

## Peer Review File

---

Chronometric dating and stratigraphic data support discontinuity between Neanderthals and early Homo sapiens in the Italian Peninsula



**Open Access** This 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 the cases where the authors are anonymous, such as is the case for the reports of anonymous peer reviewers, author attribution should be to 'Anonymous Referee' followed by a clear attribution to the source work. The images or other third party material in this 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 <http://creativecommons.org/licenses/by/4.0/>.

## Reviewers' Comments:

Reviewer #1:

Remarks to the Author:

The paper is about the demise of Neanderthal and the diffusion of anatomically modern human (AMH) in Europe during MIS 3, the so called “transition” that is since decades at the center of the scientific debate. Thanks to the application of new dating methods the quality of the data is constantly improving, allowing to fine-tune the chronology of the transition. The authors are all well-known and reputed researchers in this field.

The paper presents a novel dataset about the chronology of some Italian Uluzzian, Protoaurignacian and Mousterian key-contexts, obtained with different dating methods: 14C, and OSL and their Bayesian modeling.

This study is technically sound, and new data obtained with appropriate techniques, analyzed and interpreted carefully, and presented in sufficient detail. Moreover, the level of support for the conclusions provided for the authors’ claims is well-founded and the conclusions well-argued and convincing.

The results will be important to the field, advancing our understanding about the spatiotemporal dimension of the interaction between the last Neanderthals and first AMH in a key-region of S Europe. They will certainly contribute to our understanding of the coexistence of the two human species in the various regions of Europe, their possible interbreeding/interaction, and the duration of their co-presence.

However, the authors are missing to discuss the chronological evidence provided by some important central-Italian sites, such as La Fabbrica (Tuscany) and Colle Rotondo (Latium) that have been recently OSL dated and seem to corroborate previous suggestions that the Middle-to-UpperPaleolithic transition occurred as steps of rapid changes and geographically uneven rates of spread: see Villa P, Pollarolo L, Conforti J, Marra F, Biagioni C, Degano I, et al. (2018) From Neandertals to modern humans: New data on the Uluzzian. PLoS ONE 13(5): e0196786.

<https://doi.org/10.1371/journal.pone.0196786>

Another site to incorporate in the Discussion is Grotta Reali, whose chronological attribution to the MIS 3 allows placing the settlement in the last phase of Neanderthals presence in Europe, see: Peretto C, M Arzarello, M Coltorti, M Bertolini, Qiao-Yu Cui, O De Curtis, V Lebreton, G Lembo, L Marquer, P Pieruccini,

S Ravani, E Rufo, B Sala, S Talamo, U Thun Hohenstein, Grotta Reali, the first multilayered mousterian evidences (sic) in the Upper Volturno Basin (Rocchetta a Volturno, Molise, Italy) *Archaeological and Anthropological Sciences* (2020) 12: 67 <https://doi.org/10.1007/s12520-020-01015-3>

This site is dated from between 50,940 and 40,370 cal BP and provides detailed information on the human occupation in southern Italy, and contributes to the international debate on technical behaviour at the end of the Mousterian.

Furthermore, the very recent paper by Nicolas Teyssandier reviewing the Middle-to-Upper Paleolithic transition in Europe, and the Neanderthals/Homo sapiens technocomplexes among which the Uluzzian should be integrated in the discussion of the data too: *Us and Them: How to Reconcile Archaeological and Biological Data at the Middle-to-Upper Palaeolithic Transition in Europe?* *Journal of Paleolithic Archaeology* (2024) 7:1, <https://doi.org/10.1007/s41982-023-00166-8>

Finally, a few typos in the Tables of the Supplementary materials:

Line 1169, Table 1, first row: *Ilex aquifolium* in Italic font

Line 1164: Tables for Italian paper / what does it mean?

Line 1183: inside the Table, sample Cvl 8, *Dentalium* sp. Italic font

Line 1202: Table 4 in the table Sampler of charcoal instead of Sample

Line 1203: *Oscurusciuto* not *Oscuruscuito*

Reviewer #2:

Remarks to the Author:

This paper presents a large, robust set of dates and chronological models that will be essential to future research on Neanderthal and *Homo sapiens* occupations of Europe. This is the right venue for top-level research, and I recommend this paper be published after revisions. The data and models and conclusions are very strong, but the organization of the text stands to be improved. As is, a broad readership will have difficulty following the argument. Here I list some suggestions for improvement.

