

The Genetic Origin of the Indo-Europeans

Corresponding Author: Dr David Reich

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 1:

Reviewer comments:

Referee #1

(Remarks to the Author)

A. Summary of key results

The paper studies the genetic prehistory of the population groups that make up the Yamnaya archaeological complex, and identifies three new genetic clines that are relevant. The most important one of these is the so-called Caucasus-Lower Volga cline (CLV), since CLV people contribute 4/5 of the ancestry of the Yamnaya people. Moreover, CLV people contribute 1/10 of the ancestry of Bronze Age Anatolians. Since Yamnaya people are generally identified with the speakers of Core Proto-Indo-European, and Bronze Age Anatolians are probably speakers of Anatolian languages, the CLV people may be identified as speakers of Proto-Indo-Anatolian, dating to ca. 4400-4000 BCE.

B. Originality and significance

The results of this paper are highly original and significant.

C. Data & Methodology: validity of approach, quality of data, quality of presentation

As far as I am able to judge, the approach seems largely valid, and the quality of data is good.

However, I have problems with the sections that deal with Anatolia, and especially the question whether CLV ancestry arrived into Central Anatolia from the West ("Hypothesis A-West") or from the East ("Hypothesis A-East" / "Hypothesis B" / "Hypothesis C"). cf. suppl.materials p. 196f. In the suppl. materials it is stated that "Hypothesis A-West is not the focus of this study" (p. 197), but nevertheless the paper claims that the question about the entry of CLV ancestry into Anatolia can be answered: it claims that the data provide "a strong case" in favor of the Eastern route (line 901). This seems methodologically wrong: in a paper that contrasts two hypotheses (East vs. West), one cannot state that one of these was not the focus of the study (West), and at the same time claim that "a strong case" can be made for the other (East). Moreover, the arguments that are used against the one and in favor of the other do not seem to hold up. In suppl.mat. p. 197 it is stated that Hypothesis A-West is contradicted by the fact that, within that hypothesis, we would expect steppe migrants to have admixed with the farmers of Southeastern Europe who are "of largely Anatolian Neolithic ancestry with local Balkan hunter-gatherer admixture", but that "in Anatolia during the Chalcolithic and Bronze Age we see a decrease of Anatolian Neolithic ancestry and a lack of Balkan hunter-gatherer admixture". The first part of this argument seems a non sequitur: if a population with non-admixed Anatolian Neolithic ancestry would admix with people with an admixture of steppe ancestry and Anatolian Neolithic ancestry (as in Hypothesis A-West), the expected net result would be a decrease of Anatolian Neolithic ancestry, too. Moreover, the decrease of Anatolian Neolithic ancestry seems to have taken place at the end of the Neolithic already, i.e. too early to connect it to the influx of CLV cline genes. The second part of this argument (the lack of Balkan hunter-gatherer admixture in Anatolia) largely depends on percentages. The CLV cline signal found in Bronze Age Central Anatolians is ca. 10%. If the migrants responsible for this influx contained Balkan hunter-gatherer ancestry with a max. of, let's say, 10%, the expected percentage in the Bronze Age Central Anatolians would be below the threshold of significance any way. If the authors really want to dismiss Hypothesis A-West on archaeogenetic grounds, and thus challenge the linguistic arguments according to which Hypothesis A-West is the preferred one (as they are aware of), they have to really expand their discussion. They cannot get away with stating that this hypothesis was not the focus of their study.

Another thing that has surprised me in this regard is the shallow treatment of the samples from Western Anatolia, and specifically from Yassitepe. In the suppl.materials, Yassitepe samples are grouped with Ulucak samples under a single TUR_Aegean_BA label, and the only analysis given is that population of this label "can also be modeled in terms of Neolithic people of Western Anatolia with eastern ancestry". Yet, in the Southern Arc paper, two samples from Yassitepe were reported to have EHG ancestry (I5735, I5737), and are therefore extremely important. Would these two samples also fit a model with CLV ancestry? This needs to be treated separately. If yes, they are extremely important possible witnesses in

favor of Hypothesis A-West. (Note that on p. 213 of suppl. materials it is stated that “steppe ancestry is not seen in Western Anatolia”: this seems to be contradicted by these Yassitepe samples).

The main archaeogenetic argument in favor of an Eastern entry is that it “goes “with the grain” of the transformation of Chalcolithic and Bronze Age Anatolia” (suppl. mat. p. 197), and that the Central Anatolian Bronze Age people with CLV-admixture have “the majority of their ancestry from Mesopotamian Neolithic farmers” (main article, line 906-7). The “going with the grain” refers to the fact that in Central Anatolia there was a “major discontinuity between the Neolithic and Chalcolithic periods mediated by migrants from the east” (suppl. mat. p. 202), but this is much earlier than the spread of CLV genes can have been (which entered Anatolia after 4400-4000 BCE). There is thus no reason to connect the influx of CLV genes with this earlier gene flow from the East: they are temporally completely distinct. In order to substantiate the argument regarding the Mesopotamian Neolithic farmers, we need a much clearer and fine-grained dating of which samples show this ancestry when and where.

A puzzling aspect is the fact that this paper does not really treat the question to what extent the early west-ward migration of CLV cline people (samples in Csongrad, the making of Usatove) may be relevant to the question from where CLV cline genes entered into Anatolia. In the twin paper (“A genomic history of the North Pontic Region”) these early migrations are one of the main discoveries, but in the present paper they play hardly any role. (They are only mentioned in passing in line 678 of the main paper.)

When it comes to quality of presentation, I feel that the text of the main paper is too dense: there is an abundant use of abbreviations which makes it very difficult and slow to read. A big point of criticism is the absence of clear maps in which the archaeological complexes are presented per different time frames. It is therefore often very difficult for the reader to follow the line of thought.

E. Conclusions: robustness, validity, reliability

Overall, the robustness, validity and reliability of conclusions seem OK. I disagree, however, with the suggestion that the paper can conclude (or make a strong case for) an Eastern entry of CLV cline genes into Anatolia (and thus of an Eastern entry of the Anatolian languages into Anatolia): cf. the comments under point D above.

F. Suggested improvements: experiments, data for possible revision

The paper should include a much more detailed discussion of the “Hypothesis A-West” regarding the point of entry of the Anatolian languages into Anatolia (through the lens of the point of entry of CLV cline genes into Central Anatolia). The paper should: give a broader and clearer analysis of Western Anatolian samples; make a clearer distinction in the different time periods in which different gene flows took place in Anatolia; make clearer analyses for answering the question how and in what way it can be excluded that there was a Western Entry; make a clearer link with the early westward migrations of CLV cline people (towards Hungary, Usatove) as discussed in its twin paper; see point D above for all these points.

G. References: appropriate credit to previous work?

As far as I am aware, this is largely OK.

H. Clarity and context: lucidity of abstract/summary, appropriateness of abstract, introduction and conclusions.

The text is extremely dense, and thus lacks lucidity. The number of abbreviations is very high, which hampers the clarity. It would be extremely helpful if the article or suppl. materials would contain a clear map with the locations of different archaeological complexes and sites per time period. I often also miss the datings of specific samples that are mentioned (e.g. “the individual from Vonyucka-1 in the North Caucasus”, line 864: without giving a date for this sample, it is difficult to judge its significance). To my mind, the title of the paper is not well chosen: the term “Indo-Europeans” imply some sort of ethnicity that cannot be substantiated. Better would be “the speakers of Indo-European languages”. Moreover, in the article itself, the term “Indo-Anatolian” is used to refer to the highest ancestral language of the family, and the article in fact deals with this “Indo-Anatolian” homeland. I object to the statement “entering Anatolia from the east” in the abstract (line 24) since an eastern entry has, to my mind, not been sufficiently substantiated. This specific point therefore cannot be represented as a fact in the abstract: in that sense it is not appropriate. The last sentences of the paper are of a semi-philosophical nature, and diminish the strength of the paper.

Referee #2

(Remarks to the Author)

Key results: This study represents a massive new effort to understand the complex genetic, archaeological, and thus also linguistic prehistory of early Indo-Europeans. 372 new genomes, and 198 new C14 dates, allow a level of genetic, chronological and archaeological resolution never achieved before. It integrates all data reported in the North Pontic paper, plus the new expansive dataset reported here. And it resolves a long- standing enigma about the origin, formation, and divergence of Indo-Anatolian and Indo-European by documenting the constituting genetic and archaeological clines that contributed to their formation.

Validity: There are of course still lacunas that needs to be filled especially concerning Mesopotamia, and its interaction with Caucasus and Anatolia in the formation of Indo-Anatolian, but even with that reservation the Indo-Anatolian part of the paper looks convincing as the best interpretation so far. If substantiated it resolves a long-standing debate about a western versus northeastern/Caucasian origin/entry of Indo- Anatolian. What is truly remarkable, however, is the documentation of three new Eneolithic genetic clines, and their contribution to the formation of genetic, archaeological and linguistic formations of Indo-Anatolia/Indo-European. All summarized in a well presented/pedagogical Figure 1 and subsequent figures. However, the concrete discussion of each cline is really the strength of the paper, as it details and integrates the genetic variation along

the cline with archaeological sites/cultures in a convincing way. And it discusses alternative admixture models/interpretations, and leaves interpretations open if needed – e.g. the proposal that there is still an unsampled population to form Core Yamnaya. But the discussion also evolves with each cline, adding new evidence in a cumulative way to the final interpretation.

The uniqueness of the Yamnaya expansion from their core area, is illustrated both by demographic growth curves, as well as IBD and family sharing, which supports the maintenance of their genetic and social homogeneity. This must have been an essential part of their linguistic dominance.

The discussion of the origin and spread of Indo-Anatolian is very well argued both from a genetic/archaeological point of view and from a linguistic point of view. Many details will be sorted out by future research, but the main picture now presented looks robust, and thus solve 150 years of debate, at least in the main, as well as refutes a recent Hegarty et al. paper in Science.

I only have a critical comment on the dates of Seredny Stiih, and several other important samples, which are not corrected for marine reservoir effect, and which could affect their role in the formation of Core Yamnaya. These dates should be calibrated for marine reservoir effect, when possible.

Originality and significance: the study is truly groundbreaking, in that it allows for the first time a convincing, interpretive integration of genetics, archaeology and linguistics in the formation of early IE languages, whose 'roots' and early formation has remained in the dark. Not any longer. And even if the last word is never said, we now have a new foundation to build on.

Summary, conclusions fine

Referee #3

(Remarks to the Author)

Comments

We cannot escape from feeling uncomfortable reading a paper with massive contribution from Russian scientists at the time when Russia's aggression war puts Ukrainian scientists into the drenches or fleeing from their homeland. As a minimum, a statement by the authors, condemning Russia's aggression war, should be included. We understand this is just our opinion and such editorial decisions are with the journal.

We also understand that the authors have split what would have been one paper into two so that Ukrainian and Russian authors would not have to share authorship. We understand that it would be unacceptable for the Ukrainian researchers to appear on the same paper with Russian institutions, many of which, as per by a statement by Russian university rectors, have officially supported Russia's war in Ukraine.

The question of the genetic origins of the Indo-European speaking people has been a hot topic for a long time. Here the authors sum up a massive project which has (according to the lead author's public posting) been 10 years in the making. It is indeed an enormous effort. The two combined studies double the number Eneolithic samples (79 in addition to 82 previously published) from the steppe region and add 286 Yamnaya/Afanasievo individuals compared to 75 in the literature. This major advancement in the sampling enables this paper to set out to identify the genetic ancestors of the Yamnaya people who played the central role in reshaping the genetic and linguistic landscape of Western Eurasia in the late Bronze age. Methodologically the paper offers a new two-step "tournament" approach to filter/rank fitting qpAdm models in order to choose between competing models for the same population. This is presented as a logical thought process but not tested as a method. Not sure if there is a sensible way to do it though. Nevertheless, some additional discussion on the matter would perhaps be of help. For example, the suggestion that the fortitude criterion helps to reduce the non-falsifiability bias toward source populations with few data could be tested and shown? In this regard we were somewhat confused by the discussion around line 493 where the choice of GK2 over UNHG seems to not benefit from the fortitude criterion. Perhaps Tournament was only held for two-way models?

The paper suggests modelling Eneolithic population mixtures that led to the formation of the Yamnaya group in terms of three clines. For clarity, it could be a good idea to follow the same order of describing the clines both in the abstract (currently: CVL-Volga-Dnipro) and the main text (currently: Volga-Dnipro-CLV).

Furthermore, a cline is intuitively perceived as a linear representation of mixing of two differentiated sources. In this regard the Volga cline, as it is suggested, does not seem like a cline. Specifically, the "bend" in the cline on the PCA does not make sense. Rather the cline seems to be between BP and EHG. Moreover, panel C of Figure 2 depicts what we suggest. Volga cline is between BP and EHG and another cline is between UNHG and EHG. Also, text around line 185 seems to tell, that Upper Volga (Murzikha) groups do not fit well with BP-EHG mixing model. Looking back at the PCA, and anticipating the position of WHG, used in PC space calculation but not shown on the plot, and Baltic Mesolithic HG, for some reason not used at all, we tentatively propose that we could be seeing an "out-of-BP/PV" FAN – an array of (sampled and unsampled) radiating clines into a WHG -Siberia cline that includes Baltic HG, UNHG, Karelia EHG and so on (see also line 317). In this setting the Dnipro cline would be one of these reaching toward UNHG.

One impressive result of the massive sampling is the ability to see the temporal dynamics of admixture. For example, along the Volga cline – line 200. These results could be visualized perhaps? Especially if similar temporal patterns are observed more systematically. Much like the representation of temporal change of the Y chromosomes.

The principal result, that the paper sets out to reach, presented around line 500, deserves a summarizing simple figure, or at least a clear panel in the summarizing figure. Moreover, wouldn't it be a good idea to include a qpGraph-like summary figure of the complex interactions presented?

Overall, the level of novelty and implications of the results warrant publication in Nature. However, the manuscript would benefit from increased clarity and shortening (of the main text). The clear and cool focus of the paper is the genetic origin of the Yamnaya people. Other lines of enquiry/discussion could be condensed in the main text and opened in more detail in the supplement. The authors extend the discussion from origins of the Yamnaya to the origins of IE/IA languages half-way through the paper. This comes little bit out of the blue. Maybe structure the whole paper more clearly with clear subtitles? – origins of Yamnaya, origins of IE.

The weakest aspect of the manuscript are the figures. It is undoubtedly very challenging to come up with visual representations conveying the complexity and yet not drown in it. For example, it has been quite difficult to connect the (not always consistent) labels in the legends and figures. For some figures we have more specific comments below but overall, we feel there are ways to increase the “readability” of the visual narrative.

More specific comments and suggestions:

General comment on samples: The authors report a few new samples from Ukraine. However, the institution responsible for these samples is in Moscow, Russia. We wonder if there was any contribution from Ukrainian archaeologists in those excavations?

Summary Figure: A few of the “other sites of interest” don't have site names. Also, it's quite difficult to determine from the figure, which period(s) the samples of this study are from. Finally, the placement of the “Roots of Caucasus-Lower Volga people” text is extremely awkward – maybe it would fit better on the bottom of the figure on top of Southeastern Europe and Anatolia? Overall the Summary figure does not “tell the story” of the manuscript too well. Lots of text. Probably there should be visual and spatial representations of the clines and a schematic of the genetic makeup of the Yamnaya.

Line 43: “The first formal study of the origins of the Yamnaya identified two disparate sources of ancestry” this wording omits the contribution of Allentoft et al from the same year, who also test for admixture sources in the Yamnaya and with the absence of EHG in their data identify Mal'ta in addition to BA Armenia.

Lines 49–52: Currently, the question after the second point came very unexpectedly so maybe it would be better to unify the style and rewrite the first as an actual question as well.

Lines 71–74: The numbers of newly reported samples and those with increased data are somewhat confusing. The same samples should not be mentioned as newly reported in two manuscripts so the first sentence should be removed and the data from 299 + 55 individuals presented as newly reported and then mentioned that it's co-analysed with data from 73 + 5 individuals from Nikitin et al.

Lines 108 and 120: The use of pronouns (“these questions”, “this note of caution”) creates confusion given that the questions/note of caution were not in the previous paragraph, but earlier in the text.

Line 145: Should be “Fig. 1ab” since it's a co-description of geography and a PCA cline.

Line 152: It would be clearer and more precise to write “genetically unsampled”.

Line 161 and elsewhere: There are no references to the tables where qpAdm results are presented.

Lines 164–166: The sentence “currently dated earliest in the Samara region at Lebyazhinka-5 and in a few graves at Ekaterinovka dated before 4500 BCE” is not very clear – is Lebyazhinka-5 the earliest in the Samara region and the custom also exists in some Ekaterinovka graves dated before 4500 BCE or are Lebyazhinka-5 and Ekaterinovka both in the Samara region and the graves in both are dated before 4500 BCE (which is the earliest)? Also, is the Samara region the earliest with this custom or are these graves the earliest in the Samara region with this custom (but not elsewhere)?

Lines 169–171: It seems strange to refer to results that are presented later in the manuscript.

Lines 184–185: “Seven Volga cline populations fit this model (p-values of 0.04 to 0.72)” – the commonly used p-value threshold for a fitting model is 0.05 so it would be good to at least specify how many models have a p-value higher than 0.05.

