

## Cultural and climatic changes shape the evolutionary history of the Uralic languages

T. HONKOLA\*, O. VESAKOSKI\*, K. KORHONEN†, J. LEHTINEN‡, K. SYRJÄNEN§  
& N. WAHLBERG\*

\*Department of Biology, University of Turku, Turku, Finland

†Department of World Cultures, University of Helsinki, Helsinki, Finland

‡Department of Modern Languages, University of Helsinki, Helsinki, Finland

§School of Language, Translation and Literary Studies, University of Tampere, Tampere, Finland

### Keywords:

Bayesian phylogenetics;  
climate change;  
Court Jester model;  
divergence time estimates;  
Finno-Ugric languages;  
language evolution;  
language lineage change;  
Red Queen model;  
relaxed-clock model;  
speciation.

### Abstract

Quantitative phylogenetic methods have been used to study the evolutionary relationships and divergence times of biological species, and recently, these have also been applied to linguistic data to elucidate the evolutionary history of language families. In biology, the factors driving macroevolutionary processes are assumed to be either mainly biotic (the Red Queen model) or mainly abiotic (the Court Jester model) or a combination of both. The applicability of these models is assumed to depend on the temporal and spatial scale observed as biotic factors act on species divergence faster and in smaller spatial scale than the abiotic factors. Here, we used the Uralic language family to investigate whether both ‘biotic’ interactions (i.e. cultural interactions) and abiotic changes (i.e. climatic fluctuations) are also connected to language diversification. We estimated the times of divergence using Bayesian phylogenetics with a relaxed-clock method and related our results to climatic, historical and archaeological information. Our timing results paralleled the previous linguistic studies but suggested a later divergence of Finno-Ugric, Finnic and Saami languages. Some of the divergences co-occurred with climatic fluctuation and some with cultural interaction and migrations of populations. Thus, we suggest that both ‘biotic’ and abiotic factors contribute either directly or indirectly to the diversification of languages and that both models can be applied when studying language evolution.

### Introduction

In biological macroevolutionary studies, species divergences can be studied by inferring the phylogenetic relationships of taxa and the temporal order and range of the divergences. Such approaches also allow the investigation of biotic and abiotic forces behind the process of speciation. According to the Red Queen model, biotic interactions are the main drivers of evolutionary change (Van Valen, 1973), while the Court Jester model suggests that evolutionary changes are mainly

induced by changes in the physical environment, for example, changes in climate (Barnosky, 2001). These two models are not mutually exclusive as they seem to act on different geographical and temporal scales: the Red Queen model acts more on the local level and at short temporal scales, while the Court Jester model tends to be more dominant at larger scales in terms of both space and time (Barnosky, 2001), and is more suited for explaining extinctions than speciation (Ezard *et al.*, 2011).

The divergence of languages, in a manner similar to the divergence of biological species, has been shown to be a process which can be analysed using phylogenetic methods (e.g. Atkinson & Gray, 2005; Kitchen *et al.*, 2009). Languages provide an interesting study object for estimating the effect of ‘biotic’ and abiotic factors on

Correspondence: Terhi Honkola, Section of Ecology, Department of Biology, University of Turku, FIN-20014, Turku, Finland.  
Tel.: +358 (0)2 333 6394; fax: +358 (0)2 333 6598;  
e-mail: terhi.honkola@utu.fi

macroevolutionary processes. Although there are no strictly biotic characteristics in language, an analogy can be drawn between cultural interactions affecting language divergences and biotic interactions affecting biological speciation. As languages evolve over a relatively short-time interval (Anttila, 1989) and in a restricted geographical area, the macroevolutionary changes in languages should follow the Red Queen model. The influence of cultural changes ('biotic' factors) on language divergences is also widely attested in historical linguistics (Anttila, 1989; Dixon, 1997), and suggested also by Atkinson *et al.* (2008) in a quantitative study, which further suggests the applicability of the Red Queen model to the study of language evolution. On the other hand, human populations and their cultures can be influenced by changes in climate, as paleoclimatic, archaeological and anthropological studies have shown (deMenocal, 2001; Weiss & Bradley, 2001; Patterson *et al.*, 2010; Büntgen *et al.*, 2011; Zhang *et al.*, 2011): a moderate rise in temperature can evoke an increase in primary production (Zheng *et al.*, 2004) which in turn can lead to an increase in the population density of hunter-gatherers (Tallavaara & Seppä, 2011). When population growth approaches the carrying capacity of the environment, the individuals may be forced to migrate (Fenggui *et al.*, 2010). Similarly, deteriorating environmental conditions, for example, due to changes in climate, may lead to migrations and cultural collapses (deMenocal, 2001; Zhang *et al.*, 2011). Coupling these findings together provides us with a scenario where language divergence is triggered by changes in cultural ('biotic') factors, but where the driving forces behind the cultural changes are abiotic, such as fluctuations in mean temperature of the environment.

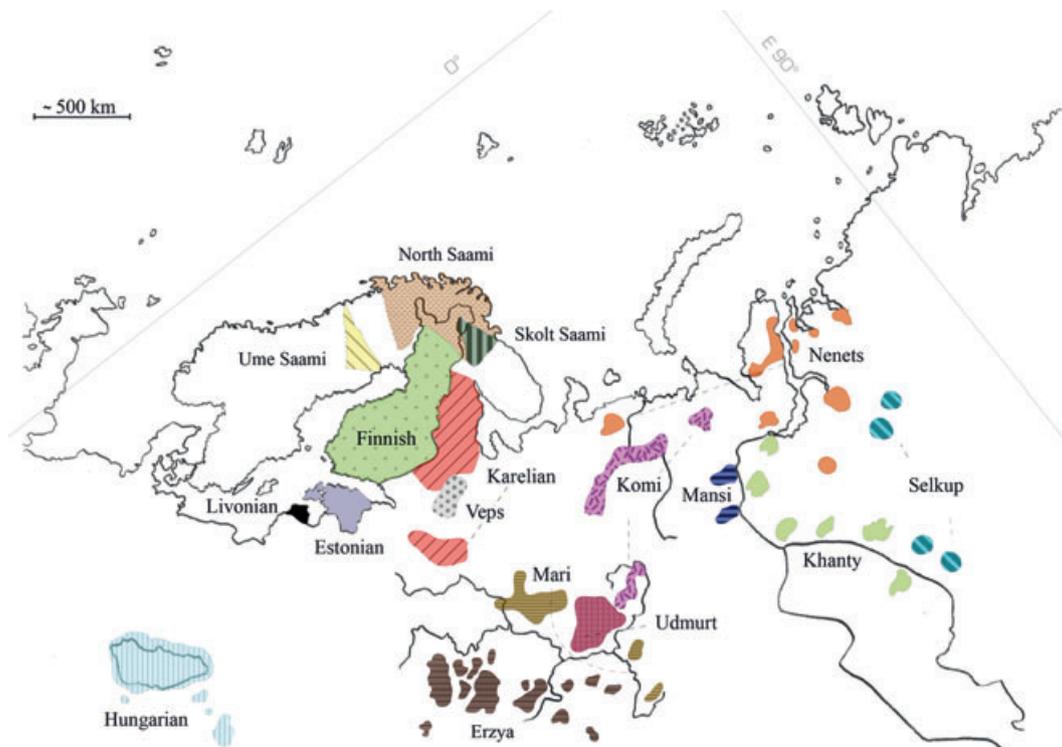
In general, the processes related to language divergences are complex. Population size and migration are affected by several factors directly or indirectly (Zhang *et al.*, 2011, Fig. 2) and once migration occurs, factors such as political complexity (Currie & Mace, 2009) and size and isolation of the speaker area as in the case of islands (Gavin & Sibanda, 2012) can have an effect on the resulting diversity of the languages. Despite the complexity of these processes, change in climate can be the principal factor starting the process leading to changes in population size (Zhang *et al.*, 2011, Fig. 2.), which further may lead to language divergences. This causality link suggests the importance of changes in climate to language divergences and calls for studies on the applicability of Court Jester model in studies of language evolution.