**Abstract.** The abstract should not be an introduction, but summarize the conclusions and results, including the temporal range for the three lithic industries in three conclusions. I would emphasize that models are consistent and reliable across three independent types of chronological markers (with almost no outliers): 14C, OSL, and regional tephras. This is a remarkable result for a chronology paper. Also, the models produce much narrower error ranges than before.

As a non-regional specialist, I find the first paragraph quite confusing, with a soup of names of periods and industries, regional debates about definitions, details about excavations at a single site, transitional lithics, a history of ideas, and comments on the latest studies. I think some of this is relevant, but in the discussion and/or Supplementary Material (SM).

The opening of paper needs to be written for a much more general audience, with a clearer statement of what we know now and the research questions and motivations. I think it's important to clarify that for this paper, the data cannot speak to the species that made each lithic industry, but they can provide temporal limits for the industries. In this paper, if I've understood correctly, we are dealing with dates associated with three industries: Mousterian, Uluzzian, and Protoaurignacian.

1. Please define these three lithic industries and differences between them, with photographs. Are there any other material differences, for example, are shell ornaments only found in association with Uluzzian lithics?

2. I would highlight this fundamental point: "Stratigraphically, [Uluzzian lithics] are always found above Mousterian [lithics], and a sedimentological unconformity is present in most cases between the two" (line 84). What is stratigraphic situation for Protoaurignacian lithics? On the map, please include the undated sites that have this stratigraphic setup – this will make your regional

generalizations much stronger, and you can hypothesize that chronologies are broadly similar at undated sites.

3. It seems to me that the consistent stratigraphic setup is what motivates this study – it suggests there is a regional pattern we need to track in time. Hence, the study's goal is to make a reliable, quantitative estimate for the final regional use of Mousterian lithics, the first use of Uluzzian lithics, and the first use of Protoaurignacian lithics. And make comparisons between northern and southern Italy. Another possible goal: define the gap between the three industries, which could be quantified with Interval query (across all site-specific boundaries). This is mentioned (line 474, 644), but I think it needs to be part of the study's goals, since it could tell us if the makers of the three industries possibly interacted (as noted in abstract).

4. I recommend emphasizing that we can now build a much more reliable chronology than ever before, not just because of new dates, but because:

- Significant IntCal20 updates to this age range. This is lost in the conclusion, but to me is a motivation to do this the study now.
- State-of-the-art sample prep for both 14C and OSL and reduced AMS error ranges can make significant improvements over existing chronologies. These are key contributions from the list of authors.
- Models have strong agreement indices even when cross-referenced with three independent time markers: 14C, TL, and tephra layers.

Alignments with the NGRIP record (Figs 3, 10) and genetic studies appear later, but are not in the introduction. They are interesting, but I would cut them, since the paper's data do not directly address these issues.

The intro and conclusion mention Uluzzian lithics in Greece, but these are not mapped or discussed. I recommend a map of these sites, which would give the paper much more regional impact. Why don't you include sites from other parts of Europe? Are there no similar lithics outside of Italy and Greece? Limiting the study area and comparisons to modern national boundaries places reflects a modern bias on data patterns created long before these countries existed.

The authors seem to be writing for regional specialists, which makes word choice confusing. It seems the authors are mixing terms for cultures, people, layers, phase, and occupations, when in fact we are talking about dates associated with three lithic industries: Mousterian, Uluzzian, and Protoaurignacian. For example:

Uluzzian and Protoaurignacian archaeology (line 440), populations (529), Protoaurignacian gic phase (331), Mousterian layer, Uluzzian occupation (476), the Uluzzian, the Mousterian, Uluzzian technocomplex (41), the identity of the Uluzzian (495) / Ulizzians (645).