Lines 188–189: It would be clearer to write “in the upriver portion of the Volga cline, between EHG and UNHG”.

Lines 196–197 and elsewhere: There are a lot of sites and ancestry sources so it would be better to stick with one label for each (BPgroup vs Berezhnovka).

Figure 1: It would be much clearer if the legends were grouped and ordered based on geography not listed alphabetically. It is not clear which groups are from this study, which are from Nikitin et al. and which previously published. The symbols for Murzikha and Ekaterinovka are extremely similar. The group labels don't fully match between different panels.

Panel a: The labels on panel a are too small and too many. Information is lost in complexity. Maybe put the full sample location map in supplementary as a larger high-resolution image and figure out how to group labels (keeping all sample locations) for the main figure?

Panel b: It's unclear what determines the font size differences on panel b. It appears that the samples used for PC space calculation are not on the plot? Would be very helpful if they were there. Also, the choice of populations for the calculation of PCs on not immediately clear and should be explained.

Panel c: The legend of panel c, with individual sample names, is suitable for a supplementary figure. The clines in panel c could for example be presented on a map (three panels if they would otherwise become too overlapping spatially).

Overall, this figure is extremely hard to follow and needs substantial work.

Line 261: Does “the Eneolithic individual from the area” mean Krivyansky? If so, then it's only possible to compare the Don Yamnaya individuals to this one on the PCA which doesn't allow for conclusions like “bear no affinity”.

Lines 292–309: It is unclear why the modelling switches between EHG and GK2.

Line 322: It would be clearer if it said “three Golubaya Krinitza individuals”.

Line 354: It's not completely clear what “their” refers to here.

Lines 390–392: Again, it seems strange to “spoil” the result of the next section.

Line 398: “Unplotted individual all give fits at $P < 0.05$ ” – should be “individuals” and “ $p < 0.05$ ”.

Figure 2: The placement of the panel labels is a bit strange. Again, it would be much clearer if the legends were grouped and ordered based on geography not listed alphabetically.

Line 424: The Dnipro cline individuals already have more than one ancestry source.

Lines 446–447: We usually know the geographic origin of a sample, but not the genetic ancestry without DNA analyses so this sentence would make more sense the other way around.

Lines 453–463: The jump from Meshoko and Svobodnoe to Unakozovskaya is confusing and unclear.

Lines 469–470: It seems more appropriate to put the full “high/low steppe ancestry” in quotation marks.

Line 476: It would be more precise to use “ends” instead of “edges”.

Lines 505–506: It's unclear what's meant by “the remaining four fifths of their ancestry”.

Line 535: There's a typo – should be “Armenia”.

Lines 538–539: Where is the Mesopotamian-Anatolian modelling shown? Also, what are the references after the end of the sentence/paragraph?

Line 554: As it is also stressed above in the manuscript, it is not possible to determine the exact source of any ancestry precisely with aDNA methods.

Lines 559–566: It is unclear which specific models are tested in this section. Are the first ones 3-way models of steppe source 1 + steppe source 2 + EHG/UNHG/GK2? But then how is the unknown admixing population's placement on the CLV cline tested?

Line 569: There's a typo – should be “Anatolian”.

Lines 587–588: There's no Extended Data Figure 2f and it is unclear which exact populations this admixture date is calculated for.

Line 596: The significance of the site of Armi is unclear to a reader not specialised in Anatolia.

Figure 3: The haplogroup labels should be moved (for example placed vertically at the left side of the plot) since they're currently difficult to see (especially G-P15). It is unclear which regions “EarlyMiddleBronzeAge” and “MiddleLateBronzeAge” comprise. Does EarlyMiddleBronzeAge include all samples from the listed cultures during the relevant time period? Which regions are the MiddleLateBronzeAge samples from? If the geographic groups in the legend don't include all samples from the region (like in the case of North Caucasus, for example), a specification should be added.

Line 630: For clarity, it would be better to write “the Caucasus Neolithic represented by Aknashen was formed by the admixture of ...”.

Lines 640–653: This paragraph is confusing as it jumps from the EHG/UNHG vs West Asia/Caucasus differentiation to the Dnipro cline and then back to the EHG/UNHG + West Asia/Caucasus admixture date.

Lines 655–657: It doesn’t seem contrasting that while it’s not possible to pinpoint the geographical origin of a Yamnaya individual based on their genetic profile, it’s also not clear exactly where their ancestors lived.

Lines 669–670: It should be stated accordingly in case this an assumption or cited if not.

Line 673: The word “there” is not necessary here and creates confusion.

Lines 679–680: It would be helpful to state how this was tested.

Lines 720–721: Allentoft et al. 2015 should also be cited here.

Line 724: For clarity, it would be good to replace “it” with “the Corded Ware population”.

Figure 4: It is stated that the individuals are projected in the same space as on Figure 1, but the PC2 axis seems to be flipped. The Core Yamnaya symbols on panel a are very similar (especially given their small size), rendering them meaningless. UNHG is not on panel a so it’s not clear directly from here that the Don Yamnaya tend towards it.

Lines 748–751: This sentence is a bit complicated and doesn’t seem fully grammatically correct.

Lines 756–758: It would be clearer if Yamnaya was mentioned here.

Line 763: Please specify what “shading shows uncertainty intervals” means.

Figure 5: Panels a and b make it seem like the Ne trajectories of the two sub-populations are very different due to the different y axis scales. It is understandably difficult to present these, but maybe having a gap in the y axis would be clearer. Also, the y axis scale of panel c is quite unusual and difficult to interpret. In sum it seems that panels a and b are perhaps redundant, and it would suffice to clarify the y axis scale on panel c. It seems that this is the only analysis where the Core Yamnaya is split into two temporal groups. It would be great to see the geographic distribution of these two groups. Furthermore, we are wondering what would happen if one would sample an expanding population in two time points, so that the later sample has seen more of the expansion. Wouldn’t the later population show a steeper Ne increase simply because it is a temporal subset not a population subset of the original population? Simply because also in an expanding population not all members of the population take part of the expansion?

Line 772: It would be helpful to be a bit more specific about what “before the Yamnaya” means.

Line 777: Please specify what “close genetic relatives” means here in terms of degrees of relatedness.

Figure 6: The panel d label is missing from the figure legend. The fonts are too small on panels c–f and on the x axis of panel g.

Line 826: Would be useful to also explain how Yamnaya can be linked to Balkan languages.

Lines 839–840: This sentence could be written in a more neutral way.

Lines 868–870: Please add references for the horse-wagon technology and for Proto-Anatolian languages being recorded in writing.

Lines 911–912: There seems to be a grammatical issue (“have” and “had”), but the sentence is also quite complex and difficult to understand so might be worth trying to rephrase.

Lines 914–916: It’s not clear how/why the Kura-Araxes expansion would’ve “driven a wedge” between steppe and West Asian Indo-Anatolian speakers.

Lines 934–938: These two questions are fully left in the air without any reference to them in the concluding sentences.

Extended Data Figure 1: It should be specified that panel a includes Eastern models, together with an explanation of what that means. It should be added that panel b models also include Cayönü next to western sources. The p-value threshold labels should be the same size on panel c as on previous panels. It should be added, which model the BPgroup proportion in panel d comes from. It is not clear how the pairs of populations mentioned form a route on panel f.

Extended Data Figure 2: Presumably DATES was not run using all European hunter-gatherers and all West Asian populations? In that case it should be specified which populations were used. The same applies to other panels where the

description is not specific.

Extended Data Figure 3: The label "Lebyazhinka_HG" is placed on top of the barplot, blocking data from view.

Supplementary Fig 2.7: There is what seems a state border between Ukraine and Crimean Peninsula. This should be removed.

Archaeological site information (Archaeological Supplement and Online Table 1):

Progress-2 and Vonyuchka are described in the Archaeological Supplement but are not in Online Table 1 – should be made clearer why they're described in the Archaeological Supplement.

Remontnoye Late Steppe Eneolithic site names are not in the table of contents of the Archaeological Supplement which makes it hard to link it with Online Table 1.

Kurmanaevka-3 site is not included in the Locality column of Online Table 1.

Krivyansky-7 sample is listed as Krivyansky-9 in the Locality column of Online Table 1.

Archaeological Supplement descriptions are missing for many sites.

Referee #4

(Remarks to the Author)

We cannot escape from feeling uncomfortable reading a paper with massive contribution from Russian scientists at the time when Russia's aggression war puts Ukrainian scientists into the drenches or fleeing from their homeland. As a minimum, a statement by the authors, condemning Russia's aggression war, should be included. We understand this is just our opinion and such editorial decisions are with the journal.

We also understand that the authors have split what would have been one paper into two so that Ukrainian and Russian authors would not have to share authorship. We understand that it would be unacceptable for the Ukrainian researchers to appear on the same paper with Russian institutions, many of which, as per by a statement by Russian university rectors, have officially supported Russia's war in Ukraine.

The question of the genetic origins of the Indo-European speaking people has been a hot topic for a long time. Here the authors sum up a massive project which has (according to the lead author's public posting) been 10 years in the making. It is indeed an enormous effort. The two combined studies double the number Eneolithic samples (79 in addition to 82 previously published) from the steppe region and add 286 Yamnaya/Afanasievo individuals compared to 75 in the literature. This major advancement in the sampling enables this paper to set out to identify the genetic ancestors of the Yamnaya people who played the central role in reshaping the genetic and linguistic landscape of Western Eurasia in the late Bronze age.

Methodologically the paper offers a new two-step "tournament" approach to filter/rank fitting qpAdm models in order to choose between competing models for the same population. This is presented as a logical thought process but not tested as a method. Not sure if there is a sensible way to do it though. Nevertheless, some additional discussion on the matter would perhaps be of help. For example, the suggestion that the fortitude criterion helps to reduce the non-falsifiability bias toward source populations with few data could be tested and shown? In this regard we were somewhat confused by the discussion around line 493 where the choice of GK2 over UNHG seems to not benefit from the fortitude criterion. Perhaps Tournament was only held for two-way models?

The paper suggests modelling Eneolithic population mixtures that led to the formation of the Yamnaya group in terms of three clines. For clarity, it could be a good idea to follow the same order of describing the clines both in the abstract (currently: CVL-Volga-Dnipro) and the main text (currently: Volga-Dnipro-CLV).

Furthermore, a cline is intuitively perceived as a linear representation of mixing of two differentiated sources. In this regard the Volga cline, as it is suggested, does not seem like a cline. Specifically, the "bend" in the cline on the PCA does not make sense. Rather the cline seems to be between BP and EHG. Moreover, panel C of Figure 2 depicts what we suggest. Volga cline is between BP and EHG and another cline is between UNHG and EHG. Also, text around line 185 seems to tell, that Upper Volga (Murzikha) groups do not fit well with BP-EHG mixing model. Looking back at the PCA, and anticipating the position of WHG, used in PC space calculation but not shown on the plot, and Baltic Mesolithic HG, for some reason not used at all, we tentatively propose that we could be seeing an "out-of-BP/PV" FAN – an array of (sampled and unsampled) radiating clines into a WHG -Siberia cline that includes Baltic HG, UNHG, Karelia EHG and so on (see also line 317). In this setting the Dnipro cline would be one of these reaching toward UNHG.

One impressive result of the massive sampling is the ability to see the temporal dynamics of admixture. For example, along the Volga cline – line 200. These results could be visualized perhaps? Especially if similar temporal patterns are observed more systematically. Much like the representation of temporal change of the Y chromosomes.

The principal result, that the paper sets out to reach, presented around line 500, deserves a summarizing simple figure, or at least a clear panel in the summarizing figure. Moreover, wouldn't it be a good idea to include a qpGraph-like summary figure of the complex interactions presented?

Overall, the level of novelty and implications of the results warrant publication in Nature. However, the manuscript would benefit from increased clarity and shortening (of the main text). The clear and cool focus of the paper is the genetic origin of the Yamnaya people. Other lines of enquiry/discussion could be condensed in the main text and opened in more detail in

the supplement. The authors extend the discussion from origins of the Yamnaya to the origins of IE/IA languages half-way through the paper. This comes little bit out of the blue. Maybe structure the whole paper more clearly with clear subtitles? – origins of Yamnaya, origins of IE.

The weakest aspect of the manuscript are the figures. It is undoubtedly very challenging to come up with visual representations conveying the complexity and yet not drown in it. For example, it has been quite difficult to connect the (not always consistent) labels in the legends and figures. For some figures we have more specific comments below but overall, we feel there are ways to increase the “readability” of the visual narrative.

More specific comments and suggestions:

General comment on samples: The authors report a few new samples from Ukraine. However, the institution responsible for these samples is in Moscow, Russia. We wonder if there was any contribution from Ukrainian archaeologists in those excavations?

Summary Figure: A few of the “other sites of interest” don’t have site names. Also, it’s quite difficult to determine from the figure, which period(s) the samples of this study are from. Finally, the placement of the “Roots of Caucasus-Lower Volga people” text is extremely awkward – maybe it would fit better on the bottom of the figure on top of Southeastern Europe and Anatolia? Overall the Summary figure does not “tell the story” of the manuscript too well. Lots of text. Probably there should be visual and spatial representations of the clines and a schematic of the genetic makeup of the Yamnaya.

Line 43: “The first formal study of the origins of the Yamnaya identified two disparate sources of ancestry” this wording omits the contribution of Allentoft et al from the same year, who also test for admixture sources in the Yamnaya and with the absence of EHG in their data identify Mal’ta in addition to BA Armenia.

Lines 49–52: Currently, the question after the second point came very unexpectedly so maybe it would be better to unify the style and rewrite the first as an actual question as well.

Lines 71–74: The numbers of newly reported samples and those with increased data are somewhat confusing. The same samples should not be mentioned as newly reported in two manuscripts so the first sentence should be removed and the data from 299 + 55 individuals presented as newly reported and then mentioned that it’s co-analysed with data from 73 + 5 individuals from Nikitin et al.

Lines 108 and 120: The use of pronouns (“these questions”, “this note of caution”) creates confusion given that the questions/note of caution were not in the previous paragraph, but earlier in the text.

Line 145: Should be “Fig. 1ab” since it’s a co-description of geography and a PCA cline.

Line 152: It would be clearer and more precise to write “genetically unsampled”.

Line 161 and elsewhere: There are no references to the tables where qpAdm results are presented.

Lines 164–166: The sentence “currently dated earliest in the Samara region at Lebyazhinka-5 and in a few graves at Ekaterinovka dated before 4500 BCE” is not very clear – is Lebyazhinka-5 the earliest in the Samara region and the custom also exists in some Ekaterinovka graves dated before 4500 BCE or are Lebyazhinka-5 and Ekaterinovka both in the Samara region and the graves in both are dated before 4500 BCE (which is the earliest)? Also, is the Samara region the earliest with this custom or are these graves the earliest in the Samara region with this custom (but not elsewhere)?

Lines 169–171: It seems strange to refer to results that are presented later in the manuscript.

Lines 184–185: “Seven Volga cline populations fit this model (p-values of 0.04 to 0.72)” – the commonly used p-value threshold for a fitting model is 0.05 so it would be good to at least specify how many models have a p-value higher than 0.05.

Lines 188–189: It would be clearer to write “in the upriver portion of the Volga cline, between EHG and UNHG”.

Lines 196–197 and elsewhere: There are a lot of sites and ancestry sources so it would be better to stick with one label for each (BPgroup vs Berezhnovka).

Figure 1: It would be much clearer if the legends were grouped and ordered based on geography not listed alphabetically. It is not clear which groups are from this study, which are from Nikitin et al. and which previously published. The symbols for Murzikha and Ekaterinovka are extremely similar. The group labels don’t fully match between different panels.

Panel a: The labels on panel a are too small and too many. Information is lost in complexity. Maybe put the full sample location map in supplementary as a larger high-resolution image and figure out how to group labels (keeping all sample locations) for the main figure?

Panel b: It’s unclear what determines the font size differences on panel b. It appears that the samples used for PC space calculation are not on the plot? Would be very helpful if they were there. Also, the choice of populations for the calculation of PCs on not immediately clear and should be explained.

Panel c: The legend of panel c, with individual sample names, is suitable for a supplementary figure. The clines in panel c

could for example be presented on a map (three panels if they would otherwise become too overlapping spatially). Overall, this figure is extremely hard to follow and needs substantial work.

Line 261: Does “the Eneolithic individual from the area” mean Kriviansky? If so, then it’s only possible to compare the Don Yamnaya individuals to this one on the PCA which doesn’t allow for conclusions like “bear no affinity”.

Lines 292–309: It is unclear why the modelling switches between EHG and GK2.

Line 322: It would be clearer if it said “three Golubaya Krinitza individuals”.

Line 354: It’s not completely clear what “their” refers to here.

Lines 390–392: Again, it seems strange to “spoil” the result of the next section.

Line 398: “Unplotted individual all give fits at $P < 0.05$ ” – should be “individuals” and “ $p < 0.05$ ”.

Figure 2: The placement of the panel labels is a bit strange. Again, it would be much clearer if the legends were grouped and ordered based on geography not listed alphabetically.

Line 424: The Dnipro cline individuals already have more than one ancestry source.

Lines 446–447: We usually know the geographic origin of a sample, but not the genetic ancestry without DNA analyses so this sentence would make more sense the other way around.

