RESEARCH ARTICLE

Migration of Alpine Slavs and machine learning: Space-time pattern mining of an archaeological data set

Benjamin Štular[1]ʘ*, Edisa Lozić[1,2]ʘ, Mateja Belak[1]‡, Jernej Rihter[1]‡, Iris Koch[2]‡, Zvezdana Modrijan[1]‡, Andrej Magdič[3]‡, Stephan Karl[2]‡, Manfred Lehner[2]‡, Christoph Gutjahr[2]‡

1 Znanstvenoraziskovalni Center Slovenske Akademije Znanosti in Umetnosti, Ljubljana, Slovenia, **2** Institute of Classics, University of Graz, Graz, Austria, **3** Območna Enota Maribor, Javni Zavod Republike Slovenije za Varstvo Kulturne Dediščine, Maribor, Slovenia

ʘ These authors contributed equally to this work.
‡ MB, JR, IK, ZM, AM, SK, ML and CG also contributed equally to this work.
* benjamin.stular@zrc-sazu.si

Abstract

The rapid expansion of the Slavic speakers in the second half of the first millennium CE remains a controversial topic in archaeology, and academic passions on the issue have long run high. Currently, there are three main hypotheses for this expansion. The aim of this paper was to test the so-called "hybrid hypothesis," which states that the movement of people, cultural diffusion and language diffusion all occurred simultaneously. For this purpose, we examined an archaeological Deep Data set with a machine learning method termed time series clustering and with emerging hot spot analysis. The latter required two archaeology-specific modifications: The archaeological trend map and the multiscale emerging hot spot analysis. As a result, we were able to detect two migrations in the Eastern Alps between c. 500 and c. 700 CE. Based on the convergence of evidence from archaeology, linguistics, and population genetics, we have identified the migrants as Alpine Slavs, i.e., people who spoke Slavic and shared specific common ancestry.

1. Introduction

On Easter Monday, 2 April 568 CE, Alboin, son of Audoin and king of the Lombards, led his people on an arduous journey. They were to "abandon the barren fields. . . and come and take possession of Italy." Because from the "west and north [Italy] is shut in by the range of Alps, they reached it from the eastern side by which it is joined to Pannonia" (Paul the Deacon, Book II, Ch. V. and IX.; [1]).

Historiography records Alboin's success: He founded the Kingdom of Italy the next year and his successors ruled there for almost two centuries, e.g., [2]. However, historiography is less clear about events in the lands that Alboin and his people left behind. These lands include the Eastern Alps; we know little more than that this region was sparsely populated by Romans

in the sixth century and that several "peoples" or *gentes*, e.g. Ostrogoths, Gepids, and Slavs, attempted to make it their home, e.g., [3, 4]. Of these, the Slavs were the most successful.

The rapid spread of the Slavic language in the second half of the first millennium CE remains a controversial topic. There are two main reasons for this. First, the lack of first-hand, written sources before the end of the ninth century. Unlike the Lombards, the Slavs had no Paul the Deacon to recount their early history. Second, archaeological evidence on this subject is sparse compared to many other Early Medieval "peoples". As a result, there is a "propensity for sweeping explanations" [5].

Currently, there are three main hypotheses for the spread of Slavic between about 400 and 850 CE, e.g., [6, 7]. The first hypothesis assumes that speakers moved in all directions from their small original habitat, the so-called *Urheimat*, e.g., [8–10]. The second hypothesis assumes the diffusion of the Slavic cultural model among non-Slavic populations or, in its extreme form, the diffusion of language alone, e.g., [7, 11–14]. Many archaeologists adhere to the third, hybrid hypothesis. The hybrid hypothesis states that movement, cultural diffusion, and language diffusion occurred simultaneously [15–18]. This is supported by recent research in population genetics and linguistics. It seems that the language spread in the West Slavic zone mainly by migration to sparsely populated areas, and in the East Slavic zone by a combination of migration and language shift. The spread in the South Slavic region was triggered by migration, but the main mechanism for further spread was a language shift from local Balkan idioms to Slavic [19].

We adhere to the hybrid hypothesis in its most recent form [19], which is based primarily on population genetics and language studies. The aim of this paper was to test this hypothesis with archaeological data from the Eastern Alps. The specific objective of the paper was to elucidate the settlement of the Alpine Slavs, as the Slavic-speaking Early Mediaeval population of the Eastern Alps is known in historiography [4, 20–22]. To this end, we applied the technique called "space-time pattern mining" to examine a large archaeological data set from the period 400 to 1100 CE. In doing so, we have developed two archaeology-specific methodological innovations that can be applied to archaeological studies of any period.

2. Material and methods

2.1. Data set and study area

Our data set is Zbiva [23], which is an open access research database for the archaeology of the Eastern Alps in the Early Middle Ages (in our study 600 to 1100 CE). The inception of the database in the early 1980s was deeply rooted in the scientific research context of that time, which determined both the geographical and chronological focus of the data set [24].

Currently, Zbiva contains data on 3,379 sites and 11,596 related bibliographic units in more than half a million database fields. Because the data set is the result of four decades of deliberate scholarly work and attentive curation, it is best described by the concept of Deep Data. The Deep Data approach is one in which we make full use of all the information available in the data to gain the knowledge [25], *cf.* [26]. As far as the authors are aware, Zbiva is unmatched in Slavic archaeology, and the only comparable data set for Early Medieval archaeology is Open-Atlas [27], with its affiliate, THANADOS [28].

Recently, the team behind Zbiva has enriched the database, adding new information with a focus on chronology and location. The chronology of each site was re-evaluated by an expert, using modern typochronologies based on C14 dates [29–31]. The locations were also improved using maps (historical and modern) and satellite imagery available through open access web GIS applications. In addition, the data set was enriched with metadata (e.g., the

confidence level for chronology and location) and paradata (e.g., sources for dating). Furthermore, it was expanded to include Late Antiquity sites (400 to 600 CE in our study).

However, this data enrichment focused on a geographically limited subset of data. This subset of 1,105 archaeological sites constitutes the study area of this article. It includes present-day Slovenia, southern Austria (Carinthia, Styria, East Tyrol, parts of Salzburg and Upper Austria) and a small part of northern Italy (the Trieste region) (Fig 1).
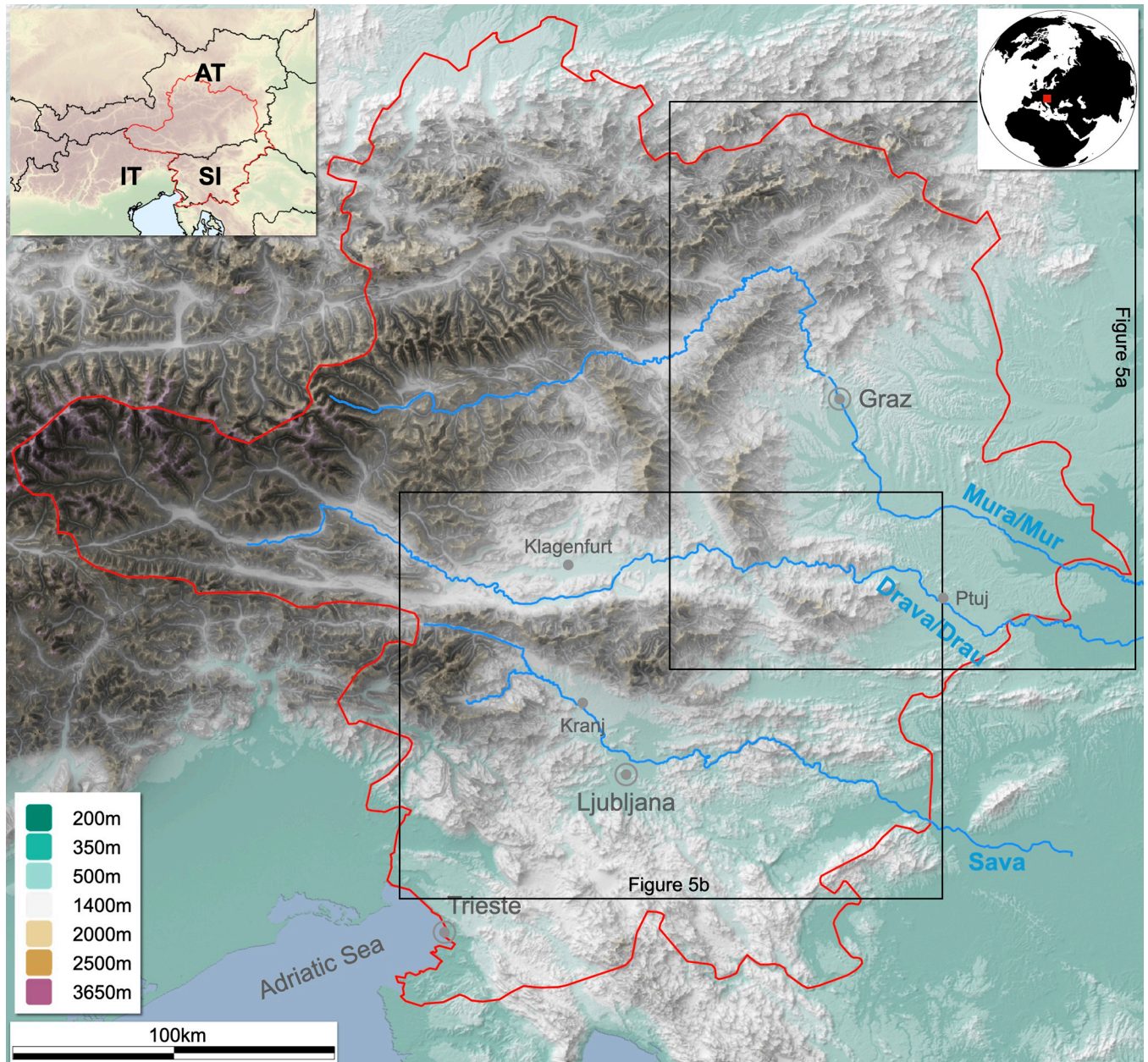
Fig 1. Map of the study area (upper left corner Lat. 48.22015, Lon. 12.35667; lower right corner Lat. 45.29785, Lon. 16.41784). The study area is marked in red and locations of Fig 5A and 5B in black (authors E.L. and B.Š; contains information from OpenStreetMap and OpenStreetMap Foundation, which is made available under the Open Database License; contains information adapted and modified from Copernicus Land Monitoring Service product EU-DEM25, which was produced with funding by the European Union).

https://doi.org/10.1371/journal.pone.0274687.g001

The data are described in more detail in the data paper [25] and are openly accessible via the Zenodo repository [32].