We studied whether the two models are suitable for studies of language evolution with the Uralic language family, which consists of 47 languages spoken mostly in the northern boreal forest zone stretching from Northeastern Europe to Western Siberia (Fig. 1) (Abondolo, 1998b; Salminen, 2007). With over a million speakers each, Hungarian, Finnish and Estonian are the

biggest languages of this family and have the status of a national language (Korhonen, 1991). The rest of the Uralic languages are minority languages spoken in Russia, Finland, Sweden, Norway and Estonia (Fig. 1), and are significantly smaller with some tens or hundreds of thousands of speakers or even fewer (Korhonen, 1991). Many are also close to extinction or already extinct, such as Livonian, which was spoken in Latvia. Many Uralic language communities have been characterized by a traditional hunter-gatherer economy which later switched to reindeer herding in the northern parts of the area. Reindeer herding has been maintained in the north to this day, while agriculture was adopted in the south and west ca. 4000–3500 YBP (years before present; Hajdú, 1975).

The phylogenetic relationships and the divergence times of the Uralic languages have been studied with traditional linguistic methods (Toivonen, 1953; Korhonen, 1981; Sinor, 1988; Abondolo, 1998c; Kallio, 2006; Häkkinen, 2009; Janhunen, 2009), but these previous estimates of the branching pattern and the estimated divergence times vary widely. For example, some scholars suggest that the Proto-Uralic language (i.e. the ancestral language of all Uralic languages) diverged into its immediate daughter branches 7000–6000 YBP (Korhonen, 1981; Sammallahti, 1988; Janhunen, 2000), while others advocate a more recent common origin, 5000–4000 YBP (Kallio, 2006; Häkkinen, 2009; Janhunen, 2009). Similarly, there are conflicting suggestions about the timings of younger divergences and about the phylogenetic relationships of the languages. During the last decade, the use of quantitative methods in linguistics has increased, and they have been successfully applied to, for example, the Austronesian (Gray *et al.*, 2009; Greenhill *et al.*, 2010), Indo-European (Gray & Atkinson, 2003), Arawak (Walker & Ribeiro, 2011), Semitic (Kitchen *et al.*, 2009) and Japonic (Lee & Hasegawa, 2011) language families (for a review, see Pagel, 2009). Also a comprehensive study on the phylogenetic relationships of the Uralic languages has recently been conducted in a Bayesian framework (Syrjänen *et al.*, 2013), but the temporal scale of the divergences has not yet been studied using computational phylogenetic methods.

Here, we study the temporal scale at which the Uralic languages have evolved by using Bayesian clock models that simultaneously infers the times of divergence and the phylogenetic relationships of the languages using linguistic data. Our aims are twofold: we (i) compare our results with the previous divergence time estimates and (ii) investigate the language divergences in a framework traditionally used in biological macroevolutionary studies. The latter is done by studying whether the language divergences coincide with abiotic and 'biotic' phenomena, that is, whether the Court Jester and the Red Queen frameworks are applicable for the study of language divergences. There are earlier studies on



**Fig. 1** A map of the Uralic languages used in this study: Estonian, Finnish, Karelian, Livonian, Veps, Eastern Khanty, Northern Mansi, Komi, Udmurt, Skolt Saami, North Saami, Ume Saami, Selkup, Tundra Nenets, Meadow Mari, Erzya, Hungarian. Map compiled from Abondolo (1998c).

the influence of climatic changes on human population dynamics and on the influence of cultural changes on language divergences, but to our knowledge, this is the first time when the effect of changes in climate and culture is studied jointly throughout a language family within a single methodological framework.

## Materials and methods

Three conventional core vocabulary lists, Swadesh 100 (Swadesh, 1955), Swadesh 200 (Swadesh, 1952) and Leipzig-Jakarta list (Tadmor, 2009) adding up to 226 different meanings, were collected from each of the 17 Uralic languages used in this study (Fig. 1). As it has been noted that the core vocabulary tends to change at a varying pace in different language families (McMahon *et al.*, 2005), these universal core vocabulary lists are not necessarily optimal for a study of a single language family, which is the case here. In fact, the standardized lists are not often applied as such, without eliminating problematic meanings from the list. Syrjänen *et al.* (2013) excluded reported loanwords in any of the 17 Uralic languages from the 226 combined core vocabulary meanings (about the literature used, see Appendix S1), which resulted in a 100-item list, 'Ura100', for which ca. 1/3 of the meanings were different from those of the conventional 100-item lists (Swadesh 100

and Leipzig-Jakarta). Thus, the Ura100 list includes fewer loanwords in the Uralic languages than any of the conventional lists and is therefore more suitable for this study. The data were encoded in binary (0, 1) format according to etymological relationships ('cognacy judgements') (words that belong to a given cognate set = 1; words that do not belong in given cognate set = 0) by using etymological dictionaries of the Uralic languages (for more details, see Syrjänen *et al.*, 2013).

We used the BEAST v.1.5.4 software package to infer phylogenetic relationships along with timing estimates (Drummond & Rambaut, 2007). This programme uses a Bayesian framework with a Markov Chain Monte Carlo (MCMC) simulation to estimate the posterior distribution of parameters from given priors. We set a simple substitution model and Yule process as the tree prior for the analysis and allowed the state frequencies to be estimated from the data, allowing rate variation according to a four-category gamma distribution. The data set was analysed both with an uncorrelated lognormal relaxed clock in which the rate on each branch of the tree is drawn independently and identically from an underlying rate distribution (Drummond *et al.*, 2006), and with a strict clock model in which the rate of change is equal. Bayes factor was calculated pairwise to relaxed and strict clock in Tracer v. 1.5 (Rambaut & Drummond, 2007) to estimate which model justified

the data better (Suchard *et al.*, 2001).  $\log_{10}$  Bayes factor 5.7 indicated decisive support for the relaxed-clock model (Jeffreys, 1961; Kass & Raftery, 1995), which was used in subsequent analysis. Times of divergence were calibrated as described below. We set the analysis to run for ten million generations with every 1000th generation sampled. Furthermore, the first 1000 trees (burn-in) were discarded in Tree Annotator after checking from Tracer v. 1.5 (Rambaut & Drummond, 2007) that all chains had reached convergence by the end of the burn-in period.