Lines 453–463: The jump from Meshoko and Svobodnoe to Unakozovskaya is confusing and unclear.

Lines 469–470: It seems more appropriate to put the full “high/low steppe ancestry” in quotation marks.

Line 476: It would be more precise to use “ends” instead of “edges”.

Lines 505–506: It’s unclear what’s meant by “the remaining four fifths of their ancestry”.

Line 535: There’s a typo – should be “Armenia”.

Lines 538–539: Where is the Mesopotamian-Anatolian modelling shown? Also, what are the references after the end of the sentence/paragraph?

Line 554: As it is also stressed above in the manuscript, it is not possible to determine the exact source of any ancestry precisely with aDNA methods.

Lines 559–566: It is unclear which specific models are tested in this section. Are the first ones 3-way models of steppe source 1 + steppe source 2 + EHG/UNHG/GK2? But then how is the unknown admixing population’s placement on the CLV cline tested?

Line 569: There’s a typo – should be “Anatolian”.

Lines 587–588: There’s no Extended Data Figure 2f and it is unclear which exact populations this admixture date is calculated for.

Line 596: The significance of the site of Armi is unclear to a reader not specialised in Anatolia.

Figure 3: The haplogroup labels should be moved (for example placed vertically at the left side of the plot) since they’re currently difficult to see (especially G-P15). It is unclear which regions “EarlyMiddleBronzeAge” and “MiddleLateBronzeAge” comprise. Does EarlyMiddleBronzeAge include all samples from the listed cultures during the relevant time period? Which regions are the MiddleLateBronzeAge samples from? If the geographic groups in the legend don’t include all samples from the region (like in the case of North Caucasus, for example), a specification should be added.

Line 630: For clarity, it would be better to write “the Caucasus Neolithic represented by Aknashen was formed by the admixture of ...”.

Lines 640–653: This paragraph is confusing as it jumps from the EHG/UNHG vs West Asia/Caucasus differentiation to the Dnipro cline and then back to the EHG/UNHG + West Asia/Caucasus admixture date.

Lines 655–657: It doesn’t seem contrasting that while it’s not possible to pinpoint the geographical origin of a Yamnaya individual based on their genetic profile, it’s also not clear exactly where their ancestors lived.

Lines 669–670: It should be stated accordingly in case this an assumption or cited if not.

Line 673: The word “there” is not necessary here and creates confusion.

Lines 679–680: It would be helpful to state how this was tested.

Lines 720–721: Allentoft et al. 2015 should also be cited here.

Line 724: For clarity, it would be good to replace “it” with “the Corded Ware population”.

Figure 4: It is stated that the individuals are projected in the same space as on Figure 1, but the PC2 axis seems to be flipped. The Core Yamnaya symbols on panel a are very similar (especially given their small size), rendering them meaningless. UNHG is not on panel a so it's not clear directly from here that the Don Yamnaya tend towards it.

Lines 748–751: This sentence is a bit complicated and doesn't seem fully grammatically correct.

Lines 756–758: It would be clearer if Yamnaya was mentioned here.

Line 763: Please specify what “shading shows uncertainty intervals” means.

Figure 5: Panels a and b make it seem like the N_e trajectories of the two sub-populations are very different due to the different y axis scales. It is understandably difficult to present these, but maybe having a gap in the y axis would be clearer. Also, the y axis scale of panel c is quite unusual and difficult to interpret. In sum it seems that panels a and b are perhaps redundant, and it would suffice to clarify the y axis scale on panel c. It seems that this is the only analysis where the Core Yamnaya is split into two temporal groups. It would be great to see the geographic distribution of these two groups. Furthermore, we are wondering what would happen if one would sample an expanding population in two time points, so that the later sample has seen more of the expansion. Wouldn't the later population show a steeper N_e increase simply because it is a temporal subset not a population subset of the original population? Simply because also in an expanding population not all members of the population take part of the expansion?

Line 772: It would be helpful to be a bit more specific about what “before the Yamnaya” means.

Line 777: Please specify what “close genetic relatives” means here in terms of degrees of relatedness.

Figure 6: The panel d label is missing from the figure legend. The fonts are too small on panels c–f and on the x axis of panel g.

Line 826: Would be useful to also explain how Yamnaya can be linked to Balkan languages.

Lines 839–840: This sentence could be written in a more neutral way.

Lines 868–870: Please add references for the horse-wagon technology and for Proto-Anatolian languages being recorded in writing.

Lines 911–912: There seems to be a grammatical issue (“have” and “had”), but the sentence is also quite complex and difficult to understand so might be worth trying to rephrase.

Lines 914–916: It's not clear how/why the Kura-Araxes expansion would've “driven a wedge” between steppe and West Asian Indo-Anatolian speakers.

Lines 934–938: These two questions are fully left in the air without any reference to them in the concluding sentences.

Extended Data Figure 1: It should be specified that panel a includes Eastern models, together with an explanation of what that means. It should be added that panel b models also include Cayönü next to western sources. The p-value threshold labels should be the same size on panel c as on previous panels. It should be added, which model the BPgroup proportion in panel d comes from. It is not clear how the pairs of populations mentioned form a route on panel f.

Extended Data Figure 2: Presumably DATES was not run using all European hunter-gatherers and all West Asian populations? In that case it should be specified which populations were used. The same applies to other panels where the description is not specific.

Extended Data Figure 3: The label “Lebyazhinka_HG” is placed on top of the barplot, blocking data from view.

Supplementary Fig 2.7: There is what seems a state border between Ukraine and Crimean Peninsula. This should be removed.

Archaeological site information (Archaeological Supplement and Online Table 1):

Progress-2 and Vonyuchka are described in the Archaeological Supplement but are not in Online Table 1 – should be made clearer why they're described in the Archaeological Supplement.

Remontnoye Late Steppe Eneolithic site names are not in the table of contents of the Archaeological Supplement which makes it hard to link it with Online Table 1.

Kurmanaevka-3 site is not included in the Locality column of Online Table 1.

Krivyansky-7 sample is listed as Krivyansky-9 in the Locality column of Online Table 1. Archaeological Supplement descriptions are missing for many sites.

Version 2:

Reviewer comments:

Referee #1

(Remarks to the Author)

I am glad to see that the authors have made several adjustments to the paper based on the review, and in that sense I am over all satisfied with the new version. However, there is one things that still needs to be adapted: the statement "entering Anatolia from the east" in line 24. In the response to the review, the authors themselves state that "People with CLV ancestry may have taken any route to reach Anatolia and from the genetic point of view we cannot distinguish between CLV people in the Caucasus-Lower Volga region and similar people who reached Moldova and Hungary", and that a western origin of CLV ancestry into Central Anatolia cannot be excluded. It therefore follows that the paper cannot simply state that CLV arrived in Central Anatolia "entering Anatolia from the east" as a fact. I therefore maintain that this statement should be adapted to, for instance, "entering Anatolia probably from the east".

Referee #2

(Remarks to the Author)

My concern about PRE correction has been adressed in away way that I find acceptable. Their position is now made clear. Thus, no more objections, and recommendation for publication without any further corrections.

Referee #3

(Remarks to the Author)

The authors have revised the manuscript largely to our satisfaction.

We do see that the authors have considered our suggestions regarding the figures but have reached the conclusion that the original ones were in fact best options. We have to agree to disagree on that. However, we don't think this warrants further delay in publication.

Beyond figure design, we noticed (Regarding Figure 1) non-matching labels not symbols. For example, Maykop in panels a and b versus Maikop in c; Karelia in the top portion of panel c and Russia_Karelia in the middle portion. Also, we would like to re-iterate that it would be beneficial to include a distinction between previously published, Nikitin et al. and your groups.

Referee #4

(Remarks to the Author)

The authors have revised the manuscript largely to our satisfaction.

We do see that the authors have considered our suggestions regarding the figures but have reached the conclusion that the original ones were in fact best options. We have to agree to disagree on that. However, we don't think this warrants further delay in publication.

Beyond figure design, we noticed (Regarding Figure 1) non-matching labels not symbols. For example, Maykop in panels a and b versus Maikop in c; Karelia in the top portion of panel c and Russia_Karelia in the middle portion. Also, we would like to re-iterate that it would be beneficial to include a distinction between previously published, Nikitin et al. and your groups.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Referees' comments:

Referee #1 (Remarks to the Author):

A. Summary of key results

The paper studies the genetic prehistory of the population groups that make up the Yamnaya archaeological complex, and identifies three new genetic clines that are relevant. The most important one of these is the so-coined Caucasus-Lower Volga cline (CLV), since CLV people contribute 4/5 of the ancestry of the Yamnaya people. Moreover, CLV people contribute 1/10 of the ancestry of Bronze Age Anatolians. Since Yamnaya people are generally identified with the speakers of Core Proto-Indo-European, and Bronze Age Anatolians are probably speakers of Anatolian languages, the CLV people may be identified as speakers of Proto-Indo-Anatolian, dating to ca. 4400-4000 BCE.

This is an accurate summary of our research findings.

B. Originality and significance

The results of this paper are highly original and significant.

C. Data & Methodology: validity of approach, quality of data, quality of presentation

As far as I am able to judge, the approach seems largely valid, and the quality of data is good. However, I have problems with the sections that deal with Anatolia, and especially the question whether CLV ancestry arrived into Central Anatolia from the West ("Hypothesis A-West") or from the East ("Hypothesis A-East" / "Hypothesis B" / "Hypothesis C"). cf. suppl.materials p. 196f. In the suppl. materials it is stated that "Hypothesis A-West is not the focus of this study" (p. 197), but nevertheless the paper claims that the question about the entry of CLV ancestry into Anatolia can be answered: it claims that the data provide "a strong case" in favor of the Eastern route (line 901).

We are grateful for the reviewer's careful reading and on reflection can see that the phrase "Hypothesis A-West is not the the focus of the study and has the difficulties we outlined above" could be misread; we have changed this to "Hypothesis A-West has the difficulties we outlined above".

We do think that our paper adds major weight to the case for an Eastern route with multiple lines of evidence which we summarize further down in our response. We think this is an original contribution of our study that should be emphasized.

By "strong case" we do not imply that the question is now settled definitively, and we are clear about this in our manuscript and revision. Indeed, can the question of the spread of language among pre-literate people ever be really definitively proven?

Our manuscript tries to be clear that our results are important not in settling the issue, but rather in shifting the balance of evidence, "greatly increasing the plausibility of scenarios of an eastern entry of Proto-Anatolian speaking ancestors into Anatolia". We think this phrasing makes it clear that our results make such scenarios more plausible rather than

definitively established. For example and for the sake of argument, if the reviewer started out with a view that the balance of evidence favored a western entry (say by 10:1 odds), and then the genetic evidence pushes in the other direction of an eastern entry as it in fact does (say 1:10), then after combining prior views based on linguistics and archaeology, with new genetic data, the reviewer would have to agree that the scenario of an eastern entry would become much more plausible in light of the genetic data (in this scenario 1:1 odds).

In our study we explicitly highlight that our genetic data cannot rule out a western scenario – the scenario is substantially less likely in light of our genetic data, but cannot be ruled out. Indeed in the Supplementary Information we write explicitly that hypothesis A-West should continue to be explored (*added emphasis*):

“Hypothesis A-West has the difficulties we outlined above. However, we think that it must continue to be explored – provided that a mechanism can be found to explain how language spread may have been effected despite twin dilutions en route from the steppe to Anatolia: first in the Balkans—to explain the paucity of EHG ancestry in Anatolia—and second in Anatolia itself, to explain the decrease of Anatolian ancestry there during the crucial Chalcolithic and Bronze Age periods. Hypothesis A-West must also provide an explanation for why the influx of substantial Mesopotamian ancestry into Central Anatolia from the east was not a harbinger of linguistic change: why did the numerically minor influence from the west -under this hypothesis- have an outsized linguistic impact compared to the numerically major influence from the east?”

This seems methodologically wrong: in a paper that contrasts two hypotheses (East vs. West), one cannot state that one of these was not the focus of the study (West), and at the same time claim that “a strong case” can be made for the other (East).

As described in the revision, we have removed the phrase “is not the focus of the study”

What we had meant to say with that phrase was simply that our study newly reported individuals from the steppe and not the putative route ancestral Anatolian speakers would have taken under the western route hypothesis. So, this statement was merely meant to describes the new data we reported. In fact, our analysis did use the extensive and highly relevant published data presented by other researchers relevant to western route, and exploration of “Western” models is included in Extended Data Fig. 1 and in the section “A Western route via Southeastern Europe into Anatolia?” of the Supplementary Information.

In our study we modeled Anatolians without presupposing any particular route and discovered the model of CLV+Mesopotamian ancestry for Central Anatolian Bronze Age people which adds weight to an eastern entry hypothesis. However, this does not exclude western origin hypotheses as we discuss in the Supplementary Information and in the manuscript; these hypotheses simply become more difficult to accommodate.

Moreover, the arguments that are used against the one and in favor of the other do not seem to hold up. In suppl.mat. p. 197 it is stated that Hypothesis A-West is contradicted by the fact that, within that hypothesis, we would expect steppe migrants to have admixed with the farmers of Shoutheastern Europe who are “of largely Anatolian Neolithic ancestry with local Balkan hunter-

gatherer admixture”, but that “in Anatolia during the Chalcolithic and Bronze Age we see a decrease of Anatolian Neolithic ancestry and a lack of Balkan hunter-gatherer admixture”. The first part of this argument seems a non sequitur: if a population with non-admixed Anatolian Neolithic ancestry would admix with people with an admixture of steppe ancestry and Anatolian Neolithic ancestry (as in Hypothesis A-West), the expected netto result would be a decrease of Anatolian Neolithic ancestry, too.

The reviewer is right that populations of mixed Anatolian Neolithic-steppe ancestry from Southeastern Europe would also decrease the Anatolian Neolithic ancestry in Anatolia. To test the feasibility of such a model, we have now tested 3-source models for TUR_C_BA, allowing the 3rd source to be Bulgaria_C (Chalcolithic Bulgaria) and thus any combination of Balkan farmers with steppe ancestry. We list below the feasible models.

A	B	C	P-value	Proportions			Std. errors		
				A	B	C	A	B	C
Armenia_C	Bulgaria_C	TUR_SE_Çayönü_PPN	0.067	33.1%	-0.7%	67.7%	6.0%	5.1%	9.6%
BPgroup	Bulgaria_C	TUR_SE_Çayönü_PPN	0.088	10.6%	-1.4%	90.8%	1.9%	5.8%	6.8%
Bulgaria_C	CoreYamnaya	TUR_SE_Çayönü_PPN	0.079	-4.7%	11.7%	93.0%	5.5%	2.2%	6.5%
Bulgaria_C	PVgroup	TUR_SE_Çayönü_PPN	0.074	-0.7%	11.5%	89.2%	5.8%	2.2%	7.1%
Bulgaria_C	Remontnoye	TUR_SE_Çayönü_PPN	0.140	2.6%	16.9%	80.5%	5.7%	2.9%	7.7%
Bulgaria_C	SShi	TUR_SE_Çayönü_PPN	0.074	-6.1%	10.5%	95.7%	5.7%	2.0%	6.5%
Bulgaria_C	SSmed	TUR_SE_Çayönü_PPN	0.067	-10.5%	9.8%	100.7%	5.6%	2.0%	6.2%

The amount of Balkan Chalcolithic ancestry is statistically not different from zero (and nominally negative), showing that when we generalize the CLV+Mesopotamian model of our paper to include the possibility of Balkan Chalcolithic ancestry, such ancestry is not inferred.

Thus the model of eastern entry is supported by the finding that a large fraction of the ancestry of Bronze Age Central Anatolia is eastern (Mesopotamian) and not western (Balkan).

We have added this content to the revised manuscript and supplement.

Where could CLV+Mesopotamian admixture have taken place?

One possibility, consistent with a “fast route” scenario also outlined in an answer below, is that CLV (BPgroup-related) ancestry arrived in Anatolia with hardly any admixture from Balkan farmers and admixed with Mesopotamian-related people who moved westward into Central Anatolia. While this scenario would be consistent with the genetic evidence it seems forced as it requires the CLV people to undertake a counter-clockwise migration route around the Black Sea and eastward into Anatolia with no admixture with locals, rather than the much shorter route via the Caucasus and Upper Mesopotamia which is genetically documented (with CLV admixtures in the North Caucasus and South Caucasus at Areni-1) and which would have brought them into contact with a population of Mesopotamian ancestry, explaining the CLV+Mesopotamian admixture more naturally, followed by their entry into Eastern Anatolia at Arslantepe as we highlight further down in this response, and finally Central Anatolia. There is thus a complete path for the transmission of CLV ancestry along the eastern route.

Moreover, the decrease of Anatolian Neolithic ancestry seems to have taken place at the end of the Neolithic already, i.e. too early to connect it to the influx of CLV cline genes.

The reviewer is right that the reduction of Anatolian Neolithic ancestry was evidenced prior to the influx of CLV cline genes. However, that does not mean that it was a single event that occurred in a short of period of time; rather, it could have been a drawn-out process in which several groups -including the CLV bearing pre-Anatolian language speakers- may have participated at different times.