I suggest the authors be more careful and consistent in referring to lithic industries, layers with these industries, dates associated with the lithics, and people who made and used these lithics. Mixing terms for cultures and artifacts recalls a bad culture history habit and can be quite confusing

for non-archaeologists. For example: "... a reduction in Neanderthal populations prior to the spread of Uluzzians in the region (line 644). What you mean, I think, is that there a potential occupation hiatus between Hominids using Mousterian lithics and those using Uluzzian lithics.

For the purposes of this paper, you seem to be following the following simplification: Mousterian lithics = Neanderthals and Uluzzian or Protoaurignacian lithics = Homo sapiens. This is fine, and would help guide the reader, but state this generalization more clearly in the introduction, since a general reader will become quickly confused. The discussion is the place to mention possible problems with this simplification.

## Figures

Figs 7–10. It would be clearer to see all of this in a single figure. I am imagining a single line for each site, with stacked HPDs for Mousterian end, Uluzzian start, and Protoaurignacian start (and if they fit, KDEs as well). This would make it clear which sites have which data. The same figure could include the summary HPDs based on the phases-of-boundaries approach. Northern Italy at the top; Southern at the bottom. This single figure could 1) summarize all the important results of the paper, and 2) make it visually clear how similar (or different) boundaries are at each site and 3) make it clear how many sites have data that inform regional conclusions, for example, the final use of Mousterian lithics is more robust since more sites have similar boundaries; Uluzzian and Protoaurignacian regional boundary trends are only based on three sites.

Fig 7. Clarify that the inset is based on a Phase of the ending Boundaries. Is this HPD a Date query from within this phase? A KDE is preferable (though will be nearly the same result, probably). I agree with your approach of grouping these boundaries in a phase, but I think a broader readership needs the clarification that this collapses intra-regional variability, for example, Mousterian lithics were last used at Rio Secco well before their final use at other sites.

Fig 8. Comparisons of starting boundaries for the Uluzzian. The authors suggest there is no statistically significant difference between these boundaries based on the command Difference. However, if we queried them with Order, we would see a strong likelihood that Uluzzian lithics appeared at Castelcivita before Cavallo (as noted in line 609), both sites in the south. While the HPDs do overlap, three boundaries are not a strong basis for regionally synchronous changes. Perhaps we don't need to force them into a phase? I would say three boundaries are too few to group, as in Figure 10, which has three Protoaurignacian starting boundaries that are not grouped. It seems quite plausible that these industries first appeared at different times; there may not be a regional trend here (at this smaller temporal scale).

Fig 9. Here I have the same question as the inset in Fig 7. All three boundaries seem to be the main conclusions of the paper. I would find it very useful to see a KDE summary of the dates from each lithic industry next to the boundaries, as in Bronk Ramsey (2017:Fig 3). With these KDEs, you could speak to the spans that people used these lithics, not only the boundaries between them.

Figs 11–12. Clarify if these maps are based on the results from Figs 7 and 8 (which seems to be the

case). Also, clarify that these circles showing ending and starting boundaries are probabilities, not the durations of when these lithic industries were in use. I would suggest caution with these figures, which have a tendency to be misread. In this case, the largest circles that appear on the most maps are in fact the least precise boundaries, which is often confused for duration or significance. I suggest the authors consider a more intuitive figure: circles representing KDEs for each lithic industry. In this case, change circle sizes at each site would better show declining Mousterian dates and rising Uluzzian and Protoaurignacian dates, perhaps in a single figure.

Details to clarify in the Supplementary Material:

- Anything that does not directly deal with the research questions should be moved to the supplementary material, for example: model sensitivity to Firenze dates (line 379), dates with caveats (paragraph starting on line 383), and model variants. If these are relevant to the paper's questions, clarify how.
- Clarify that in Europe, OSL and 14C have been robustly correlated. TL and OSL remain less reliable in regions with younger geology, including the Andes (Marsh et al. 2021). This paper's results provide further confirmation of this correlation.
- Please clarify the basis for the Delta\_R on shell dates ( $4\pm 189$  years). Are there paired dates?
- Add Cavallo's unpublished field notes here (line 474) – this is too fine a detail for the main text. The discussion should focus on inter-site comparisons.
- The authors should not assume readers know about site details, as this is not a regional journal, for example, line 325: "As is widely known, the Protoaurignacian layers are capped by a series of flowstones...". This seems like a detail for the Supplementary Material.