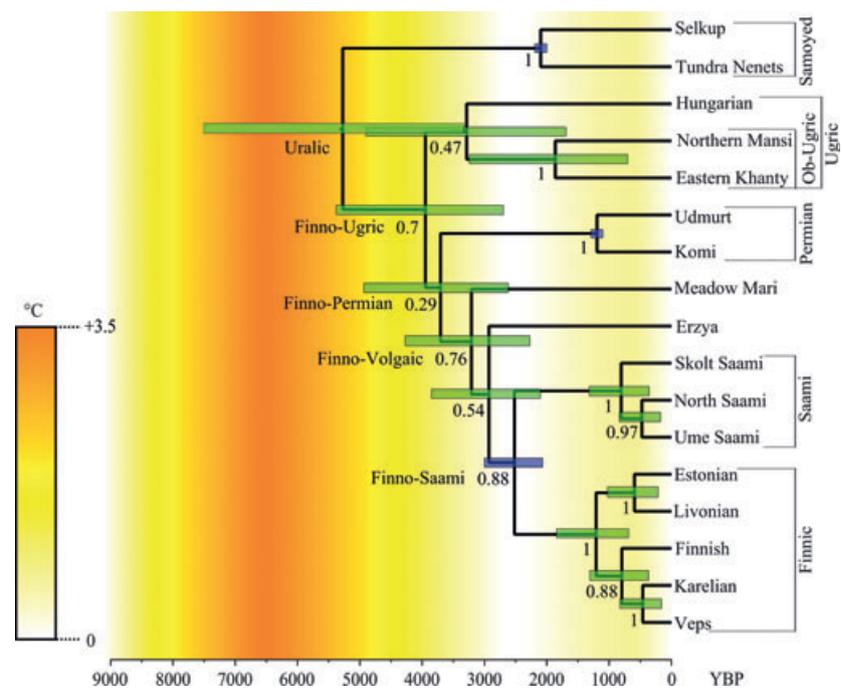
We calibrated our topology based on the estimated times of divergence of three subgroups of the Uralic languages. These dates can be considered fairly well supported by several fields of study, that is, loanword studies, history and archaeology. The calibrated nodes are the Permian, Samoyed and Finno-Saami divergences. The divergence time of the Permian languages (Komi and Udmurt) has been traced to 1300–1100 YBP on the basis of Volga Bulgarian loanwords that are more frequent in Udmurt than in Komi (Hajdú, 1975; Kokkonen, 1991). This timing is also supported by archaeological findings that suggest the separation of these two populations at this time (Hajdú, 1975). Similarly, the dating of the Samoyed split to around 2200–2000 YBP is based on historical data on the movements of Turkic tribes in the Proto-Samoyed homeland and on Yeniseian loanwords in Samoyed (Hajdú, 1975; Korhonen & Kulonen, 1991; Janhunen, 1998). The divergence time estimates for the Finno-Saami branch vary more; in general, scholars who have studied Baltic and Germanic loanwords have proposed dates that vary

between 3000 and 2000 YBP (Hajdú, 1975; Kallio, 2006). We ran the analysis with both uniform and normal prior distributions for the age constraints. The range of the uniform prior covers the values between the maximum and minimum estimates given above. For calibration points using a normal distribution prior, the mean and standard deviation were calculated from divergence time estimates taken from the literature (Toivonen, 1953; Décsy, 1965; Hajdú, 1975; Korhonen, 1981, 1988; Kokkonen, 1991; Korhonen & Kulonen, 1991; Laakso, 1991; Janhunen, 1998, 2009; Riese, 1998; i.e. Finno-Saami 2630 YBP  $\pm$  390 years, Permian 1225 YBP  $\pm$  60 years, Samoyed 2030 YBP  $\pm$  60 years). When, instead of a single value, a time range was given in the literature, we used the midpoint as our value estimate, for example, 2500 YBP when a range of '3000–2000 YBP' was given.

The stability of the results and the quality of the calibration points was investigated by repeating the analysis four times for each calibration point combination: Permian + Samoyed + Finno-Saami, Permian + Finno-Saami, Permian + Samoyed, Samoyed + Finno-Saami for both the uniform and normal prior distributions (in total 32 analyses). In addition, posterior probability values indicating the robustness of the branching order were obtained from the analysis.

The temperature data were compiled from the studies by Kremenetski *et al.* (1997), Davis *et al.* (2003), Väiliranta *et al.* (2003) and Heikkilä & Seppä (2010), which were based on lake sediment, pollen and peat data collected from several locations in Northeastern Europe. The colour gradient in Fig. 2 is a generalization

**Fig. 2** Timing analyses of the Uralic languages. Green bars represent the 95% highest probability density (HPD) for the divergence times. Scale values represent years before present (YBP). Values outside the nodes represent posterior probabilities. Calibration points (Samoyed, Permian and Finno-Saami) are labelled with blue bars indicating the uniform prior of the calibration points. Names of different protolanguages are marked on the nodes of the tree, and the names of different subclasses are on the right margins. The colour scale of the picture describes the temperature changes with relation to current temperature (+3.5–0 °C red-white) of the Northeastern Europe/East European tundra (west side of the Ural Mountains) (compiled from Kremenetski *et al.*, 1997; Davis *et al.*, 2003; Väiliranta *et al.*, 2003; Heikkilä & Seppä, 2010).



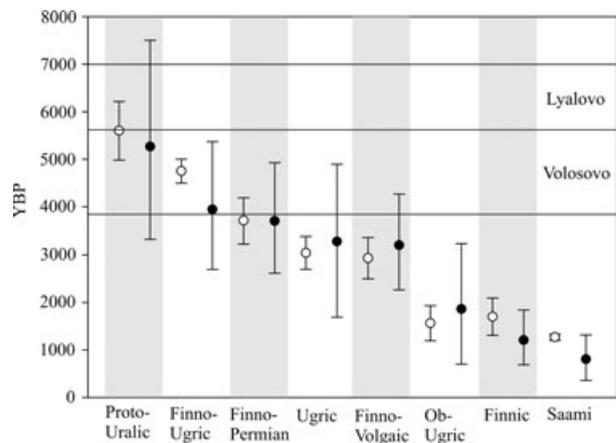
of the changes in temperature drawn over the western side of the Ural Mountains. This generalization was considered feasible as the Holocene thermal maximum seems to have had a remarkably similar pattern and timing throughout boreal Northeastern Europe (Heikkilä & Seppä, 2010) and, in general, it was followed by gradual cooling (Kremenetski *et al.*, 1997; Davis *et al.*, 2003; Väiliranta *et al.*, 2003; Heikkilä & Seppä, 2010). Archaeological and historical information was also compiled from scholarly literature as cited in the Discussion.

