Indo-European demic diffusion model

A dissertation by
Carlos Quiles M.D., M.Sc.

Ph.D. candidate
Department of Anatomy, Cell Biology, and Zoology
University of Extremadura

Badajoz
February 2017
Abstract

Introduction.-

Previous archaeological and linguistic theories have based the expansion of Indo-European languages on the expansion of the Corded Ware culture from the Yamna horizon.

Aim.-

To obtain a more precise general picture of the evolution of Eurasian cultures, peoples, and languages related to Indo-European languages.

Materials and Methods.-

Analysis of ancient and modern DNA samples together with available archaeological and linguistic data.

Results.-

The Indo-European demic diffusion model proposed advances the theory that the expansion of Indo-European languages from the steppe was linked to the expansion of peoples belonging to haplogroup R1b in Eurasia.

Discussion.-

Late Indo-European most likely expanded directly in its western migration into the Bell Beaker culture, and the Corded Ware culture was probably not linked to that expansion. That challenges previous archaeological and linguistic theories concerning the dialectal evolution of Indo-European languages and peoples.

Conclusion.-
Careful cross-disciplinary investigation of ancient DNA samples recently published supports a demic diffusion model for the expansion of Indo-European languages.
# Table of Contents

Abstract............................................................................................................................. 3

Table of Contents ............................................................................................................. 5

List of Figures ................................................................................................................... 7

I. Introduction ................................................................................................................... 9

II. Aim ............................................................................................................................ 11

III. Materials and Method ............................................................................................... 13

IV. Results ...................................................................................................................... 19

IV.1. Before Indo-European ....................................................................................... 19

IV.1.1. West Eurasian ancestry .............................................................................. 19

IV.1.2. East Eurasian ancestry ................................................................................ 20

IV.1.3. Languages of Eastern Europe in the Mesolithic ......................................... 21

IV.2. Early Indo-European ........................................................................................... 25

IV.3. Middle Indo-European ...................................................................................... 28

IV.4. Late Indo-European ........................................................................................... 31

IV.4.1. Contacts with the Caucasus ........................................................................ 32

IV.5. The Forest Zone ................................................................................................ 35

IV.5.1. Expansion of the Corded Ware culture ...................................................... 36

IV.6. Tocharian expansion ......................................................................................... 38

IV.7. Late Indo-European expansion .......................................................................... 41
List of Figures

Figure 1. Example of diagram of expansion and relationships of Indo-European languages, based solely on linguistic considerations. Adapted from Adrados (1998)......................... 14
Figure 2. Diachronic map of Paleolithic migrations (continued in the next page)........ 23
Figure 3. Diachronic map of Mesolithic migrations ca. 6500-5000 (continued in the next page)............................................................................................................................... 26
Figure 4. Modern distribution of LPCS ancestry (R1b-M269 (xL23)), adapted from Richard Rocca (2012)........................................................................................................................ 29
Figure 5. Diachronic map of Early Neolithic migrations ca. 5000-4000 BC .......... 30
Figure 6. Diachronic map of Neolithic migrations ca. 4000-3000 BC.................... 34
Figure 7. Diachronic map of Copper Age migrations ca. 3000-2250 BC ............... 37
Figure 8. Diachronic map of migrations in Asia ca. 3000-2800 BC ......................... 40
Figure 9. Diachronic map of migrations in Europe ca. 3000-2250 BC................. 42
Figure 10. Modern distribution of WYP ancestry (L-51*). Adapted from Richard Rocca (2012). ............................................................ 45
Figure 11. Modern distribution of EYP ancestry (R1b-Z2103). Adapted from Myres et al. (2011). ............................................................... 48
Figure 12. Modern distribution of Atlantic WEP ancestry (R1b-L21). Adapted from Myres et al. (2011) ............................................................... 51
Figure 13. Modern distribution of Gallic WEP ancestry (R1b-DF27). Adapted from Myres et al. (2011) ............................................................... 52
Figure 14. Diachronic map of migrations in Europe ca. 2250-1750 BC............ 54
Figure 15. Diachronic map of migrations in Europe ca. 1750-1250 BC............. 58
Figure 16. Modern distribution of Northern WEP ancestry (R1b-U106). Adapted from Myres et al. (2011). ............................................................... 59
Figure 17. Diachronic map of migrations in Europe ca. 1250-750 BC. ........................................ 62
Figure 18. Modern distribution of Alpine WEP ancestry (R1b-DF27). Adapted from Myres et al. (2011). ........................................................................................................................................... 64
Figure 19. Diachronic map of migrations in Europe ca. 750-250 BC. . . . .................................. 65
Figure 20. Diachronic map of migrations in Europe ca. 250 BC – 250 AD. .................. 67
Figure 21. Diachronic map of migrations in Europe ca. 250-750 AD. ......................... 74
Figure 22. Modern distribution of ECW ancestry (R1a-Z282). Adapted from Underhill et al. (2015). ........................................................................................................................................... 75
Figure 23. Diachronic map of migrations in Asia ca. 2250-1750 BC. ......................... 78
Figure 24. Diachronic map of migrations in Asia ca. 1750-1250 BC. ....................... 80
Figure 25. Diachronic map of migrations in Asia ca. 1250-750 BC. ................................ 81
Figure 26. Diachronic map of migrations in Asia ca. 750-250 BC. ................................. 82
Figure 27. Diachronic map of migrations in Asia ca. 250 BC - 250 AD. ...................... 83
Figure 28. Diachronic map of migrations in Asia ca. 250-750 AD. ................................. 84
Figure 29. Modern distribution of ACW ancestry (R1a-Z93). Adapted from Underhill et al. (2015). ........................................................................................................................................... 86
Figure 30. Diachronic map of migrations ca. 750 – 1300 AD (continued in the next page). ........................................................................................................................................... 88
I. Introduction

Language and culture expansion is explained by two main alternative models: the demic diffusion model, which involves mass movement of people; and the cultural diffusion model, which refers to cultural impact between populations, and involves limited genetic exchange between them. Language transfer since ancient times seems to be associated with an expansion of people (Mikhailova 2015), which is demonstrated, in most cases, by a significant replacement of patrilineal Y-DNA. Investigation of Y-DNA haplogroups help demonstrate e.g. the expansion of Han people in Northern and Southern China (Wen et al. 2004; Zhao et al. 2015), and the expansion of Arabs in the Arab peninsula (Chiaroni et al. 2010), and into Southern Levant and North Africa (Nebel et al. 2002). Recently, the genetic history of Europe – including the expansion of hunter-gatherers and farmers – has been more precisely shaped thanks to ancient DNA research (Fu et al. 2016).

The recent expansion into Europe and Asia of Eurasian pastoralists, commonly identified with Indo-European speakers in mainstream diffusion models (Gimbutas 1993; Mallory 2014), was linked to haplogroup R1a (Semino 2000; Wells et al. 2001; Zerjal et al. 1999) due to the correlation of its modern geographic distribution with the ancient Corded Ware culture, and modern Balto-Slavic, Germanic, and Indo-Iranian speaking areas (Mirabal et al. 2009; Underhill et al. 2010).

Haplogroup R1b, which shows a modern Western European distribution peaking in the British Isles and around historically Basque-speaking regions (Myres et al. 2011; Lucotte 2015), was until recently associated with a Paleolithic Western European origin (Morelli et al. 2010; Semino 2000). With decreased age estimates of haplogroup R1b in Europe, a more recent spread with farming has been suggested (Myres et al. 2011; Chiaroni, Underhill, and Cavalli-Sforza 2009; Cruciani et al. 2011; Balaresque et al. 2010).
Following these genetic frameworks, Indo-European languages would have spread with an Indo-European-speaking, R1a-dominated, invasive, Eastern (Corded Ware) population into a non-Indo-European-speaking, R1b-dominated, Western Atlantic (Bell Beaker) population. This connection was the weakest link between the supposed archaeological and the attested historical European linguistic landscapes, needing explanatory models that included some kind of cultural diffusion model – e.g. technologically- or economically-based (Brandt et al. 2015).

Ancient DNA (aDNA) investigation allows us to disentangle complex human history (Slatkin and Racimo 2016). The most recent research of ancient genetics (Haak et al. 2015; Allentoft et al. 2015; Mathieson et al. 2015), concerned with general population movements of Eurasians westwards from the steppe, has shown with their published data that haplogroup R1b was almost absent from Western Europe until after the expansion of Eurasian pastoralists, that the origin of most of its modern descendants in Western Europe is probably to be traced to the Pontic-Caspian steppes, and therefore that its expansion into central Europe happened at nearly the same time as haplogroup R1a, i.e. from the East and after ca. 3000 BC (Haak et al. 2015). In these studies, R1a was absent from samples of the Yamna horizon, most of which belonged to haplogroup R1b-M269.

The earliest linguistic link between haplogroups R1b and R1a, deemed until recently a cultural diffusion along the Corded Ware – Bell Beaker contact area (and later among Bell Beaker groups), seems thus to be contested by the latest genetic research. However, alternative explanations are being sought to adapt older paradigms to the newest research, e.g. suggesting a direct connection of the expansion of Indo-European languages to the Corded Ware culture (Allentoft et al. 2015), and thus R1a as the genetic marker of the expansion of Proto-Indo-European speakers in Europe (Horvath 2015).
II. Aim

To obtain a more precise general picture of the evolution of Eurasian cultures, peoples, and languages since prehistoric times, and especially of the expansion of Indo-European-speaking peoples.
III. Materials and Methods

The theory presented here offers an alternative population expansion model that seems to better fit the recent genetic research (involving ancient as well as modern DNA investigations) with mainstream archaeological, anthropological and linguistic models.

Linguistic models of Indo-European (IE) dialectal differentiation based on comparative grammar and internal reconstruction (Figure 1) will be used to illustrate this theory, most of which currently follow the mainstream three-stage migration model (Meid 1975). The most common nomenclature of Early, Middle, and Late Indo-European periods is used (Dunkel 1997).

Such linguistic models consider Proto-Indo-European (PIE) as a product of a long historical development, formed gradually – like most natural languages –, and having thus stages of development (Lehmann 1992). This theory is therefore in contrast with the ‘constellation analogy’ (Clackson 2007) and similar negations of a concrete community of speakers – defined in time and space – of PIE or any of its later dialects. Historical linguistics can only provide a relative historical framework for individual Indo-European languages and proto-languages, though (Mallory and Adams 2007).

Archaeology works with the concept of culture, and as such it is able to determine timelines. When these timelines complement linguistics beautifully both are able to provide a contextualized historical explanation of linguistic frameworks (Vander Linden 2015; Hänsel and Zimmer 1994). The model set forth by Marija Gimbutas (1953), impressively expanded recently by Anthony (Anthony 2007; Anthony and Brown 2011; Anthony 2013), of potential cultures where Indo-European was spoken, is used as framework for the potential expansion of Indo-European peoples.
Even though Anthony links his theory to a linguistic model developed by phylogenetics (Anthony and Ringe 2015; Ringe, Warnow, and Taylor 2002), it seems more reasonable to avoid such methods, due to their controversial nature and labile results (Pereltsvaig and Lewis 2015).

The theory laid in this paper takes dialectal evolution – lying at the core of any IE expansion model – as its stable framework, and uses genetic investigation (of ancient and modern DNA samples) and its potential relationship with archaeological cultures to establish an expansion model step by step, taking into account that there are complex problems found in correlations of languages with archaeological cultures (Meier-Brügger 2003) and human genetics (Campbell 2015).

Ancestry of any selected population is likely to be a mixture of several ancient groups, which is reflected on the genetic structure (Haak et al. 2010; Skoglund et al. 2012; Malmström et al. 2009; Lazaridis et al. 2014). However, the genetic landscape for ancient populations is limited by the number of ancient DNA samples and ancient populations studied (Hellenthal et al. 2014). For simplicity purposes, then, none of the new haplotype-
based techniques are used to obtain admixture analyses. Only SNPs of ancient Y-DNA samples will be taken into account in this study, since mtDNA samples involve a more complex analysis in demic diffusion models – where the paternal ancestry of the invaded territory is believed to be replaced or displaced to a certain extent. According to this simplification, meaningful names are given to ancestries associated with relevant Y-DNA SNPs, and not to actual communities.