That the population of Bronze Age Central Anatolia represents a different mix than what was previously there can be seen unambiguously by the fact that TUR_C_BA can be modeled as BPgroup+TUR_SE_Çayönü_PPN ($p=0.14$) while the Chalcolithic population of Central Anatolia could not ($p=0.0005$; Extended Data Fig. 1c). The mix of CLV+Mesopotamian ancestry that appears during the Bronze Age (after 3000 BCE) thus does not describe the population at Çamlıbel Tarlası (Skourtanioti et al., Cell, 2020) from mid-4th millennium BCE Chalcolithic Central Anatolia within the bend of the Kızılırmak River. The genetic change is dated to perhaps the late 4th millennium BCE, broadly consistent with the inferred late 4th millennium split of Anatolian languages on linguistic grounds (Kloekhorst <https://doi.org/10.1017/9781009261753.007>, Fig. 4.2). We add content to spell out these issues in the revised section of the supplement discussing this.

Note also that while the Chalcolithic population of Central Anatolia as a whole has “eastern” ancestry compared to the earlier Neolithic population, this is not of Mesopotamian origin (Table S33: all eastern sources are from the Caucasus, not Mesopotamia). This underscores the fact that the Bronze Age population was of a different origin, a different “mix”. The “eastern” ancestry was transformative of Anatolia during the Chalcolithic-Bronze Age periods but was not a singular event emanating from a single population of the east migrating at one time across the entirety of the Anatolian peninsula. Thus, the fact that “eastern” ancestry flowed westward *earlier* than the putative arrival of Anatolian speakers does not mean that it *only* arrived earlier. In fact, we have direct evidence (via the Mesopotamian admixture) that it also arrived later.

The second part of this argument (the lack of Balkan hunter gatherer admixture in Anatolia) largely depends on percentages. The CLV cline signal found in Bronze Age Central Anatolians is ca. 10%. If the migrants responsible for this influx contained Balkan hunter-gatherer ancestry with a max. of, let's say, 10%, the expected percentage in the Bronze Age Central Anatolians would be below the threshold of significance any way.

We agree that the absence of Balkan hunter-gatherer ancestry is not definitive as it was variable in Balkan Neolithic-Chalcolithic populations and a small amount of ancestry could be below noise. We are explicit about this in our manuscript: “The absence of hunter-gatherer ancestry could arguably be explained if little of it was present in the Chalcolithic Southeastern European populations encountered by migrants en route to Anatolia.” When we add a 3rd source to the BPgroup+Mesopotamian model to be Serbia_IronGates_Mesolithic, we estimate a contribution of $0.3\pm1.5\%$ ($Z=0.2$). This, like

the $-1.4 \pm 1.9\%$ Bulgaria_C ancestry in the Table above is not significantly different from zero. We include the Balkan hunter-gatherer ancestry as a potential tracer dye of a southeastern European route for the sake of completeness lest we overlook any trace of such a migration route.

What we can say is that we simply see no evidence for ancestry of *either* farmer (Bulgaria_C) *or* hunter-gatherer (Serbia_IronGates_Mesolithic) origin in Bronze Age Central Anatolia. This does not preclude the possibility of its existence at low levels beyond our ability to detect it.

Hypothesis A-West has the follow challenges:

1. It must explain how, as we write in the supplement: *“language spread may have been effected despite twin dilutions en route from the steppe to Anatolia”* That is a question that genetics cannot answer on the basis of its data and is a question that linguists and archaeologists need to address, proposing a plausible model for how a Chalcolithic group – prior to the technological advances associated with the Yamnaya spread – may have been able to effect major linguistic change in Anatolia without making any genetic impact there above the level of statistical detectability.
2. It needs to propose a new mechanism such as a “fast route” to Anatolia during Wave 1 which as we argue in the linked manuscript on the North Pontic Region saw migrants of CLV ancestry penetrating deep into Europe, at least as isolated individuals, as far west as Hungary and Moldova without any discernible ancestry from locals. Could such Wave 1 migrants have gone much further distances without any admixture, south into the Balkans and deep into Central Anatolia via the western route? We see no genetic evidence for such a movement, and, if it happened it must be also complemented with a westward movement of Mesopotamian ancestry to account for the ancestral mix in Bronze Age Central Anatolia, a process that provides evidence for an eastern entry event at exactly the right time and place.

If the authors really want to dismiss Hypothesis A-West on archaeogenetic grounds, and thus challenge the linguistic arguments according to which Hypothesis A-West is the preferred one (as they are aware of), they have to really expand their discussion. They cannot get away with stating that this hypothesis was not the focus of their study.

We do not dismiss hypothesis A-West on archaeogenetic grounds; we simply find no statistical evidence for it on archaeogenetic grounds, and we are clearer about this in our revision. It is entirely possible that future studies will find missing links and early representatives of Proto-Anatolian speakers in western Turkey or the Balkans. As we describe above, hypothesis A-West could be reconciled with the archaeogenetic results in two ways that we spell out explicitly in our revised supplement. While the genetic evidence for it is lacking, we think that it should continue to be explored by either (i) exploring archaeological or linguistic models in which linguistic change could be driven with minimal genetic impact (was it a stochastic fluke or due to an underappreciated advantage?), or (ii) continuing to explore the genetic prehistory of the Balkans and Anatolia to uncover traces of the proposed movement prior to its dilution.

Another thing that has surprised me in this regard is the shallow treatment of the samples from Western Anatolia, and specifically from Yassitepe. In the suppl.materials, Yassitepe samples are grouped with Ulucak samples under a single TUR_Aegean_BA label, and the only analysis given is that population of this label “can also be modeled in terms of Neolithic people of Western Anatolia with eastern ancestry”. Yet, in the Southern Arc paper, two samples from Yassitepe were reported to have EHG ancestry (I5735, I5737), and are therefore extremely important. Would these two samples also fit a model with CLV ancestry? This needs to be treated separately. If yes, they are extremely important possible witnesses in favor of Hypothesis A-West. (Note that on p. 213 of suppl. materials it is stated that “steppe ancestry is not seen in Western Anatolia”: this seems to be contradicted by these Yassitepe samples).

We did not treat the listed individuals in any special way as their ancestry is not exceptional. The EHG ancestry in the three Yassitepe individuals was estimated in the Southern Arc paper as $2.9 \pm 2.6\%$ (I5737 in the MBA), and $-1.5 \pm 1.0\%$ (I5733) and $1.8 \pm 2.3\%$ (I5735) in the EBA. These results are consistent with zero EHG ancestry for all three individuals given the standard errors. If there is steppe ancestry here it is at levels below our ability to detect it definitively. The presence of steppe ancestry here – if it were statistically established – would also not be definitive about the direction of the entry of the entry of Anatolian speakers. This is because such ancestry could be alternatively interpreted as: (i) a legacy of Proto-Anatolian speakers coming from the Balkans, (ii) a legacy of Proto-Anatolian speakers coming from the east, or (iii) unrelated to Proto-Anatolian speakers and indicative of contact between Bronze Age Western Anatolians and steppe admixed people from the west via the Balkans (where steppe ancestry was copious during the EBA and later periods).

In response to the reviewer’s comment, we also tested the two individuals (with nominally positive EHG ancestry albeit not significantly different from zero) separately to see if any of them could be modeled with CLV ancestry as a source, using the same set of Sources as in other analysis of Anatolian populations and trying all combinations of them.

Individual I5735 (of the EBA) can be modeled with a single source (Çayönü; $p=0.50$) and possibly Masis Blur ($p=0.029$); all remaining models do not fit at $p<0.001$. When we add BPgroup as a 2nd source (thus modeling it precisely as the Central Anatolians of the Bronze Age), the model also fits ($p=0.419$) with an estimated proportion of $1.9 \pm 3.1\%$ BPgroup ancestry. Thus, there is no statistical evidence for CLV ancestry here and, if anything, the individual could be interpreted as an “easterner”.

No single-source model fits individual I5737 ($p<1e-5$). We list 2-source models in the table below:

A	B	P-value	A	B	Std. error
Armenia_Aknashen_N	Bulgaria_C	0.061	30.7%	69.3%	4.2%
Armenia_Aknashen_N	TUR_C_Çatalhöyük_N	0.052	27.2%	72.8%	5.1%
Armenia_C	TUR_C_Çatalhöyük_N	0.078	30.1%	69.9%	5.3%
Armenia_C	TUR Marmara Barcın N	0.060	51.7%	48.3%	4.4%
Armenia_KuraAraxes_EBA	TUR_C_Çatalhöyük_N	0.089	25.8%	74.2%	4.6%

Armenia_MasisBlur_N	Bulgaria_C	0.265	46.0%	54.0%	6.3%
Armenia_MasisBlur_N	TUR_C_Boncuklu_PPN	0.112	52.2%	47.8%	6.0%
Armenia_MasisBlur_N	TUR_C_Çatalhöyük_N	0.145	44.0%	56.0%	8.6%
Azerbaijan_C	Bulgaria_C	0.169	35.1%	64.9%	4.7%
Azerbaijan_C	TUR_C_Boncuklu_PPN	0.109	41.1%	58.9%	4.6%
Azerbaijan_C	TUR_C_Çatalhöyük_N	0.124	32.0%	68.0%	5.8%
Azerbaijan_N	Bulgaria_C	0.171	39.4%	60.6%	4.9%
Azerbaijan_N	TUR_C_Boncuklu_PPN	0.069	46.2%	53.8%	5.1%
Azerbaijan_N	TUR_C_Çatalhöyük_N	0.356	36.8%	63.2%	5.6%
Azerbaijan_N	TUR Marmara Barcın N	0.072	55.2%	44.8%	5.0%
Iran_GanjDareh_N	TUR Marmara Barcın N	0.083	29.6%	70.4%	2.7%
Maikop	TUR_C_Çatalhöyük_N	0.103	25.8%	74.2%	4.5%

Thus, individual I5737 fits models with a mixture of Anatolian farmer and Caucasus Neolithic ancestry (without steppe ancestry). We can also look at more complex models with three sources; 166 of them fit this individual. In the table below we show the feasible subset with Bulgaria_C and Masis Blur as fixed sources, thus allowing the 3rd source to be any other source (including those from the steppe):

A	B	C	P-value	Proportions			Std. errors		
				A	B	C	A	B	C
Armenia_MasisBlur_N	BPgroup	Bulgaria_C	0.162	45.2%	0.1%	54.7%	8.7%	3.2%	7.0%
Armenia_MasisBlur_N	Bulgaria_C	CoreYamnaya	0.161	46.0%	54.4%	-0.3%	8.0%	6.5%	3.4%
Armenia_MasisBlur_N	Bulgaria_C	Levant_N	0.199	44.8%	52.3%	2.9%	7.5%	8.0%	8.5%
Armenia_MasisBlur_N	Bulgaria_C	PVgroup	0.170	46.5%	54.0%	-0.5%	9.2%	7.1%	3.7%
Armenia_MasisBlur_N	Bulgaria_C	Remontnoye	0.142	46.2%	54.1%	-0.3%	11.1%	7.3%	5.5%
Armenia_MasisBlur_N	Bulgaria_C	SShi	0.170	46.1%	54.4%	-0.4%	7.9%	6.5%	3.1%
Armenia_MasisBlur_N	Bulgaria_C	SSlo	0.201	45.9%	54.2%	-0.1%	6.6%	6.3%	2.6%
Armenia_MasisBlur_N	Bulgaria_C	SSmed	0.185	46.3%	54.0%	-0.3%	7.6%	6.6%	2.9%
Armenia_MasisBlur_N	Bulgaria_C	TUR_C_AşıklıHöyük_PPN	0.559	47.6%	66.3%	-13.9%	6.9%	11.4%	10.3%

All proportions from the 3rd source (including steppe ones) are consistent with zero, so the analysis of this individual by itself is consistent with the analysis in the Southern Arc paper that failed to establish statistically the presence of EHG ancestry.

We added this additional analysis in the Supplementary Information.

The main archaeogenetic argument in favor of an Eastern entry is that it “goes “with the grain” of the transformation of Chalcolithic and Bronze Age Anatolia” (suppl. mat. p. 197), and that the Central Anatolian Bronze Age people with CLV-admixture have “the majority of their ancestry from Mesopotamian Neolithic farmers” (main article, line 906-7). The “going with the grain” refers to the fact that in Central Anatolia there was a “major discontinuity between the Neolithic and Chalcolithic periods mediated by migrants from the east” (suppl. mat. p. 202), but this is much earlier than the spread of CLV genes can have been (which entered Anatolia after 4400-4000 BCE). There is thus no reason to connect the influx of CLV genes with this earlier gene flow from the East: they are temporally completely distinct. In order to substantiate the argument regarding the Mesopotamian Neolithic farmers, we need a much clearer and fine-grained dating of which samples show this ancestry when and where.

We have argued above that the eastern gene flow was not a singular event that should be ascribed only to a period much earlier than the spread of CLV gene flow and thus does not preclude a migration into Central Anatolia in a time frame consistent with the arrival of Anatolian speakers.

In Extended Data Fig. 1d we show all Chalcolithic/Bronze Age Anatolians who fit ($p > 0.05$) the CLV+Mesopotamian model. In the revised manuscript—to highlight this important evidence—we now also highlight in the text individual ART027: “In our analysis, the R-V1636 individual (ART038) from Arslantepe does not provably have BPgroup ancestry ($3.6 \pm 3.1\%$) but individual ART027 from the site (3370-3100 BCE) does have a substantial proportion of such ancestry ($16.7 \pm 3.5\%$, $p = 0.171$ for the model fit), a few centuries before the appearance of this mix in the Early Bronze Age of Central Anatolia.”

A puzzling aspect is the fact that this paper does not really treat the question to what extent the early west-ward migration of CLV cline people (samples in Csongrad, the making of Usatove) may be relevant to the question from where CLV cline genes entered into Anatolia. In the twin paper (“A genomic history of the North Pontic Region”) these early migrations are one of the main discoveries, but in the present paper they play hardly any role. (They are only mentioned in passing in line 678 of the main paper.)

People with CLV ancestry may have taken any route to reach Anatolia and from the genetic point of view we cannot distinguish between CLV people in the Caucasus-Lower Volga region and similar people who reached Moldova and Hungary. While in principle CLV migrants may have taken a roundabout way to reach Central Anatolia via Europe, it seems simpler to link them to the east when they are evidenced in the North Caucasus (Maikop and Unakozovskaya), South Caucasus (Areni-1), Eastern Anatolia (Arslantepe), and eventually Central Anatolia during the Bronze Age. We have added a line to the revised manuscript to be more explicit that such CLV cline people with no detectable European farmer or hunter-gatherer admixture did penetrate deep into Europe: “CLV ancestry entered Anatolia from the west without bringing along with it substantial southeastern European ancestry, as a southern extension of the westward wave of Lower Volga-related people that penetrated deep into Europe at Giurgiulești and Csongrád, or even directly from the steppe via a maritime route, bypassing both Balkans and Caucasus”

When it comes to quality of presentation, I feel that the text of the main paper is too dense: there is an abundant use of abbreviations which makes it very difficult and slow to read.

We feel that the abbreviations are helpful, but if the editor would prefer that we spell them out we would be happy to do so.

A big point of criticism is the absence of clear maps in which the archaeological complexes are presented per different time frames. It is therefore often very difficult for the reader to follow the line of thought.

We have added several maps organized by chronological period to the end of SI1 which will allow the reader to locate all key sites and view the distribution of archaeological cultures.

E. Conclusions: robustness, validity, reliability

Overall, the robustness, validity and reliability of conclusions seem OK. I disagree, however, with the suggestion that the paper can conclude (or make a strong case for) an Eastern entry of CLV cline genes into Anatolia (and thus of an Eastern entry of the Anatolian languages into Anatolia): cf. the comments under point D above.

As discussed above, we think the paper greatly increases the plausibility of scenarios of an eastward entry of Anatolian languages into Anatolia, and we are also clear in our manuscript that this is not a proof. We enumerate several lines of evidence:

- **Central Anatolians of the Bronze Age are modeled as CLV+Mesopotamian, i.e., with sources to their east. Hypothesis A-West proposes a more complex scenario in which CLV ancestry must first go west, around the Black Sea and into Anatolia from the west, hardly admixing with the Balkan farmers *en route* to their Anatolian destination and meeting Mesopotamians moving in the opposite direction.**
- **The shortest route from the Caucasus-Lower Volga to Central Anatolia is from the east and not around the Black Sea. Hypothesis A-West proposes a much longer route around the Black Sea.**
- **The eastern entry hypothesis cleanly describes why Anatolian and Indo-European languages would split as the consequence of their ancestors' migrations in opposite directions (into the NPR and the Serebnii Stih out of which Proto-Indo-European would emerge and southward into the Caucasus and Mesopotamia out of which Proto-Anatolian would emerge). In the western entry hypothesis CLV people must move westward, prolonging the proximity of putative pre-Indo-European and pre-Anatolian groups.**
- **There is Y-chromosome evidence for the CLV dominant haplogroup R-V1636 in the proposed geographical route with examples in EBA Armenia and Eastern Anatolia (at Arslantepe).**
- **There is direct evidence of CLV ancestry in the North Caucasus (Maikop) and South Caucasus (Areni-1) along the proposed migratory path; further south along that path is Upper Mesopotamia whose ancestry could very plausibly have mixed with the southward migrating CLV populations to form the mix we detected in Central Anatolians.**
- **In our revision we now newly highlight that there is direct evidence for CLV ancestry also at Arslantepe in Eastern Anatolia where there is also Y-chromosome evidence of CLV ancestry.**
- **There is evidence of genetic discontinuity between Chalcolithic and EBA Central Anatolia. The Chalcolithic population is neither Mesopotamian ($p=0.0003$) nor a Mesopotamian+CLV mix ($p=0.0005$).**
- **Hypothesis A-West predicts an increase of western ancestry in Bronze Age Central Anatolia, yet we do not see evidence of either Balkan farmer or hunter-gatherer ancestry. Hypothesis A-West can be reconciled with the genetic data if either the western ancestry is diluted to trace levels (necessitating an explanation of why it**

could nonetheless effect linguistic change) or that CLV ancestry itself is the western ancestry but arrives in Anatolia (from the west) with no discernible admixture of Balkan farmers and hunter-gatherers and in synchrony with Mesopotamian ancestry moving in the opposite direction.