## Results

The Proto-Uralic phase diverged ca. 5300 YBP [HPD (= 95% highest probability density): 7503–3330 YBP] with the separation of the Finno-Ugric and Samoyed clades (Fig. 2). This divergence was followed by the split of Finno-Ugric ca. 3900 YBP (HPD: 5371–2695 YBP) into the Ugric and Finno-Permian clades. The date obtained for the divergence of the Ugric clade into the Hungarian and the Ob-Ugric groups was ca. 3300 YBP (HPD: 4895–1690 YBP). This dichotomous branching was, however, only weakly supported as the posterior probability value was below 0.5. The Ob-Ugric languages diverged ca. 1900 YBP (HPD: 3239–702 YBP). The date obtained for the divergence of Permian from the rest of the Finno-Permian clade was ca. 3700 YBP (HPD: 4928–2615 YBP), but this node was also only weakly supported. The rest of the languages form the Finno-Volgaic group, from which Mari diverged ca. 3200 YBP (HPD: 4267–2267 YBP), soon followed by Erzya ca. 2900 YBP (HPD: 3846–2102 YBP).

The first divergence of Finnic languages into northern (Finnish, Karelian, Veps) and southern (Estonian, Livonian) groups occurred ca. 1200 YBP (HPD: 1842–687 YBP) after which Finnish diverged from Karelian and Veps ca. 800 YBP (HPD: 1307–370 YBP). Karelian and Veps diverged from each other ca. 450 YBP (HPD: 826–160 YBP). Finally, the divergence of Estonian and Livonian occurred ca. 600 YBP (HPD: 1024–213 YBP). The first divergence of the Saami languages, separating Skolt Saami from a branch that includes North Saami and Ume Saami (i.e. Eastern Saami languages from Western Saami), occurred ca. 800 YBP (HPD: 1318–358 YBP). Finally, North Saami and Ume Saami diverged ca. 450 YBP (HPD: 837–173 YBP).

The HPD values of the divergences overlap to at least some extent. Exceptions to this are the HPDs of Proto-Uralic and Ob-Ugric divergences which do not overlap and the HPD values of the Finnic and Saami divergences which do not overlap with any of the HPD values of the preceding divergences (Figs 2 and 3). The results did not vary to any considerable degree between the uniform and normal prior and between the four replicates made with each calibration point combination (see Fig. S1), which supported the robustness of our results.



**Fig. 3** Average divergence time estimates of different protolanguages. Open circles indicate previous results (compiled from Kettunen & Vaula, 1943; Toivonen, 1953; Décsy, 1965; Korenchy, 1972; Hajdú, 1975; Korhonen, 1981, 1988; Sammallahti, 1988; Korhonen, 1991; Kulonen, 1991; Laakso, 1991; Lehtiranta & Seurujärvi-Kari, 1991; Abondolo, 1998a; Honti, 1998; Keresztes, 1998; Janhunen, 2000; Kallio, 2006; Häkkinen, 2009; Janhunen, 2009) with 95% confidence intervals. Solid circles indicate our results with a 95% HPD interval. The archaeological periods of Lyalovo and Volosovo cultures are included in the figure as described by Carpelan & Parpola (2001).

## Discussion

### Comparisons with previous divergence time estimates

Our phylogenetic hypothesis follows the traditional view of the Uralic language divergences in that the first split occurs between Samoyed and Finno-Ugric languages and is followed by further divergences of the Finno-Ugric group leading eventually to smaller groups such as Finnic, Saami, Permian and Ob-Ugric (e.g. Korhonen, 1981). Thus, our results do not support the highly polytomous views of the Uralic language tree (Häkkinen, 1983; Salminen, 1999). However, our phylogenetic hypothesis does not specify a completely clear branching pattern; that is, the ambiguity in the order of the intermediate divergences seen in some previous linguistic studies (e.g. Kulonen, 2002; Michalove, 2002) can also be seen in our results.

Our results suggest that the first divergence of the Uralic language occurred ca. 5300 YBP, which is close to the average calculated from previous linguistic studies (5600 YBP) (Fig. 3; Kettunen & Vaula, 1943; Toivonen, 1953; Décsy, 1965; Korhonen, 1981, 1991; Sammallahti, 1988; Janhunen, 2000, 2009; Kallio, 2006; Häkkinen, 2009). The intermediate level divergences of Finno-Ugric, Finno-Permian and Finno-Volgaic occurred during ca. 3900–2900 YBP, which is largely what previous linguistic studies also suggest (from Finno-Ugric to

Finno-Volgaic ca. 4750–3000 YBP) (Fig. 3). Within this period of multiple divergences, our results propose a slightly more recent divergence of Finno-Ugric languages than what is suggested in the earlier studies. Instead, the divergence time estimates proposed in this study and in earlier studies are rather similar for Finno-Permian and Finno-Volgaic protolanguages (Fig. 3). It is noteworthy that our results do not support the entity of Finno-Permian as the posterior probability value for this group is low (0.29; Fig. 2). Similarly, the results do not support an Ugric branch containing Hungarian and Ob-Ugric exclusive of other languages (posterior probability 0.47) (Fig. 3). Due to these low posterior probability values for the Ugric and Finno-Permian protolanguages, a polytomous branching of the Finno-Ugric clade to Hungarian, Ob-Ugric, Permian and Finno-Volgaic can be posited. Similar polytomous branching is also suggested by Syrjänen *et al.* (2013) based on the same data as here, but analysed with a different algorithm. Various views about the degree of polytomous branching in Uralic languages have been suggested, varying from highly polytomous (Häkkinen, 1983; Salminen, 1999) to strictly binary (Korhonen, 1981) trees, along with different intermediate forms (Kulonen, 2002; Michalove, 2002). This ambiguous nature of the intermediate branchings can also be seen in our results: they show the lowest posterior probability values proposing an uncertain pattern of binary branchings.

Our results suggest a rather ancient divergence for the Ob-Ugric languages (ca. 1900 YBP), as our estimate is close to the upper limit of what previous linguistic studies have estimated, namely between 2000 and 1000 YBP with an average of 1550 YBP (Décsy, 1965; Kálmán, 1988; Honti, 1998; Keresztes, 1998) (Fig. 3). The divergence of Finnic languages into northern (Finnish, Karelian, Veps) and southern (Estonian, Livonian) groups occurred around 1200 YBP according to our results. Our estimate places this divergence some 500 years later than has been obtained on average in earlier research (1700 YBP; Kettunen & Vaula, 1943; Décsy, 1965; Hajdú, 1975; Korhonen, 1981; Janhunen, 2009) (Fig. 3). Similarly, the average divergence time of Skolt Saami from North and Ume Saami is also ca. 500 years more recent in our results (ca. 800 YBP) than suggested by previous scholars (1250 YBP, Korhonen, 1981, 1988; Lehtiranta & Seurujärvi-Kari, 1991) (Fig. 3).

The overlapping HPD values and broad HPD ranges, especially those close to the root of the tree, are partly due to the relatively recent dates of the calibration points used in the analyses. Rather recent divergences were chosen as calibration points as these time estimates were more reliable than the estimates for earlier divergences. Another reason for the overlapping of HPDs lies in the short intermediate branches of the tree (Fig. 2).