Compared to the entirety of Slavic territories ours is a small study region. But this region is an excellent (if not pivotal) case study for understanding the general processes of the spread of Slavic speakers in Early Middle Ages for three reasons. (i) Archaeologically, this is the only region where data is readily available for advanced spatial analysis, including machine learning (see above). (ii) The historiographical sources are second to none and include the oldest permanent Slavic political entity (Carantania, after 650 CE; e.g., [4, 22, 33–35]), the oldest Slavic text other than the canonical Old Church Slavic (ninth-century *Monumenta Frisingensia*; [36]), and the oldest mention of a member of a specifically Slavic social elite, a *župan* (iopan Physso, 777 CE [21, 22, 37]). (iii) Linguistically, the area is on the southwestern periphery of the spread of Slavic, bordering Germanic and Romance languages; this is important because peripheries typically preserve archaisms better than centres.

2.2. Space-time cube

Artificial Intelligence (AI) is becoming an ever more integral part of the digital humanities. Here, we focus on machine learning, which is a subset of AI. It refers to a set of data-driven algorithms and techniques that automate data prediction, classification, and clustering. Machine learning is rapidly being adopted by archaeologists interested in the analysis of geospatial data, material culture, texts, and artistic data [38, 39] and is most often used in combination with Big Data. Among the most prolific fields for machine learning within archaeology is airborne LiDAR data, e.g. overview in [40].

Currently, most machine learning processing takes place in the Phyton and R programming environments. This means that it requires coding and is therefore not readily accessible to most archaeologists. In this paper, we have used the only off-the-shelf pipeline that can be readily applied to archaeology: The toolset for space-time pattern mining in ArcGIS Pro v. 2.9 (ESRI, Redlands CA, USA). A similar pipeline has already been demonstrated on archaeological data [41], but to our knowledge ours is the first time this method was used to test a specific archaeological hypothesis.

Some comments on terminology are in order. "Space Time Pattern Mining" is a commercial name for the toolset used in this work [42]. We classify our method as machine learning, following [39]. They list the methods used in the cultural heritage domain, which are similar to the time series clustering, under unsupervised machine learning. In addition, machine learning is used in the background of this software to enable, for example, intelligent data-driven defaults [43].

Why use space-time pattern mining? Whenever archaeologists (or anyone, for that matter) look at a map, we inherently begin to turn that map into information by finding patterns and assessing trends. However, sometimes the patterns in the data are too complex to be clustered, observed and aggregated by the human eye. In such cases, we can use space-time pattern mining tools to answer our questions confidently, objectively, and repeatably.

The first step in space-time pattern mining is to organize the data. In this case, the data is incorporated into what is known as the "space-time cube model." This model is based on Torsten Hägerstand's time geography, which introduces the time axis into the traditional Cartesian coordinate system, e.g., [44]. It can be thought of as a three-dimensional cube consisting of space-time fields, where the dimensions x and y represent space and the dimension z represents time (Fig 2A). The data is stored in the so-called "space-time NetCDF cube." The process of constructing the space-time cube data model is complex and beyond the scope of this article, but it is well documented, e.g., [41].

The concept of the space-time cube is familiar to archaeologists. At its core, it is an application of the way we perceive (or have perceived in the past) archaeological excavations: In
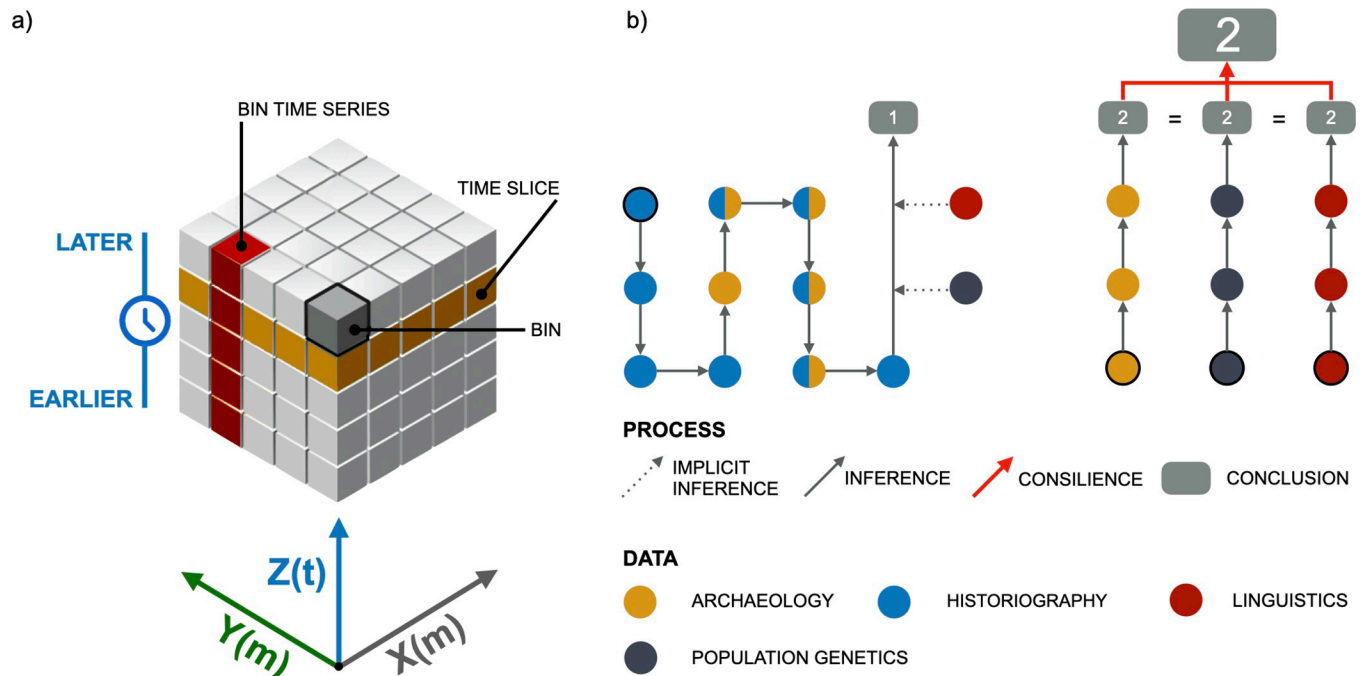
Fig 2. Graphical representations of selected methods: a) space-time cube model (after ESRI); b) comparison of two types of inference: On the left, data from different fields are compiled to draw a unified conclusion (analysis of Pohl's [17] interpretation of our study region as an example), and on the right, consilience.

https://doi.org/10.1371/journal.pone.0274687.g002

spatial quadrants and time phases. The quadrants in the xy grid are constant throughout the excavation, and the phases stack on top of each other, with the earliest at the bottom and the latest at the top.

For our case study, we aggregated data into 5 km big (y dimension) and 25 years long (z dimension) hexagonal bins. The size of the hexagon was chosen as the largest in which the relevant processes can be observed; in Early Medieval archaeology, it roughly corresponds to the site catchment area of a single settlement, e.g., [45]. The time interval of 25 years was determined on the basis of the data properties. The time sensitivity of relevant archaeological dating is about half a century, i.e., ± 25 years. However, start and end dates can often be a quarter of a century, rarely as brief as decades or even years. The accuracy of our dates is therefore 50 years, but the precision is approximately 25 years. Accordingly, 25-year intervals were chosen for analysis, but the accuracy of the data requires that archaeological interpretation be limited to 50-year intervals.

This method is very sensitive to the difference between no data (areas not analyzed) and null data (areas analyzed, but not found to contain any known sites). To account for this, we limited the space-time cube to the area used for data collection. Additionally, we excluded areas higher than 1,400 m above sea level (Fig 1: shades of brown). In our region, this altitude delineates the highest valley settlements from the lowest high-mountain pastures. The latter were excluded from the analysis because they are specialized seasonal settlements that were always dependent on valley settlements.