The following is a list of haplogroups relevant to this work, their formation date and the time to most recent common ancestor (TMRCA), as well as the common names and abbreviations used to refer to them in this paper:

- **I** (formed ca. 40900 BC, TMRCA ca. 25500 BC): European hunter-gatherer (EHG).
  - **I1** (formed ca. 25500 BC, TMRCA ca. 2600 BC): Northern European hunter-gatherer (NEHG).
  - **I2** (formed ca. 25500 BC, TMRCA ca. 21600 BC): Western European hunter-gatherer (WEHG).
    - **I2-L621** (formed ca. 9300 BC, TMRCA ca. 4500 BC): Balkan Neolithic.
    - **I2-M223** (formed ca. 15500 BC, TMRCA ca. 12100 BC): Central European Neolithic.
  - **I2-M284** (formed ca. 8500 BC, TMRCA ca. 5200 BC): Western Central European Neolithic.
- **G2-P15** (formed ca. 18600 BC, TMRCA ca. 16200 BC): Neolithic farmer (NF).
- **N1c1** (formed ca. 13000 BC, TMRCA ca. 12000 BC): East Asian hunter-gatherer (EAHG).
  - **N1c1a1** (formed ca. 11800 BC, TMRCA ca. 5100 BC): Siberian hunter-gatherer (SHG).
    - **N-Z1936** (formed ca. 2300 BC, TMRCA ca. 2300 BC): Uralic hunter-gatherer (UHG).
    - **N-VL29** (formed ca. 2300 BC, TMRCA ca. 1500 BC): Baltic hunter-gatherer (BFHG).
  - **N-Y4706** (formed ca. 600 BC, TMRCA ca 500 BC): North Baltic hunter-gatherer.
• **N-M2783** (formed ca. 600 BC, TMRCA ca. 600 BC): South Baltic hunter-gatherer.

• **R*** (formed ca. 29,900 BC, TMRCA ca. 26200 BC): Ancient West Eurasian (AWE).
  
  o **R1** (formed ca. 26200 BC, TMRCA ca. 20800 BC): Northern Ancient West Eurasian (NAWE).
    
    ▪ **R1a** (formed ca. 20800 BC, TMRCA ca. 16400 BC): East Eurasian (EE).

• **R1a-M17** (formed ca. 12100 BC, TMRCA ca. 6500 BC): Late EE.
  
  o **R1a-M417** (formed ca. 6500 BC, TMRCA ca. 3500 BC): Eastern European hunter-gatherer (EEHG).

    ▪ **R1a-Z283, R1a-Z282** (formed ca. 3000 BC, TMRCA ca. 3000 BC): European Corded Ware (ECW).
      
      ▪ **R1a-Z284** (formed ca. 2700 BC, TMRCA ca. 2300 BC): Northern European Corded Ware (NECW).
      
      ▪ **R1a-M458** (formed ca. 2700 BC, TMRCA ca. 2700 BC): Central European Corded Ware (CECW).
      
      ▪ **R1a-M558** (formed ca. 2600 BC, TMRCA ca. 2300 BC): Eastern European Corded Ware (EECW).

    ▪ **R1a-Z93** (formed ca. 3000 BC, TMRCA ca. 2700 BC): Asian Corded Ware (ACW).
      
      ▪ **R1a-L657** (formed ca. 2700 BC, TMRCA ca. 2200 BC): West ACW.
      
      ▪ **R1a-Z2125** (formed ca. 2700 BC, TMRCA ca. 2100 BC): East ACW.
- **R1b** (formed ca. 20800 BC, TMRCA ca. 18400 BC) West Eurasian (WE).
  - **R1b-L278** (formed ca. 18400 BC, TMRCA ca. 16800 BC): European West Eurasian (EWE).
    - **R1b-V88** (formed ca. 9700 BC, TMRCA ca. 9700 BC): Southern European West Eurasian (SEWE).
    - **R1b-P297** (formed ca. 14800 BC, TMRCA ca. 13300 BC): Early Pontic-Caspian steppe (EPCS).
  - **R1b-M269** (formed ca. 11300 BC, TMRCA ca. 4500 BC): Middle Pontic-Caspian steppe (MPCS).
  - **R1b-L23** (formed ca. 4500 BC, TMRCA ca. 4200 BC): Late Pontic-Caspian steppe (LPCS).

The following lineages derived from **R1b-L23**:

- **R1b-Z2103** (formed ca. 4200 BC, TMRCA ca. 4100 BC): Eastern Yamna pastoralist (EYP).
  - **R1b-Z2106** (formed ca. 4100 BC, TMRCA ca. 4100 BC).
  - **R1b-Z2108** (formed ca. 4100 BC, TMRCA ca. 4100 BC): Balkan EYP.
  - **R1b-2110** (formed ca. 4100 BC, TMRCA ca. 4100 BC): Central EYP.
    - **R1b-L584** (formed ca. 4100 BC, TMRCA ca. 2700 BC): Anatolian EYP.
    - **R1b-L277** (formed ca. 4100 BC, TMRCA ca. 2100 BC): Northern EYP.
- **R1b-L51** (formed ca. 3800 BC, TMRCA ca. 3600 BC): Western Yamna pastoralist (WYP).
  - **R1b-L151** (formed ca. 3600 BC, TMRCA ca. 2900 BC), R1b-Z2111 (formed ca. 3600 BC, TMRCA ca. 2800 BC): Western European pastoralist (WEP).
  - **R1b-U106** (formed ca. 2900 BC, TMRCA ca. 2900 BC): Northern WEP (NWEP).
  - **R1b-P312** (formed ca. 2900 BC, TMRCA ca. 2400 BC): Southern WEP (SWEP).
From **R1b-P312**, the following lineages derived:

- **R1b-DF27** (formed ca. 2400 BC, TMRCA ca. 2400 BC):
  Gallic WEP (GWEP).
- **R1b-U152** (formed ca. 2400 BC, TMRCA ca. 2400 BC):
  Alpine WEP (AWEP).
- **R1b-L21** (formed ca. 2400 BC, TMRCA ca. 2400 BC):
  Atlantic WEP (AtWEP).

Modern physical maps are used to illustrate potential expansion routes of ancient cultures, peoples, and languages, even though they pose a significant danger to the development of a sound model, since they almost invariably involve “a concatenation of weakly supported links that corporately form an “arrow” of dispersion” (Mallory 2014). Map routes are only depicted as a visual help to add movement to the otherwise stationary maps of ancient cultures, peoples, languages, and ancient DNA obtained from scattered burials.
IV. Results

IV.1. Before Indo-European

Ancient West Eurasian (AWE) ancestry is linked to Ancient North Eurasian (ANE) genetic lineages – itself descending from Ancient South Eurasian (Lazaridis et al. 2016) –, and represents a patrilineal line traceable to the Mal’ta-Buret’ culture, or a population closely related to them (Flegontov et al. 2016).

Western Eurasian (WE) and Eastern Eurasian (EE) lineages split from a common ancestor that emerged shortly after the Mal’ta-Buret’ culture (Raghavan et al. 2014).

IV.1.1. West Eurasian ancestry

WE lineage was deemed to have originated ca. 16400 BC in Western Asia (Karafet et al. 2008), and it has been proposed that it survived the Last Glacial Maximum in refugia near the southern Ural Mountains and the Aegean Sea (Lobov 2009). Its latest westward migration has been proposed to have happened during the Late Neolithic (Myres et al. 2011).

However, an Epigravettian individual with European West Eurasian (EWE) ancestry from Villabruna is dated ca. 12000 BC (Fu et al. 2016). Based on the most recent data of modern populations an origin of EWE ca. 20800 BC is suggested, with a time to MRCA of ca. 16800 BC. Early Pontic Caspian Steppe (EPCS) ancestry (formed ca. 14800 BC, TMRCA ca. 13300) has been found in two individuals of the Narva culture dated ca. 5700 and ca. 5000 BC (Jones et al. 2017), and another EPCS (but potentially intermediate stage EWE-EPCS) individual was found in Mesolithic Samara (Mathieson et al. 2015). EWE examples have been found as Southern European West Eurasian (SEWE) ancestry in Els Trocs ca. 5311-5068 BC, and as non-SEWE, non-EPCS EWE ancestry in Quedlinburg
ca. 3645-3537 BC (Haak et al. 2015). These facts may suggest an early migration of EWE ancestry into Europe, and a spread from Europe in different directions.

The finding of basal WE ancestry in modern populations of southern Kazakhstan (Myres et al. 2011) and Iran (Grugni et al. 2012) may suggest a southern migration route into Europe. Basal EWE ancestry was found in five cases out of 5,326 cases: 3 Italians, 1 West Asian, 1 East Asian (Cruciani et al. 2010), which also point to a potential ancestral European migration of EWE lineages (Figure 2).

**IV.1.2. East Eurasian ancestry**

EE lineage originated ca. 25000 BC, and has been proposed to diverge initially in the vicinity of present-day Iran based on the study of modern populations (Underhill et al. 2015). Hunter-gatherers of EE ancestry were proposed to have migrated from the Iranian area to the forests of Eastern Europe, since early samples were found in the Narva culture (Horvath 2015).

While the oldest aDNA sample for EE is found in an individual from a post-Swiderian culture dated ca. 6850-6000 BC (Lazaridis et al. 2016), two samples of Late EE ancestry (ca. 6125-4885 BC) found in Irkutsk – near the zone where the ancient Mal’ta-Buret’ culture was located – and the TMRCA of ca. 6500 BC for Late EE ancestry (based on modern populations) point to a more recent westward migration into Europe compared to WE lineages. It is tempting to find the origin of this late migration in the traditional association of hunters of the Kelteminar culture (which began ca. 5500 BC) to the steppe: ancestors of this population are supposed to have originally migrated from the Hissar range ca 6000 BC.

The end of the last Ice Age ca. 14000-12000 BC brought instability to the Pontic-Caspian zone: meltwater flew torrentially from the northern glaciers and the permafrost into the Khvalynian Sea (the Caspian Sea is a small remaining part of it), and a shoreline between the middle Volga and the Ural River restricted east-west movements south of the Ural Mountains (Anthony 2007). By 11000-9000 BC a violent flood may have poured into the Black Sea, enlarging it and creating the Sea of Azov. Although the magnitude and rapidity of this flow remains controversial, it is agreed that meltwater created unstable shores in this area. After about 8000 BC, the Pontic-Caspian steppe was stable, and Mesolithic hunters settled there.
IV.1.3. Languages of Eastern Europe in the Mesolithic

The different migration times and paths of WE and EE groups puts the Ural-Caspian frontier as a linguistic and cultural barrier: their languages and cultures (if they were once related), must have evolved differently – and they might have acquired different languages from other peoples along the way. It does not seem reasonable to assume a single language for both populations by the time period when a common Proto-Indo-European is believed to have been spoken.

ECPS formation date and TMRCA point to a slow spread of hunters in Eastern Europe, coinciding with the geographic changes associated with the last deglaciation. The population of the final Paleolithic Swiderian culture, which developed in Poland on the sand dunes left behind by retreating glaciers, migrated during the Paleolithic-Mesolithic transition (ca. 9500 BC) to the northeast following the retreating tundra, which is evidenced by a 300-year-long settlement break before a new population arrived (Kobusiewicz 2002).

Post-Swiderian cultures developed in the Baltic and in the Forest Zone north of the unstable Pontic-Caspian zone, and it is reasonable to assume a migration of hunters from the north into the newly created steppes ca. 8000 BC.

Indo-European has been described as “a branch of Indo-Uralic which was transformed under the influence of a Caucasian substratum” (Kortlandt 2002), which would imply an invasion of ECPS ancestry in a territory of previous Caucasian hunter-gatherers. Such Caucasian influence has found support recently by the finding of a genetic contribution of a pocket of Caucasus hunter-gatherers (ca. 11000–8000 BC), who seem to have weathered much of the last Ice Age in apparent isolation (Jones et al. 2015).

The invasion of EE ancestry must have disrupted the pre-Indo-European WE community thriving in Eastern Europe. In this context, Uralic – as the language of EE migrants that came to the Forest Zone from the East – could be considered an Uralo-Yukagir superstratum over a Pre-Indo-European substratum. A Uralo-Yukagir community spread over Eurasia is supported by the finding of maximum Ancient North Eurasian ancestry in modern-day Kets, Mansi, Native Americans, Nganasans and Yukagirs (Flegontov et al. 2016).
Proto-Uralic is therefore assumed to be the language spoken by hunter-gatherers that lived in the Forest Zone, north of the Pontic Caspian steppe, at the same time as Middle Indo-European was spoken in the steppes (Parpola 2012).

The arrival of East Asian hunter-gatherer (EAHG) ancestry into Northern Europe has been dated after 5000 BC (Ilumae et al. 2016), and the still more recent formation and TMRCA for common European ancestries point a late and stepped spread of these hunter-gatherer groups into the Forest Zone, hence cultural assimilation remains the best explanation at the moment for the shared Uralic languages of modern EE and EAHG communities. EAHG groups may have brought with them the Altaic traits found in Uralic languages (Kortlandt 2010), and even though an aDNA sample of EAHG ancestry is found in the Forest Zone dated ca. 2500 BC, it is tempting to place the mass migration of Siberian hunter-gatherer communities around the Urals with the expansion of the poorly understood Seima-Turbino phenomenon (which began ca. 2000 BC in East Asia), since it connected cultures from Mongolia to Finland.