### Comparisons with abiotic and 'biotic' changes

The first divergence of the Uralic language tree occurred ca. 5300 YBP according to our results. Based on earlier divergence time estimates, Carpelan & Parpola (2001) connected Proto-Uralic to the Lyalovo culture (7000–5650 YBP). Our results, however, also allow the possibility that the initial divergence of Proto-Uralic happened during Volosovo culture which followed the Lyalovo culture. At the time of Lyalovo culture period, the climate was relatively warm (Holocene thermal maximum ca. 7500–5000 YBP; Kremenetski *et al.*, 1997; Davis *et al.*, 2003; Väiliranta *et al.*, 2003; Heikkilä & Seppä, 2010). A moderate increase in temperature leading to the Holocene thermal maximum may have induced a rise in population size (Tallavaara & Seppä, 2011) and led to migrations when the carrying capacity of the area was approached or reached (Lidicker, 1962; Zubrow, 1971). Thus, it is possible that increasing population size may have been indirectly associated with cultural alterations (from Lyalovo culture to Volosovo) which further could have had an impact on the diversification of Proto-Uralic. This idea of possible migratory movements among the Proto-Uralic speakers is also supported by suggestions that the Volga-Oka area – the area of the earliest ceramic finds of this part of Europe and the hypothesized speaking area of Proto-Uralic (Toivonen, 1953; Salminen, 1999; Häkkinen, 2009) – continuously produced a demographic surplus beginning in the Mesolithic (ca. 10 000 YBP). This caused migrations especially towards the north and the northwest (Carpelan & Parpola, 2001). Thus, it can be suggested that a significant underlying cause for the divergence of Proto-Uralic was an ongoing climate change, which led to migrations due to a high population density, which in turn induced cultural changes.

The divergences of Finno-Ugric, Finno-Permian, Ugric, Finno-Volgaic and Finno-Saami occurred ca. 3900–2500 YBP. Carpelan & Parpola (2001) connect the Proto-Finno-Ugric phase to Volosovo culture (5650–3900 YBP; Fig. 3), but also suggest that during this period a cultural border was formed between the Proto-Permian, Proto-Volgaic and Proto-Ugric linguistic communities. Our results also connect the Proto-Finno-Ugric phase to Volosovo culture but instead of showing further divergences during Volosovo these divergences occur rapidly after the end of Volosovo period (Fig. 3). Therefore, we argue on the basis of our results, suggested also by archaeological evidence, that the Proto-Finno-Ugric started diverging into several different clades during Volosovo period after which several divergences occurred during a rather short period. Moreover, the divergences of Finno-Ugric, Finno-Permian, Ugric, Finno-Volgaic and Finno-Saami coincide with a period of cooling after the Holocene thermal maximum (for the variation in temperature, see Kremenetski *et al.*, 1997; Davis *et al.*, 2003; Väiliranta *et al.*, 2003; Heikkilä & Seppä, 2010,

Fig. 2). In hunter-gatherer populations, declining temperature can lead to lowered carrying capacity through lower food abundance (Tallavaara & Seppä, 2011), which forces inflicting populations to either gradual extinction or migrations (Lidicker, 1962; Zubrow, 1971). Thus again, the ultimate cause for these language divergences may have been a change in temperature which led to migrations, this time through lowered carrying capacity, which again in turn induced cultural changes. It seems reasonable to conclude that it is the change in temperature which matters, not the direction of the change. The study of causal linkages from changes in climate to changes in population size and migrations through different pathways remains out of scope for this study, but for potential pathways of causal linkages in preindustrial agricultural societies, see Zhang *et al.* (2011).

According to our results, the diversification of Ob-Ugric into different languages occurred ca. 1900 YBP. The divergence of Samoyed has been dated to ca. 2000 YBP on the basis of alleged contacts with Turkic tribes (Hajdú, 1975; Korhonen & Kulonen, 1991; Janhunen, 1998). As the assumed living areas of the speakers of Proto-Ob-Ugric and Proto-Samoyed were rather close to each other, the migrations of the Turkic tribes affecting the Samoyeds may also have affected the Ob-Ugric speakers, further impacting on the divergence of Khanty and Mansi. These migrations also co-occurred with declining temperatures around 2000 YBP (note a difference to Fig. 2 in which the temperature increases around 2000 YBP due to the location on the other side of the Ural Mountains), which changed the vegetation from steppe to forest on the eastern side of the Urals (Zakh *et al.*, 2010), and which most likely had an impact on the local human populations as well. In sum, we consider it possible that the migrations of Turkic tribes affected both Samoyed and Ob-Ugric speakers, while the subsequent migrations of the speakers of Khanty and Mansi could have been intensified by changes in their living environment and its carrying capacity.

Our results suggest that the Finnic languages diverged into northern (Finnish, Karelian, Veps) and southern (Estonian, Livonian) groups around 1200 YBP. It is assumed that the temperature was slightly warmer again after the gradual cooling and the short cold period (Fig. 2). The density of human settlements is known to have increased in all coastal Baltic areas between 1600 and 1200 YBP (Meinander, 2006). This increase in population density was probably induced by the intensification of agriculture around 2500 YBP in the area what is now Latvia (Heikkilä & Seppä, 2010). Correlation between temperature and population size is weaker in agricultural than in hunter-gatherer societies (Tallavaara & Seppä, 2011). Instead, population growth is high soon after transition to agriculture (Bocquet-Appel, 2002; Bandy, 2005), which could explain why

the density of human settlements increased despite only a slight rise in temperature. Migrations across the Gulf of Finland had been occurring reciprocally and irregularly for thousands of years (Miettinen, 1996), and it can be suggested that the increase in population density in the southern coast of the Gulf of Finland would have led to one of these migratory waves to the north. In effect, similar archaeological finds from the northern and southern coasts of the Gulf of Finland dated to 2000–1600 YBP indicate gradual northward migrations across the gulf to the current area of Finland (Kivikoski, 1961). This migration period precedes our divergence time estimate for the Finnic languages. Nevertheless, it could have been the starting point for the process of language divergence which was possibly intensified by different economies on the opposite sides of the Gulf of Finland (agriculture in Estonia; slash-and-burn agriculture, fishing and hunting in Finland) and by the political split in early mediaeval times (ca. 900–800 YBP; Carpelan & Parpola, 2001). Thus, we suggest that the division of the Finnic languages into southern and northern groups was at least partly driven by agriculture-induced increase in population density, which led to migration from the southern coast of the Gulf of Finland and was further deepened by nascent cultural differences.

The divergence of Finnish from Karelian and Veps occurred around 800 YBP (Fig. 2). This happened soon after the time when East Slavic tribes arrived in the Baltic area (Bjørnflaten, 2006). These new arrivals introduced many loanwords, especially into the eastern Finnic languages, and brought the Karelians and Vepsians under Eastern Slavic cultural influence and the Greek Catholic Church around 900 YBP (Hajdú, 1975; Laakso, 1991). At that time, the speakers of Finnish on the other hand were brought into close contact with Sweden and the Roman Catholic Church (Hajdú, 1975; Meinander, 2006). Therefore, it can be suggested that this divergence occurred mainly due to cultural reasons.