- The strongest evidence for a western entry (Kloekhorst <https://doi.org/10.1017/9781009261753.007>) is the central-western distribution of Anatolian languages in Anatolia. However, that does not definitively establish the geographical route of transmission as there is a geographical gap between Anatolian languages and the steppe on both the west and east sides of their historical distribution: (Southeastern Europe and Eastern Anatolia/Caucasus respectively). In these gaps there may have once lived Anatolian-related speakers whose languages were later replaced (for example, by the well-attested Yamnaya and Kura-Araxes phenomena respectively). Or, the gaps exist because they were migratory paths and not places of settlement for pre-Anatolian speakers. We are clear about this in our revised manuscript.

In summary, while we do think both hypotheses should continue to be pursued, and we do think there is a strong case for the eastern entry hypothesis from the archaeogenetic perspective and no definitive rejection of that hypothesis from either archaeology or linguistics.

F. Suggested improvements: experiments, data for possible revision

The paper should include a much more detailed discussion of the “Hypothesis A-West” regarding the point of entry of the Anatolian languages into Anatolia (through the lens of the point of entry of CLV cline genes into Central Anatolia). The paper should: give a broader and clearer analysis of Western Anatolian samples; make a clearer distinction in the different time periods in which different gene flows took place in Anatolia; make clearer analyses for answering the question how and in what way it can be excluded that there was a Western Entry; make a clearer link with the early westward migrations of CLV cline people (towards Hungary, Usatove) as discussed in its twin paper; see point D above for all these points.

We appreciate the feedback on these issues, and revision of the manuscript to address these points has in our opinion made it more compelling. We have added some additional discussion to the section “A Western route via Southeastern Europe into Anatolia?” of the Supplementary Information. We also added the two scenarios outlined above in the main text as possibilities consistent with A-West. We do not think more attention to that hypothesis can be given in the text as our analysis does not reveal any new genetic facts in favor of it and we think that the text should be devoted to new discoveries, while noting appropriately, as we do, the continued need to consider alternatives and the fact that our results only change the weight of evidence and do not support a definitive conclusion.

G. References: appropriate credit to previous work?
As far as I am aware, this is largely OK.

H. Clarity and context: lucidity of abstract/summary, appropriateness of abstract, introduction

and conclusions.

The text is extremely dense, and thus lacks lucidity. The number of abbreviations is very high, which hampers the clarity. It would be extremely helpful if the article or suppl. materials would contain a clear map with the locations of different archaeological complexes and sites per time period. I often also miss the datings of specific samples that are mentioned (e.g. “the individual from Vonyucka-1 in the North Caucasus”, line 864: without giving a date for this sample, it is difficult to judge its significance). To my mind, the title of the paper is not well chosen: the term “Indo-Europeans” imply some sort of ethnicity that cannot be substantiated. Better would be “the speakers of Indo-European languages”. Moreover, in the article itself, the term “Indo-Anatolian” is used to refer to the highest ancestral language of the family, and the article in fact deals with this “Indo-Anatolian” homeland. I object to the statement “entering Anatolia from the east” in the abstract (line 24) since an eastern entry has, to my mind, not been sufficiently substantiated. This specific point therefore cannot be represented as a fact in the abstract: in that sense it is not appropriate. The last sentences of the paper are of a semi-philosophical nature, and diminish the strength of the paper.

We have added date information about the date of Vonyuchka-1. We think that “Indo-Europeans” does not necessarily imply an ethnicity and is an appropriate short name rather than the longer “speakers of Indo-European Languages”. It has been used before in reputable standard works such as JP Mallory’s *In Search of the Indo-Europeans*. We are explicit about our word usage choice in our manuscript.

The article deals with the origin of the Proto-Indo-European speakers (who we identify with the Core Yamnaya, propose that they were formed on the basis of the Serednii Stih culture in the North Pontic Region, and trace further back to CLV precursors). We also show that the CLV can be tied to Anatolian speakers as well which has implications for the Proto-Indo-Anatolian homeland question. We decided to keep the title succinct rather than write “Proto-Indo-Europeans and Proto-Indo-Anatolians”, focusing on the Indo-European aspect. The term “Proto-Indo-Anatolian” is mentioned in the abstract in which we have more space to summarize our reconstruction and relate the two terms.

The reviewer writes that they ‘object to the statement “entering Anatolia from the east” in the abstract (line 24) since an eastern entry has, to my mind, not been sufficiently substantiated’. However, we are careful in this clause to specify that the statement is about genes not language, and here the evidence for an eastern origin is in fact definitive: “CLV cline people with Lower Volga ancestry contributed four fifths of the ancestry of the Yamnaya, but also, entering Anatolia from the east, contributed at least a tenth of the ancestry of Bronze Age Central Anatolians”. Nowhere in the abstract do we say that Indo-European/Indo-Anatolian languages definitively came into Anatolia from the east. We discuss the subtleties in the main text where we have more space to unpack the issues.

Referee #2 (Remarks to the Author):

Key results: This study represents a massive new effort to understand the complex genetic, archaeological, and thus also linguistic prehistory of early Indo-Europeans. 372 new genomes, and 198 new C14 dates, allow a level of genetic, chronological and archaeological resolution never achieved before. It integrates all data reported in the North Pontic paper, plus the new expansive dataset reported here. And it resolves a long- standing enigma about the origin, formation, and divergence of Indo-Anatolian and Indo-European by documenting the constituting genetic and archaeological clines that contributed to their formation.

Validity: There are of course still lacunas that needs to be filled especially concerning Mesopotamia, and its interaction with Caucasus and Anatolia in the formation of Indo-Anatolian, but even with that reservation the Indo-Anatolian part of the paper looks convincing as the best interpretation so far. If substantiated it resolves a long-standing debate about a western versus northeastern/Caucasian origin/entry of Indo- Anatolian. What is truly remarkable, however, is the documentation of three new Eneolithic genetic clines, and their contribution to the formation of genetic, archaeological and linguistic formations of Indo-Anatolia/Indo-European. All summarized in a well presented/pedagogical Figure 1 and subsequent figures. However, the concrete discussion of each cline is really the strength of the paper, as it details and integrates the genetic variation along the cline with archaeological sites/cultures in a convincing way. And it discusses alternative admixture models/interpretations, and leaves interpretations open if needed – e.g. the proposal that there is still an unsampled population to form Core Yamnaya. But the discussion also evolves with each cline, adding new evidence in a cumulative way to the final interpretation.

The uniqueness of the Yamnaya expansion from their core area, is illustrated both by demographic growth curves, as well as IBD and family sharing, which supports the maintenance of their genetic and social homogeneity. This must have been an essential part of their linguistic dominance.

The discussion of the origin and spread of Indo-Anatolian is very well argued both from a genetic/archaeological point of view and from a linguistic point of view. Many details will be sorted out by future research, but the main picture now presented looks robust, and thus solve 150 years of debate, at least in the main, as well as refutes a recent Hegarty et al. paper in Science.

I only have a critical comment on the dates of Seredny Stiih, and several other important samples, which are not corrected for marine reservoir effect, and which could affect their role in the formation of Core Yamnaya. These dates should be calibrated for marine reservoir effect, when possible.

We thank the reviewer for this important comment.

In our manuscript, we address these concerns by providing faunal dates unaffected by Freshwater Reservoir Effect (FRE) for cemeteries where we have such dates. All specialists agree that this is the best and most accurate method to correct for FRE. In other cases, we

identify outlier dates that appear too old. For example for the important site of Berezhnovka on the Volga, the Archaeological Supplement states: "This difference in FRE in between grave 5 and the others implies that one Eneolithic individual at Berezhnovka II (I22199, with Y-haplogroup Ia2 like Serednii Stih) had a more riverine diet that created a FRE, while others buried beside him (I22202, I22201, and I22200 all with Y-haplogroup R1b) had a more terrestrial diet and therefore yielded accurate radiocarbon dates."

To clarify our approach including addressing questions about the reliability of the Serednii Stih chronology, in our revised manuscript we make three changes to the main text:

(1) In the section on the new data reported, we now include the following revised sentence:

"Technical details of the 803 ancient DNA libraries that are the basis for the newly reported data (and an additional 195 libraries that failed our screening) are presented in Online Table 2, while details of 198 newly generated radiocarbon dates on these individuals are presented in Online Table 3 (we make no attempt to correct radiocarbon dates for freshwater reservoir effects (FRE); while old carbon in diets can make dates hundreds of years too old ¹, there is no reliable formula for correcting for this bias based on stable isotope measurements in the Eneolithic Steppe², and so here we take the approach of presenting uncorrected dates and noting the degree of consistency or lack of consistency with archaeological context dates not affected by FRE)".

(2) In the section on the Dnipro Cline, we write:

"The Dnipro Cline is formed at one end by Neolithic individuals living along the Dnipro River rapids whose union of calibrated radiocarbon date ranges is 6242-4542 BCE (UNHG), and at the other end by the Serednii Stih population represented by 13 individuals with good quality data whose union of radiocarbon date ranges is 4996-3372 BCE (both date ranges span the conventional archaeological dates for these cultures and hence are unlikely to be substantially distorted by Freshwater Reservoir Effects)."

(3) We added a Methods section on our Radiocarbon Dating approach:

"Radiocarbon dates: We report 198 accelerator mass spectrometry (AMS)-based radiocarbon dates, in almost every case using the same skeletal elements employed for ancient DNA analysis (Online Table 3). Nearly all dates were performed at the Pennsylvania State University Radiocarbon Laboratory using standard methods^{98,99}, and also include measurements of dietary stable isotopes (carbon $\delta^{13}\text{C}$ and nitrogen $\delta^{15}\text{N}$) (Online Table 3). We calibrated dates using OxCal (v4.4)¹⁰⁰ and the IntCal20 calibration curve¹⁰¹. We make no attempt to correct radiocarbon dates reported in this study for freshwater reservoir effect (FRE) using stable isotope measurements from the same samples, since formulae for such corrections are unreliable for Eneolithic steppe contexts such as we study². For example at Eneolithic Khvalynsk, there is no correlation between radiocarbon dates and the $\delta^{15}\text{N}$ levels conventionally used for correction²; and for the father-son (order unknown) pair I6408 [5209-5006 calBCE (6150 $\pm$ 25 BP, PSUAMS-4155, $\delta^{15}\text{N}$ =15.1)] and I6737 [4783-4555 calBCE (5820 $\pm$ 25 BP, PSUAMS-4545, $\delta^{15}\text{N}$ =15.6)], the 95% confidence intervals are non-overlapping by more than two hundred years, and the discrepancy compared to the genetic inference that they must have lived at very similar times would be even larger if we corrected based on $\delta^{15}\text{N}$ isotope levels (which generally

imply a larger downward correction for larger $\delta^{15}\text{N}$). Because an isotopic correction in this context would likely add uncertainty to our results, we take the approach of not applying a correction. Instead, when relevant we note the degree of consistency with faunal and plant remains not affected by FRE as well as material culture typology (as we do in the case of faunal dates in graves we assembled for Khvalynsk and Ekaterinovka). We also note comparison to conventional archaeological dating (as at Serednii Stih culture graves where our radiocarbon dates are broadly consistent with archaeological dating, and Golubaya Krinitsa where some radiocarbon dates may be too old based on FRE). Finally, we identify individuals whose dates are outliers compared to others in the same context likely due to FRE specific to the diets of some individuals (like I22199 at Berzhnovka II, discussed in Supplementary Information section 1).”

Originality and significance: the study is truly groundbreaking, in that it allows for the first time a convincing, interpretive integration of genetics, archaeology and linguistics in the formation of early IE languages, whose ‘roots’ and early formation has remained in the dark. Not any longer. And even if the last word is never said, we now have a new foundation to build on.

Summary, conclusions fine

We thank the reviewer and we are glad they found our presentation of the genetic results accessible as that is always a challenge when trying to integrate a large number of genetic findings in conjunction with the archaeology.

Referees #3/4 (Remarks to the Author):

Comments

We cannot escape from feeling uncomfortable reading a paper with massive contribution from Russian scientists at the time when Russia's aggression war puts Ukrainian scientists into the drenches or fleeing from their homeland. As a minimum, a statement by the authors, condemning Russia's aggression war, should be included. We understand this is just our opinion and such editorial decisions are with the journal.

We also understand that the authors have split what would have been one paper into two so that Ukrainian and Russian authors would not have to share authorship. We understand that it would be unacceptable for the Ukrainian researchers to appear on the same paper with Russian institutions, many of which, as per by a statement by Russian university rectors, have officially supported Russia's war in Ukraine.

In our opinion and that of all of the co-authors of both papers it is positive for scientific research to be able to move forward even in this terrible time. Our formula of two coordinated papers that shared data was one co-authors of both papers supported.

The question of the genetic origins of the Indo-European speaking people has been a hot topic for a long time. Here the authors sum up a massive project which has (according to the lead author's public posting) been 10 years in the making. It is indeed an enormous effort. The two combined studies double the number Eneolithic samples (79 in addition to 82 previously published) from the steppe region and add 286 Yamnaya/Afanasievo individuals compared to 75 in the literature. This major advancement in the sampling enables this paper to set out to identify the genetic ancestors of the Yamnaya people who played the central role in reshaping the genetic and linguistic landscape of Western Eurasia in the late Bronze age.

Methodologically the paper offers a new two-step “tournament” approach to filter/rank fitting qpAdm models in order to choose between competing models for the same population. This is presented as a logical thought process but not tested as a method. Not sure if there is a sensible way to do it though. Nevertheless, some additional discussion on the matter would perhaps be of help. For example, the suggestion that the fortitude criterion helps to reduce the non-falsifiability bias toward source populations with few data could be tested and shown? In this regard we were somewhat confused by the discussion around line 493 where the choice of GK2 over UNHG seems to not benefit from the fortitude criterion. Perhaps Tournament was only held for two-way models?

We employed the model tournament approach to present a transparent account of the models we considered and to compare them against each other in a systematic and symmetric manner. We added the following discussion to the Supplementary Information:

“The model tournament approach was motivated by the tension between (a) the desire to make the Right set consist of populations that are maximally discriminative between the Sources and are thus close (in time and place) to the population history of the Test population and (b) the desire to make the Right set an impartial arbiter among the competing models

which suggests that its populations should be chosen to be as distant (in time and place) to the population history.

The inclusion in the Right of spatio-temporally “close” populations in furtherance of goal (a) has the promise that they and the Sources will share substantial genetic drift which is what allows the technology of qpAdm to distinguish between them. This, however, comes at the danger that these “close” populations will share genetic drift with the Test population as well to the exclusion of the Sources, which is a violation of the qpAdm phylogenetic assumptions.³

Conversely, the inclusion of spatio-temporally “distant” populations in furtherance of goal (b) excludes the possibility of some model violations: for example, Afontova Gora 3, an Upper Paleolithic Siberian⁴ axiomatically did not experience any post-18kya history, making it a safe and impartial choice for distinguishing between sources that differ in their pre-18kya affinity to it, but of little use to distinguish between sources that differentiated from each other, accumulating genetic drift in a more recent time frame.

The “model tournament” approach is a formal protocol that allows us to pursue both goals, and is qualitatively different from protocols like “model rotation” that has been used in previous publications. In pursuance of (a) the Right set is formed with a conservative set of distant and early sources so that the combinatorial number of possible models with K sources can be reduced to a much smaller set of candidates. This smaller set of models cannot be rejected in terms of deep history alone, and their competition adds spatio-temporally “close” sources to the Right set in a systematic and symmetric (among models) manner, in pursuance of goal (b).”

We do compete all feasible 3-source models against each other in the Supplementary Information section 2 in the “Full exploration of 3-way models for the core Yamnaya” section. The model with Ukraine_N as the source was not included in that competition as it was not feasible ($p=0.04$) according to our set criteria, but we competed against the model with GK2 as the source as described in our reply below.

The match between Ukraine_N and GK2 as the 3rd source in the Aknashen+BPgroup+X model is not decisive. The p-value with GK2 as the source is 0.93 and with Ukraine_N it is 0.04 as we write in the text. These values become correspondingly 0.69 and 0.07 when we add Ukraine_N or GK2 to the Right respectively. So, the model tournament cannot give us a clear winner among the two models. This is why we try to consistently refer to Dnipro-Don hunter-gatherers or North Pontic hunter-gatherers rather than prefer either the Dnipro (Ukraine_N) or Don (GK2) as the source, as this question cannot be decided to the limits of our resolution. What complicates matters further is the presence of Igren_o (which is geographically from the Dnipro area but a clade with GK2 from the Don).