2.3. Time series clustering

Clustering is one of the most widely used machine learning techniques in the field of cultural heritage [39]. Its goal is to organize similar data into homogeneous groups or clusters. Clusters

are formed by grouping objects that have maximum similarity with other objects within the cluster and minimum similarity with objects in other clusters. For large and complex data sets, unsupervised approaches offer the best solution. Time series clustering is a type of unsupervised clustering used for data with a temporal component [46, 47].

The concept of time- series clustering is deeply familiar to archaeologists having been used since the nineteenth century. At that time, for example, the three-age system, which divides the development of human civilization into the Stone Age, Bronze Age, and Iron Age, was defined by clustering similarly dated stone/bronze/iron artefacts. As most archaeologists know from experience, such clustering is relatively easy for a few artefacts or sites but becomes daunting when the numbers run into the hundreds or thousands of objects or sites. In such cases, unsupervised time series clustering can be used.

In this article, we applied time series clustering to classify sites into chronological groups. In each group, the chronology (start date, end date) of the sites is more similar to each other than that of the sites outside the group.

The similarity between the clusters is measured by the so-called "pseudo-F statistic." The larger the pseudo-F value, the more different each cluster is from the other clusters [48]. There are several ways to calculate the pseudo-F statistic, each depending on which characteristics of the time series are considered important. In our experience, the most appropriate for archaeology is the "Profile (Fourier)" method, i.e., method based on Fourier series periodic function. It is used to cluster time series that have similar, smooth, and periodic patterns over time [42].

This method lends itself to the analysis of archaeological processes because they usually follow a consistent pattern: A gradually introduced innovation is followed by a peak of use and a steady decline. Thus, archaeological processes can be compared to seasons, where temperature follows a consistent annual pattern, with higher temperatures in summer and lower temperatures in winter. The "Profile (Fourier)" method is best suited to finding locations that have the most similar annual temperature patterns, for example, to distinguish between locations with mild and severe winters. A season in this example represents an archaeological phase or period.

In our case, we opted to ignore the range, i.e., the magnitude of the values in each period. To extend the analogy above, ignoring the range causes the change of seasons in two places occurring at the same times to be considered similar, even though the actual temperatures are different.

2.4. Modified emerging hot spot analysis

Spatial analysis is often called upon to determine the density of observed phenomena, and one of the most common tools to do this in archaeology is the so-called "hot spot analysis." It uses the Getis-Ord G^* statistic to calculate z-scores and p-values within a given spatial neighborhood. These indicate whether the observed spatial clustering of high and low values is more (hot spot) or less (cold spot) pronounced than would be expected from a random distribution, e.g., [49].

In this article, we have used an emerging hot spot analysis that examines the clustering of high and low values over time, in addition to spatial trends. The space-time cube is evaluated bin-by-bin, and each bin is analyzed relative to its space-time neighbors. Thus, each site is related not only spatially but also temporally to neighboring sites. The result is similar to the traditional hot spot analysis, except that it is in 3D (where the z dimension represents time).

Such a result can deliver an overwhelming amount of information. Therefore, the tool evaluates the trends of hot spots and cold spots over time using the Mann-Kendall trend test, e.g., [50] and categorizes each location in the study area accordingly. For example, a location is

considered a consecutive hot spot if it has an uninterrupted series of statistically significant hot spot bins over the latest time step intervals, but less than 90% of the total [42]. The resulting 2D representation of trends can be termed a "trend map."

However, the trend map provided by the tool is not suitable for archaeology for two reasons. First, it assumes that the latest records are the focus of analysis. Second, it was designed for data sets much larger than ours and those of most archaeological studies.

We therefore modified the trend map by focusing on chronological periods previously calculated by the time series clustering method. For example, a location was considered a first period consecutive hot spot if it had an uninterrupted hot spot series of at least 100 years within the first period (detailed description in S1 Table). We term this an "archaeological trend map".

In emerging hot spot analysis, the spatial and temporal neighborhoods have a significant influence on the results. We found, through empirical observation, that the best results for hot spots and cold spots were obtained with different settings. For cold spots: fixed distance method with 20 km neighborhood, and time step three. For hot spots: k-nearest neighbors (kNN) method with six spatial neighbors, time step one [42].

Therefore, we have introduced another archaeology-specific modification to the tool. For the purposes of this article, we superimposed the cold spots derived using the first settings with the hot spots derived using the second settings in a single visualization. We refer to this method as "multiscale emerging hot spot analysis."

To ensure the highest level of methodological transparency, reproducibility, and transparency and to reduce the time researchers spend replicating the work of other research groups, we provide the ready-to-re-use data in GIS format and the GIS protocol (S1 Appendix).

3. Theory

3.1. Consilience

The specifics of archaeological inference, e.g., [51], are not often outlined in articles such as ours. In this case, however, it is necessary because academic passions on the question of the migration of the Slavs have long been running high, and the methods of inference are often scrutinized. Moreover, this topic is invariably interdisciplinary, but the interdisciplinarity is achieved through a variety of approaches.

In order to enrich this discussion with the most objective archaeological information possible, we have chosen to base our inference on consilience of induction. Consilience, also known as "convergence of evidence," is a scientific principle that states that the same conclusion is much stronger when drawn from independent and unrelated sources. Confidence is strongest when evidence from different fields is considered because the methods and/or data are different [52].

Although it is rarely referred to by its name, this principle is popular in archaeology. For example, consilience is applied whenever radiocarbon dating is invoked to support archaeological dating.

It is important to distinguish between consilience, where conclusions are drawn independently before being correlated, and the more common interdisciplinary approach to the study of Slavic migrations, where data from different fields is compiled to draw a unified conclusion with a mix and match approach (Fig 2Bb). To this end, we have been careful to consider only information from each field that has not been influenced by findings from another field. For example, in the Discussion we consider linguistic information [53], but disregard the conclusions drawn on the same subject matter using supporting evidence from archaeology [54]. We

also take care to include only interpretations reached by domain specialists, as reinterpretations by non-specialists can be problematic [6].

3.2. Material culture as ethnicity, identity, and habitus?

The main archaeological argument for the migrations of the Slavs is based on the association of the Slavs with various archaeological cultures or *habitus*. For instance, the archaeological assemblages of the so-called Prague Culture are associated with the Early Slavs, e.g., [55–61].

From a modern theoretical perspective, this argument draws on Pierre Bourdieu's [62] notion of *habitus*; its basic premise is that practical knowledge is embodied in daily practices and that material culture, including pottery, expresses these practices, e.g., [63]. However, equating a *habitus* (e.g., the archaeological assemblages of Prague culture) with a people/tribe/ethnicity (e.g., the Early Slavs) is an additional step that Bourdieu did not anticipate. In much of the literature after the mid-1960s, the notion that material culture is more or less directly related to cognition of peoples was questioned by many; the acceptance that archaeological cultures simply cannot be directly correlated with ethnicity took hold, e.g., [64, 65].

Rather than engage in this discourse, we based our argument only on the categories of archaeological data that are indisputable: Location and chronology of the site. Our inference thus eclipses theoretical issues about the associations of material culture with ethnicity, e.g., [66], identity, e.g., [67], or habitus, e.g., [63]. Whether or not the Prague Culture archaeological assemblages are associated with the Early Slavs was immaterial to our conclusions.

4. Results

4.1. Archaeological periodization

The result of our time series clustering is the archaeological periodization of sites. The most archaeologically meaningful result is when the data is clustered into three periods (pseudo-F statistic value 421.595). These are the Late Antiquity and two Early Middle Ages periods (Fig 3). The latter two correspond to the traditional periodization of the jewelry into the Carantanian and the Köttlach phases, e.g., [68] or groups A/B and C of Eichert [69].
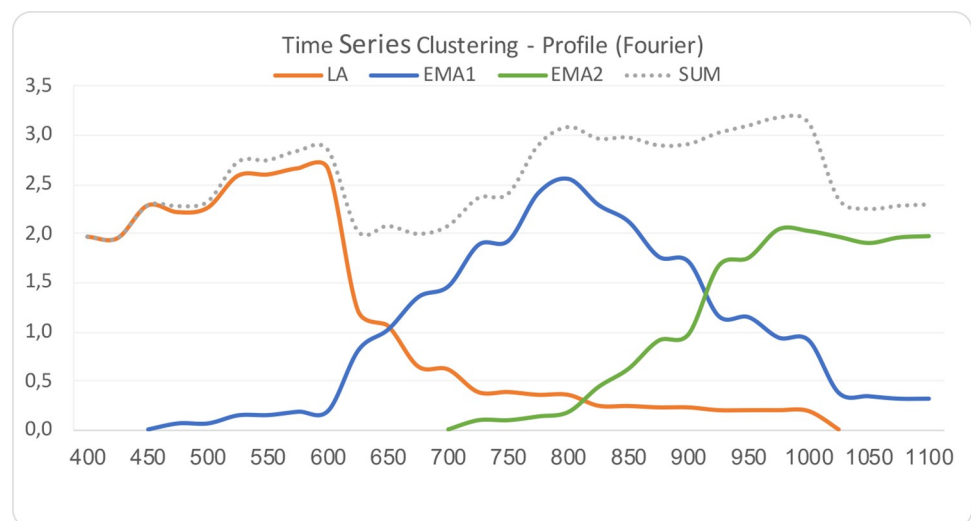

Fig 3. Archaeological periodization with time series clustering: LA—Period 1, Late Antiquity; EMA1—Period 2, Early Middle Ages 1; EMA2—Period 3, Early Middle Ages 2. Values on x axis are years CE, values on y axis are unitless and relative.

https://doi.org/10.1371/journal.pone.0274687.g003

Three important conclusions may be drawn from these results. The first is, that the general increase in sites between 400 and 500 CE and the decrease after 1000 CE does not reflect reality, as we know it from numerous sources, e.g., [22, 70]. Rather, it reveals the weakness of the underlying data set: The period before 500 CE has not been collated systematically, and data for the period after 1000 CE (which, until recently, was not considered relevant to archaeology in the region, e.g., [71]) is lacking. Regardless, the present data set is suitable for the study of the half millennium between 500 and 1000 CE, which was the aim.