The influence of Novgorod (the centre of the Eastern Slavic culture) extended also to northern areas (Hajdú, 1975), coinciding with our results on the divergence time of Skolt Saami from North and Ume Saami (ca. 800 YBP). Reindeer herding, which is still practised by the Saami tribes, may have also had an influence on the diversification of the Saami languages. Kortessalmi (2008) proposed that reindeer herding and nomadism existed around the area of Ume Saami already around 1200–800 YBP. From there, it moved towards the north reaching the North Saami area in 400 YBP at the latest (Kortessalmi, 2008) but without ever fully reaching the more Eastern Saami areas (to which Skolt Saami belongs to). As the availability of draught animals facilitates long-distance travel, it is possible that reindeer herding in the western areas strengthened the connection between Ume and North Saami. This economical

border separating the areas with and without reindeer herding and most recently the state border in the east may have encouraged closer contacts between Ume and North Saami leaving Skolt Saami to develop more in the eastern sphere of influence. This probably affected the divergence of languages as well.

We show here that both 'biotic' and abiotic processes are involved in language diversification and that this may happen primarily through climatic changes triggering secondary forces, such as cultural alterations. The connection between cultural changes ('biotic' processes) and the diversification of languages can also be seen in the Indo-European language family where major processes of diversification within the Germanic and Romance subgroups of Indo-European co-occur with the significant cultural changes known as the Migration period (Gray & Atkinson, 2003; Büntgen *et al.*, 2011). These involved, besides migratory movements, the disintegration of Roman administrative structures in many areas (Gray & Atkinson, 2003; Halsall, 2007; Büntgen *et al.*, 2011), and which also interestingly coincided with exceptional variability in climate (Büntgen *et al.*, 2011).

## Conclusion

Our divergence time estimates are largely comparable with the divergence time estimates suggested by earlier linguistic studies. Some differences, however, do exist: our result suggest more ancient divergence for Ob-Ugric, and more recent divergence for the Finno-Ugric, Finnic and Saami unities. For the shape of the Uralic language phylogeny, our results support the hypothesis of polytomous branching of the Finno-Ugric clade.

The divergence times of the Uralic languages coincide both with known processes of climatic fluctuation and with changes in cultural spheres of influence. Climatic changes and other abiotic factors are thought to have an effect on the diversity, the distribution and the macroevolutionary processes of biological species (Hewitt, 2000; Benton, 2009), but mostly over large temporal scales (Barnosky, 2001). However, given that, on one hand, climatic changes can have significant effects on human populations (Tallavaara & Seppä, 2011) also on a short-time scale (deMenocal, 2001; Büntgen *et al.*, 2011; Zhang *et al.*, 2011) and, on the other hand, that languages change fast (Anttila, 1989), the connection between changes in climate and language divergence is reasonable and supported also by our study. Thus, both abiotic and 'biotic' events contribute to language divergences, and what is considered a cultural effect on language change may be primarily caused by abiotic fluctuation. However, there are occasions where the diversification of languages divergence seems to be caused solely by 'biotic' interaction, that is, cultural turnovers that have no clear cause in the natural environment. We suggest that 'biotic' processes involved in

the Red Queen model can explain language divergences, but that they are often triggered by abiotic factors indicating the importance of including the Court Jester model when discussing the evolution of languages.

## Acknowledgments

We thank U. Määttä, T. Suutari and A. Leino from the BEDLAN project and E. Salmela, J. Häkkinen, B. Planade, R. Kubitzka, E. Vesterinen, M. Pagel and anonymous reviewers for their helpful comments. We also thank J. Ruohonen, M. Tallavaara and T. Saarinen for providing insights into the fields of archaeology and geology. Institute for the Languages of Finland provided facilities and expertise for collecting the lexical list. Professor S. Saarinen helped us to find experts to double-check the vocabulary lists. The experts we thank are M. Torikka (Karelian); S. Jantunen (Veps, Livonian); E. Mikone (Estonian); K. Vuolab-Lohi (North Saami); E. Koponen (Ume Saami, Skolt Saami); N. Chinaeva (Erzya); A. Moisio (Meadow Mari); S. Lumme (Komi); O. Titova (Udmurt); J. Varga and Z. Balogh (Hungarian); S. Virtanen (Northern Mansi); M. Salo (Eastern Khanty); T. Salminen (Tundra Nenets, Selkup). We also thank I. Tammi for the map and D. Tringham for the language revision. This research was funded by Kone Foundation for TH, JL, KS and OV, by the Academy of Finland for NW and by the University of Helsinki for KK.

## References

- Abondolo, D. 1998a. Hungarian. In: *The Uralic Languages* (D. Abondolo, ed.), pp. 428–456. Routledge, London.
- Abondolo, D. 1998b. Introduction. In: *The Uralic Languages* (D. Abondolo, ed.), pp. 1–42. Routledge, London.
- Abondolo, D. (ed.) 1998c. *The Uralic languages*. Routledge, London.
- Anttila, R. 1989. *Historical and Comparative Linguistics*, 2nd rev. edn. John Benjamins, Amsterdam, Philadelphia.
- Atkinson, Q.D. & Gray, R.D. 2005. Curious parallels and curious connections – phylogenetic thinking in biology and historical linguistics. *Syst. Biol.* **54**: 513–526.
- Atkinson, Q.D., Meade, A., Venditti, C., Greenhill, S.J. & Pagel, M. 2008. Languages evolve in punctuational bursts. *Science* **319**: 588.
- Bandy, M.S. 2005. New world settlement evidence for a two-stage neolithic demographic transition. *Curr. Anthropol.* **46**: 109–115.
- Barnosky, A.D. 2001. Distinguishing the effects of the Red Queen and Court Jester on miocene mammal evolution in the northern rocky mountains. *J. Vertebr. Paleontol.* **21**: 172–185.
- Benton, M.J. 2009. The Red Queen and the Court Jester: species diversity and the role of biotic and abiotic factors through time. *Science* **323**: 728–732.
- Bjørnflaten, J.I. 2006. Chronologies of the Slavization of Northern Russia Mirrored by Slavic Loanwords in Finnic and Baltic. In: *The Slavization of the Russian North. Mechanisms*