The paper suggests modelling Eneolithic population mixtures that led to the formation of the Yamnaya group in terms of three clines. For clarity, it could be a good idea to follow the same order of describing the clines both in the abstract (currently: CVL-Volga-Dnipro) and the main text (currently: Volga-Dnipro-CLV).

We have chosen to introduce the CLV cline first in the abstract because it is a major finding of our study that ties all other clines together, so we need to have introduced it

before we say that, e.g., “CLV people bearing both Caucasus Neolithic and Lower Volga ancestry” when discussing the Dnipro Cline. We experimented with different orders and found that this is the most succinct in the abstract.

In the main text we briefly introduce the three clines in the context of the PCA and take a geographical approach to their discussion, first discussing the two “northern” riverine clines and then introducing the CLV cline just after them (as it makes sense of both origins), and just before the discussion of Anatolia (to which CLV people also contributed).

Furthermore, a cline is intuitively perceived as a linear representation of mixing of two differentiated sources. In this regard the Volga cline, as it is suggested, does not seem like a cline. Specifically, the “bend” in the cline on the PCA does not make sense. Rather the cline seems to be between BP and EHG.

We agree with the reviewer that our use of “Volga Cline” to refer to the non-linear interaction zone along the Volga is unconventional as this is not linear. In our revised manuscript, we clarify that the Volga region harbors at least two ancestry gradients and reserve the use of the name “Volga Cline” to refer only to the downriver portion (EHG-BPgroup). We include the upriver portion of the Volga region in the “European Hunter-Gatherer Cline”.

Moreover, panel C of Figure 2 depicts what we suggest. Volga cline is between BP and EHG and another cline is between UNHG and EHG. Also, text around line 185 seems to tell, that Upper Volga (Murzikha) groups do not fit well with BP-EHG mixing model. Looking back at the PCA, and anticipating the position of WHG, used in PC space calculation but not shown on the plot, and Baltic Mesolithic HG, for some reason not used at all, we tentatively propose that we could be seeing an “out-of-BP/PV” FAN – an array of (sampled and unsampled) radiating clines into a WHG -Siberia cline that includes Baltic HG, UNHG, Karelia EHG and so on (see also line 317). In this setting the Dnipro cline would be one of these reaching toward UNHG.

We agree with the reviewer, and indeed this “fan” is a central theme of the sister North Pontic Region paper.

The reason we do not plot Baltic, Balkan, and Western European hunter-gatherer is that they are further along UNHG on the EHG-UNHG-BHG-WHG cline (see e.g., the PCA from Mathieson et al. 2018) reproduced below. This has the effect of packing data points closer together along PC1, thus making them less distinct. So, we chose to include the EHG-UNHG portion of the cline to make better use of the plot area to bring out distinctions within the steppe region. In Fig. 2a of the accompanying “A genomic history of the North Pontic Region from the Neolithic to the Bronze Age” paper, Balkan hunter-gatherers are plotted as they are very important for the discussion of that paper.

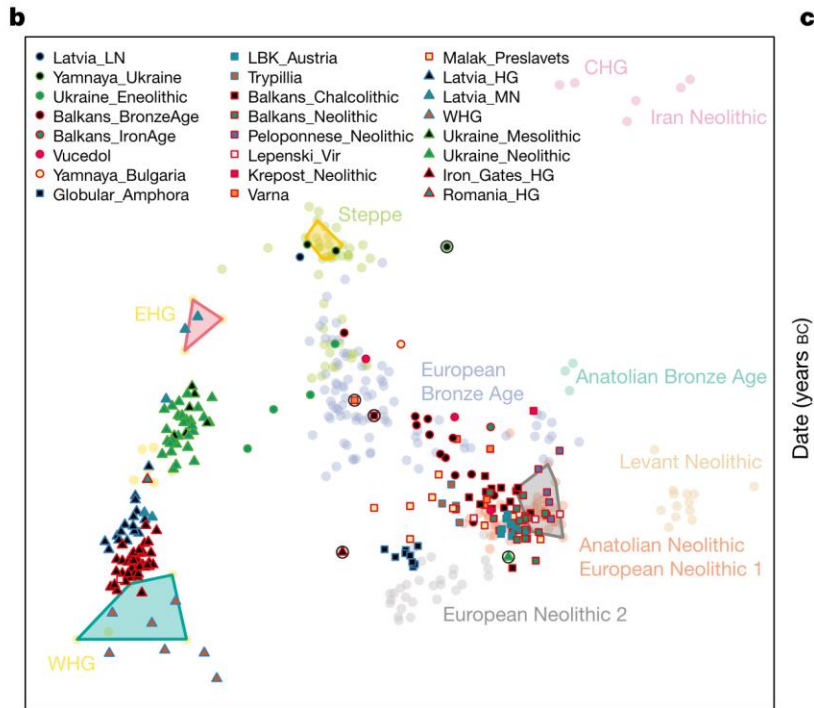


Figure 1 | Geographic and genetic structure of 216 analysed individuals. a, Locations of newly reported individuals. b, Ancient individuals projected onto principal components defined by 777 present-day west Eurasians (shown in Extended Data Fig. 1); data include selected published individuals (faded circles, labelled) and newly reported individuals (other symbols, outliers enclosed in black circles). Coloured polygons cover individuals that had cluster memberships fixed at 100% for supervised ADMIXTURE analysis. c, Direct or contextual dates

One impressive result of the massive sampling is the ability to see the temporal dynamics of admixture. For example, along the Volga cline – line 200. These results could be visualized perhaps? Especially if similar temporal patterns are observed more systematically. Much like the representation of temporal change of the Y chromosomes.

We explored trying to create such a visualization, but we do not see a clear pattern among the large number of individuals of similar and overlapping dates (based on C14 dates) from Khvalynsk where ancestry variation is extreme. The situation is further complicated by possible reservoir effects in the dates of humans from the area. For these reasons we decided not to make any claims or visual representation about temporal patterns other than the observation in the text of the contrast between the Lebyazhinka HG, Ekaterinovka, and Khvalynsk populations as a whole.

The principal result, that the paper sets out to reach, presented around line 500, deserves a summarizing simple figure, or at least a clear panel in the summarizing figure. Moreover, wouldn't it be a good idea to include a qpGraph-like summary figure of the complex interactions presented?

We think qpGraph is not well-suited in this setting. For example, we infer Core Yamnaya to have three proximate sources of ancestry (Aknashen-BPgroup-UNHG/GK2) each of which is a complex mix of earlier sources (BPgroup is EHG-CHG-Central Asian) and thus the graph would need to encompass at fine detail the entangled history of the entirety of Eurasia, and doubtless would be wrong in many respects. We do not feel confident in our ability to fit such very complex graphs with many admixture events uniquely (see e.g., Maier et al.: <https://elifesciences.org/articles/85492>)

Overall, the level of novelty and implications of the results warrant publication in Nature. However, the manuscript would benefit from increased clarity and shortening (of the main text). The clear and cool focus of the paper is the genetic origin of the Yamnaya people. Other lines of enquiry/discussion could be condensed in the main text and opened in more detail in the supplement. The authors extend the discussion from origins of the Yamnaya to the origins of IE/IA languages half-way through the paper. This comes little bit out of the blue. Maybe structure the whole paper more clearly with clear subtitles? – origins of Yamnaya, origins of IE. The weakest aspect of the manuscript are the figures. It is undoubtably very challenging to come up with visual representations conveying the complexity and yet not drown in it. For example, it has been quite difficult to connect the (not always consistent) labels in the legends and figures. For some figures we have more specific comments below but overall, we feel there are ways to increase the “readability” of the visual narrative.

We thank the reviewer for the very close reading of the text and the many helpful comments which we individually address below. We think responding to these comments has increased readability and manuscript quality.

More specific comments and suggestions:

General comment on samples: The authors report a few new samples from Ukraine. However, the institution responsible for these samples is in Moscow, Russia. We wonder if there was any contribution from Ukrainian archaeologists in those excavations?

The eight samples from Ukraine that in the original submission of our manuscript were marked as reported for the first time here are from a collection of teeth from Deriivka-1 and Olexandria sampled a decade ago at the Russian Academy of Sciences in Moscow. We reported the screen for ancient DNA data for these two collections in Mathieson et al. Nature 2018, and in the present paper cite the archaeological contextualization given in Mathieston et al. 2018, prepared by a team overseen by Ukrainian-based coauthor of that study Inna Potekhina at the National Academy of Sciences of Ukraine in Kyiv (we do not present a new contextualization here). The ancient DNA data for the 8 specific samples in question was poor in quality and hence not listed in the main table of newly analyzed data from these two sites present in Mathieson et al 2018 although we would have listed them if we had included in the paper a table of samples with low quality data as is now our standard practice. Since Mathieson et al. 2018 we made 6-34 libraries from each of the eight samples, and after merging data had enough to be useful. In our revision, we mark these samples as improved quality data from samples first processed in Mathieson et al.

2018. Thus, we thus no longer mark any samples from Ukraine as processed for the first time here.

Summary Figure: A few of the “other sites of interest” don’t have site names. Also, it’s quite difficult to determine from the figure, which period(s) the samples of this study are from. Finally, the placement of the “Roots of Caucasus-Lower Volga people” text is extremely awkward – maybe it would fit better on the bottom of the figure on top of Southeastern Europe and Anatolia? Overall the Summary figure does not “tell the story” of the manuscript too well. Lots of text. Probably there should be visual and spatial representations of the clines and a schematic of the genetic makeup of the Yamnaya.

Thanks for these presentation suggestions.

We added the labels of the missing “other sites of interest”. We decided to place the two blocks of text in the largely empty areas of the map so as not to interfere with the plotted points. The uneven geographical landscape of the Balkans and Anatolia makes the text less readable as a background and there is also less space in that area of the plot, so we decided to put “Roots of Caucasus-Lower Volga people” against the even “greenish” background of eastern Europe where there was enough room for it.

We could not place both blocks of text dealing with the origins/expansion of CLV people close to the actual CLV area because of the high density of plotted points in this area, so we decided to put them in chronological order with “origins” on the left and “expansions” on the right.

Line 43: “The first formal study of the origins of the Yamnaya identified two disparate sources of ancestry” this wording omits the contribution of Allentoft et al from the same year, who also test for admixture sources in the Yamnaya and with the absence of EHG in their data identify Mal’ta in addition to BA Armenia.

Allentoft et al. (2015) did not infer an Armenian-related component in the Yamnaya which is the statement referenced here so it would not be appropriate to cite it here as supportive of the two sources of ancestry in the sentence “and a southern, West Asian source related to present-day Armenians”. Allentoft et al. wrote: “Using D-statistics, we find that Corded Ware and Yamnaya individuals form a clade to the exclusion of Bronze Age Armenians ([Extended Data Table 1](#)) showing that the genetic ‘Caucasus component’ present in Bronze Age Europe has a steppe origin rather than a southern Caucasus origin.” In Haak et al. (2015) it was written: “This was likely due to admixture of EHG with a population related to present-day Near Easterners, as the most negative f_3 statistic in the Yamnaya (giving unambiguous evidence of admixture) is observed when we model them as a mixture of EHG and present-day Near Eastern populations like Armenians (Z 5 26.3; Supplementary Information section 7).”

The two papers are often cited together but we think it is appropriate to cite only Haak et al. (2015) in this sentence for that particular assertion. We believe that the Allentoft et al. manuscript is exceedingly important in its scientific contributions in many areas, and in our revised manuscript we now also cite Allentoft et al. (2015) earlier, at the first group of

citations (*The spread of Indo-European language and culture transformed all these regions.*) as they certainly were one of the first papers to study the Yamnaya and we had omitted that citation from that block. We also rephrased our sentence to “The formal study ... first identified” instead of “The first formal study” so as to make clear that Haak et al. first identified the two sources of ancestry, not that it studied the origins of the Yamnaya first.

Lines 49–52: Currently, the question after the second point came very unexpectedly so maybe it would be better to unify the style and rewrite the first as an actual question as well.

We re-wrote this in the form of a question.

Lines 71–74: The numbers of newly reported samples and those with increased data are somewhat confusing. The same samples should not be mentioned as newly reported in two manuscripts so the first sentence should be removed and the data from 299 + 55 individuals presented as newly reported and then mentioned that it’s co-analysed with data from 73 + 5 individuals from Nikitin et al.

We agree this sentence was confusing. We added “formally reported in the present study and a parallel one⁶” so readers will be aware that the numbers of the first sentence are the joint dataset. We then continue to break down the samples into the ones reported in the present one and in the ones reported in Nikitin et al. (to which an additional 3 individuals were added, so a total of 78+3=81 in the revision).

Lines 108 and 120: The use of pronouns (“these questions”, “this note of caution”) creates confusion given that the questions/note of caution were not in the previous paragraph, but earlier in the text.

We were referring to the immediately preceding question. We have rephrased to “this question”. We removed “this note of caution”.

Line 145: Should be “Fig. 1ab” since it’s a co-description of geography and a PCA cline.

We changed this.

Line 152: It would be clearer and more precise to write “genetically unsampled”.

We changed this.

Line 161 and elsewhere: There are no references to the tables where qpAdm results are presented.

We followed what we believe is standard practice at the journal in not references individual tables and figures included in-line in the Supplementary Information sections but we can certainly add these more detailed references if the editor requests them. The material in the Supplementary Information is copious and various points are supported by either statements in the text or particular tables and figures, so we think it would add a large

number of such items burdening the text. We think that the Table of Contents of the Supplementary Information aids the reader in locating the discussion of particular populations and analyses. We will abide with whatever style the editor suggests is most appropriate.

Lines 164–166; The sentence “currently dated earliest in the Samara region at Lebyazhinka-5 and in a few graves at Ekaterinovka dated before 4500 BCE” is not very clear – is Lebyazhinka-5 the earliest in the Samara region and the custom also exists in some Ekaterinovka graves dated before 4500 BCE or are Lebyazhinka-5 and Ekaterinovka both in the Samara region and the graves in both are dated before 4500 BCE (which is the earliest)? Also, is the Samara region the earliest with this custom or are these graves the earliest in the Samara region with this custom (but not elsewhere)?

We have expanded this text to refer to the dates of both Ekaterinovka and Lebyazhinka-5 both of which furnish evidence for this custom before 4500 BCE.

Lines 169–171: It seems strange to refer to results that are presented later in the manuscript.

We thought it would be best not to break the flow of discussing the variation in the Volga Cline for an aside about the Central Asian component in BPgroup. So, we chose to alert the reader that this will be discussed further down where the Central Asian signal is explored in detail.

Lines 184–185: “Seven Volga cline populations fit this model (p-values of 0.04 to 0.72)” – the commonly used p-value threshold for a fitting model is 0.05 so it would be good to at least specify how many models have a p-value higher than 0.05.

All but one have p-value>0.05. To clarify this, we expanded this to “p-values of 0.04 for Ekaterinovka and 0.12-0.72 for the rest”

Lines 188–189: It would be clearer to write “in the upriver portion of the Volga cline, between EHG and UNHG”.

We now write: “in the upriver segment of the Volga zone, between EHG and UNHG”

Lines 196–197 and elsewhere: There are a lot of sites and ancestry sources so it would be better to stick with one label for each (BPgroup vs Berezhnovka).

We changed to Berezhnovka-related ancestry here as we think it is more appropriate when discussing the Volga as both its genetic and geographical end is Berezhnovka (which is the genetically major component of BPgroup). When discussing this type of ancestry flowing elsewhere we prefer to use BPgroup-related so that the reader will not think that the ancestry comes necessarily from Berezhnovka itself (e.g., to the Dnipro or the Caucasus), as BPgroup is geographically distributed in both the Lower Volga (Berezhnovka) and North Caucasus (Progress-2).

Figure 1: It would be much clearer if the legends were grouped and ordered based on geography not listed alphabetically. It is not clear which groups are from this study, which are from Nikitin et al. and which previously published. The symbols for Murzikha and Ekaterinovka are extremely similar. The group labels don't fully match between different panels.

We thank the reviewer for the careful inspection of the plotting points. We aimed to have a consistent plotting scheme across the different figures, so we inspected panels a and b to identify any discrepancies:

- **Four populations (Russia_Karelia, Russia_AfontovaGora3, TTK, WSHG) are only plotted in the PCA of panel b; these are previously published data well outside the geographical scope of the map so we did not plot them**
- **Eighteen populations (mainly regional subsets of Turkey and Iran) are only plotted on the map of panel a and not in the PCA of panel b. We avoided adding all these different symbols on the PCA as the PCA aims to maximally discriminate between steppe populations and all West Asian populations cluster in the relatively small area at the bottom of the PCA. We only included a few key groups from West Asia that flesh out its overall variation without overburdening it with symbols**
- **We checked the remaining 53 populations (plotted in both panels a and b) and we think they are all plotted using the same symbols**

Panel a: The labels on panel a are too small and too many. Information is lost in complexity. Maybe put the full sample location map in supplementary as a larger high-resolution image and figure out how to group labels (keeping all sample locations) for the main figure?