Second, unlike changes in material culture, changes in landscape are more gradual and often overlap. For example, the time series of Late Antiquity does not end until 1000 CE, as some sites exhibit continuity from Late Antiquity onward (for example, the town of Kranj, e.g., [72]). The results of time series analysis in archaeology are therefore complex and must be interpreted with great care.

Third, we substantiated the long-established periodization of the Early Middle Ages into two periods by an independent source of data: The chronology of sites rather than the typology of jewelry. This, then, is the first quantitative evidence that changes in jewelry styles taking place in the second half of the 9th century were reflected in changes in the archaeological landscape. The most likely explanation is that both changes had the same underlying cause, which however is beyond the scope of this article.

4.2. Archaeological landscape

The emerging hotspot analysis revealed an astonishing quantity and quality of information (Fig 4). Most relevant to our topic are the extensive areas of cold spots in the northern part of the region and the general patchiness, i.e. activity is concentrated in enclaves. In this respect, the archaeological landscape between 500 and 1000 CE differs from both the preceding Roman period settlement and subsequent High Medieval period, which both exhibit a more regular pattern of settlement. This is important in providing a context for understanding various historical processes. For example, the reason it is so difficult for historiographers to define the exact borders of *Carniola*, e.g., [3] and *Carantania*, e.g., [4] is that in the patchy landscape precise fixed borders most likely never existed.

The main focus of this article was migration. The two tools for detecting migrations in our data were provided by Curta [13]. First, migration must have occurred if settlements and cemeteries suddenly appear in a previously sparsely populated area, i.e. cold spots are immediately followed by hot spots. Second, migration can also be detected by the sudden appearance of a material culture without local traditions or parallels in a given area.

With the first tool, migration was documented in the easternmost part of the study area. In the period between 450 and 500 CE this is a cold spot area, but after c. 500 CE hot spots appear along the river Mura (*Ger*. Mur). After a period of consolidation until c. 600 CE, a series of small-scale neighbourhood migrations upstream of the Mura and the adjacent Drava (*Ger*. Drau) rivers is documented by numerous hot spots (Fig 5A; S2 Appendix).

With the second tool, a migration upstream of the Sava river after c. 675 CE was documented. Between c. 600 and 675 CE, there was a gradual decline in hot spots along the Sava, but between c. 675 and 750 CE there was a reversal of that trend (Fig 5B; S2 Appendix). This trend reversal alone could be explained by other causes than migration. The evidence for migration was provided by the time series clustering, which showed a sudden and complete shift in material culture: the number of Late Antiquity sites diminishes dramatically and at the same time Early Medieval sites start appearing (Fig 3). This shift has long been known in archaeology as the transition from fortified hilltop settlements to unfortified lowland settlements, e.g., [70], which determines the transition from Late Antiquity to the Early Middle Ages.
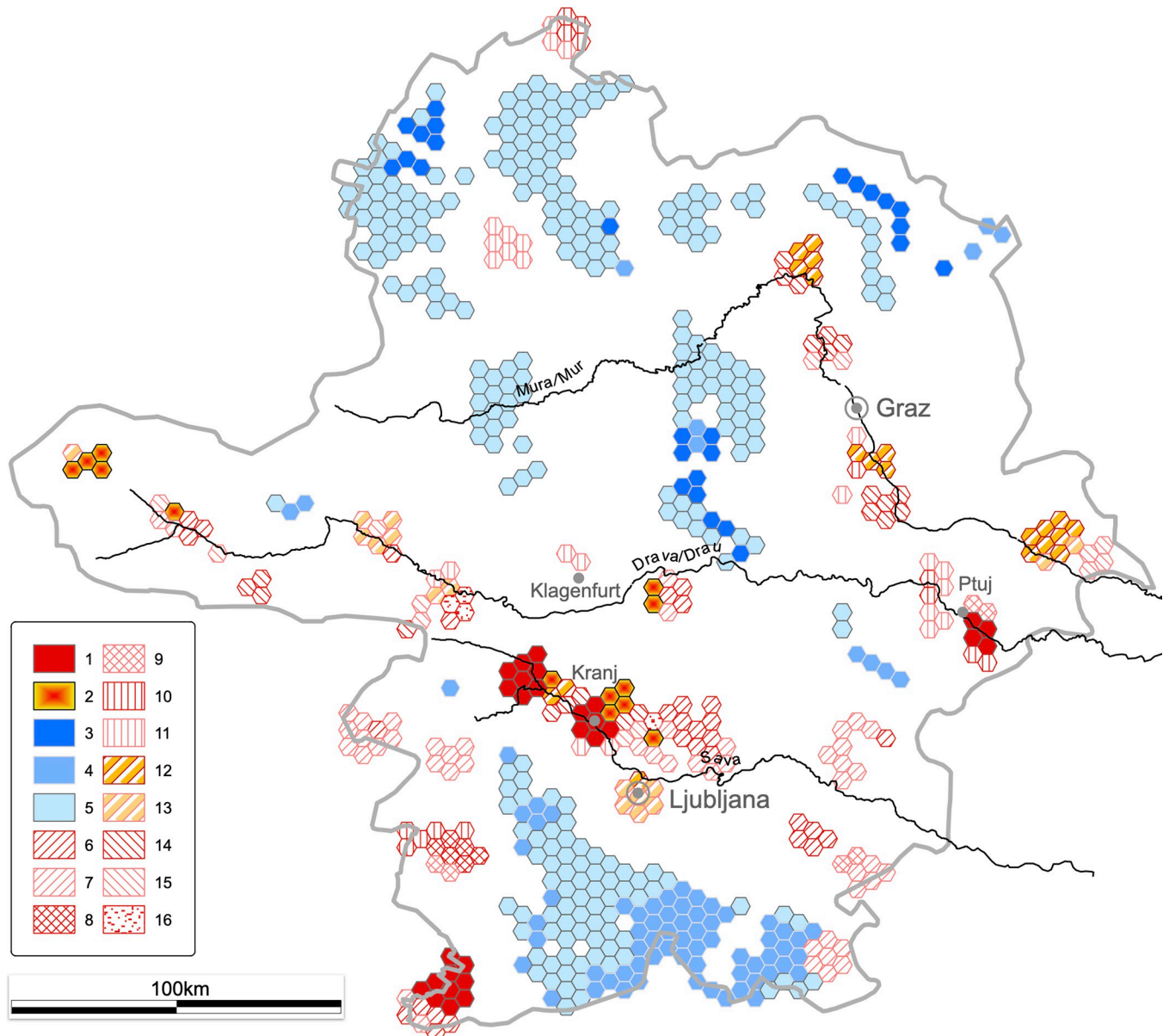
Fig 4. Archaeological trend map of the modified categorization of the multiscale emerging hot spot analysis. See S1 Table for the legend (authors E.L. and B.Š; contains information from OpenStreetMap and OpenStreetMap Foundation, which is made available under the Open Database License; contains information adapted and modified from Copernicus Land Monitoring Service product EU-DEM25, which was produced with funding by the European Union).