- and Chronology (J. Nuorluoto, ed.), pp. 50–77. Department of Slavonic and Baltic Languages and Literatures, University of Helsinki, Helsinki.
- Bocquet-Appel, J. 2002. Paleoanthropological traces of a neolithic demographic transition. *Curr. Anthropol.* **43**: 637–650.
- Büntgen, U., Tegel, W., Nicolussi, K., McCormick, M., Frank, D., Trouet, V. et al. 2011. 2500 years of European climate variability and human susceptibility. *Science* **331**: 578–582.
- Carpelan, C. & Parpola, A. 2001. Emergence, contacts and dispersal of Proto-Indo-European, Proto-Uralic and Proto-Aryan in archaeological perspective. In: *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations* (C. Carpelan, A. Parpola & P. Koskikallio, eds), pp. 55–150. Suomalais-Ugrilaisen Seuran Toimituksia 242, Helsinki.
- Currie, T.E. & Mace, R. 2009. Political complexity predicts the spread of ethnolinguistic groups. *Proc. Natl. Acad. Sci. USA*. **106**: 7339–7344.
- Davis, B.A.S., Brewer, S., Stevenson, A.C., Guiot, J. & Contributors Data. 2003. The temperature of Europe during the Holocene reconstructed from pollen data. *Quat. Sci. Rev.* **22**: 1701–1716.
- Décsy, G. 1965. *Einführung in die finnisch-ugrische Sprachwissenschaft*. Harrassowitz, Wiesbaden.
- Dixon, R.M.W. 1997. *The Rise and Fall of Languages*. Cambridge University Press, Cambridge.
- Drummond, A.J. & Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**: 214.
- Drummond, A.J., Ho, S.Y.W., Phillips, M.J. & Rambaut, A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**: 699–710.
- Ezard, T.H.G., Aze, T., Pearson, P.N. & Purvis, A. 2011. Interplay between changing climate and species' ecology drives macroevolutionary dynamics. *Science* **332**: 349–351.
- Fenggui, L., Yili, Z., Zhaodong, F., Guangliang, H., Qiang, Z. & Haiheng, Z. 2010. The impacts of climate change on the Neolithic cultures of Gansu-Qinghai region during the late Holocene Megathermal. *J. Geog. Sci.* **20**: 417–430.
- Gavin, M.C. & Sibanda, N. 2012. The island biogeography of languages. *Glob. Ecol. Biogeogr.* **21**: 958–967.
- Gray, R.D. & Atkinson, Q.D. 2003. Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* **426**: 435–439.
- Gray, R.D., Drummond, A.J. & Greenhill, S.J. 2009. Language phylogenies reveal expansion pulses and pauses in pacific settlement. *Science* **323**: 479–483.
- Greenhill, S.J., Drummond, A.J. & Gray, R.D. 2010. How accurate and robust are the phylogenetic estimates of austronesian language relationships? *PLoS ONE* **5**: e9573.
- Hajdú, P. 1975. *Finn-Ugric Languages and Peoples*. André Deutsch, London.
- Häkkinen, K. 1983. *Suomen kielen vanhimmasta sanastosta ja sen tutkimisesta. Suomalais-ugrilaisen kielten etymologisen tutkimuksen perusteita ja metodiikkaa*, Doctoral edn. University of Turku, Turku.
- Häkkinen, J. 2009. Kantauralin ajoitus ja paikannus: perustelut puntarissa. *JSFOu* **92**: 9–56.
- Halsall, G. 2007. *Barbarian Migrations and the Roman West, 376–568*. Cambridge University Press, Cambridge.
- Heikkilä, M. & Seppä, H. 2010. Holocene climate dynamics in Latvia, eastern Baltic region: a pollen-based summer temperature reconstruction and regional comparison. *Boreas* **39**: 705–719.
- Hewitt, G. 2000. The genetic legacy of the quaternary ice ages. *Nature* **405**: 907–913.
- Honti, L. 1998. Ob-Ugric. In: *The Uralic Languages* (D. Abondolo, ed.), pp. 327–357. Routledge, London.
- Janhunen, J. 1998. Samoyedic. In: *The Uralic Languages* (D. Abondolo, ed.), pp. 457–479. Routledge, London.
- Janhunen, J. 2000. Reconstructing Pre-Proto-Uralic typology spanning the millennia of linguistic evolution. In: *Congressus nonus internationalis Fenno-Ugristarum. Pars 2, Summaria Acroasium in Sectionibus et Symposiis Factarum* (A. Nurk, T. Palo & T. Seilenthal, eds), pp. 59–76. University of Tartu, Estonian Finno-Ugric Committee, Tartu.
- Janhunen, J. 2009. Proto-Uralic – what, where, and when? *Mémoires de la Société Finno-Ougrienne* **258**: 57–78.
- Jeffreys, H. 1961. *Theory of Probability*, 3rd edn. Clarendon press, Oxford.
- Kallio, P. 2006. Suomen kantakielen absoluuttista kronologiaa. *Virittäjä* **110**: 2–25.
- Kálmán, B. 1988. The history of Ob-Ugric languages. In: *The Uralic languages: Description, History and Foreign Influences* (D. Sinor, ed.), pp. 395–412. E. J. Brill, Leiden.
- Kass, R. & Raftery, A. 1995. Bayes factors. *J. Am. Stat. Assoc.* **90**: 773–795.
- Keresztes, L. 1998. Mansi. In: *The Uralic Languages* (D. Abondolo, ed.), pp. 387–427. Routledge, London.
- Kettunen, L. & Vaula, M. 1943. *Suomen kielioppi sekä tyyl- ja runo-opin alkeet oppikouluille ja seminaareille*, 3rd edn. WSOY, Porvoo.
- Kitchen, A., Ehret, C., Assefa, S. & Mulligan, C.J. 2009. Bayesian phylogenetic analysis of semitic languages identifies an early bronze age origin of semitic in the near east. *Proc. Biol. Sci.* **276**: 2703–2710.
- Kivikoski, E. 1961. *Suomen Historia I: Suomen Esihistoria*. Werner-Söderström Oy, Porvoo.
- Kokkonen, P. 1991. Komit eli syrjäänit. In: *Uralilaiset kansat: Tietoa suomen sukukielistä ja niiden puhujista* (J. Laakso, ed.), pp. 236–253. WSOY, Juva.
- Korenchy, E. 1972. *Iranische Lehnwörter in den obugrischen Sprachen*. Akadémiai Kiadó, Budapest.
- Korhonen, M. 1981. *Johdatus Lapin Kielen Historiaan*. Suomalaisen Kirjallisuuden Seura, Helsinki.
- Korhonen, M. 1988. The history of the Lapp language. In: *The Uralic languages: Description, History and Foreign Influences* (D. Sinor, ed.), pp. 264–287. E. J. Brill, Leiden.
- Korhonen, M. 1991. Uralin tällä ja tuolla puolen. In: *Uralilaiset kansat: Tietoa suomen sukukielistä ja niiden puhujista* (J. Laakso, ed.), pp. 20–48. WSOY, Juva.
- Korhonen, M. & Kulonen, U. 1991. Samojedit. In: *Uralilaiset kansat: Tietoa suomen sukukielistä ja niiden puhujista* (J. Laakso, ed.), pp. 302–317. WSOY, Juva.
- Kortessalmi, J.J. 2008. *Poronhoidon Synty Ja Kehitys Suomessa*. Suomen Kirjallisuuden Seura, Helsinki.
- Kremenetski, C., Vaschalova, T., Goriachkin, S., Cherkinsky, A. & Sulerzhitsky, L. 1997. Holocene pollen stratigraphy and bog development in the western part of the Kola Peninsula, Russia. *Boreas* **26**: 91–102.
- Kulonen, U. 1991. Obinugrilaiset. In: *Uralilaiset kansat: Tietoa suomen sukukielistä ja niiden puhujista* (J. Laakso, ed.), pp. 254–278. WSOY, Juva.
- Kulonen, U. 2002. Kielitiede ja Suomen väestön juuret. In: *Ennen, minoin: Miten menneisyyttämme tutkitaan* (R. Grünthal, ed.), pp. 102–116. Suomalaisen Kirjallisuuden Seura, Helsinki.