Long-ranging language relationships are difficult to prove. If Uralic and Indo-European shared a common ancestor – Indo-Uralic (Kloekhorst 2008) –, it should be associated to the post-Swderian European WE communities. It would then be conceivable to give some credit to the nature of Proto-Basque as of Pre-Indo-Uralic substratum (Blevins 2015), beyond pre- and post-Roman IE superstrata (Koch 2013), based on the presence of SEWE ancestry in Els Trocs ca. 5295-5066 BC.
Figure 2. Diachronic map of Paleolithic migrations (continued in the next page)
IV.2. Early Indo-European

Early Pontic-Caspian steppe ancestry is found in a male from a Mesolithic pre-Yamna culture buried ca. 5650-5555 BC in the Samara region (Mathieson et al. 2015), north of the Caspian Sea.

Hunters from the Pontic-Caspian steppe – as European Mesolithic hunter-gatherers in general – possessed no domesticated animals before the arrival of stockbreeding, spread with Neolithic farmers from Anatolia after about 6100 BC.

However, before the arrival of southern farmers, pottery had begun in the Elshanian culture ca. 7000-6500 BC, and it had spread south- and westward first into the Northern Caspian culture ca. 6500 BC, and then into North Pontic societies ca. 6200-6000 BC (Zaitseva et al. 2009), which suggests a close cultural link among Mesolithic Pontic-Caspian cultures, or maybe a population movement from the Samara region.

Contacts of North Pontic cultures with Criş settlers from the Starčevo–Körös–Criş culture about 5800 BC introduced domesticated cattle to the Bug-Dniester culture, but no signs of cultural assimilation has been found, with the later invasion of Linear Pottery sites ca. 5500-5200 BC respecting a similar cultural frontier, geographically coincident with the Dniester (Anthony 2007). Hence the language of western Neolithic settlers – assumed to come from the Middle East – was probably not transferred to North Pontic herders. From the Bug-Dniester culture domesticated cattle, sheep, and goats spread quickly from about 5200 BC east- and northward into Pontic-Caspian sites, reaching Khvalynsk and the Samara region about 5100 BC (Figure 3).

A population expansion during this time – eastward or westward from any small community within the Pontic-Caspian steppe – is not warranted by the available data, and seems not necessary to explain subsequent developments, since different Mesolithic communities of the Pontic-Caspian steppe may well have spoken Early Indo-European sister languages.
IV.3. Middle Indo-European

The Copper Age began in Bulgaria ca. 5200-5000 BC, and Old European copper-trade network included the Pontic-Caspian steppe societies after ca. 4600 BC. The revolution of herding, travel, and raiding – and thus the change in the steppe – came with horseback riding, appearing ca. 4800 in Early Khvalynsk, and spreading south- and eastward. The early Sredni Stog culture began about 4400 BC, and it seems that people from the east Pontic Caspian steppe (related to Early Khvalynsk) brought a new culture (Anthony 2007), and probably also their Middle Indo-European language.

Within this new culture, a new elite group associated with the Suvorovo-Novodanilovka complex (Anthony 2007) was involved in raiding and trading with the lower Danube valley during the Tripolye B1 period, before and during the collapse of Old Europe. Settlements of Suvorovo-Novodanilovka chiefs have been found along the lower Danube (Figure 5), and sites of the posterior Cernavodă I culture seem to represent the assimilation of migrants from the steppes, therefore linked to Anthony’s first expansion from the Pontic-Caspian steppes into southern Europe ca. 4200-4000 BC (Anthony 2007; Anthony 2013).

Middle Pontic-Caspian steppe (MPCS) ancestry seems to have expanded westward from the Pontic-Caspian steppe horizon early, coinciding with the successfully spread of basal Late Pontic-Caspian steppe (LPCS) ancestry, since both populations are found in the Balkans, Central Europe, and Armenia (Myres et al. 2011; Herrera et al. 2012).

The old origin of MPCS (ca. 11300 BC) compared to a later TMRCA (ca. 4500 BC) points to a survival of subclades in the modern population related to an expansion occurred exactly around this time period, probably from some eastern clans of Pontic-Caspian herders that developed the Sredni Stog culture in the west, and turned into Suvorovo-Novodanilovka chiefs and southeastern European settlers. This is in turn to be associated with the split of the ancestor of Anatolian (Kortlandt 1990; Ringe 2006) from a common Middle Indo-European (Tischler and Oettinger 1989; Lehrman 1996; Melchert 1998).

The modern distribution of MPCS ancestry in the Balkans and Anatolia (not reaching the Armenian highlands) points to the posterior migration of MPCS lineages with Anatolian languages (Figure 4). Its modern peak around Kosovo can be explained by a later founder effect that might have happened during any expansion of peoples in the region in the past.
four thousand years, and which can tentatively be assigned to a recent Albanian expansion. Its modern distribution in the Alps and in ancient Tyrrhenia might point to an eastern route of the Suvorovo-Novodanilovka settlers of eastern Hungary, hence giving support to the theories describing Etruscan as an Anatolian branch (Adrados 1989, 1994). On the other hand, it could well be a sign of independent back and forth migrations between the Adriatic Sea and the Italian Peninsula.
Figure 5. Diachronic map of Early Neolithic migrations ca. 5000-4000 BC.
IV.4. Late Indo-European

After 4000 BC, different groups were formed in the steppes. In the west, Late Sredni Stog and “Post-Mariupol” (“Extended-Position-Grave”) communities, the heirs of the western Early Sredni Stog clans, remained in contact with Tripolye villagers, and some assimilation seems to have happened east of the Dnieper ca. 3700-3500 BC.

In the east, Early Khvalynsk gave way to late Khvalynsk and Repin societies in the Volga-Don region, whose language is to be associated with a common Late Indo-European (Anthony 2007). The split of LPCS into Eastern Yamna pastoralist (EYP) ancestry must have happened early, possibly during the previous westward expansion of eastern clans (of MPCS and LPCS ancestry) in and outside of the Pontic-Caspian steppes – given the similar forming (ca. 4200 BC) and TMRCA (ca. 4100 BC) dates. The earliest aDNA samples of EYP ancestry are three individuals found in the Late Khvalynsk zone in Lopatino I (3339-2917 BC), Peshany V (3334-2635 BC), and Ishkinovka I (3305-2925 BC) (Haak et al. 2015).

Western Yamna pastoralist (WYP) ancestry split later than EYP from the common LPCS stem (formed ca. 3800 BC, TMRCA ca. 3600), and given its later successful expansion into western Europe it is hypothesized to have successfully expanded to a certain extent during the common Yamna (“Pit Grave”) period of the Pontic-Caspian steppe cultures. Successful groups of the Repin culture might have first expanded coinciding with the common date of ca. 3600 for TMRCA of WYP and formation of WEP ancestry, but no early sample has been found to date in the Yamna horizon.

Given the lack of aDNA from the Western Yamna horizon, and the later westward expansion of EYP lineages, it is probably safest to assume a western location of the group within Yamna. It would have formed a community with EYP, but somehow separated culturally from it, and thus the two main dialects of Late Indo-European may have developed separately.

Graeco-Aryan (probably including at least Greek, Armenian, and Indo-Iranian) has been argued as a dialect continuum or a linguistic community where a number of common innovations were shared at an early time (Mallory and Adams 2007; West 2007). North-West Indo-European – including Italic, Celtic, Germanic, Balto-Slavic, and Tocharian – has been proposed as a group of closely related dialects (Adrados 1998; Oettinger 1997; Mallory and Adams 2007).
Both linguistic communities remained thus in close contact, and are probably to be located to the eastern Don-Volga steppes, spreading across the Pontic-Caspian steppes after about 3300 BC (Anthony 2007). Because of their later expansion, their division could be speculatively traced back to the early division of Volga-Don groups: the western, Don-based Repin culture, and the eastern, Volga-based Late Khvalynsk culture (Figure 6).

The westward and eastward expansion of the Repin culture about 3300 BC is associated to the rapid diffusion of the Yamna horizon across the Pontic-Caspian steppes, and a common, disintegrating Late Indo-European may have been spoken in this common period, where laryngeals were already unstable, and had probably already undergone the first common phase of laryngeal loss (Quiles 2012).

IV.4.1. Contacts with the Caucasus

The Uruk expansion in Mesopotamia after about 3700 BC intensified during the Late Uruk period ca. 3350-3100 BC, and its expansion reached toward the gold, silver, and copper sources in the Caucasus Mountains. The Maykop culture of rich chieftains’ graves with Mesopotamian ornaments developed from this trade network in the North Caucasus Piedmont, and a western and probably also a later eastern southern routes through the shores of the Black and Caspian seas respectively have been proposed (Anthony 2007).

Steppe-Caucasian trade is supported by Maykop imports found in the North Pontic steppes from the Dniester to the lower Volga in the east, but no Caucasian imports have been found in the Volga-Ural region. Late Maykop peoples – most likely speaking languages ancestral to modern Caucasian languages – probably interacted with individuals from Repin and Late Khvalynsk cultures, and the contact was most direct on the lower Don. Late Maykop graves incorporated carved stone stelae like those of western Yamna. The trading of drugs, wool, and horses has been proposed as main steppe imports into Maykop (Anthony 2007).

Horse trade – including wheels, carts, and the possibility of a quicker transport of metals into Uruk – are proof of an indirect contact between steppe herders and Mesopotamia. The association of exported domesticated horses with experienced breeders and riders of the lower Don offers a solid framework to support the hypothesis of the presence of Late-Indo-European-speaking peoples in Mesopotamia, and thus allow for IE borrowings in Sumerian (Sahala 2009-2013). The condition of North-West Indo-European as an
Euphratic superstratum of Sumerian (Whittaker 2008, 2012) would require a more detailed explanation of internal and external influence, and reasons for potential language replacement and expansion in Mesopotamia.
Figure 6. Diachronic map of Neolithic migrations ca. 4000-3000 BC.
IV.5. The Forest Zone

The oldest remains of EE ancestry are found in the forests north of the Pontic-Caspian steppe: ca. 5500-5000 BC in Yzhhnyy Oleni Ostrov (Haak et al. 2015), and ca. 4000 BC in Serteya VIII (Chekunova et al. 2014). During this stage of Rudnyayan culture there is continuity in relation to the previous stage, and contacts are made with eastern Baltic area and through the Western Dvina (Mazarkevich et al. 2009).

The oldest pottery of the Forest Zone – from the older Early Comb Ware culture – was created about 5200-4500 BC in southern Finland and Karelia. At the end of this period, ca. 4500-4300 BC, Pit-Comb Ware replaces pottery in the southern sites, but in southern Finland and Karelia variants of the older types were still in use (Nordqvist and Mökkönen 2016).

The introduction of Typical Cord Ware, which heralded the appearance of Neolithic traits in the Forest Zone, is dated to around 3900 BC, and it was discontinued ca. 3400 BC. It was a relatively uniform culture that covered a vast area ranging from the Urals to the Baltic Sea, and from Northern Ukraine to the Arctic Ocean.

Recent research points to the split of Eastern European hunter-gatherer (EEHG) ancestry from EE ca. 3800 BC (Underhill et al. 2015). EEHG foragers were found not to have received genetic influx from Anatolian-farmer-related genes during the Mesolithic or Neolithic, and therefore an inner cultural diffusion of pottery, farming and metallurgy is assumed for the population of the Baltic and Dnieper Rapids (Jones et al. 2017).

The disintegration of the Comb Ware phase began ca. 3500 BC, coinciding with the influence of the Volga-Kama region and the birth of several variants of Asbestos- and Organic-tempered Wares, although no break has been observed in cultural development (Nordqvist et al. 2012). These groups also maintained vast and varying intra- and inter-regional contact networks.

During this period of 3500-3000 BC a shift to drier and cooler conditions is found in the steppes, with steppes expanding, and therefore also Yamna pastoralists and their cattle following them. The emergence of Volosovo and Garino-Vor metallurgy in the fourth millennium has been attributed to external influences from Yamna.
IV.5.1. Expansion of the Corded Ware culture

Between 3500-2000 BC an interruption in cultural continuity in the Forest Zone is found, coinciding with a major change in the environment, with selective felling and subsequent regeneration of forests in the Pit-Comb Ware area (Mazurkevich et al. 2009; Poska and Saarse 2002), which could have been caused by the complex movement of peoples in this period, as reflected by the interaction or “checkerboard of regional cultures covering the rolling hills and valleys of the forest steppe zone” (Anthony 2007), and a complex set of cultures is found in the East European Forest Zone, different from Central European cultures (Czebreszuk and Szmyt 2004).

Corded Ware (CW) ancestry separated from EEHG ca. 3000 BC, nearly at the same time as the westward migration of the Corded Ware culture is supposed to have begun, and at the same time as European (ECW) and Asian Corded Ware (ACW) ancestries splintered from it. Corded Ware groups migrated east- and westward, reaching the Middle Elbe-Saale region – where most aDNA samples analyzed come from – about 2750 BC (Figure 7).