https://doi.org/10.1371/journal.pone.0274687.g004

On the basis of this data, a comment can be made on the size of the migrations. Overall, hot spots interpreted as resulting from the first migration account for only 4% of all hot spots in c. 500 CE, which indicates a relatively small founder population. However, by c. 700 CE, 59% of hot spots can be interpreted as resulting directly or indirectly from both migrations. Although this is a very rough estimate, far from giving a direct indication of the actual number of people involved, it is the best available and by far the most tangible to date, cf., [3, 4, 37, 73–76]. As such, it is an invaluable foundation for explaining acculturation processes following migrations, which, however, are not the subject of this article.
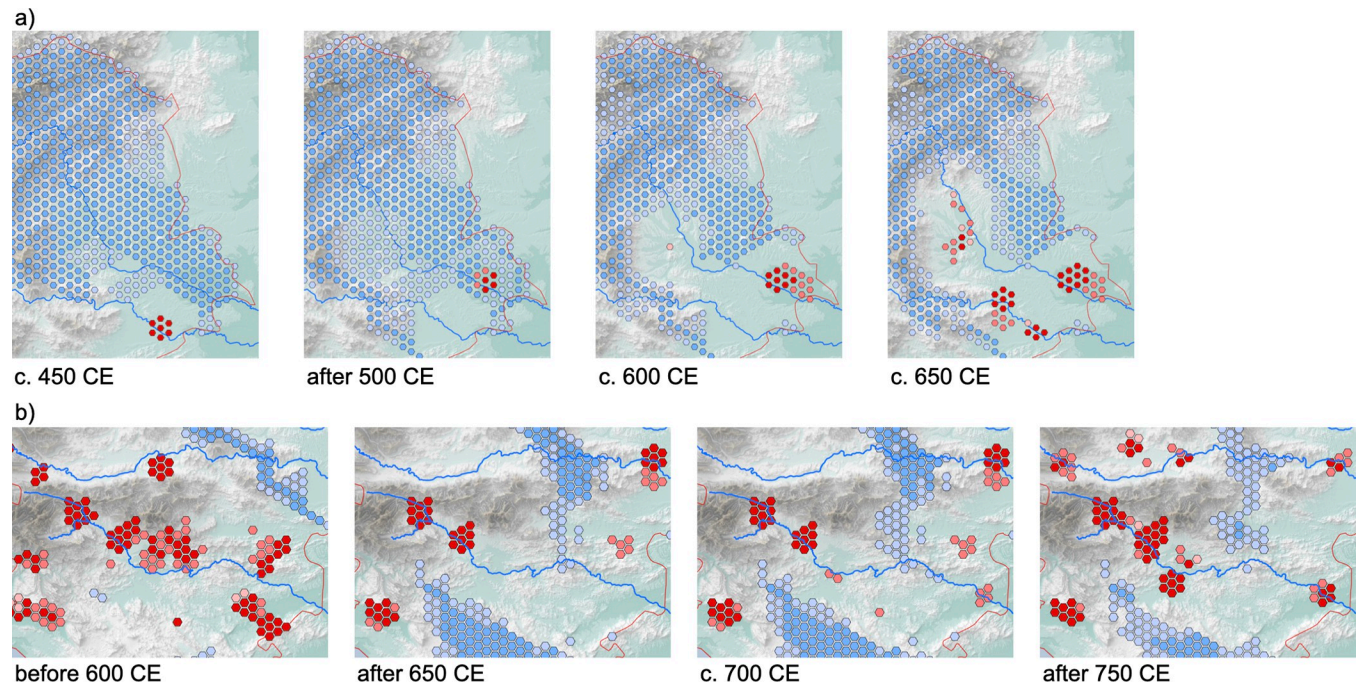
Fig 5. Time slices from the emerging hot spot analysis for selected areas: a) Eastern part of the study area from 450 CE to 650 CE; b) central part of the study area from before 600 CE to after 750 CE (authors E.L. and B.Š; open access raw data sources used: EU-DEM v1.1, https://land.copernicus.eu; OpenStreetMap, https://www.openstreetmap.org).

https://doi.org/10.1371/journal.pone.0274687.g005

5. Discussion

5.1. Archaeology

Our data thus testifies to two migrations: The first upstream of the Mura and Drava rivers after c. 500 CE, and the second before c. 700 CE upstream of the Sava river. This is an important discovery, but it does not shed light on who the migrants were. Based on the archaeological [77–80]) and historiographical [17, 20, 22] context, we can hypothesize that they were Early Slavs. But this hypothesis inherits all the weaknesses of the existing ones, which are based on the controversial presence of archaeological assemblages of Prague Culture and scant written sources.

The hypothesis may, however, be considered the null hypothesis that can be tested with the consilience principle. The new archaeological evidence for two separate migrations allows us to correlate it with interpretations from the linguistics and genetics of modern populations. These two fields of science use completely different data sources and methods than archaeology and have recently made significant advances in understanding Slavic migrations.

5.2. Linguistics

Let us first take a look at linguistics. While a language or dialect may be tied to any number of identities within a given period, a shared linguistic innovation requires a linguistic community, for which the term "founder population" has been proposed [6, 54].

Modern Slovenian, which is spoken today in the southern part of the research area, belongs to the South Slavic clade, according to the traditional classification of the Slavic languages [81]. There are, however, considerable linguistic similarities between the Slovenian and the West

Slavic lects. These similarities were explained either by the existence of a specific link between Slovenian and West Slavic or by a mixed South and West Slavic origin of Slovenian [82–86].

Specific ties between Slovenian and South Slavic on the one hand and West Slavic on the other have recently been demonstrated with a series of phylogenetic NeighborNet networks. The analysis concluded that Slovenian seems to be almost equally close to the West and South Slavic, but distant from the East Slavic, "thus supporting the putative mixed nature of Modern Slovenian" [87]. It is the latter interpretation that interests us.

In conclusion of the above cited analysis further studies of Slovenian dialects are proposed in order to clarify the position of Slovenian among the Slavic languages. One of such study examined the diatopic distribution and semantic development of *glčěti as the primary neutral verb meaning 'to speak'. It was carried from an emergent dialect of Slavic and is now widespread in present-day central Russia, central Bulgaria, and in Slovenia along the Mura and Drava rivers. Of interest is the hypothesis of possible relationships between the "early Slavic speakers who spoke dialects in which *glčěti played a central role as a verb of speech" and those who did not within modern Slovenia, i.e. the southern part of our study area. The hypothesis states that this dichotomy, together with the -ny- ‖ -nǫ- isogloss, "can be viewed as inherited pre-migration cleavages" [53], that is, "the dialects of Slavic brought to the subalpine area. . . differed (amongst themselves)" ([6]; translated from the Slovenian by B.Š.; the subalpine area mentioned corresponds approximately to our study region). Since "shared linguistic innovation presupposes a community" ([6]; translated from the Slovenian by B.Š.), it follows that heterogeneous dialects presuppose heterogeneous communities or founder populations.

Therefore, the linguistic interpretations imply that the southern part of the region under study was originally populated by two founder populations that spoke two heterogeneous Slavic dialects. One, using *glčěti, populated areas along the rivers Mura and Drava, the other one populated areas further west.

5.3. Genetic history

Second, let us turn to genetic history. This scientific field attempts to reconstruct human evolution and the history of human populations using genetic information obtained from either modern or ancient DNA [88]. DNA has been described as a document containing "messages from the past" [89] and is a proven tool in prehistoric archaeology, e.g., [90–93]. However, there are significant obstacles to the use of modern DNA when it comes to Late Antiquity and the Middle Ages. For example, the historical population-level information that this method reflects is complex and overlapping and should not be understood as representing a direct correspondence between population history and social history [94, 95]. For the study of this time period, ancient DNA or ancient genomic DNA data are more appropriate, e.g., [96, 97]. However, ancient DNA data are not available in sufficient quantity to study the expansion of Slavic speakers.

Regardless of the methodological shortcomings, it is agreed that the results of genetic studies on modern DNA are indisputable in terms of providing information on genetic proximity and can contribute to hypotheses about human population history, including Late Antiquity and Early Mediaeval migrations [94, 97–99].

The most complete study of modern DNA pertaining to the expansion of Slavic speakers examined all ethnic groups living today who speak Balto-Slavic languages, as well as their neighbors. It concluded that the genetic diversity of today's Slavs was predominantly formed *in situ* (i.e., the substrate genetic components in the settled areas prevail), with marked differences between West and East Slavs on the one hand and South Slavs on the other. However, there is genetic affinity showing a common ancestry (i.e., a homogeneous genetic substrate

inherited from the ancestral population) among the Slavs, which probably demonstrates the historical dispersion of a once uniform population [87, 100]. This was recently confirmed in a review article, which concluded, that the migration of Slavs was accompanied by active assimilation of indigenous European populations [101].

Looking more closely at the region under study, the variability of the microsatellite loci of the Y chromosome is telling. It shows that the inhabitants of present-day Slovenia are far removed from all other South Slavic populations [87]. When this was first discovered in an earlier study it was interpreted to "suggests that at least two different migration waves of the Slavs may have reached the Balkans in the early Middle Ages" [102].

Considering only the modern Slovenian population, there is one possible ancestral haplogroup for all Slovenian populations which has the highest frequency along the Mura and Drava rivers. This could indicate that the origin of this ancestral haplogroup was in this area and that it later spread westward [103].

Thus, population genetic studies show that the southern part of the study region was possibly settled by two separate migrations. The earlier one took place along the Mura and Drava rivers; the later one was to the west of that area.

5.4. Consilience

Interpretations from three scientific fields, using completely different data sets and methods, shows a consilience or convergence of evidence. Archaeological, linguistic, and genetic evidence suggests, with varying degrees of certainty, that there were two separate migrations to the southern part of the region under study: The earlier one along the Mura and Drava rivers, and the later one, which archaeology can locate along the Sava river. Archaeology and genetics suggest that acculturation was the predominant post-migration process. Linguistics and genetics indicate that the migrants were Slavs. In particular, linguistics indicates that the migrants were speakers of Slavic, and genetics confirms that they had a homogeneous genetic substrate inherited from a single ancestral population common only to ethnic groups speaking Slavic today.

Based on consilience, we can define the immigrants who arrived in the Eastern Alps between c. 500 and c. 700 CE with by far the highest reliability to date. They were speakers of Slavic and shared a specific "Slavic" ancestry. Our archaeological analysis places these migrations in space and time with some precision (Fig 5).

6. Conclusions

The aim of this article was to test the hybrid hypothesis of Slavic migration using archaeological data, with the Alpine Slavs as a case study. The term "migration" in the title was deliberately chosen to be somewhat provocative, as modern historiography and archaeology of the Early Middle Ages tend to downplay the role of physical migrations.