- Laakso, J. 1991. Itämerensuomalaiset sukukielemme ja niiden puhujat. In: *Uralilaiset kansat: Tietoa suomen sukukielistä ja niiden puhujista* (J. Laakso, ed.), pp. 49–122. WSOY, Juva.
- Lee, S. & Hasegawa, T. 2011. Bayesian phylogenetic analysis supports an agricultural origin of Japonic languages. *Proc. Biol. Sci.* **278**: 3662–3669.
- Lehtiranta, J. & Seurujärvi-Kari, I. 1991. Saamelaiset. In: *Uralilaiset kansat: Tietoa suomen sukukielistä ja niiden puhujista* (J. Laakso, ed.), pp. 123–155. WSOY, Juva.
- Lidicker, W.Z. Jr 1962. Emigration as a possible mechanism permitting the regulation of population density below carrying capacity. *Am. Nat.* **96**: 29–33.
- McMahon, A., Heggarty, P., McMahon, R. & Slaska, N. 2005. Swadesh sublists and the benefits of borrowing: an Andean case study. *Trans. Philol. Soc.* **103**: 147–170.
- Meinander, H. 2006. *Suomen historia: Linjat, rakenteet, käännekohtat*. WSOY, Helsinki.
- deMenocal, P.B. 2001. Cultural responses to climate change during the late holocene. *Science* **292**: 667–673.
- Michalove, P.A. 2002. The classification of the Uralic languages: lexical evidence from Finno-Ugric. *Finnisch-ugrische Forschungen* **57**: 58–67.
- Miettinen, T. 1996. Suomenlahden ulkosaarten esihistoriaa. In: *Suomenlahden ulkosaaret: Lavansaari, Seiskari, Suursaari, Tytärsaari* (R. Hamari, M. Korhonen, Miettinen Timo & I. Talve, eds), pp. 49–68. Suomalaisen Kirjallisuuden Seura, Helsinki.
- Pagel, M. 2009. Human language as a culturally transmitted replicator. *Nat. Rev. Genet.* **10**: 405–415.
- Patterson, W.P., Dietrich, K.A., Holmden, C. & Andrews, J.T. 2010. Two millennia of North Atlantic seasonality and implications for Norse colonies. *Proc. Natl. Acad. Sci. USA* **107**: 5306–5310.
- Rambaut, A. & Drummond, A.J. 2007. *Tracer*. URL <http://tree.bio.ed.ac.uk/software/tracer>.
- Riese, T. 1998. Permian. In: *The Uralic Languages* (D. Abondolo, ed.), pp. 249–275. Routledge, London.
- Salminen, T. 1999. Euroopan kielet muinoin ja nykyisin. In: *Pohjan poluilla: Suomalaisten juuret nykytutkimuksen mukaan* (P. Fogelberg, ed.), pp. 13–26. Ekenäs Tryckeri Ab, Ekenäs.
- Salminen, T. 2007. Europe and North Asia. In: *Encyclopedia of the World's Endangered Languages* (C. Moseley, ed.), pp. 211–282. Routledge, New York.
- Sammallahti, P. 1988. Historical phonology of the Uralic languages. In: *The Uralic languages: Description, History and Foreign Influences* (D. Sinor, ed.), pp. 478–554. E. J. Brill, Leiden.
- Sinor, D. (ed.) 1988. *The Uralic Languages: Description, History and Foreign Influences*. E. J. Brill, Leiden.
- Suchard, M.A., Weiss, R.E. & Sinsheimer, J.S. 2001. Bayesian selection of continuous-time markov chain evolutionary models. *Mol. Biol. Evol.* **18**: 1001–1013.
- Swadesh, M. 1952. Lexicostatistic dating of prehistoric ethnic contacts. *Proc. Am. Philos. Soc.* **96**: 452–463.
- Swadesh, M. 1955. Towards greater accuracy in lexicostatistic dating. *Int. J. Am. Linguist.* **21**: 121–137.
- Syrjänen, K., Honkola, T., Korhonen, K., Lehtinen, J., Vesakoski, O. & Wahlberg, N. 2013. Shedding more light on language classification using basic vocabularies and phylogenetic methods: a case study of Uralic. *Diachronica* **30**: in press.
- Tadmor, U. 2009. Loanwords in the world's languages: findings and results. In: *Loanwords in the World's Languages: A Comparative Handbook* (M. Haspelmath & U. Tadmor, eds), pp. 55–75. Walter de Gruyter, Berlin.
- Tallavaara, M. & Seppä, H. 2011. Did the mid-Holocene environmental changes cause the boom and bust of hunter-gatherer population size in eastern Fennoscandia? *Holocene* **22**: 215–225.
- Toivonen, Y.H. 1953. Suomalais-ugrilaisesta alkukodista. *Virtittäjä* **57**: 5–35.
- Väliranta, M., Kaakinen, A. & Kuhry, P. 2003. Holocene climate and landscape evolution East of the Pechora Delta, East-European Russian Arctic. *Quatern. Res.* **59**: 335–344.
- Van Valen, L. 1973. A new evolutionary law. *Evol. Theor.* **1**: 1–30.
- Walker, R.S. & Ribeiro, L.A. 2011. Bayesian phylogeography of the Arawak expansion in lowland South America. *Proc. Biol. Sci.* **278**: 2562–2567.
- Weiss, H. & Bradley, R.S. 2001. What drives societal collapse? *Science* **291**: 609–610.
- Zakh, V.A., Ryabogina, N.E. & Chlachula, J. 2010. Climate and environmental dynamics of the mid- to late Holocene settlement in the Tobol-Ishim forest-steppe region, West Siberia. *Quatern. Int.* **220**: 95–101.
- Zhang, D.D., Lee, H.F., Wang, C., Li, B., Pei, Q., Zhang, J. et al. 2011. The causality analysis of climate change and large-scale human crisis. *Proc. Natl. Acad. Sci. USA* **108**: 17296–17301.
- Zheng, D., Prince, S. & Hame, T. 2004. Estimating net primary production of boreal forests in Finland and Sweden from field data and remote sensing. *J. Veg. Sci.* **15**: 161–170.
- Zubrow, E.B.W. 1971. Carrying capacity and dynamic equilibrium in the prehistoric southwest. *Am. Antiquity* **36**: 127–138.

## Supporting information

Additional Supporting Information may be found in the online version of this article:

**Appendix S1** Data set sources.

**Figure S1** Phylogenies made with different calibration point and prior combinations repeated four times each.

Received 27 March 2012; revised 11 December 2012; accepted 16 December 2012