The earliest ECW lineages are found in Globular Amphorae and early Corded Ware cultures, while ancient DNA from Neolithic Linear Pottery culture (ca. 5500–3500 BC) have been found to correspond to European farmer and hunter-gatherer lineages (Haak et al. 2015), which raises the possibility of a wide and rapid spread of ECW ancestry in Europe (Underhill et al. 2015), associated with the expansion of the Corded Ware culture.

The diversification of ACW in the Middle East and South Asia remains more obscure (Underhill et al. 2015).

Ancient DNA samples point to a western expansion of EEHG foragers into Central Europe as part of the spread of the Corded Ware culture: basal EEHG (ca. 2473-2348 BC) in Erperstedt (Haak et al. 2015); individuals with EEHG ancestry (ca. 2286-2048 BC) in Leki Male, and in the Nordic Late Neolithic (ca. 2191-1972 BC) in Marjberg (Allentoft et al. 2015). Two aDNA samples indicate that a population of EE ancestry also remained in this area during the Usvyatyan culture (ca. 2500 BC) in Naumovo and Sertaya II (Chekunova et al. 2014), and were thus not displaced from the region.
Figure 7. Diachronic map of Copper Age migrations ca. 3000-2250 BC.
IV.6. Tocharian expansion

The Early Bronze Age Afanasievo culture (ca. 3500-2600) in the Altai-Sayan region has been found to be genetically indistinguishable from Yamna in admixture analyses using exclusively mtDNA (Allentoft et al. 2015).

Linguists have placed Proto-Tocharian within the North-West Indo-European dialectal zone, but with innovations compatible with an isolated development (Mallory and Adams 2007). The findings of Anthony (2007) regarding the spread of a group from the Repin society into Afanasievo (ca. 3000-2800) supports the eastward expansion of a western Yamna group, and therefore the position of Tocharian as part of an early North-West Indo-European split.

The only non-EYP found in the Yamna horizon comes from Lopatino II (ca. 3300-2700 BC), in the Samara region (Haak et al. 2015), which might be speculatively interpreted as part of an expanding western group of WYP ancestry in the Volga-Ural zone, where previous (Late Khvalynsk) and posterior (Poltavka) aDNA samples show an overwhelming majority of EYP ancestry.

In the Copper Age, one sample of the Bolshemys culture (ca. 4th millennium BC) and three individuals from the succeeding Afanasievo culture (ca. 3000 BC) were found to belong to MPCS lineages, and three more to EWE lineages (Holland 2014), which more clearly points to the expansion of Yamna migrants (Figure 8). A division between a Mongol Altai and a Siberian area of Indo-European influence was clearly seen in ancient DNA samples from a later period, where only one sample from Okunevo culture (ca. 2300-1800 BC) was of MPCS ancestry, the rest from Asian hunter-gatherer ancestry – and another probably related to Siberian hunter-gatherer –, all of which suggests a change in the genetic pool of the paternal line in the region (Holland 2014).

Affinity with a phylogenetic network created in the study by Holland (2014) placed males of MPCS ancestry most likely within the LPCS lineages.

However, the main lineages found in Tarim Basin mummies of the Xiaohe necropolis (dated ca. 2000 BC), 11 out of 12 male remains, were of EEHG (Li et al. 2010). If these were actually ancestors of later Tocharian speakers, this could support a cultural assimilation of Pre-Tocharian into peoples of EEHG ancestry, and also that the migration of the Pre-Tocharian language to the Tarim Basin was coincident with the Andronovo expansion, which “had transformed the steppes from a series of isolated cultural ponds to
a corridor of communication” (Anthony 2007). That change is clearly attested in aDNA samples from the final Bronze Age, where no WE ancestry is found, and one sample of ACW ancestry is found in the area. This period marks also the appearance of Paleo-Siberian lineages in south Siberia (Holland 2014).

From all paternal lineages found in modern Uyghurs (Zhong et al. 2013), only that of MPCS ancestry cannot be explained by earlier or later population expansions.
Figure 8. Diachronic map of migrations in Asia ca. 3000-2800 BC.
IV.7. Late Indo-European expansion

Anthony’s third migration wave of ca. 3000–2800 BC (Anthony 2013) must include the expansion of Eastern and Western Yamna pastoralists into Europe.

The most obvious material division within the early Yamna horizon was between east and west (see above Figure 7). According to forming and TMRCA dates of EYP subclades, communities with different EYP ancestry might have already developed different communities within the Volga–Ural–North Caucasian zone, a part of the more mobile eastern Yamna pastoral economy (Anthony 2007).

The remaining North-West Indo-European community – separated from Pre-Tocharian speakers – lived more likely around the South Bug – Lower Don steppe, and it is possible that their genetic ancestry was already dominated by WYP lineages. It expanded west possibly early within the southern stream of the third migration wave, from the Bug-Dnieper-Azov steppes into the lower Danube valley and Bulgaria, pushing farther up the Danube to the middle Danube valley in eastern Hungary through an Old Europe in crisis (Figure 8) – contemporary with late Baden / Cernavodă III (Anthony 2013; Anthony 2007).

This period represents the end of the second horizon of the Corded Ware culture, and the last corded pottery in the narrower territory of the central Balkans, but steppe elements are found spread into central and southern Balkans, and corded ware is found in the third horizon into southern Balkan groups and into Globular Amphorae, Bell Beaker and Corded Ware cultures to the north (Bulatović 2014). The spread of the Corded Ware culture in Northern Europe has been connected to the expansion of IE dialects via the northern route and through the Coţofeni culture, and more precisely through Usatovo (Anthony 2007; Gimbutas 1977); however, no ancient population exchange of WYP or EYP has been found in this area, and neither has LPCS ancestry been found in early Corded Ware cultures.

The expansion of peoples of WYP ancestry must have happened later, during the migration along the Danube, in the Yamna settlements in Central Europe, only some generations before the time when earliest West European pastoralist (WEP) paternal lines appeared and simultaneously spread – the Northern WEP, ca. 2900 BC, and the Southern WEP, ca. 2900 – 2800 BC.
Figure 9. Diachronic map of migrations in Europe ca. 3000-2250 BC
IV.7.1. The western migration path

Yamnaya migrations spread rapidly into the Danube valley, beginning ca. 3100 and continuing up to ca. 2800 BC (Figure 8). Probably the earliest settlement appeared in Varna bay in Bulgaria. A large settlement appeared to the west in southwestern Romania divided by the Danube river (Tarnava-Rast group). Migrants pushed west, appearing west of the Iron Gates in Jabuke, but the largest number of migrants ended up in East Hungarian settlements. Another settlement appeared south of the Varna bay, in the Balkan uplands (Kovachevo-Troyanovo), within the Ezero culture (Anthony 2007).

This expansion of Yamna into Late Neolithic / Bronze Age central European population has been shown to be strongly male-biased, suggesting 14 migrating males for every migrating female (Goldberg et al. 2017). This is in contrast with the migrations from Anatolia during the spread of farming, which has not been found to be sex-biased, and suggests a different sociocultural phenomenon.

Two alternative migration paths have been proposed for the western expansion of WYP lineages. A northern path, following the Corded Ware culture expansion west, has been proposed based on the analysis of modern populations (Valverde et al. 2016), and could be supported by the presence of an individual with EPCS ancestry in a Corded Ware burial (ca. 2865-2578 BC) from Oblaczkowo (Allentoft et al. 2015).

A southern path, compatible with Anthony’s (2007) account, may be supported the finding of WE ancestry in Lánycsők (ca. 2860-2620 BC) of an unidentified culture (Szecsenyi-Nagy et al. 2015). Stronger support is probably to be found in the modern distribution of basal WYP ancestry (not belonging to any of the most common subclades) in Central Europe, which suggests a western migration of WYP along the Danube (Figure 10).

A rapid decline in human activities peaked in Central Europe between 4000–3000 BC and recovered only after 3000 BC, accelerating after 2500 BC. This decline has been related to adaptation processes during climatic changes (Kolář et al. 2016; Gardner 2002) – which might have helped the expansion of Yamna settlers into scarcely populated areas –, and it recovered after 3000 BC, accelerating after 2500 BC, which is compatible with the expansion of the horse, the wheel, and pastoralist societies into these areas.

Large stone anthropomorphic stelae seem to have first appeared in the Mikhailovka I culture in the second half of the 4th millennium. Mikhailovka I was later replaced by
Usatovo, and its culture continued in the Kemi-Oba culture of Crimea. Carved stone stelae appear to have expanded in frequency and elaboration in both territories, and in part of the North Pontic steppes, after about 3300 BC (Anthony 2007), coinciding with the western expansion of Repin societies within Yamna. Strikingly similar stone stelae appeared later in the Caucasus, Troy, and also in Central and Western Europe, and with special frequency in the Swiss Alps and in the Provence. A maritime route for such cultural expansion has been proposed, which would justify e.g. its early presence in Troy (Anthony 2007).

Mainly associated with funerary customs in the Yamna horizon, the use of other carved anthropomorphic stones (like figurines) in some local traditions of Western Europe seem to predate that of the Yamna horizon. Also, large stone stelae might have been used quite differently, or for different purposes, in certain local cultures (Robb 2009).

It is therefore speculative to assign any value to the potential terrestrial expansion of stone stelae, but the time of appearance of similarly carved large slabs of anthropomorphic stones in Europe, as well as its association with funerary rites – especially burial sites –, might point to a link with the expansion of certain communities of WYP ancestry from the South Bug – Lower Don steppe (and their traditions) into Western Europe, and the southern route would therefore be more likely.

East Hungarian Yamna settlers would then have migrated west into central Europe. WYP lineages in Europe are mainly found in Bell Beaker peoples, with a sample of WYP not-Northern-WEP ancestry (i.e. probably South WEP) found in a Bell Beaker burial in Kromsdorf ca. 2600-2500 BC (Lee et al. 2012), and later South WYP subclades found exclusively in Bell Beaker groups.

Together, these data suggest an expansion of North-West Indo-European – separated from the Tocharian branch – into Central European Bell Beaker groups, partially coincident with the contact zones of Bell Beaker and Corded Ware cultures. Similar economic practices – the continuation of an agricultural way of life – are found in Central Europe, with no difference in mobility between Bell Beaker and Corded Ware folk (Sjogren, Price, and Kristiansen 2016), which suggest an adaptation of both migrant groups to the newly settled areas.
IV.7.2. The expansion of Balkan Indo-European

Later examples of EYP subclades suggest a genetic continuity of Early Khvalynsk population in the pastoral groups within Volga–Ural–North Caucasian steppe: Early Yamna in Ekaterinovka ca. 2910-2875, Late Yamna in Temrta IV ca. 2887-2634, and two samples from the Poltavka culture in Kutuluk ca. 2867-2484, and in Lopatino II ca. 2885-2665 (Mathieson et al. 2015).

Eastern groups – if only because of their location within Yamna – may have followed the southern stream somehow later compared to WYP, possibly through the settlement of Kovachevo-Troyanovo.

EYP lineages appear in modern populations as a Balkan group, an Anatolian group distributed in the Balkans and Anatolia, an Armenian group found mainly in the Armenian highlands, and a Central group spread in Greek and in Central Europe (Figure 11).
The modern distribution of EYP subclades around the Balkans gives therefore support to the existence of a Paleo-Balkan Sprachbund or dialect continuum (Kortlandt 2003).

**IV.7.2.1. Greek**

The southern route of the southern stream of Anthony’s third migration wave has been described as the later expansion of the Yamna culture from the Lower Danube and Tisza rivers to the southern tip of the Balkan peninsula, where its population are probably represented by the transition of Early Helladic II to Early Helladic III period ca. 2200 BC (Gimbutas 1977), coincident with the arrival of Minyan pottery style. Others have proposed a later date, the beginning of the Middle Helladic culture ca. 2000-1900 BC (Beekes 2011).

The complexity of Y-DNA haplogroups found in the modern population of Greece bears witness to the thousands of years of European and Asian interaction in the formation of its peoples. While it is clear that Yamna ancestry does not represent a big part of its modern population in any study published to date, it is unclear how the ancient population was affected by the migration of peoples of EYP lineages.

Analysis of modern Greek and Cretan lineages point to a Neolithic expansion of MPCS ancestry in the region, which were found nearer to Italian than to Balkan lineages (which in turn might be related to Central EYP), but no subclades were given in the study (King et al. 2008). Analysis of Greek-Cypriot modern populations revealed the presence of EYP lineages in easternmost and westernmost sides of the island, with Central EYP appearing only in the east (Voskarides et al. 2016). The early attestation of Mycenaean Greek in the island point to an early expansion of Central EYP, but the early Anatolian influence over the island preclude a precise identification of their origin.

The potential invasion and assimilation of EYP settlers in Proto-Anatolian-speaking populations in Cernavodă III and Ezero cultures further confounds modern genetic studies, and aDNA samples are needed to more clearly depict the expansion of both populations. Such a contact may have happened early during the southward migration, as suggested by Anatolian loanwords found in Greek.