We used selected machine learning methods to analyze an archaeological data set that can be described as Deep Data. Specifically, we used two methods: Time series clustering and a modified emerging hot spot analysis. The former method is directly suitable for archaeology without modification, whereas the latter required two archaeology-specific modifications: The archaeological trend map and the multiscale emerging hot spot analysis.

The results have provided us with an overwhelming quality and quantity of new information. In this article, we have focused on confirming two separate migrations of the Alpine Slavs. Based on the convergence of evidence from archaeology, linguistics, and population genetics, we define the immigrants as Alpine Slavs who were speakers of Slavic and shared

specific "Slavic" ancestry. Two founder populations migrated to the Eastern Alps: The first after c. 500 and the second before c. 700 CE.

The identities and ethnicity of the migrants (as defined by modern historiography) are, in our view, beyond the scope of archaeology. The acculturation processes that took place after the migration will be discussed elsewhere. From the available evidence, however, it is clear that the crucial process was cultural spread *sensu* Heather [18]. We envisage that the number of migrating people was relatively small and more akin to a small group infiltration than a mass migration. The movement itself was a part of it, but the processes that took place afterwards were historically the most important.

Thus, we have achieved the aim of the article, which was to prove the validity of the hybrid hypothesis of Slavic migration with archaeological data. The migration of the Alpine Slavs was a combination of movement of people, cultural diffusion, and language diffusion, all occurring simultaneously. This clearly refutes the hypothesis that only the cultural model or even only the language spread.

While this paper focused on a specific question related to the migration of Slavs, the methods we developed, borrowing and adapting from a variety of disciplines, can be applied to archaeological studies of any period, anywhere that suitable data is available. We hope that these advances will be used beneficially by other scholars and establish a new, practical approach to add to the archaeological arsenal of methodologies.

In the field of machine learning in archaeology and the digital humanities in general, we hope to have shown that, in addition to Big Data, Deep Data also holds great potential.

Supporting information

S1 Table. Archaeological trend map, detailed description of 16 classes.
(DOCX)

S1 Appendix. GIS Protocol for multy-scale emerging hot spot analysis.
(DOCX)

S2 Appendix. Animation of the emerging hot spot analysis time slices from 400 CE to 1100 CE (authors E.L. and B.Š; contains information from OpenStreetMap and OpenStreetMap Foundation, which is made available under the Open Database License; contains information adapted and modified from Copernicus Land Monitoring Service product EU-DEM25, which was produced with funding by the European Union).
(GIF)

Acknowledgments

Author Contributions

Conceptualization: Benjamin Štular, Edisa Lozić.

Data curation: Mateja Belak.

Formal analysis: Edisa Lozić.

Funding acquisition: Benjamin Štular.

Investigation: Mateja Belak, Jernej Rihter, Iris Koch, Zvezdana Modrijan, Andrej Magdič, Stephan Karl.

Methodology: Benjamin Štular, Edisa Lozić.

Project administration: Benjamin Štular, Andrej Magdič, Manfred Lehner, Christoph Gutjahr.

Software: Benjamin Štular.

Supervision: Benjamin Štular, Manfred Lehner.

Validation: Edisa Lozić.

Visualization: Edisa Lozić.

Writing – original draft: Benjamin Štular, Edisa Lozić.

Writing – review & editing: Benjamin Štular, Edisa Lozić, Andrej Magdič, Stephan Karl, Manfred Lehner.

References

1. Deacon P, Foulke WD. History of the Lombards. Peters E, editor. Philadelphia: University of Pennsylvania Press; 1974.

2. Bradač F, Grafenauer B, Gantar K. Pavel Diakon (Paulus Diaconus), Zgodovina Langobardov (Historia Langobardorum). Maribor: Obzorja; 1988.

3. Bratož R. Med Italijo in Ilirikom. Slovenski prostor in njegovo sosedstvo v pozni antiki. Ljubljana: Filozofska fakulteta; 2014. (Zbirka ZČ 46).

4. Kahl H-D. Der Staat der Karantanen. Fakten, Thesen und Fragen zu einer frühen slawischen Machtbildung im Ostalpenraum (7.-9. Jh.). Bratož R, editor. Ljubljana: Narodni muzej Slovenije; 2002.

5. Pohl W. The Early Slavs: Culture and Society in Early Medieval Eastern Europe. P. M. Barford. Speculum. 2004; 79: 448–450. https://doi.org/10.1017/s0038713400087996

6. Snoj M, Greenberg ML. O jeziku slovanskih prebivalcev med Donavo in Jadranom v srednjem veku (pogled jezikoslovcev). Zgodovinski časopis. 2012; 66: 276–305.

7. Curta F. Slavs in the Making: History, Linguistics, and Archaeology in Eastern Europe (ca. 500-ca. 700). London, New York: Routledge; 2020. https://doi.org/10.4324/9780203701256

8. Herrmann J. Urheimat und Herkunft der Slawen. In: Herrmann J, editor. Welt der Slawen. Geschichte, Gesellschaft, Kultur. München: Verlag C. H. Beck; 1986. pp. 11–18.

9. Dolukhanov P. The early Slavs: Eastern Europe from the initial settlement to the Kievan Rus. London, New York: Routledge; 1996.

10. Timberlake A. Culture and the spread of Slavic. In: Bickel B, Grenoble LA, Peterson DA, Timberlake A, editors. Language Typology and Historical Contingency: In honor of Johanna Nichols. Amsterdam, Ph: John Benjamin's Publishing Company; 2013. pp. 331–356.

11. Lunt HG. Common Slavic, Proto-Slavic, Pan-Slavic: What Are We Talking About? International Journal of Slavic Linguistics and Poetics. 1997; 41: 7–67.

12. Curta F. The Making of the Slavs: History and Archaeology of the Lower Danube Region, c. 500–700. Cambridge: Cambridge University Press; 2001.

13. Curta F. Migrations in the Archaeology of Eastern and Southeastern Europe in the Early Middle Ages (Some Comments on the Current State of Research). In: Preise-Kapeller J, Reinfandt L, Stouraitis Y, editors. Migration Histories of the Medieval Afro-Eurasian Transition Zone: Aspects of mobility between Africa, Asia and Europe, 300–1500 CE. London, New York: Brill; 2020. pp. 101–140.

14. Pritsak O. The Slavs and the Avars. In: Gli slavi occidentali e meridionali nell'alto medioevo 15–21 Aprile 1982 Settimane di Studio del Centro Italiano di Studi Sull'Alto Medioevo, XXX. Fondazione Centro italiano di studi sull'alto Medioevo; 1983. pp. 353–435.

15. Kazanski M. Archaeology of the Slavic Migrations. In: Greenberg ML, Grenoble LA, editors. Encyclopedia of Slavic Languages and Linguistics. London, New York: Brill; 2020. https://doi.org/10.1163/2589-6229_ESLO_COM_035967

16. Pleterski A. Etnogeneza slavena—metode i proces. Starohrvatska prosvjeta. 2013; 40: 8–32.

17. Pohl W. The Avars: a Steppe Empire in Europe, 567–822. Ithaca, NY: Cornell University Press; 2018. https://doi.org/10.7591/9781501729409

18. Heather P. Empires and barbarians: The fall of Rome and the birth of Europe. Oxford: Oxford University Press; 2010.

19. Lindstedt J, Salmela E. Language Contact and the Early Slavs. In: Klír T, Boček V, Jansens N, editors. New perspectives on the Early Slavs and the rise of Slavic. Universitätsverlag Winter GmbH; 2020. pp. 275–299.

20. Grafenauer B. Zgodovina Slovenskega naroda, I. zvezek: Od naselitve do uveljavljenja frankovskega fevdalnega reda. Ljubljana: Kmečka knjiga; 1954.

21. Friesinger H. Alpenslawen und Bayern. In: Herrmann J, editor. Welt der Slawen. Geschichte, Gesellschaft, Kultur. München: Verlag C. H. Beck; 1986. pp. 109–122.

22. Štih P. The Middle Ages between the Eastern Alps and the Northern Adriatic. Select Papers on Slovene Historiography and Medieval History. London, New York: Brill; 2010.

23. Pleterski A. Zbiva v3.08; 2016. [cited 2021 Dec 30]. Database: Sites and Monuments [Internet]. Available from: http://zbiva.zrc-sazu.si

24. Štular B. The Zbiva Web Application: a tool for Early Medieval archaeology of the Eastern Alps. In: Richards JD, Niccolucci F, editors. The ARIADNE Impact. Budapest: Archaeolingua; 2019. pp. 69–82. https://doi.org/10.5281/zenodo.3476712

25. Štular B, Belak M. Deep Data Example: Zbiva, Early Medieval Data Set for the Eastern Alps. Research Data Journal for the Humanities and Social Sciences. 2022; In press.

26. Huggett J. Is Less More? Slow Data and Datafication in Archaeology. In: Garstki K, editor. Critical Archaeology in the Digital Age. Los Angeles, CA: Cotsen Insititute of Archaeology; 2022. pp. 156–184.