## References

Bronk Ramsey, C., 2017. Methods for Summarizing Radiocarbon Datasets. *Radiocarbon* 59, 1809–1833. <https://doi.org/10.1017/RDC.2017.108>

Marsh, E.J., Korpisaari, A., Mundt, S.P., Gasco, A., Durán, V., 2021. Radiocarbon vs. luminescence dating of archaeological ceramics in the southern Andes: a review of paired dates, bayesian models, and a pilot study. *Radiocarbon* 63, 1471–1501. <https://doi.org/10.1017/RDC.2021.82>

Reviewer #3:

Remarks to the Author:

The authors of this manuscript present results of an impressive project to explore whether or not there was overlap between Neanderthals (associated with the Mousterian tool industry) and

modern humans (associated with Uluzzian tool industry) in four cave sites in southern Italy. They then compare the timing of the Uluzzian tool industry to that of another modern human tool industry (Protoaurignacian) using spatial-temporal mapping to interpret the results. The project shows the power of careful radiocarbon and optical stimulated luminescence (OSL) dating combined with stratigraphy in Bayesian models. The radiocarbon dating uses the most advanced pretreatment techniques as well as using ZooMS to identify the species of bone fragments.

However, I feel that the interpretation of the model results regarding the Mousterian ending prior to the beginning of the Uluzzian may be slightly overstretched. At Grotta del Cavalo there is a clear end to the Mousterian below a tephra layer and apparently some time elapsed before for the start of the Uluzzian but there appears to be potential overlap between the end of the Mousterian and the start of the Uluzzian in the Grotta di Castelcivita age model (Figure 3). The individual model boundaries for the start of the Uluzzian at these two sites are used as priors in a phase model along with that from Riparo Brion (Figure 8) to produce a single HPD (Figure 9). With only three sites, does the later start and tight HPD of the Cavolo Uluzzian disproportionately shift the combined HPD to age younger age? If so, that may be forcing the start of the Uluzzian to be later than the end of the Mousterian rather than overlapping with it. I think the authors should acknowledge that this conclusion is somewhat uncertain.

Also, in the conclusions the authors state: 'We also see evidence that there may have been reduction in Neanderthal populations prior to the spread of Uluzzians in the region'. As far as I can tell this evidence is not presented in the paper.

Otherwise, I think the paper is a significant contribution to the field and should be published with minor revisions.

Specific comments:

Line 137 and elsewhere with regards to Castelcivita: What does Taglio refer to? It is confusing since it doesn't seem to correspond to stratigraphic units and this is not shown on Fig S2.

Line 192: 'Shell carbonates' Are the shell carbonates all marine shells?

Line 360: 'We calculated and applied a new local marine reservoir correction (or Delta\_R) for the shell dates in the model of  $4 \pm 189$  years, along with the Marine20 calibration' In the supplementary info it is stated that Delta\_R has been updated to Marine20 but the authors doesn't state where the original reservoir age data is from. Is the data from the [calib.org/marine](http://calib.org/marine) website or is it based on previously published marine/terrestrial pairs?

Line 496-498: 'There appears to be a slight delay in the Cavallo site by comparison, as hinted at above, although it is not statistically significant compared with Castelcivita.' How was the significance tested?

Figure 1 or on Figures 11-12: Please label Oronto Chanel, Adriatic Sea, Conero Promontory, Uluzzo Bay etc. on the figure as these locations are discussed with regards to the spatial-temporal discussion and in the conclusions.

Figures 11-12: It would make more sense to reverse the order of these so that the oldest plots were first. Also in captions, please explain what the size of the circle means.

## REVIEWER COMMENTS and replies in red

Reviewer #1 (Remarks to the Author):

The paper is about the demise of Neanderthal and the diffusion of anatomically modern human (AMH) in Europe during MIS 3, the so called “transition” that is since decades at the center of the scientific debate. Thanks to the application of new dating methods the quality of the data is constantly improving, allowing to fine-tune the chronology of the transition. The authors are all well-known and reputed researchers in this field.