**IV.7.2.2. Other Balkan languages**

The language ancestral to Armenian is – like Phrygian – believed to have belonged to the peoples that came from the west and overran the Hittite empire in the 12th century BC (Beekes 2011). The language ancestral to Albanian, sometimes identified with Illyrian,
should therefore be also originally located in the Balkans early during the west migration of Balkan Indo-European.

Both the Albanian and Armenian languages are spoken by modern populations where the majority of WE ancestry is formed by basal MPCS and LPCS ancestry, which point to a resurge of a Proto-Anatolian genetic component (together with European hunter-gatherer ancestry) after the expansion of EYP groups.

In the case of Armenian, this has been explained by a history of genetic isolation from their surroundings (Haber et al. 2016). The oldest sample found in the region is dated ca. 2619-2465 from Kura-Araxes, from EPCS (not MPCS) ancestry (Lazaridis et al. 2016)¹, suggesting the presence of previous EPCS lineages in the region, unrelated to the later Armenian migration.

Also, populations of the western part of the Armenian Highland, Van, Turkey, and Lebanon show genetic affinity with European populations, and their absence in previous studies “should be considered a consequence of the absence in their Armenian datasets of populations from the western region of the Armenian highland” (Hovhannisyan et al. 2014), which is also hindered by the loss of data in modern populations due to the effects of the Armenian Genocide.

¹ Lazaridis, Twitter, 18 juni 2016: "I1635 (Armenia_EBA) is R1b1-M415(xM269). We'll be sure to include in the revision. Thanks to the person who noticed! #ILovePreprints."
Figure 11. Modern distribution of EYP ancestry (R1b-Z2103). Adapted from Myres et al. (2011).
IV.7.3. The expansion of North-West Indo-European in Europe

IV.7.3.1. Bell Beaker

The Bell Beaker phenomenon is defined by groups that show a common know-how in technology, especially regarding pottery, copper metallurgy (Amzallag 2009), and flint. No single unified network of know-how transmission can be reconstructed, only local or regional networks (Linden 2015), but despite this a supra-local homogeneity can be observed in the whole of Europe from 2500 BC “in similar funerary rituals, in the way of interacting with territory, in the way of representing iconography and decorating pottery, and in the way of representing social differences” (Martínez and Salanova 2015). The Bell Beaker phenomenon made thus the previous regional networks of Western Europe uniform with identical social codes.

The Bell Beaker migrations (ca. 2500-2100 BC) between Central Europe, the British Isles, and the Iberian Peninsula have been associated with the expansion over central and western Europe of Yamna migrants through the Vučedol complex in the Carpathian basin, with specific correspondences found in burial rites, armament, costume, ornaments, technology in general, and also in ranked society, funerary rites, belief in life after death, and in general symbolism (Gimbutas 1993). Bell Beaker sites of Csepel around Budapest, west of the East Hungarian settlement region of Yamna migrants, and dated ca. 2800-2600, could have been a bridge between Yamna on the east and Moravia and Bavaria to the west, through which Yamna dialects spread to Southern Germany, where decorated cup styles, domestic pot types, and grave dagger types from the Middle Danube were adopted ca. 2600 BC (Anthony 2007).

Central WEP split ca. 2600 BC into its main subclades Alpine, Gallic, and Atlantic. The first individuals of Central WEP ancestry attested are both of a Bell Beaker culture, one of Gallic WEP ancestry (ca. 2431-2150 BC) in Quedlingburg (Lazaridis et al. 2016), and one of Alpine WEP ancestry in Osterhofen (Allentoft et al. 2015). The presence of individuals of Atlantic WEP ancestry (ca. 2026-1885 BC and ca. 2024-1741 BC) in Rathlin, Ireland (Cassidy et al. 2016), suggests that the East-West migration of pastoralists continued for centuries and reached the westernmost areas within the third millennium.

While the regional substrate for eastern and northern Bell Beaker groups is in many cases formed by Late Corded Ware culture groups – with some pottery types persisting in later
times, and with individual burials being also used by later settlers—, in southern and western Bell Beaker territory previous regional substrates do not herald the Bell Beaker groups, with newer settlements using locations different to Late Neolithic sites, and collective graves being reused or substituted by individual graves (Besse 2014).

Demic diffusion of Atlantic WEP ancestry (Figure 12) accompanying Bell Beaker expansion in the British Isles is supported by ancient DNA analysis (Cassidy et al. 2016). Recent research supports a considerable degree of mobility with little difference between male and female migration in Britain (Parker Pearson et al. 2016), which do not support exchange of female marriage partners (Brodie 2001) or inter-cultural contact consolidation (Vander Linden 2007), as previously proposed.

The presence of Gallic WEP ancestry in the Bell Beaker group of Central Europe might be linked to this lineage’s expansion to the west and south during the Bell Beaker phenomenon (Figure 13).

Studies of ancient Indo-European hydronymy (Krahe 1964; Krahe 1949; Nicolaisen 1957) have revealed a quasi-uniform name-giving system for water courses that shows Indo-European water-words and suffixes following rules of Late Indo-European word formation (Adrados 1998), pointing to an ancient wave of Late Indo-European speakers spread over Western and Central Europe before the Celtic and Germanic expansions, including the British Isles, the Italian and Iberian peninsulas.

The expansion of both the Gallic and Atlantic WEP groups could then be linked to the first introduction of Indo-European languages in Western Europe (Cassidy et al. 2016), which could have left some traces of their existence. However, the later expansion of Celtic languages, and an apparent resurgence of the probably indigenous Proto-Iberian and Proto-Basque languages – possibly the descendant of the languages of early farmers, similar to Paleo-Sardinian (Terradas et al. 2014) – over an expanding Iberian branch of the Gallic WEB ancestry (Gunther et al. 2015) have left scarce data on the older situation. The only certain Indo-European language of Iberia that can be considered of a non-Celtic nature is Lusitanian (which has been linked to a potential Galaico-Lusitanian group of the north-western Iberian Peninsula), and there has been some discussion on the pre-Celtic nature of the languages of Cantabri, Astures, Pellendones, Carpetani, and Vettones. Also,
while the position of Tartessian as Indo-European (Koch 2009) is highly doubted\(^2\), there is some support for a borrowing of names from a “lost Indo-European language” over the course of long-term contacts (Mikhailova 2015).

The other region where modern Gallic ancestry peaks in the modern population corresponds to the old Nordwestblock cultural region, where a non-Celtic, non-Germanic Indo-European language might have been spoken (Kuhn, Hachmann, and Kossack 1986). The contended Pre-Celtic Irish and Pictish substrates might have been of Late Indo-European nature, imported by peoples of Atlantic WEP ancestry.

\(^2\) It was criticized extensively in a special section of Vol. 42 of The Journal of Indo-European Studies (No. 3 & 4, Fall/Winter 2014)
Figure 13. Modern distribution of Gallic WEP ancestry (R1b-DF27). Adapted from Myres et al. (2011)
**IV.7.3.2. Únětice culture**

The Únětice culture (ca. 2300-1700 BC) has been cited as a pan-European cultural phenomenon (Kristiansen and Larsson 2005), whose influence covered large areas due to intensive exchange (Pokutta 2013), with Únětice pottery and bronze artifacts found from Ireland to Scandinavia, the Italian Peninsula, and the Balkans (Figure 14). As such, it is the best candidate for the last community of a West Indo-European language ancestral to Italic, Celtic, Germanic, and perhaps to Balto-Slavic (Beekes 2011).

Thought to have evolved from Bell Beaker cultures, the scarce ancient DNA analyzed is however formed by four individuals from pre-Yamna lineages: one individual from Corded Ware/Proto-Únětice culture at Łęki Małe (ca. 2286-2048), and three from Únětice proper, one ca. 2137-1965 BC from Eulau, and two from Erperstedt dated ca. 2131-1979 BC and 2134-1939 BC (Mathieson et al. 2015). On its periphery MPCS lineages are found in Gata/Wieselburg (ca. 1770-1760 BC) and Untermeitingen (Allentoft et al. 2015).

Ancient DNA samples suggest at least a partial resurge of hunter-gatherer ancestry in Únětice, although only a slightly lesser genetic affinity to Yamna than in Bell Beaker groups (Haak et al. 2015). Úněticean genetic melting pot strengthens its origin as the vector of cultural diffusion of West Indo-European – a late dialect of the North-West Indo-European spoken by Bell Beaker groups – into Late Corded Ware and other groups of central Europe.

Differences found in West Indo-European dialects in the last phase of the laryngeal loss (Quiles 2012) points to a time coincident with the spread of Únětice culture.

Bell Beakers and early Únětice represented the first prospectors and metallurgists, travelling and sharing their skills, with Adlerberg and Straubing cultures being small local centers (Kristiansen 1987).
Figure 14. Diachronic map of migrations in Europe ca. 2250-1750 BC.
IV.7.3.3. Nordic Bronze Age

In Scandinavia, a migration of Bell Beaker groups to Jutland (ca. 2300-1700) seems to have brought skills in mining and sailing, introducing mass production of flint daggers, as well as the first metal daggers (an imitation of copper and bronze prototypes). This Dagger Period of the Late Nordic Neolithic also represents the introduction of a more ranked social organization, and large chiefly houses similar to Únětice appeared in south Scandinavia, and point to “a radical reorganization of economy and social organization which wiped out or integrated the diverse cultural traditions of the previous period into a single south Scandinavian cultural sphere”, which suggests “the formation of a shared Nordic language based upon the frequent interaction that followed from the distribution of flint daggers” (Kristiansen 2009).

The best candidate for an expansion of the Pre-Germanic dialect of West Indo-European into Scandinavia from Únětice is the Barbed Wire Beaker culture of the Low Countries and Northern Lowland, which would later show a period of change (Figure 15) starting ca. 1850 until its complete cultural change evident after ca. 1500 BC (Fokkens and Harding 2013), into the Elp culture (ca. 1800-800 BC). These cultures are situated close to areas where Northern WEP ancestry forms the majority of the present genetic make-up.

The expansion of Germanic is thus to be linked to the expansion of Northern WEP lineages in southern Scandinavia. Investigation of aDNA shows that the first individual with Northern WEP ancestry is found in Lilla Beddings (2275-2032 BC), belonging to either the Battle Axe group of the Corded Ware culture or to the Nordic Bronze Age culture (Allentoft et al. 2015). While the date is compatible with the expansion of Beaker groups into Southern Scandinavia – and their early contacts along the Øresund strait –, it is a member of an extinct subclade not found in modern populations.

West European hunter-gatherer (WEHG) ancestry was formed ca. 25000 BC, and the modern European population has a TMRCA ca. 21600 BC. Many ancient DNA samples are found since the Paleolithic, and two main branches – a Balkan and a Central European ones – seem to have divided early, the former associated with the Balkans, the latter apparently following the expansion of Italo-Celtic and Germanic, and therefore possibly integrated with WEP ancestry since Únětice. Samples from Sweden since the Mesolithic show exclusively samples of WEHG ancestry before the Bronze Age.
Northern European hunter-gatherer (NEHG) ancestry was formed at the same time as WEHG. The first example is found in Neolithic Linear Pottery culture in Hungary (Szecsenyi-Nagy et al. 2015), which suggests its distribution in central Europe before the Corded Ware and Yamna expansion. The next sample found in aDNA records is in the Nordic Bronze Age in Angmollan, ca. 1493-1302 BC (Allentoft et al. 2015).

MPCS ancestry in Denmark from Early (1499-1324 BC) and Late Nordic Bronze Age (794-547 BC), and European hunter-gatherer ancestry from Sweden (Angmollan ca. 1493-1302 BC, ca. 1432-1292; Abekas ca. 1395-1132), point to the suggested south-north culture and population cline in Scandinavia during these times of ethnolinguistic change.

The modern population with NEHG ancestry is centered in northern Scandinavia near Skagerrak strait and Kattegat sea area, and shows a TMRCA ca. 2600 BC. On the other hand, Northern European Corded Ware ancestry (formed ca. 2700 BC, TMRCA ca. 2300) distribution in modern Scandinavian populations is located further to the north and west of that zone. If invasions from southern to northern Scandinavia are supposed to have happened in a south-to-north route, through the Øresund strait into the Skåne region, NEHG ancestry should be assumed to have migrated during the proposed expansion of Barbed Wire Beaker peoples into Scandinavia, and thrived after that.

Whether both populations, of Northern WEP and NEHG ancestry, were already mixed in northern Germany before their northern migration into Jutland, or remained separated is difficult to assess, but it seems likely that in a common NWEP-NEHG society with a common language its diffusion north might have happened through different clans that might have already integrated different lineages, and that later founder effects simplified to a greater extent the situation east and west of the Øresund strait.

After an obscure period of internal development, the situation in Northern Germany and Scandinavia before the Iron Age would have probably corresponded loosely to the present situation, with the NWEP-NEHG divide possibly located to the east of the current cline, at the Øresund strait, given the quite late invasion of Jutland by Danes.