27. Filzwieser R, Eichert S. Towards an Online Database for Archaeological Landscapes. Using the Web Based, Open Source Software OpenAtlas for the Acquisition, Analysis and Dissemination of Archaeological and Historical Data on a Landscape Basis. Heritage. 2020; 3: 1385–1401. https://doi.org/10.3390/heritage3040077

28. Eichert S. Digital Mapping of Medieval Cemeteries. Journal on Computing and Cultural Heritage. 2021; 14: 1–15. https://doi.org/10.1145/3406535

29. Pleterski A. Datiranje zgodnjesrednjeveške naselbine Lehen pri Mitterkirchnu v Zgornji Avstriji kot kontrola nove datacijske metode s pomočjo referenčne tabele in korelacijske formule ustij loncev. Vjesnik Arheološkog muzeja u Zagrebu. 2010; 43: 309–324.

30. Pleterski A. Zgodnjesrednjeveška naselbina na blejski Pristavi: tafonomija, predmeti in čas. Ljubljana: Založba ZRC; 2010.

31. Pleterski A. A step towards the chronology of early medieval head ornaments in the Eastern Alps. Arheološki vestnik. 2013; 64: 299–334.

32. Štular B, Pleterski A, Belak M. Zbiva, Early Medieval Data Set for the Eastern Alps. Data sub-set. 2021. [cited 2021 Dec 29]. Repository: Data Set [Internet]. Available from: https://doi.org/10.5281/ZENODO.5761811

33. Gleirscher P. Karantanien: slawisches Fürstentum und bairische Grafschaft. Klagenfurt/Celovec: Hermagoras Verlag., Mohorjeva založba; 2018.

34. Eichert S. Karantanische Slawen—slawische Karantanen. Überlegungen zu ethnischen und sozialen Strukturen im Ostalpenraum des frühen Mittelalters. In: Biermann F, Kersting T, Klammt A Der Wandel um 1000 Beiträge der Sektion zur slawischen Frühgeschichte der 18. Jahrestagung des Mittel- und Ostdeutschen Verbandes für Altertumsforschung in Greifswald, 23. bis 27. März 2009. Langenweissbach: Beier & Beran; 2011. pp. 433–440.

35. Pleterski A. Die Kärntner Fürstensteine in der Struktur dreier Kultstätten. In: Huber A, editor. Der Kärntner Fürstenstein im europäischen Vergleich (Symposium Gmünd 1996). Gmünd: Die Stadt; 1997. pp. 43–119.

36. Ogrin M, Grdina I, Erjavec T, Bojadžiev D. Brižinski spomeniki: Monumenta Frisingensia: elektronska znanstvenokritična izdaja. Ljubljana: Inštitut za slovensko literaturo in literarne vede ZRC SAZU; 2007. Available from: http://nl.ijs.si/e-zrc/bs/html/bs.html.

37. Štih P. Strukture današnjega slovenskega prostora v zgodnjem srednjem veku. In: Bratož R, editor. Slowenien und die Nachbarländer zwischen Antike und karolingischer Epoche. Anfänge der slowenischen Ethnogenese I. Ljubljana: Narodni muzej Slovenije; 2000. pp. 355–394.

38. Bickler SH. Machine Learning Arrives in Archaeology. Advances in Archaeological Practice. 2021; 9: 186–191. https://doi.org/10.1017/aap.2021.6

39. Fiorucci M, Khoroshiltseva M, Pontil M, Traviglia A, Del Bue A, James S. Machine Learning for Cultural Heritage: A Survey. Pattern Recognition Letters. 2020; 133: 102–108. https://doi.org/10.1016/j.patrec.2020.02.017

40. Lozić E, Štular B. Documentation of Archaeology-Specific Workflow for Airborne LiDAR Data Processing. Geosciences. 2021; 11: 26. https://doi.org/10.3390/geosciences11010026

41. Cui J, Liu Y, Sun J, Hu D, He H. G-STC-M Spatio-Temporal Analysis Method for Archaeological Sites. ISPRS International Journal of Geo-Information. 2021; 10: 312. https://doi.org/10.3390/ijgi10050312

42. ESRI. How Create Space Time Cube works. 2021. [cited 29 December 2021]. In: ESRI Help [Internet]. Available from: https://pro.arcgis.com/en/pro-app/latest/tool-reference/space-time-pattern-mining/learnmorecreatecube.htm.

43. Bennett L. Machine Learning in ArcGIS. ArcUser, the Magazine for Esri Software Users. 2018; 21(2): 8–9. Available from: https://www.esri.com/about/newsroom/arcuser/machine-learning-in-arcgis/.

44. Kraak M-J. Geovisualization and time: new opportunities for the space-time cube. In: Dodge M, McDerby M, Turner M, editors. Geographic visualization: concepts, tools and applications. Chichester, England; Hoboken, NJ: John Wiley & Sons, Ltd; 2008. pp. 293–306. https://doi.org/10.1002/9780470987643.ch15

45. Lozić E. Application of Airborne LiDAR Data to the Archaeology of Agrarian Land Use: The Case Study of the Early Medieval Microregion of Bled (Slovenia). Remote Sensing. 2021; 13(16): 3228. https://doi.org/10.3390/rs13163228

46. Montero P, Vilar JA. TSclust: An R package for time series clustering. Journal of Statistical Software. 2014; 62: 1–43. https://doi.org/10.18637/jss.v062.i01

47. Aghabozorgi S, Shirkhorshidi AS, Wah TY. Time-series clustering—A decade review. Information Systems. 2015; 53: 16–38. https://doi.org/10.1016/j.is.2015.04.007

48. Vogel MA, Wong AKC. PFS clustering method. IEEE transactions on pattern analysis and machine intelligence. 1979; 237–245. https://doi.org/10.1109/tpami.1979.4766919 PMID: 21868854

49. Achino KF, Štular B, Rihter J, Rihter J. Assessing the intentionality of spatial organization. Cemetery of Župna Cerkev (Kranj, Slovenia) case study. Arheološki vestnik. 2019; 70: 297–313.

50. Hamed KH, Rao AR. A modified Mann-Kendall trend test for autocorrelated data. Journal of hydrology. 1998; 204: 182–196.

51. Hodder I. The Archaeological Process an Introduction. Oxford: Wiley-Blackwell; 1999.

52. Wilson EO. Consilience. The unity of knowledge. New York: Vintage Books; 1998.

53. Schallert J, Greenberg ML. The Prehistory and Areal Distribution of Slavic *glčěti 'Speak'. Slovenski jezik / Slovene Linguistic Studies. 2007; 6: 9–76.

54. Greenberg ML. Slavs as migrants: mapping prehistoric language variation. In: Genis R, de Haard E, Lučić R, editors. Definitely Perfect Festschrift for Janneke Kalsbeek. Amsterdam: Uitgeverij Pegasus; 2017. pp. 169–183.

55. Kazanski M. Les Slaves. Les origines (Ier-VIIe siècle après J.-C.). Paris: Editions errance; 1999.

56. Barford PM. The early Slavs: culture and society in early medieval Eastern Europe. Ithaca: Cornell University Press; 2001.

57. Gojda M. The Ancient Slavs; Settlement and Society. Edinburgh: Edinburgh University Press; 1991.

58. Parczewski M. Origins of early Slav culture in Poland. Antiquity. 1991; 65: 676–683. https://doi.org/10.1017/S0003598X00080303

59. Biermann F, 2016 New archaeological evidence from the Late Migration and Early Slavic period in the north-east German region. In: Chudzinśka B, Wojenka M, Wołoszyn M, editors. Od Bachórza Do Swiatowida Ze Zbrucza: Tworzenie Się Słowianskiej Europy W Ujęciu Zródłoznawczym: Księga Jubileuszowa Profesor a Michała Parczewskiego. Kraków- Rzeszów: Wydawnictwo Uniwersytetu Rzeszowskiego; 2016. pp. 113–123.

60. Pavlovič D, Vojakovič P, Toškan B. Cerklje ob Krki: novosti v poselitvi Dolenjske v zgodnjem srednjem veku. Arheološki vestnik. 2021; 72: 137–186. https://doi.org/10.3986/av.72.06

61. Macháček J, Nedoma R, Dresler P, Schulz I, Lagonik E, Johnson SM, et al. Runes from Lány (Czech Republic)—The oldest inscription among Slavs. A new standard for multidisciplinary analysis of runic bones. Journal of Archaeological Science. 2021; 127: 105333. https://doi.org/10.1016/j.jas.2021.105333

62. Bourdieu P. Outline of a Theory of Practice. Cambridge: Cambridge University Press; 1977.

63. Skibo JM, Schiffer MB. People and Things. A Behavioral Approach to Material Culture. New York: Springer; 2008.

64. Jones S. The Archaeology of Ethnicity: A Theoretical Perspective. London, New York: Routledge; 1997.

65. Brather S. Ethnische Identitäten als Konstrukte der frühgeschichtlichen Archäologie. Germania: Anzeiger der Römisch-Germanischen Kommission des Deutschen Archäologischen Instituts. 2000; 78: 139–177.

66. Hu D. Approaches to the Archaeology of Ethnogenesis: Past and Emergent Perspectives. Journal of Archaeological Research. 2012; 371–402.

67. Dzino D. Becoming Slav, Becoming Croat. Identity Transformations in Post-Roman and Early Medieval Dalmatia. Leiden, Boston: Brill; 2010.

68. Korošec P. Zgodnjesrednjeveška arheološka slika karantanskih Slovanov. Ljubljana: SAZU; 1979.

69. Eichert S. Die frühmittelalterlichen Grabfunde Kärntens: Die materielle Kultur Karantaniens anhand der Grabfunde vom Ende der Spätantike bis ins 11. Jahrhundert. Klagenfurt: Verlag des Geschichtsvereines für Kärnten; 2010.