We think that the plotted points are necessary in the main figure. We already removed many (such as the different BHG/WHG and subsets from Turkey and Iran mentioned above) to maximize clarity, and the remaining ones are crucial for describing the genetic structure of the steppe and adjacent regions.

Panel b: It's unclear what determines the font size differences on panel b. It appears that the samples used for PC space calculation are not on the plot? Would be very helpful if they were there. Also, the choice of populations for the calculation of PCs is not immediately clear and should be explained.

We added Supplementary Information Fig. S2 to add the samples used to compute the PCA and explain our choice more fully there: "It includes populations surrounding the steppe but does not include the Test/Sources populations so as to avoid any systematic difference between populations used to compute the PCA and those projected on it. We show in Fig. S 2 the populations used to compute the PCA, dominated by the Siberian-European distinction along PC1 contrasting Firsovo with Oberkassel Cluster and contrasting both these boreal populations with those of the Near East and derived populations (such as those from Iran and LBK) along PC2."

We do not show this version of the PCA in the main manuscript as we want to maximize the use of space to focus on the samples of interest rather than the Siberian-WHG-related individuals that dominate the PCA when all are included.

Panel c: The legend of panel c, with individual sample names, is suitable for a supplementary figure. The clines in panel c could for example be presented on a map (three panels if they would otherwise become too overlapping spatially).

Overall, this figure is extremely hard to follow and needs substantial work.

We chose not to depict the data points on a map because there is considerable spatial overlap between points that are very different in their ancestry proportions (for example the SShi SSmed SSlo of Serednii Stih and Khi Klo Med of Khvalynsk) and also large number of individuals are sampled at particular sites (such as Khvalynsk). We think a barplot in which individuals are sorted by their proportions of ancestry is more appropriate.

Line 261: Does “the Eneolithic individual from the area” mean Kriviansky? If so, then it’s only possible to compare the Don Yamnaya individuals to this one on the PCA which doesn’t allow for conclusions like “bear no affinity”.

Indeed we meant the Eneolithic individual from Kriviansky. We changed the vague “bear no affinity” to a precise statement: “cluster distinctly from the Eneolithic individual from Kriviansky in PCA (Fig. 1) and are not a clade with it using qpWave ($p=8e-15$)”

Lines 292–309: It is unclear why the modelling switches between EHG and GK2.

GK2 lies on a particular position along the broader UNHG-EHG cline. By modeling with a point on the UNHG-EHG cline we allow this to vary freely rather than presuppose that the admixing population corresponds exactly to the sampled GK2.

We have clarified this in the text: “We therefore model individuals of the Cline with ancestry from any population along the extent of the entire UNHG-EHG cline (Fig. 1c), without presupposing that this population corresponds exactly to the UNHG or GK2 positions along this cline.”

Line 322: It would be clearer if it said “three Golubaya Krinitza individuals”.

We rephrased this as suggested.

Line 354: It’s not completely clear what “their” refers to here.

We rephrased to “the shared genetic drift of the Core Yamnaya”

Lines 390–392: Again, it seems strange to “spoil” the result of the next section.

We rephrased this so that it prepares the reader to anticipate that the answer to the problem of the ancestral blend of the Yamnaya will be answered in the next section: “Crucially, the Core Yamnaya fail all models of Fig. 2a-d ($p<1e-8$), and thus its origins must include not only the CHG-EHG-UNHG-Tutkaul sources involved in these models, but

a different blend of ancestries documented in the third major cline of ancestry to which we now turn.”

Line 398: “Unplotted individual all give fits at $P < 0.05$ ” – should be “individuals” and “ $p < 0.05$ ”.

We changed this

Figure 2: The placement of the panel labels is a bit strange. Again, it would be much clearer if the legends were grouped and ordered based on geography not listed alphabetically.

We anticipate that sometimes the reader may want to look up data points based on their plotting symbol and color, sometimes they will want to look up sites based on their geographic proximity, and sometimes they will want to look up sites based on their name that is mentioned in the text. The legend order cannot simultaneously accommodate all three needs, so we decided to use the alphabetical order which allows the reader to quickly find the symbol for a population of interest named in the text.

Line 424: The Dnipro cline individuals already have more than one ancestry source.

We re-wrote this sentence to better convey the idea that the Yamnaya cannot be modeled in terms of the Dnipro cline populations as they all have more UNHG/GK2-related ancestry than them, and therefore they must have ancestry from another source that has even lower or no UNHG/GK2-related ancestry:

“The Yamnaya are on the end of the Dnipro cline, having less UNHG/GK2-related ancestry than all other cline populations; thus, they cannot be modeled in terms of them alone (Fig. 1), but must have possessed more of an unknown second source with even lower or even no UNHG/GK2-related ancestry”

Further down we show that Remontnoye does indeed fit the criteria of being the unknown second source:

“We estimate $-0.3 \pm 2.9\%$ UNHG or $-0.5 \pm 3.5\%$ GK2 ancestry when we add either population as a 3rd source added to the Aknashen+BPgroup model, so Remontnoye is indeed a population of no discernible UNHG/GK2-related ancestry as anticipated for the unknown second source for the Yamnaya.”

Lines 446–447: We usually know the geographic origin of a sample, but not the genetic ancestry without DNA analyses so this sentence would make more sense the other way around.

We added “or vice versa” to account for that possibility as well. What we were trying to do was to warn readers against that erroneous inference that a population of either high or low Berezhnovka-related ancestry must come from the northern end of the Caucasus-Lower Volga geography; the fact that BPgroup (consisting of Berezhnovka and Progress-2 individuals) spans the geographical range cautions against that interpretation. But the converse is also problematic: we should not expect more geographically “northern”

populations to be necessarily more genetically “northern”, with e.g., Remontnoye being geographically well to the north of the North Caucasus piedmont but genetically more “southern” than the people of it at Progress-2 and Vonyuchka.

Lines 453–463: The jump from Meshoko and Svobodnoe to Unakozovskaya is confusing and unclear.

We have added a sentence to clarify: “These sites are bracketed temporally in the North Caucasus by two populations for which ancient DNA data exists: preceded by the Eneolithic Unakozovskaya (ref.⁵ 4607-4450 BCE, and this study) and succeeded by the Maikop.”

Lines 469–470: It seems more appropriate to put the full “high/low steppe ancestry” in quotation marks.

We removed the quotation marks altogether as they did not serve an obvious purpose.

Line 476: It would be more precise to use “ends” instead of “edges”.

We changed this

Lines 505–506: It’s unclear what’s meant by “the remaining four fifths of their ancestry”.

We meant the non-UNHG/GK2-related ancestry, but removed this parenthetical statement to avoid confusion.

Line 535: There’s a typo – should be “Armenia”.

We fixed this.

Lines 538–539: Where is the Mesopotamian-Anatolian modelling shown? Also, what are the references after the end of the sentence/paragraph?

We do not show modeling of Mesopotamian-Anatolian contrasts in this paper. The references pertain to relevant literature that observed this “eastern” influence in the Chalcolithic and later periods on top of the Anatolian Neolithic population substratum.

Line 554: As it is also stressed above in the manuscript, it is not possible to determine the exact source of any ancestry precisely with aDNA methods.

We agree.

Lines 559–566: It is unclear which specific models are tested in this section. Are the first ones 3-way models of steppe source 1 + steppe source 2 + EHG/UNHG/GK2? But then how is the unknown admixing population’s placement on the CLV cline tested?

As we write in the Supplement: “To this end, we fixed Çayönü as one source and considered the endpoints of the steppe clines as additional sources: BPgroup, Russia_Karelia, Armenia_Aknashen_N, Ukraine_N, GK2, CoreYamnaya. We fit models with 3 sources in total which allowed us to place the steppe component in TUR_C_BA along both the clines inferred by analysis (e.g., Volga=BPgroup+Russia_Karelia), as well as others non-attested with any genetic data (e.g., Aknashen+Ukraine_N) and unlikely on geographical grounds.”

We re-wrote the manuscript to clarify this: “Moreover, when we fix Çayönü as one source and add pairs of steppe sources”.

Line 569: There’s a typo – should be “Anatolian”.

We fixed this.

Lines 587–588: There’s no Extended Data Figure 2f and it is unclear which exact populations this admixture date is calculated for.

This was a typo. We fixed this to refer to Extended Data Figure 2e. The admixture refers to “a combined Caucasus-Anatolia population (Maikop-Armenia_C-TUR_C_BA)” consisting of populations that have CLV ancestry.

Line 596: The significance of the site of Armi is unclear to a reader not specialised in Anatolia.

Kroonen et al. cited in our paper discusses this in detail and gives references to the relevant literature. We have now added Ebla to the map of Extended Data Fig. 1f and also specified in the text that it is “just south of the proposed migratory path” and specified its early date. We think this will help the reader understand why the evidence from Ebla is significant as a potential witness to the presence of Anatolian speakers in its vicinity.

Figure 3: The haplogroup labels should be moved (for example placed vertically at the left side of the plot) since they’re currently difficult to see (especially G-P15). It is unclear which regions “EarlyMiddleBronzeAge” and “MiddleLateBronzeAge” comprise. Does EarlyMiddleBronzeAge include all samples from the listed cultures during the relevant time period? Which regions are the MiddleLateBronzeAge samples from? If the geographic groups in the legend don’t include all samples from the region (like in the case of North Caucasus, for example), a specification should be added.

We increased the font size of the labels and the vertical size of the plot. Since random jitter is added to the position of the points along the vertical axis, whether the labels overlap with the points is stochastic but the labels are all clearly visible. We prefer not to have the labels on the left as the figure legend would intervene between the labels and the distributions and horizontal space would be wasted that we prefer to use to more clearly show the distribution of points.

We have also added the individuals plotted in a new Online Table 6 so that the interested reader can identify exactly what is plotted in the broad categories defined.

Line 630: For clarity, it would be better to write “the Caucasus Neolithic represented by Aknashen was formed by the admixture of ...”.

We changed this.

Lines 640–653: This paragraph is confusing as it jumps from the EHG/UNHG vs West Asia/Caucasus differentiation to the Dnipro cline and then back to the EHG/UNHG + West Asia/Caucasus admixture date.

We removed the middle sentence about the Dnipro Cline, so the paragraph flows better now.

Lines 655–657: It doesn’t seem contrasting that while it’s not possible to pinpoint the geographical origin of a Yamnaya individual based on their genetic profile, it’s also not clear exactly where their ancestors lived.

The contrast is between their narrow geographical distribution within the North Pontic Region (in the associated North Pontic manuscript the vicinity of Mykhailivka as a candidate locale is highlighted) and their expansive later distribution. We re-wrote this to better convey that idea: “The proximate ancestors of the Core Yamnaya must have lived somewhere within the territory of the Serednii Stih culture across the North Pontic (or Dnipro-Don) region, contrasting with their expansive distribution after the formation of the Yamnaya archaeological horizon”

Lines 669–670: It should be stated accordingly in case this an assumption or cited if not.

We rewrote this: “The Don Yamnaya were formed in the late 4th millennium BCE (Extended Data Fig. 2b), a time during which, one may assume, unmixed UNHG, after a millennium or more of the Serednii Stih culture, would be rare if they existed at all.”

Line 673: The word “there” is not necessary here and creates confusion.

This was removed.

Lines 679–680: It would be helpful to state how this was tested.

We added a new piece of analysis in the supplement in the “Grouping core Yamnaya individuals” section to support the idea that the Don Yamnaya is a genetic cul-de-sac:

“The Don Yamnaya do not belong to the main core Yamnaya cluster, so we were curious if we could identify other Yamnaya-Afanasievo individuals similar to them in our dataset using a similar procedure as with the Core Yamnaya. However, when we apply the exact same criteria (with the difference that we use the p-value for being a fit with the Don Yamnaya (instead of Samara Yamnaya), we only find 4 fits (compared to 104 for the Samara Yamnaya). These four are a PCA outlier from Krivyansky in the Lower Don (I8950), two individuals marked as Yamnaya-Catacomb transitional (I23654 and I23655) also from the Don, and only a single individual belonging to the Afanasievo (I13812; p=0.201) which, however, also fits as part of the Core Yamnaya grouping

($p=0.435$). It appears that the expansive distribution of the Yamnaya stemmed primarily from the core (Samara-like) group; the Don Yamnaya have no similar evidence of wide expansion.”

In the revised manuscript we point to this analysis: “It is interesting that after the Don Yamnaya formed they participated little or not at all in the Core Yamnaya expansion to either the Altai or SE Europe, as virtually no high-quality individuals outside the Don are a clade with it (Supplementary Information, section 2) and thus the Lower Don represented a cul-de-sac for the Yamnaya expansion.”

Lines 720–721: Allentoft et al. 2015 should also be cited here.

We added this.

Line 724: For clarity, it would be good to replace “it” with “the Corded Ware population”.

We changed this.

Figure 4: It is stated that the individuals are projected in the same space as on Figure 1, but the PC2 axis seems to be flipped. The Core Yamnaya symbols on panel a are very similar (especially given their small size), rendering them meaningless. UNHG is not on panel a so it’s not clear directly from here that the Don Yamnaya tend towards it.

We rotated the orientation of Fig. 4 so it is now the same as that of the re-oriented Fig. 1 in which UNHG is top-left and EEF is bottom-left. We did not include UNHG on panel a as it would dominate the Figure and cause the Yamnaya individuals to cluster even closer together visually and be less distinguishable; that the Don Yamnaya tends towards UNHG can be better seen in Fig. 1. We also changed the Core Yamnaya symbols so as not to all be black and thus not appear “smudged” to each other; the great homogeneity of the Core Yamnaya which occupy a small area of the plot while consisting of many individuals is inherently challenging for visualization purposes, making it practically impossible to show individuals that overlap with each other in the same PCA space. To make it possible for interested readers to appreciate the very subtle differences that do exist, we added their raw co-ordinates to an Online Table 7 which we refer to in the legend.

Lines 748–751: This sentence is a bit complicated and doesn’t seem fully grammatically correct.

We re-wrote the sentence as follows: “Figure 5 shows the results separately analyzed for Core Yamnaya dating to the first three hundred ($n=25$) and later three hundred ($n=26$) years of our sampling who produce a 95% confidence interval of 3829-3374 BCE and 3642-3145 BCE for the time before growth respectively.”

Lines 756–758: It would be clearer if Yamnaya was mentioned here.

We changed “a subgroup” to “the Yamnaya subgroup”.

Line 763: Please specify what “shading shows uncertainty intervals” means.

We clarified this: “*shading shows confidence intervals provided by HapNE (dark: 50%, light: 95%)*”

Figure 5: Panels a and b make it seem like the Ne trajectories of the two sub-populations are very different due to the different y axis scales. It is understandably difficult to present these, but maybe having a gap in the y axis would be clearer. Also, the y axis scale of panel c is quite unusual and difficult to interpret. In sum it seems that panels a and b are perhaps redundant, and it would suffice to clarify the y axis scale on panel c.

We agree with the reviewer that conveying the information was difficult. We believe that panel c, where the y-axis is shown on a log-scale, is the way to proceed. However, some readers might not be used to such a scaling, and this is the reason why we also present the two plots on a linear scale.

To clarify this issue, we have edited the legends in the following way:

OLD:

We infer an extraordinary population expansion (c) after 3642-3374 BCE (intersection of 95% confidence intervals for the two analyses for the minimum), from a time when the effective size is a few thousand to an order of magnitude larger.

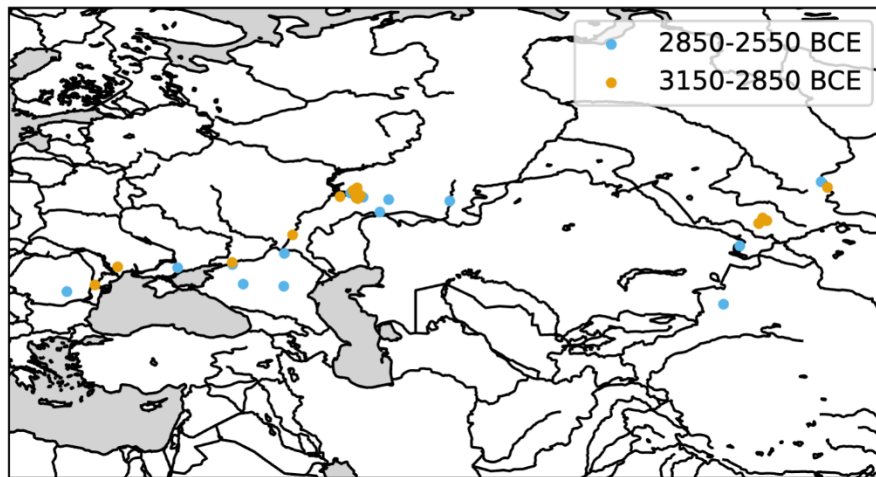
NEW:

Jointly displaying these two trajectories on a log-scale reveals an extraordinary population expansion (c) after 3642-3374 BCE (intersection of 95% confidence intervals for the two analyses for the minimum), from a time when the effective size is a few thousand to an order of magnitude larger.

It seems that this is the only analysis where the Core Yamnaya is split into two temporal groups. It would be great to see the geographic distribution of these two groups.