The irruption of Germanic peoples in central, eastern, and western Europe including the Roman Empire – the Barbarian Invasions from Classical sources, renamed the Migration Period since the Romantic era – suggest a NWEP-dominated West Germanic area, and Viking migrations point to a NEHG- and NECW-dominated North Germanic area (see below Figure 21).
The modern distribution of NWEP (Figure 16) is roughly coincident with the expansion of West Germanic with the medieval Ostsiedlung, showing a west-east cline compatible with the Germanization of Slavs to the east of the Elbe. Although modern population samples are difficult to assess without genealogical information – due to the expulsion of Germans after World War II –, medieval samples from Podlažice (ca. 1180 AD) in Czech lands and Nicolaus Copernicus’ (Bogdanowicz et al. 2009) family origin from Koperniki near Nysa in Silesia before the 14th century seem to support the expansion of NWEP lineages associated with German settlers of the Holy Roman Empire (see below Figure 30).

The question of the dialectal nature of East Germanic remains a purely linguistic one, but NEHG and NECW lineages scattered throughout Europe seem to support the classical description of East Germanic tribes migrating from Scandinavia to the east of the Elbe, and thus its connection with the Nordic branch.
Figure 15. Diachronic map of migrations in Europe ca. 1750-1250 BC.
Figure 16. Modern distribution of Northern WEP ancestry (R1b-U106). Adapted from Myres et al. (2011).
IV.7.3.4. Tumulus Culture

It was only after 2000 BC that large-scale mining operations and production which required specialized metallurgical and organizational know-how began in a few centers, and they reached distant regions as far as Northern Scandinavia. And only from 1750/1700 BC began the actual Pan-European tradition of metal work until its consolidation in 1600 BC, with different regions in Europe producing their own products, most specially the cultures of the Carpathian basin (Kristiansen and Larsson 2005).

The contacts of Únětice with Carpathian territories are constant, e.g. in the Únětice-Nitra and Únětice-Hatva horizons, where settlement micro-regions and relationships are difficult to assess. Únětice elites controlled trade routes from the Baltic Sea shores to Aegean Sea artisans, with Úněticean daggers found all over Europe and in Anatolia, and the nature of weapons and metal work suggest a chronic state of warfare and the emergence of a warrior class until its demise by the Tumulus culture, born in the area previously occupied by Únětice groups in Southern Germany.

Central European groups from southern Germany would then in this context correspond to a community with a common West Indo-European language ancestral to Italic and Celtic (Kortlandt 2007), whose continuous development and dialectal development is to be followed into the Tumulus culture (ca. 1600-1200 BC). The Tumulus culture was eminently a warrior society which expanded eastward with new chiefdoms east into the Carpathian Basin (up to the river Tisza) and northward into Polish and central European and Únětice territories, with dispersed settlements centered around fortified structures (see above Figure 15).

In the subsequent period of crisis, it developed into bands of raiders and mercenaries, and took control of peasant societies, as happened in several regions during the Urnfield and La Tène periods, and similar to the society of mercenaries and warring city states in the Celtic period (Kristiansen 2000). The European world ca. 1450-1100 BC has been compared to the Viking Age, with population pressure and lack of land for young sons with no inheritance leading to war-bands that engage in seasonal raiding, trading, and piracy; followed up by more massive colonizing ventures and migrations; and a political economy based on a chiefdom form of society where free farmers were the dominant class, with commoners and slaves as dependent groups (Kristiansen 2016).
Exogamous and endogamous strategies and variable distances of marriage exchanges to maintain alliances complicate this picture further in the Tumulus/Nordic traditions (Kristiansen 2000), which suggests that post-Únětice societies could especially benefit from more aDNA samples and the application of admixture analysis.

The Urnfield culture (ca. 1300-750 BC) is associated with the rise of a new warrior elite, and the formation of new farming settlements and their urnfields. In some areas there is continuity from Tumulus to Urnfield culture, with narrowing and concentration of settlements along the river valleys, but there is also wide-ranging migrations (Figure 17). These migrations are similar to those seen later in La Tène culture (Kristiansen 2000). Urnfield migrations south of the Pyrenees may have brought the pre-Celtic Sorothaptic language believed to be behind certain toponyms and inscriptions around the Pyrenees (Coromines 1976).

Scarce aDNA from Late Urnfield populations from its northeastern territories in Saxony – near the Lusatian culture –, show a mixture of lineages, which suggest genetic continuity with older cultures: ECW ancestry was found in Halberstadt (ca. 1193-979), and of the eight males studied from the Lichtenstein cave (ca. 1000 BC), five were of European hunter-gatherer, two of EEHG, and one of WE ancestry (Schilz 2006).

Given the modern distribution of Alpine WEP ancestry (see below Figure 18), its expansion is probably to be connected to the spread of the Urnfield culture and later offshoots Hallstatt and Villanovan cultures.
Figure 17. Diachronic map of migrations in Europe ca. 1250-750 BC.
IV.7.3.4.1. Celtic

From the early Urnfield culture expanded the Hallstatt culture (ca. 1200-500 BC), associated with Proto-Celtic (Chadwick 1970). The later expansion of La Tène culture (ca. 500-1st c. BC) from certain core Hallstatt regions – valleys of Marne and Moselle and neighboring Rhineland in the west, and a Moravian zone in the east – has been linked to the spread of Celtic languages (Figure 19). However, the Mainz research project of bio-archaeometric identification of mobility has not proven to date a mass migration of Celtic peoples in central Europe ca. 4th-3rd centuries BC, i.e. precisely in a period where textual evidence informs of large migratory movements (Scheeres 2014). La Tène material culture points to far-reaching inter-regional contacts and cultural transfers (Burmeister 2016).

Alpine WEP lineages are found today (Figure 18), scattered to the north, south, and west of the Alps, reaching the southwest corner of the Iberian Peninsula, and the British Isles. It seems to peak around the current borders between Italy, France, and Switzerland.

The expansion of Central European Neolithic (CEN) ancestry – stemming from WEHG –, assumed to be already mixed with Alpine WEP ancestry since at least the Tumulus culture, is found from the British isles to Turkey, with a West CEN ancestry concentrated in Great Britain (with mutational divergence suggesting its foundation ca. 300 BC) providing “some tentative evidence of ancient flow with eastern areas that could support the idea that the La Tène culture was accompanied by some migration” (McEvoy and Bradley 2010).

The scarcity of Alpine WEP and CEN ancestry in the modern populations of the British Isles and Iberia – where Celtic languages had clearly spread by the time of the Roman invasion – appear to suggest a successful cultural diffusion of the language from warring Celtic minorities who established new chiefdoms throughout Europe. It is also possible that a previous admixture of Atlantic and Gallic WEP ancestry in the expanding Celtic population further confounds the genetic change associated with the Celtic expansion.

Lacking ancient DNA samples and more complex genetic analyses, it may be assumed from the available data that the sociocultural phenomenon associated with the expansion of La Tène culture (and Celtic-speaking peoples) is different from the Neolithic expansion of farming – where Anatolian ancestry spread slowly with technology –, and also from the Bronze Age expansion of herding – where male-dominated Indo-European-speaking
groups of Yamna ancestry spread rapidly into western Europe and partially replaced or displaced the original population.

Figure 18. Modern distribution of Alpine WEP ancestry (R1b-DF27). Adapted from Myres et al. (2011).
Figure 19. Diachronic map of migrations in Europe ca. 750-250 BC.
IV.7.3.4.2. Italic

A spread of early eastern Urnfield cultures from Transdanubia is attested in the Po Valley (Váczi 2013).

The Villanovan (ca. 1100-700) culture, expanded from early Urnfield, has long been associated with Proto-Italic (Gimbutas 1965). However, the association of Villanovan withItalic remains controversial, since Villanovan territory is partially coincident with the later Etruscan-speaking zone, and no clear cultural break is seen between both cultures. However, a resurge of a previous language – akin to the Basque example – might explain the cultural continuity in Etruria.

Genetic analysis of the modern population show a spread of Alpine WEP ancestry south of the Alps including north and center Italy, which supports the invasion of this group from the north, through the Alps. However, as with Greek – and even more so – the complexity of the current Y-DNA maps of the region mirrors the Italian Peninsula’s historical development since the Paleolithic.

Classical sources since the Aeneid – Virgil’s epic poem linking the foundation of Rome to the flight of Aeneas from the Troy of Homer’s Iliad – have possibly influenced the rejection of famous Italian linguists of the unity between Latin and Osco-Umbrian. While both branches share common innovations, and it is therefore difficult today to reject a shared community by relating all differences to recent contacts, some linguists have tried to reconcile the obvious Italic nature of Latin and its morphological differences compared with Osco-Umbrian with a potential late Anatolian substratum, and have thus supported an eastern invasion through Apulia.

To further complicate the linguistic and archaeological discussion around Latin, a recent Anatolian connection has been found by examining mtDNA in modern populations of present day Tuscany (Brisighelli et al. 2009).

The expansion of Rome (Figure 20) seems not to have been accompanied by a massive migration of peoples, and cultural diffusion is likely to have played a bigger role in the expansion of Latin.
Figure 20. Diachronic map of migrations in Europe ca. 250 BC – 250 AD.
IV.7.3.4.3. Italo-Celto-Germanic and Celto-Germanic contacts

The initial phase of the Elp culture (1800-1200) in the Low Countries is characterized by tumuli related to the Tumulus culture and later to the Urnfield culture, with Dutch-German lowland areas found to share cultural roots with the southern Scandinavian area (Butler, Arnoldussen, and Steegstra 2012) predating technologic and economic exchanges between Urnfield and Northern Bronze Age Scandinavia (Kristiansen and Suchowska-Ducke 2015).

These complicated cultural-economic networks (see above Figure 15 and Figure 17) that preclude precise ethnic (and thus linguistic) differentiation supports the maintenance of late contacts between the languages ancestral to Germanic and Celtic, assuming a position of Proto-Celtic to the north of the Hallstatt culture – as supported by the known homelands of La Tène culture.

A careful analysis of CEN ancestry – found today distributed among Germanic and Italo-Celtic territories – might bring light to population movements and exchanges during the Bronze Age and the Iron Age in Europe.
IV.8. Indo-European in Corded Ware societies

IV.8.1. Previous models

Some Late Yamna and Usatovo migrants seem to have penetrated northward up the Dniester, South Bug and Dnieper valleys, and Globular Amphorae and Corded Ware expansion is noted to the east, forming the Middle Dnieper culture in the forest-steppe around Kiev, beginning ca. 2800-2600 BC (Anthony 2007).

Migrating Yamna pastoralists into already expanding Corded Ware groups (Wencel 2015) might have created the necessary environment for the spread of Indo-European languages, as suggested by previous mainstream models for Indo-European expansion (Gimbutas 1977).

The admixture of Yamna aDNA samples found elevated (up to 76%) in Corded Ware samples has been said to support the migration of Yamna populations into Corded Ware groups, while the lower percentage found in Bell Beaker and Únětice groups was explained by a subsequent, less profound displacement process triggered by western and central European groups (Haak et al. 2015; Allentoft et al. 2015; Mathieson et al. 2015). According to Anthony (2007), this admixture might have happened during interactions precisely through the contact zone north of the Usatovo culture.

However, data from individuals of early Pontic-Caspian Steppe and Eastern European hunter-gatherer communities (probably up to the Paleolithic Swiderian culture territory) are shown to be already similar before the Neolithic expansion\(^3\), and it is therefore more likely that genetic admixture between neighboring populations of WE (Pontic-Caspian steppe) and EE (Forest Zone) ancestry had occurred during previous Mesolithic and Neolithic contacts, including likely raids, intermarriage and minor migrations, as witnessed in the Mesolithic–Neolithic transition with samples of EPCS ancestry in the Narva culture in Latvia ca. 5791-5586, and of EEHG ancestry in the Dnieper-Donets culture ca. 4469-4293 BC (Jones et al. 2017). One of the oldest samples of Corded Ware burials is of EPCS ancestry (ca. 2865-2578 BC) from Oblaczkowo (Allentoft et al. 2015), probably suggesting the presence of this ancestry in some Corded Ware migrants as well as in the settled territories. To further complicate the situation, the absorption of Bug-Dniester and later the formation of the Usatovo culture by Old European cultures left

\(^3\) See e.g. Extended Data Figure 2 in Haak et al. (2014), Extended Data Table 2 in Mathieson et al. (2015), Figure 2 in Jones et al. (2017).
probably a mixed genetic make-up, in the most likely zone for interaction with developing Corded Ware groups.

Genetic affinity between such closely related groups have no doubt influenced genetic admixture results. In light of this similarity, the more revealing data from ancient DNA samples concern not the affinity found between Yamna and Corded Ware, but the one found between Yamna and Bell Beaker, and between Yamna and Únětice.