70. Ciglenečki S. Archaeological investigations of the decline of antiquity in Slovenia. In: Bratož R, editor. Slowenien und die Nachbarländer zwischen Antike und karolingischer Epoche. Anfänge der slowenischen Ethnogenese I. Ljubljana: Narodni muzej Slovenije; 2000. pp. 119–139.

71. Predovnik KK, Nabergoj T. Archaeological research into the periods following the Early Middle Ages in Slovenia. Arheološki vestni. 2010; 61: 245–294.

72. Kosi M. Predurbane ali zgodnjeurbane naselbine?: (Civitas Petouia, Carnium/Creina in druga centralna naselja neagrarnega značaja v zgodnjem srednjem veku). Del 2. Zgodovinski časopis. 2010; 64: 8–44.

73. Grafenauer B. Ustoličevanje koroških vojvod in država karantanskih Slovencev. Ljubljana: SAZU; 1952.

74. Grafenauer B. Zgodovina Slovenskega naroda, I. zvezek: Od naselitve do uveljavljenja frankovskega fevdalnega reda. Ljubljana: Državna založba Slovenije; 1978.

75. Vilfan S. Zur Struktur der freisingischen Herrschaften südlich der Tauern in Frühmittelalter. In: Hödl G, Grabmayer J, editors. Karantanien und der Alpen-Adria-Raum im Frühmittelalter 2 St Veiter Historikergespräche. Wien: Böhlau; 1993. pp. 209–222.

76. Szameit E. Zum archäologischen Bild der frühen Slawen in Österreich. Mit Fragen zur ethnischen Bestimmung karolingerzeitlicher Gräberfelder im Ostalpenraum. In: Bratož R, editor. Slowenien und die Nachbarländer zwischen Antike und karolingischer Epoche. Anfänge der slowenischen Ethnogenese I. Ljubljana: Narodni muzej Slovenije; 2000. pp. 507–548.

77. Pavlovič D. Začetki zgodnjeslovanske poselitve Prekmurja = Beginnings of the Early Slavic settlement in the Prekmurje region, Slovenia. Arheološki vestnik. 2017; 68: 349–386.

78. Guštin M, Tifengraber G. Oblike in kronologija zgodnjesrednjeveške lončenine na Novi tabli pri Murski Soboti = Formen und Chronologie frühmittelalterlicher Keramik in Nova tabla bei Murska Sobota. In: Guštin M, editor. Zgodnji Slovani. Zgodnjesrednjeveška lončenina na obrobju vzhodnih Alp (Die frühen Slawen. Frühmittelalterliche Keramik am Rand der Ostalpen). Narodni Muzej Slovenije; 2002. pp. 46–64.

79. Novšak M. Zgodnjesrednjeveške najdbe z najdišča Grofovsko pri Murski Soboti. In: Guštin M, editor. Zgodnji Slovani. Zgodnjesrednjeveška lončenina na obrobju vzhodnih Alp (Die frühen Slawen. Frühmittelalterliche Keramik am Rand der Ostalpen). Narodni Muzej Slovenije; 2002. pp. 27–32.

80. Kerman B. Arheološka slika slovanske poselitve Prekmurja = Archaeological picture of the settlement of the Slavs in Prekmurje. In: Lux J, Štular B, Zanier K, editors. Slovani, naša dediščina = Our heritage: the Slavs. Ljubljana: Zavod za varstvo kulturne dediščine Slovenije; 2018. pp. 55–68.

81. Sussex R, Cubberley P. The Slavic Languages (Cambridge Language Surveys). Cambridge: Cambridge University Press; 2006.

82. Sobolev A. N. Voprosy drevnejshej istorii yuzhnoslavyanskikh yazykov i areal'naya lingvistika. užnoslovenski filolog. 2000; 56: 1035–1050.

83. Bezlaj F. Položaj slovenščine v okviru slovanskih jezikov. In: Furlan M, editor. Zbrani jeziskoslovni spisi. Ljubljana: Založba ZRC; 2003. pp. 268–277.

84. Bernstein SB. 1961. Ocherk sravnitel'noj grammatiki slavyanskikh yazykov, vol. 1. Moscow: Izd-vo Akademii nauk SSSR.

85. Stieber Z. 1972. O drevnikh slovensko-zapadnoslavyanskikh svyazyakh. In: Russkoe i slavyanskoe yazykoznanie. K 70-letiyu chl.-korr. AN SSSR R. I. Avanesova. Moscow: Nauka.

86. Lekov I. 1958. Znachenieto na gramaticheskite, slovoobrazovatelni i leksikalni danni za klasi- fikatsiyata na slavyanskite ezitsi ot savremenno gledische. In: Slavyanskaya filologiya. IV Mezhdunarodnyj syezd slavistov 2. Moscow: Izd-vo Akademii nauk SSSR.

87. Kushniarevich A, Utevska O, Chuhryaeva M, Agdzhoyan A, Dibirova K, Uktveryte I, et al. Genetic Heritage of the Balto-Slavic Speaking Populations: A Synthesis of Autosomal, Mitochondrial and Y-

Chromosomal Data. Calafell F, editor. PLOS ONE. 2015; 10: e0135820. https://doi.org/10.1371/journal.pone.01358200135820 PMID: 26332464

88. Wawruschka C. Genetic History and Identity: The Case of Turkey. Medieval Worlds. 2016; 4: 123–161. https://doi.org/10.1553/medievalworlds_no4_2016s123

89. Sykes B. The Seven Daughters of Eve and Blood of the Isles: Exploring the Genetic Roots of our Tribal History. London: Transworld Corgi; 2007.

90. Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, Roodenberg SA, et al. Genome-wide patterns of selection in 230 ancient Eurasians. Nature. 2015; 528(7583): 499–503. https://doi.org/10.1038/nature16152 PMID: 26595274

91. Lazaridis I, Nadel D, Rollefson G, Merrett DC, Rohland N, Mallick S, et al. Genomic insights into the origin of farming in the ancient Near East. Nature. 2016; 536(7617): 419–424. https://doi.org/10.1038/nature19310 PMID: 27459054

92. Lipson M, Cheronet O, Mallick S, Rohland N, Oxenham M, Pietrusewsky M, et al. Ancient genomes document multiple waves of migration in Southeast Asian prehistory. Science. 2018; 361: 92–95. https://doi.org/10.1126/science.aat3188 PMID: 29773666

93. Gokcumen O, Frachetti M. The Impact of Ancient Genome Studies in Archaeology. Annual Review of Anthropology. 2020; 49: 277–298. https://doi.org/10.1146/annurev-anthro-010220-074353

94. Geary P, Veeramah K. Mapping European Population Movement through Genomic Research. Medieval Worlds. 2016; 4:: 65–78. https://doi.org/10.1553/medievalworlds_no4_2016s65

95. Brather S, editor. New Questions Instead of Old Answers: Archaeological Expectations of aDNA Analysis. Medieval Worlds. 2016; 4: 22–41. https://doi.org/10.1553/medievalworlds_no4_2016s22

96. O'Sullivan N, Posth C, Coia V, Schuenemann VJ, Price TD, Wahl J, et al. Ancient genome-wide analyses infer kinship structure in an Early Medieval Alemannic graveyard. Science Advances. 2018; 4(9): eaao1262. https://doi.org/10.1126/sciadv.aao1262 PMID: 30191172

97. Amorim CEG, Vai S, Posth C, Modi A, Koncz I, Hakenbeck S, et al. Understanding 6th-century barbarian social organization and migration through paleogenomics. Nature Communications. 2018; 9. https://doi.org/10.1038/s41467-018-06024-4 PMID: 30206220

98. Pohl W. Editor's Introduction: The Genetic Challenge to Medieval History and Archaeology. P Medieval Worlds. 2016; 4: 2–4. https://doi.org/10.1553/medievalworlds_no4_2016s2

99. Samida S, Feuchter J. Why Archaeologists, Historians and Geneticists Should Work Together–and How. Medieval Worlds. 2016; 4: 5–21. https://doi.org/10.1553/medievalworlds_no4_2016s5

100. Kushniarevich A, Kassian A. Genetics and Slavic Languages. In: Greenberg ML (editor) Encyclopedia of Slavic Languages and Linguistics Online. https://doi.org/10.1163/2F2589-6229_eslo_com_032367

101. Malyarchuk BA, Derenko MV. Diversity and Structure of Mitochondrial Gene Pools of Slavs in the Ethnogenetic Aspect. Biology Bulletin Reviews. 2021; 11: 122–133. https://doi.org/10.1134/s2079086421020067

102. Malyarchuk BA, Grzybowski T, Derenko MV, Czarny J, Drobnič K, Miścicka-Śliwka D. Mitochondrial DNA variability in Bosnians and Slovenians. Annals of human genetics. 2003; 67: 412–425. https://doi.org/10.1046/j.1469-1809.2003.00042.x PMID: 12940915

103. Zupan A. Genetska struktura Slovencev, kot jo razkrivajo polimorfizmi kromosoma Y in mitohondrijske DNA. Doctoral Thesis, The University of Ljubljana. 2014. Available at: https://repozitorij.uni-lj.si/IzpisGradiva.php?id=76713.