The two groups have similar geographic distributions which we show below:

Geographic Distribution of Core Yamnaya samples used in the HapNe analysis



Furthermore, we are wondering what would happen if one would sample an expanding population in two time points, so that the later sample has seen more of the expansion. Wouldn't the later population show a steeper N_e increase simply because it is a temporal subset not a population subset of the original population? Simply because also in an expanding population not all members of the population take part of the expansion?

The reason we split the Core Yamnaya into two temporal groups was precisely in order to be able to appreciate whether there are differences in the reconstructed curves, which could happen as the reviewer says if the earlier or later individuals included groups that did not participate fully in the expansion. The fact that the reconstructed curves overlap is thus reassuring. We now added a clause explaining this to the text. Given that the curves are overlapping over the reconstructed range, we don't think it is helpful to show the differences in the geographic distribution.

If we correctly understood this remark, the reviewer asked what would happen if we applied HapNe to two distinct sets of individuals from an expanding population. We agree that the analysis based on the most recent individuals will have a higher N_e at their first generation, because the population at that time is higher than the one of the other group at their sampling time. Note, however, that the predicted times are shifted by the time offset between the two populations so that the values of N_e at a given year should match in both experiments. Regarding the second part of the comments, our inferred value of N_e is based on the Wright-Fisher model, and potential caveats are present in the Supplementary Material.

OLD:

[...] when the effective size is a few thousand to an order of magnitude larger.

NEW:

[...] when the effective size is a few thousand to an order of magnitude larger. The offset on the x-axis is due to the difference in sampling time between the two groups.

Line 772: It would be helpful to be a bit more specific about what “before the Yamnaya” means.

We added our precise definition in Methods: “Pre-Yamnaya individuals plotted in Fig. 6 are from the 5500-3500 BCE period.”

Line 777: Please specify what “close genetic relatives” means here in terms of degrees of relatedness.

We define close genetic relatives as: “defined as sharing at least three $\geq 20\text{cM}$ segments or a total sum of IBD $\geq 100\text{cM}$ ”.

The criterion of at least three $\geq 20\text{cM}$ segments is the same one used in the Ringbauer et al. Nature Genetics 2024 ancIBD paper to define close relatives in Figure 3 of that study (excerpted below, red points in Figure 3a), and based on the simulations in Figure 3b corresponds approximately to individual related within the 5th degree and closer.

We have revised the text in our manuscript to the following “defined as sharing at least three $\geq 20\text{cM}$ segments (corresponding to relatedness to within about the 5th degree⁶) or a total sum of IBD $\geq 100\text{cM}$ ”

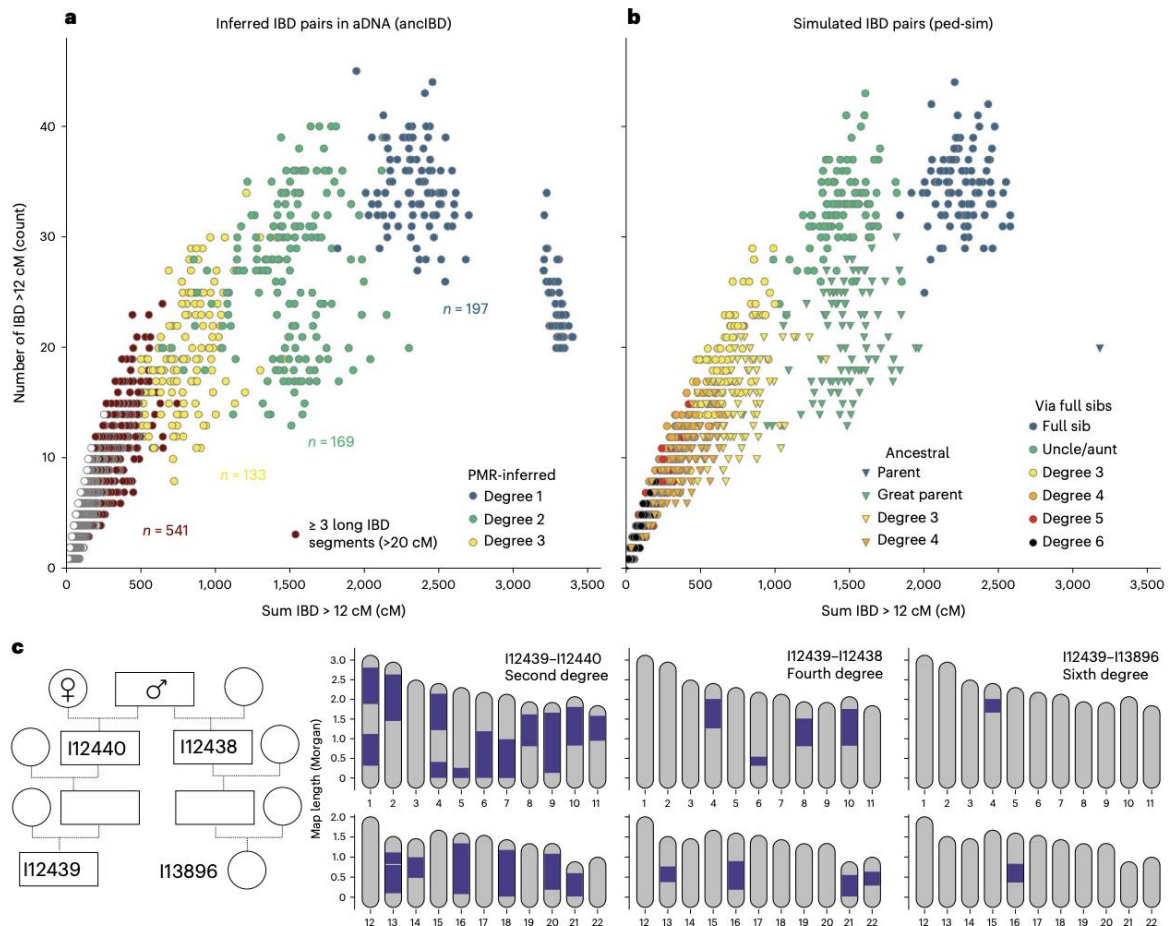


Fig. 3 | Inferring biological relatives in the aDNA record using long IBD inferred with ancIBD. a, Inferred IBD among pairs of 4,248 ancient Eurasian individuals. The plot visualizes both the count (y axis) as well as the summed length (x axis) of all IBD > 12 cM long. For comparison, we colour-code pairs on the basis of relatedness estimates from pairwise mismatch rates (PMR) that can detect up to third-degree relatives (Supplementary Note 9). We also annotate new relatives found by ancIBD, indicated by at least three very long IBD segments (> 20 cM) typical of up to sixth-degree relatives. **b**, Simulated IBD among pairs of relatives. For each relative class, we simulated 100 replicates using the software

ped-sim²⁵, as described in Supplementary Note 8. As in **a**, we depict the summed length and the count of all IBD at least 12 cM long. **c**, Inferred IBD among four ancient English Neolithic individuals, who lived about 5,700 years ago and were entombed at Hazleton North long cairn. A full pedigree was previously reconstructed using first- and second-degree relatives inferred using pairwise SNP matching rates²⁶. We depict all IBD at least 12 cM long. The four individuals were genotyped using 1240k aDNA capture (I12438, 3.7× average coverage on target; I12440, 2.1×; I13896, 1.1×; I12439, 6.7×).

Figure 6: The panel d label is missing from the figure legend. The fonts are too small on panels c–f and on the x axis of panel g.

We added (d) in the figure legend. We also increased the fonts in panels c–g and they now much more readable.

Line 826: Would be useful to also explain how Yamnaya can be linked to Balkan languages.

We now wrote: “Third, the Yamnaya can be linked to the languages of Armenia⁷ via both autosomal and Y-chromosome ancestry after ~2500 BCE, and to the Balkans⁸ where, in addition to Greek, several poorly-attested ancient Indo-European languages were spoken in antiquity, such as Illyrian and Thracian.”

Lines 839–840: This sentence could be written in a more neutral way.

We removed: “The technical reasons for these older dates will doubtlessly be debated by linguists.”

Lines 868–870: Please add references for the horse-wagon technology and for Proto-Anatolian languages being recorded in writing.

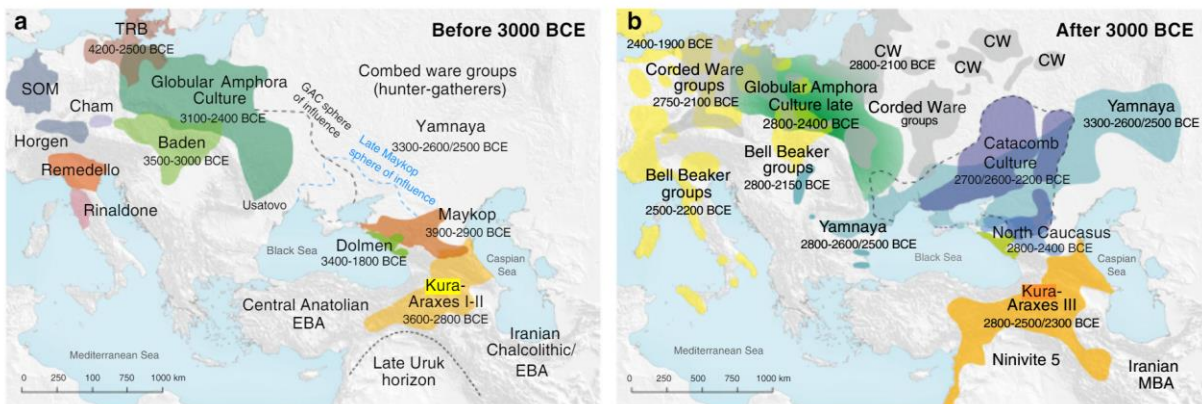
We added references and clarified that the daughter languages are recorded in writing.

Lines 911–912: There seems to be a grammatical issue (“have” and “had”), but the sentence is also quite complex and difficult to understand so might be worth trying to rephrase.

We fixed the grammar and also re-wrote the sentence.

Lines 914–916: It’s not clear how/why the Kura-Araxes expansion would’ve “driven a wedge” between steppe and West Asian Indo-Anatolian speakers.

The geographical distribution of the Kura-Araxes expansion in the South Caucasus and eastern Anatolia encompasses the route that people from the North Caucasus could have taken to reach Central-Western Anatolia where Anatolian languages were spoken. See, for example, Fig. 1a,b of Wang et al. (2019) which we reproduce below. We think the description in the text accurately captures the pattern shown in the map.



Lines 934–938: These two questions are fully left in the air without any reference to them in the concluding sentences.

We think it is of value to pose questions that we don’t currently have an answer for to highlight hypotheses raised by the new findings in this study that have not been resolved but can potentially be addressed in further work.

Extended Data Figure 1: It should be specified that panel a includes Eastern models, together

with an explanation of what that means. It should be added that panel b models also include Cayönü next to western sources. The p-value threshold labels should be the same size on panel c as on previous panels. It should be added, which model the BPgroup proportion in panel d comes from. It is not clear how the pairs of populations mentioned form a route on panel f.

We have re-written the descriptions of the panels to specify these points. We have also made the p-value labels the same size across panels a-c. Finally, we enhanced and enlarged the map to better show the set of sites with CLV ancestry along the path from the North Caucasus to Central Anatolia.

Extended Data Figure 2: Presumably DATES was not run using all European hunter-gatherers and all West Asian populations? In that case it should be specified which populations were used. The same applies to other panels where the description is not specific.

DATES and related methods benefit from large sample sizes so we added many European and West Asian populations to maximize statistical power and drive down the standard errors of the estimation. We now list all these populations in the Supplementary Information in the “When was the Yamnaya formed and when did it expand?” section.

Extended Data Figure 3: The label “Lebyazhinka_HG” is placed on top of the barplot, blocking data from view.

We moved this label to the right where it does not obstruct the data.

Supplementary Fig 2.7: There is what seems a state border between Ukraine and Crimean Peninsula. This should be removed.

We modified the map removing all lines that could be read as political borders.

Archaeological site information (Archaeological Supplement and Online Table 1): Progress-2 and Vonyuchka are described in the Archaeological Supplement but are not in Online Table 1 – should be made clearer why they're described in the Archaeological Supplement. Remontnoye Late Steppe Eneolithic site names are not in the table of contents of the Archaeological Supplement which makes it hard to link it with Online Table 1. Kurmanaevka-3 site is not included in the Locality column of Online Table 1. Krivyansky-7 sample is listed as Krivyansky-9 in the Locality column of Online Table 1. Archaeological Supplement descriptions are missing for many sites.

Progress-2 and Vonyuchka were previously published and we include detailed archaeological descriptions as background due to their extreme importance as part (together with the newly described individuals from Berezhnovka) of the BPgroup and PVgroup subsets of the Caucasus-Lower Volga cline.

We removed Kurmnaevka-3.

We corrected Krivyansky-7.

We also filled in the missing archaeological information for sites included in the dataset.

This email has been sent through the Springer Nature Manuscript Tracking System NY-610A-SN&MTS

Confidentiality Statement:

This e-mail is confidential and subject to copyright. Any unauthorised use or disclosure of its contents is prohibited. If you have received this email in error please notify our Manuscript Tracking System Helpdesk team at <http://platformsupport.nature.com>.

Details of the confidentiality and pre-publicity policy may be found here <http://www.nature.com/authors/policies/confidentiality.html>

- 1 Shishlina, N. I. *et al.* Paleoecology, Subsistence, and 14C Chronology of the Eurasian Caspian Steppe Bronze Age. *Radiocarbon* **51**, 481-499 (2009). <https://doi.org:10.1017/S0033822200055879>
- 2 Anthony, D. W. *et al.* The Eneolithic cemetery at Khvalynsk on the Volga River. *Præhistorische Zeitschrift* **97**, 22-67 (2022). <https://doi.org:doi:10.1515/pz-2022-2034>
- 3 Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207-211 (2015). <https://doi.org:10.1038/nature14317>
- 4 Fu, Q. *et al.* The genetic history of Ice Age Europe. *Nature* **534**, 200-205 (2016). <https://doi.org:10.1038/nature17993>
- 5 Wang, C.-C. *et al.* Ancient human genome-wide data from a 3000-year interval in the Caucasus corresponds with eco-geographic regions. *Nature Communications* **10**, 590 (2019). <https://doi.org:10.1038/s41467-018-08220-8>
- 6 Ringbauer, H. *et al.* Accurate detection of identity-by-descent segments in human ancient DNA. *Nature Genetics* **56**, 143-151 (2024). <https://doi.org:10.1038/s41588-023-01582-w>
- 7 Lazaridis, I. *et al.* The genetic history of the Southern Arc: A bridge between West Asia and Europe. *Science* **377**, eabm4247 (2022). <https://doi.org:10.1126/science.abm4247>
- 8 Mathieson, I. *et al.* The genomic history of southeastern Europe. *Nature* **555**, 197-203 (2018). <https://doi.org:10.1038/nature25778>

Referees' comments:

Referee #1 (Remarks to the Author):

I am glad to see that the authors have made several adjustments to the paper based on the review, and in that sense I am over all satisfied with the new version. However, there is one things that still needs to be adapted: the statement "entering Anatolia from the east" in line 24. In the response to the review, the authors themselves state that "People with CLV ancestry may have taken any route to reach Anatolia and from the genetic point of view we cannot distinguish between CLV people in the Caucasus-Lower Volga region and similar people who reached Moldova and Hungary", and that a western origin of CLV ancestry into Central Anatolia cannot be excluded. It therefore follows that the paper cannot simply state that CLV arrived in Central Anatolia "entering Anatolia from the east" as a fact. I therefore maintain that this statement should be adapted to, for instance, "entering Anatolia probably from the east".

We are glad that we were able to address the reviewer's comments. For the remaining comment, we now say "entering Anatolia likely from the east".

Referee #2 (Remarks to the Author):

My concern about PRE correction has been adressed in away way that I find acceptable. Their position is now made clear. Thus, no more objections, and recommendation for publication without any further corrections.

We thank the reviewer.

Referee #3 (Remarks to the Author):

The authors have revised the manuscript largely to our satisfaction.

We do see that the authors have considered our suggestions regarding the figures but have reached the conclusion that the original ones were in fact best options. We have to agree to disagree on that. However, we don't think this warrants further delay in publication.

Beyond figure design, we noticed (Regarding Figure 1) non-matching labels not symbols. For example, Maykop in panels a and b versus Maikop in c; Karelia in the top portion of panel c and Russia_Karelia in the middle portion. Also, we would like to re-iterate that it would be beneficial to include a distinction between previously published, Nikitin et al. and your groups.

We thank the reviewer for their careful reading of the paper. We use Maikop in all panels of Fig. 1. We also use Russia_Karelia now in both portions of panel (c) which also matches the naming in panel (b).

Referee #4 (Remarks to the Author):

The authors have revised the manuscript largely to our satisfaction.

We do see that the authors have considered our suggestions regarding the figures but have reached the conclusion that the original ones were in fact best options. We have to agree to disagree on that. However, we don't think this warrants further delay in publication.

Beyond figure design, we noticed (Regarding Figure 1) non-matching labels not symbols. For example, Maykop in panels a and b versus Maikop in c; Karelia in the top portion of panel c and Russia_Karelia in the middle portion. Also, we would like to re-iterate that it would be beneficial to include a distinction between previously published, Nikitin et al. and your groups.

We thank reviewer 4 for their joint comments with reviewer 3 addressed above.