Adding to the controversy, mtDNA – which is also included in previous admixture analyses – in European samples posterior to the migration has been demonstrated to come overwhelmingly from previous European Neolithic / Bronze Age populations (Goldberg et al. 2017). The difference in mean X-to-autosomal ancestry ratios differ among groups, with larger male bias and lower European Neolithic ancestry in later Bell Beaker and Únětice, although the trend found was not significant. This may further increase the relative importance of the affinity found between the original Yamna population and the Bell Beaker and Únětice descendants of Yamna migrants, compared to Corded Ware descendants of Forest Zone migrants.

All these data suggest that findings from recent genetic admixture studies are far more significant in assessing the admixture of Yamna in Bell Beaker and Únětice populations of central and western Europe than in eastern European Corded Ware groups, a fact that has apparently been overlooked.

Central EYP ancestry in modern central European populations does not support an expansion beyond the contact zone to the north, but rather – if traced to the Yamna migrations – a path similar to the southern route followed by WYP lineages.

Therefore, no strong cultural or genetic link has been found to date for such a generalized assumption in Indo-European studies, that Corded Ware cultures was the (or a) vector for the expansion of Indo-European languages into Europe.

---

4 Data is available in SI Appendix, Fig. S1B for Corded Ware and Únětice, and Table S7 for Bell Beaker – which shows high inter-individual variability. Even being non-significant, Corded Ware samples seem to have a slightly earlier mean date – compatible with its earlier migration. An increasing X-to-autosomal ratio would be expected in the traditional model from Corded Ware to Únětice, while in the latter it is actually found decreased compared to the former, which doesn’t fit well with the interpretation of a resurgence of old lineages compared to Corded Ware (Haak et al. 2015). However, it fits well with a previous similar ratio found in Bell Beaker, although the scarcity of aDNA samples seem to preclude the significance and thus a generalization of any conclusions.
IV.8.2. Balto-Slavic

As already stated, samples from late Corded Ware cultures to the east of Únětice show a high percentage of EEHG ancestry in Bronze Age and modern population samples.

The expansion of ECW ancestry appears therefore to be strongly linked to the spread of the Corded Ware culture, but the original location of modern population ancestries (Figure 22) is unclear. Based only on current genetic mapping (Underhill et al. 2015), basal CW seems to be distributed following a westward and eastward expansion from a Pit-Comb Ware ancestral homeland; Northern ECW seems to have expanded early to Scandinavia and expanded later from a secondary nucleus there (given its later TMRCA); Western ECW appears in Central Europe, with diffusion to the East; whereas Eastern ECW seems to be centered on Eastern Europe.

Examples of EEHG and ECW ancestry from ancient populations show an original spread east- and westward from the late Pit-Comb Ware cultures – which support an original expansion of ECW subclades unrelated to the Pre-Balto-Slavic expansion (see below Figure 23).

The current distribution and older TMRCA of Central ECW compared to the other ECW lineages could support its position as the original Pre-Balto-Slavic-speaking population. Some late Corded Ware groups in central Europe are known to have smoothly transitioned to Bell Beaker cultures (Besse 2014), and some of these formed proto-Únětice groups. Some Úněticean groups later evolved into early Lusatian Tumulus culture, (ca. 1700-1400), originally located between the Elbe and Oder basins (see above Figure 15), which later expanded east (ca. 1300-500 BC) into territories of previous Trzciniec culture (see above Figure 17).

Diffusion of West Indo-European language and culture has already been proposed to be identified with the Úněticean expansion into peoples of different ancestry, and continuity of such admixture from the region of Tumulus–Early Lusatian into Urnfield–Lusatian cultures is supported by findings of ECW lineages in the Urnfield group from Saxony-Anhalt, close to the Lusatian cultural homeland. That points to the Elbe basin as the original site of cultural breach for ECW lineages, between the older Corded Ware tradition and the new Úněticean culture.

The east and western peaks in Central ECW ancestry might support a west-east migration, as well as an east-west migration, and indeed both in different periods, which is expected.
to be found if Lusatian is indeed linked to the expansion of Pre-Balto-Slavic, and later Central ECW ancestry is linked to the West Slavic expansion to the west.

The Pomeranian and related West-Baltic culture of cairns (ca. 650-150 BC) evolved from the Lusatian culture to the east, following the expansion of the Jastorf and Hallstatt/La Tène cultures (see above Figure 19). Under pressure from Germanic migrations to the south and east from Scandinavia and the German lowlands, represented by Oksywie (2nd c. BC – 1st c. AD) and later Wielbark (1st c. AD – 4th c. AD) in eastern Pomerania.

The Przeworsk culture (3rd c. BC – 5th c. AD) shows continuity in its roots with the preceding Pomeranian culture, but its extension north from the Vistula to the Oder, and south toward the middle Danube from the Dniester to the Tisza valley was accompanied by significant influences from La Tène and Jastorf cultures, and the subsequent absorption into the Wielbark culture – related to the East Germanic expansion – make its precise association controversial, and it is sometimes viewed as an amalgam of a series of localized cultures (see above Figure 20).

East of the Przeworsk zone was the Zarubinets culture (3rd c. BC – 2nd c. AD), considered a part of the Przeworsk complex (Mallory and Douglas 1997), located between the upper and middle Dnieper and Pripyat rivers. Early Slavic hydronyms are found in the area, and the prototypical examples of Prague-type pottery later originated there (Curta 2001). It is therefore to be identified as Proto-Slavic (Kobyliński 2005).

Zarubinets came to an end with the migration of its population, linked to the increasingly arid climate. By the 3rd century western parts of Zarubinets had been integrated into the Wielbark culture, and some Zarubinets groups had moved southward into river valleys, moving closer to Sarmatian and Thracian-Celtic groups of the Don region and forming the Chernoles culture. Central late Zarubinets sites gradually turned into the Kiev culture (ca. 3rd-5th c.), widely considered the first identifiable Slavic archaeological culture, from which Prague-Penkov-Kolochin complex of cultures – identified with the expansion of Proto-Slavic (Mallory and Douglas 1997) – descended about the 5th c. (Figure 21).

The division of historical Slavic tribes in territories and cultures in the 5th-7th centuries remains a hotly debated topic (Curta 2001). Nevertheless, the expansion of the Prague-Korchak culture from its nucleus in the older Przeworsk-Zarubinets contact zone to the west – including its expansion as the Sukow-Dziedzice group to the Baltic Sea – can be
identified with the West Slavic expansion, and – at least part of – the western peak of Western ECW ancestry in modern populations.

The Kolochin material culture was a transformation of the old Kiev culture (Kobyliński 2005), but evidence of Baltic river names in the region have made some propose an original Proto-Baltic population (Mallory and Douglas 1997) before the East Slavic occupation. Indeed the Baltic populations have been found to be genetically the closest to East Slavs (Kushniarevich et al. 2015), which is compatible with an originally Baltic-speaking Eastern ECW ancestry undergoing a cultural assimilation with the East Slavic expansion. A precise analysis of Finno-Ugric and Baltic populations would be necessary to discern which ECW subclades were associated with which population migrations and expansions.

The expansion of the Penkov culture in the Danube seems related to the expansion of South Slavic. Confusing accounts of the Byzantine Empire of the raids and migrations of a federation of tribes (the Antes and the Sklavenes) in their frontiers give a general idea of the complex interaction of different groups in the Balkans (Curta 2001), which might justify a late assimilation of the language by groups of Balkan Neolithic ancestry, which is prevalent today in South Slavic territory (Kushniarevich et al. 2015). However, apart from the main peak of Balkan Neolithic ancestry in the south-east Balkan territory, a secondary peak around Bessarabia, as well as its general distribution around the same territory as the Prague-Penkov-Kolochin complex probably point to an earlier assimilation of the group, during the transition to a Proto-Slavic community and before its migration.
Figure 21. Diachronic map of migrations in Europe ca. 250-750 AD.
Figure 22. Modern distribution of ECW ancestry (R1a-Z282). Adapted from Underhill et al. (2015).
IV.8.3. Indo-Iranian

During the western expansion of Yamna herders in the Bronze Age, the Fatyanovo group emerged early at the northeastern edge of the Middle Dnieper group, still showing mixed Corded Ware / Globular Amphorae traits, substituting the Volosovo culture and occupying the Volga-Kama region. Near it the Balanovo group seemed to be its metallurgical heartland (Anthony 2007).

In the forest-steppe zone of the middle Volga and upper Don, at the easternmost aspect of the Russian forest-zone, the last cultures descended from Corded Ware ceramic tradition, the Abashevo group, emerged ca. 2500 BC or later (Anthony 2007), substituting late Volosovo culture, and reaching the Upper Ural basin. Abashevo showed a mix of Fatyanovo/Balanovo and Catacomb/Poltavka culture (Figure 23).

Genetic make-up of modern populations show a distribution of basal ECW ancestry centered on the old territory of Middle-Dnieper – Fatyanovo – Abashevo groups (Underhill et al. 2015), and an example of ECW is found later in the Potapovka culture, in Utyevka ca. 2200-1900 BC (Mathieson et al. 2015).

Early Yamna continued in the Lower Don – North Caucasian steppe as the Catacomb culture, and in the Volga-Ural region as the Poltavka culture, where cultural continuity implies that a language from the Graeco-Aryan continuum – already separated from the Paleo-Balkan group – was spoken, i.e. the ancestor of Indo-Iranian. Herders from the Poltavka culture began to move to the Ural-Tobol steppes probably about 2800-2600 BC.

Coinciding with more arid climate after ca. 2500 BC, both Poltavka and Abashevo herders settled between the Tobol and Ural River valleys. The Poltavka outlier of ACW ancestry found in Potapovka, in the Samara region, ca. 2925-2536 BC (Mathieson et al. 2015), in a date only slightly later to its formation, points to a process of population expansion and probably also to intense early regional contacts between peoples of Abashevo and Poltavka cultures.

Cultures that emerged around 2100-1800 BC in the region – Sintashta in the Ural-Tobol steppes, and Potapovka in previous Poltavka territory – seemed to continue in an early phase the previous Abashevo tradition, but retained and gradually expanded many cultural traits of Poltavka pottery, followed the same burial rites, and settled on top of or incorporated older Poltavka settlements. “It is difficult to imagine that this was accidental.
A symbolic connection with old Poltavka clans must have guided these choices” (Anthony 2007).

Both Sintashta and Potapovka were born from a time of escalating conflict and competition between rival tribal groups in the northern steppes, where raiding must have been endemic, and intensified fighting led to the invention of the light chariot (Anthony 2007).

Ancient DNA samples of Sintashta, Potapovka, and later of the Srubna culture (Figure 24) show a substitution of EYP ancestry by Corded Ware lineages, and admixture analysis suggests similar genetic sources between peoples of Corded Ware and Sintashta (Allentoft et al. 2015). However, cultural continuity with Poltavka is not only seen archaeologically in material and symbolic culture, but is also evident from the association of the Sintashta expansion with Andronovo, and therefore with the later expansion of Indo-Iranian peoples and their languages. The most likely explanation for the eastern expansion of Indo-Iranian by peoples with ACW ancestry is therefore the assimilation by Sintashta groups of the Proto-Indo-Iranian language spoken by Poltavka herders.

The process by which this evident cultural assimilation happened, given the presupposed warring nature of their contacts, remains unclear. It is conceivable, in a region of highly fortified settlements, to think about alliances of different groups against each other, akin to the situation found in Bronze Age Europe: a minority of Abashevo chiefs and their families would dominate over certain fortified settlements and wage war against other, neighboring tribes. After a certain number of generations, the most successful settlements would have replaced the paternal lineages of the region, while the majority of the population in these settlements – females, commoners and slaves – retained the original Poltavka culture. Indeed, not only EYP ancestry was replaced in the region, but also all other EEHG and Corded Ware lineages, as demonstrated by the later expansion of ACW ancestry in Andronovo and Srubna cultures, and by present day EEHG ancestry in Asia.
Figure 23. Diachronic map of migrations in Asia ca. 2250-1750 BC.
Difficult to determine is the language spoken by peoples of the Srubna (Timber Grave) culture of ca. 1800-1200 BC – heirs of the Pokrovka complex of (ca. 1900-1750 BC) created by Potapovka and late Abashevo groups –, although later (probably Indo-European-speaking) Cimmerian or Thraco-Cimmerian groups might have emerged from this remaining group of Pontic-Caspian herders, since their relationship to Scytho-Sarmatian groups migrated from south Asia is unclear.

A comprehensive description of Sintashta-Petrovka expansion east as part of the Andronovo horizon in Asia – coinciding with the western expansion of the Seima-Turbino phenomenon to the Forest Zone – is given by Anthony (2007). Chariots were probably invented in the steppes, improving warfare and likely playing a big role in Indo-Iranian expansion within the Andronovo horizon (after ca. 1900 BC) and south from the Zeravshan valley into the Bactria-Margiana Archaeological Complex (after ca. 1800 BC), creating between 1800-1600 BC a post-BMAC culture dominated by Tazabagyab-Andronovo herders.

After about 1600 BC pastoral economies spread across Iran and into Baluchistan, and ca. 1500 BC Indo-Aryan chariot warriors invaded a Hurrian-speaking kingdom of Mitanni in north Syria. At the same time, as post-BMAC herders spread to the northeast Indian subcontinent, the Rig Veda was probably being composed (see Figure 25 to Figure 28).
Figure 24. Diachronic map of migrations in Asia ca. 1750-1250 BC.
Figure 25. Diachronic map of migrations in Asia ca. 1250-750 BC.
Figure 26. Diachronic map of migrations in Asia ca. 750-250 BC.
Figure 27. Diachronic map of migrations in Asia ca. 250 BC - 250 AD.
Figure 28. Diachronic map of migrations in Asia ca. 250-750 AD.
Modern distribution of ACW lineages (Figure 29) show a clear division between western and eastern subclades – with basal ACW located east of the Andronovo horizon – (Underhill et al. 2015). While the western subclade has an expected peak in the northern part of the Indian subcontinent – broadly coincident with the spread of Indo-Aryan languages –, the eastern subclade peaks at the core of the Proto-Iranian Yaz culture and East Iranian expansion (of languages related to old Bactrians, Sogdians, and Scytho-Sarmatian peoples). Its spread west of the Iranian Plateau, however, is complicated by its condition of place of transit of innumerable cultures and peoples in prehistoric and historic times – as is the case with the genetic make-up of southern Italian and Balkan peninsulas.
Figure 29. Modern distribution of ACW ancestry (R1a-Z93). Adapted from Underhill et al. (2015).
IV.8.4. A common Corded Ware substrate

It has been argued that similarities found in Indo-Iranian and Balto-Slavic languages – like the peculiar phonetic ruki development, a similar satem trend in both groups (Meier-Brügger 2003) – suggest a sort of west-east continuum between both languages, with certain features running through them (Mallory and Adams 2007).

Since both Únětice (ca. 2300-1600) and Sintashta (ca. 2100-1800 BC) potential language expansions into populations with Corded Ware ancestry happened at the same time, it could be argued that both communities happened to speak similar dialects that could have influenced both languages – a North-West Indo-European and a Graeco-Aryan dialect already developed quite differently – in a similar manner, and thus their similarities could be explained from a common language substrate, whether non-Indo-European, Pre-Indo-European, or even Indo-European.

It has been classically proposed that the Mesolithic language of the Narva and Combed Pit Ware cultures is to be identified with a Uralic community, and dates ca. 4000 have been proposed for the common reconstructible Proto-Uralic language (Parpola 2012; Kortlandt 2002). Finno-Ugric has also been shown to have developed in close contact with Proto-Indo-Iranian (Kallio 2002).

According to the theory presented in this paper, the EEHG population of the Combed Pit Ware culture expanded with the Corded Ware culture into western and eastern Europe, so it is likely that their language was indeed Uralic.

From a linguistic point of view, the characteristic palatalization of the consonant system in Proto-Uralic is compatible with the similarly transposed velar system adopted for Late Indo-European dialects by Balto-Slavic and Indo-Iranian speakers, thus explaining the strongest phonetic connection between these dialectally diverse Indo-European languages. Differences in Baltic and Slavic satemization process might point to an early split of the West Indo-European dialect ancestral to both, before or during its assimilation by different Uralic-speaking communities.

A similar western Corded Ware substratum could be argued to be the origin of certain common isoglosses found in Germanic and Balto-Slavic.

The study of precise isoglosses connecting these languages, and their potential relation to Uralic proto-languages lies beyond the scope of this paper.
Figure 30. Diachronic map of migrations ca. 750 – 1300 AD (continued in the next page).
V. Discussion

The core problem addressed by this paper has been the inconsistency found between the prevalent theories on migration routes and recent research on the genetic make-up of peoples from the Pontic-Caspian steppe.

The Indo-European demic diffusion model proposed here advances the theory that the expansion of Indo-European languages from the steppe was linked to the expansion of peoples with MPCS ancestry in Eurasia.

V.1. Consequences of the Indo-European demic diffusion model

There is a long-held assumption, since the kurgan hypothesis was laid out (Gimbutas 1963), that Corded Ware herders had helped spread Indo-European languages into Europe and Asia. This assumption is not fully explained by recent archaeological research (Anthony 2007; Anthony 2013), and recent findings in ancient human genetics question it on the grounds of a different path for human migration from the steppes.

To reject this old tenet has wide-ranging consequences:

- The natural trend of Indo-Europeanists to date Indo-European proto-languages all separated at the same time, and usually farther back in time than is warranted by the linguistic evidence (Kortlandt 1990) is challenged, offering a more naturally stepped separation. There is no need to place all known Indo-European branches simultaneously separated in a massive expansion into Corded Ware, Middle Dnieper, Bell Beaker, and Sintashta cultures (Anthony 2007; Anthony 2013).
- Balto-Slavic and Indo-Iranian similarities can be thought of as stemming from a common Corded Ware (or Eastern European hunter-gatherer) language substrate, with implications for the still prevalent three-dorsal reconstruction of PIE, which seems to be a resilient tradition from the early days of the centum–satem division.
of Indo-European, and has long been contested with sound linguistic arguments (Adrados, Bernabé, and Mendoza 2010; Quiles and López-Mencherero 2012). The association of this language substrate with Proto-Uralic offers an elegant explanation for these developments, and is supported by linguistic, archaeological, and now also ancient genetic data.

- A shared linguistic unity of Italic, Celtic, Germanic, and probably Balto-Slavic, is likely to have existed, probably at the same time as Indo-Iranian – or slightly earlier –, and both later than a potential Paleo-Balkan community.

- Pre-Germanic is more likely to have been imported into southern Scandinavia by peoples of mixed late Northern WEP / NEHG ancestry, marking the transition to the Nordic Bronze Age. The precise pre-West-IE linguistic landscape is unknown, but the previous arrival and likely expansion of peoples of ECW ancestry might have brought with them the language of Eastern European hunter-gatherers.

- An Italo-Celtic community is compatible with this expansion model, as is their close contact with a Pre-Germanic community, during the period of intense economic exchanges during the Bronze Age.

Some linguists have used the biological foundations of phylogenetics to extrapolate questionable methods to linguistics, and have thus obtained questionable results (Gray and Atkinson 2003). Similarly, scientists are using the available statistical means to study genetic admixture in modern human populations, extrapolating admixture mapping methods to scarce ancient human samples, and deriving far-fetched conclusions. This paper demonstrates the need to include wide anthropological investigation of the historical context of the samples studied, including linguistics, archaeology, and cultural anthropology, to obtain plausible explanations for the complex data obtained in human biology.

Even with the latest methods improving ancient DNA admixture analysis (Ni et al. 2016), a clear picture of genetic introgression would be necessary to establish which differences between compared groups are due to population inbreeding, and which to later adaptations (Quach et al. 2016), before being able to derive potential migration theories.
V.2. Problems with the demic diffusion model

V.2.1. Demic diffusion, cultural diffusion, or founder effect

Ethnos and language are intimately associated, and are known to be much more resistant to change than culture and social stratification, and thus changes in material culture are not to be equated to changes in language, even if ethnicity may take on new meanings (Kristiansen 2000).

Demic diffusion shows theoretically the simplest (and thus strongest) link with ethnic and linguistic change, since it shows the predominance of a new people that displaces or absorbs the original one. This is usually accompanied by a decline in Y-DNA variation, since certain clans usually predominate in the expansion of a population.

Founder effect (or genetic introgression) seems theoretically second to demic diffusion, and has been used to explain the replacement of genetic make-up without replacement of language. To resort to a founder effect to explain population changes when enough ancient DNA samples are lacking to suggest them is dangerous: the scarcity of ancient DNA samples makes the interpretation of their meaning – in relation to actual ancient areal occupation – a matter of subjective evaluation, in conjunction with archaeological finds (Campbell 2015).

Cultural diffusion has a weaker potential explanation for linguistic change, since it implies that the language of a minority – whose genetic lineages might have survived or not – spreads to the original population by way of economic or cultural (e.g. religious) domination. Population expansion into certain territories and decline of the original population are followed in some cases by a rising of the original paternal lineages (Brandt et al. 2015), in some cases with a huge decline of the invasive paternal ancestry. A static genetic situation observed after that process is often quite simply interpreted as cultural diffusion, supposing that no (or almost no) population exchange has taken place.

Other potential models can only be weaker than these main three. It seems logical that weaker models should not be used lightly, and clear proof of their applicability (and non-applicability of the stronger models) should be given in each case.

V.2.2. Admixture analysis and simple SNP comparison

This demic diffusion model relies on the comparison of ancient and modern Y-DNA SNPs, by observing how patrilineal lineages are replaced in certain areas that belonged to certain archaeological cultures. A male-dominated expansion from Yamna (Goldberg
et al. 2017) offers precisely the right case for this kind of study, and points to many potential flaws of admixture analysis in assessing the actual impact of the migration in European Neolithic / Bronze Age populations.

Examination of SNP of the Y-chromosome of ancient individuals one by one seems more suited to the scarcity of aDNA samples and the quality of its analysis, since defects in the STR sequencing are frequent, and thus only certain SNP markers may be obtained, with less information – and higher subclades – obtained from the samples.

The date calculated for TMRCA of modern populations has been used to define when certain migrations or expansions might have occurred. While it seems a good starting point for that purpose, it relies on the survival of modern populations related to such ancient population movements, and as such it could miss different – now extinct – lineages that could give an earlier date if they were included in TMRCA calculations.

Ancient and modern mtDNA distribution analyses – although they can help more clearly determine migration paths (Brandt et al. 2013) and other interesting characteristics of ancient cultures, such as female exogamy (Sjogren, Price, and Kristiansen 2016) –, has not been included in this paper for simplicity purposes.

Potential language relationships have been used to illustrate the Indo-European diffusion model. Many long-term linguistic relationships beyond Middle Indo-European remain hypothetical at best – when not completely discarded with the current data –, and it is not the intention of this paper to support or dismiss such connections. Such a relationship must be proven by linguistic research, and archaeology and genetics can only add precision to such studies.

While the theory here presented seems rational and scientifically sound, many details and alternative explanations have been omitted for clarity purposes.

More ancient DNA samples are needed to precisely draw most details of the general theory laid out in this paper.
VI. Conclusion

Careful cross-disciplinary investigation of ancient DNA samples recently published supports a demic diffusion model for the expansion of Indo-European languages directly into central and western Europe through the Bell Beaker culture, challenging previous archaeological and linguistic theories based on the expansion through the Corded Ware culture. Potential consequences of this new model in archaeological and linguistic investigation have been outlined in this paper, among them the development of a stable framework of time and space for Indo-European dialectal classification, allowing for a more precise dating of Indo-European branches and their splits and expansions, and why and how they might have occurred.
Acknowledgements

I have benefitted from corrections and contributions by Fernando López-Menchero.

For the sake of consistency, YFull estimates for year formed and TMRCA of each subclade have been used unless other sources are expressly stated. Also for the sake of consistency, YBP dates have been approximated to BC.

Public data from FTDNA-associated groups R1b, R1b-U106, R1b-P312, R1b-DF27, and R1b-U152 have been used in assessing modern population expansions.
References


Chiaroni, Jacques, Roy J King, Natalie M Myres, Brenna M Henn, Axel Ducourneau, Michael J Mitchell, Gilles Boetsch, Issa Sheikh, Alice A Lin, Mahnoosh Nik-Ahd, Jabeen Ahmad, Francesca Lattanzi, Rene J Herrera, Muntaser E Ibrahim, Aaron Brody, Ornella Semino,


Horvath, Csaba Barnabas. 2015. R1a subclades and Bronze Age migrations on the Eurasian steppes. *ESJ* 2.


Lobov, Artem Sergeevich. 2009. Структура генофонда субпопуляций башкир, Институт биохимии и генетики, Уфимский научный центр РАН, Уфа.


Pokutta, Anna. 2013. Population Dynamics, Diet and Migrations of the Únětice culture in Poland, Gothenburg, Department of Archaeology, University of Gothenburg, Gothenburg.


Scheeres, Mirjam. 2014. High mobility rates during the period of the “Celtic migrations”? 87Sr/86Sr and δ18O evidence from Early La Tène Europe, Fachbereich Biologie, Johannes Gutenberg-Universität Mainz, Mainz.


Zhao, Yong-Bin, Ye Zhang, Quan-Chao Zhang, Hong-Jie Li, Ying-Qiu Cui, Zhi Xu, Li Jin, Hui Zhou, and Hong Zhu. 2015. Ancient DNA Reveals That the Genetic Structure of the Northern Han Chinese Was Shaped Prior to 3,000 Years Ago *PLoS One* 10 (5):e0125676.