The paper presents a novel dataset about the chronology of some Italian Uluzzian, Protoaurignacian and Mousterian key-contexts, obtained with different dating methods: 14C, and OSL and their Bayesian modeling.

This study is technically sound, and new data obtained with appropriate techniques, analyzed and interpreted carefully, and presented in sufficient detail. Moreover, the level of support for the conclusions provided for the authors’ claims is well-founded and the conclusions well-argued and convincing.

The results will be important to the field, advancing our understanding about the spatiotemporal dimension of the interaction between the last Neanderthals and first AMH in a key-region of S Europe. They will certainly contribute to our understanding of the coexistence of the two human species in the various regions of Europe, their possible interbreeding/interaction, and the duration of their co-presence.

We thank the reviewer for these positive comments.

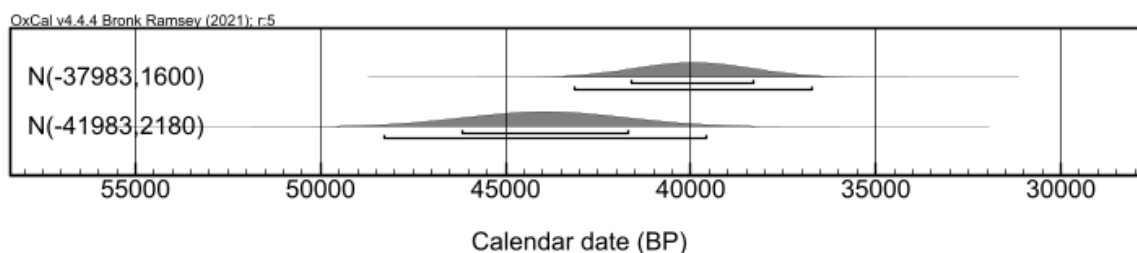
However, the authors are missing to discuss the chronological evidence provided by some important central-Italian sites, such as La Fabbrica (Tuscany) and Colle Rotondo (Latium) that have been recently OSL dated and seem to corroborate previous suggestions that the Middle-to-UpperPaleolithic transition occurred as steps of rapid changes and geographically uneven rates of spread: see Villa P, Pollarolo L, Conforti J, Marra F, Biagioni C, Degano I, et al. (2018) From Neandertals to modern humans: New data on the Uluzzian. PLoS ONE 13(5): e0196786. <https://doi.org/10.1371/journal.pone.0196786>

Another site to incorporate in the Discussion is Grotta Reali, whose chronological attribution to the MIS 3 allows placing the settlement in the last phase of Neanderthals presence in Europe, see: Peretto C, M Arzarello, M Coltorti, M Bertolini, Qiao-Yu Cui, O De Curtis, V Lebreton, G Lembo, L Marquer, P Pieruccini, S Ravani, E Rufo, B Sala, S Talamo, U Thun Hohenstein, Grotta Reali, the first multilayered mousterian evidences (sic) in the Upper Volturno Basin (Rocchetta a Volturno, Molise, Italy) Archaeological and Anthropological Sciences (2020) 12: 67 <https://doi.org/10.1007/s12520-020-01015-3>

This site is dated from between 50,940 and 40,370 cal BP and provides detailed information on the human occupation in southern Italy, and contributes to the international debate on technical behaviour at the end of the Mousterian.

We thank the reviewer very much for these additional comments. The La Fabbrica and Colle Rotondo sites are very important additions to the question of the Uluzzian/Mousterian in Italy, of that there is no doubt. At La Fabbrica an erosional surface separates the thin Uluzzian layer (layer 2) from the underlying Mousterian (layer 1a). The OSL dates also suggest a temporal interval, since layer 1a is dated to  $44 \pm 2.1$  and layer 2 is dated to  $40 \pm 1.6$  ka). The problem with Fabbrica in terms of the chronometric side, however, is the precision of the results; they are spread so widely in terms of their uncertainties that it becomes extremely difficult to reliably interpret their meaning (see image below). The GeGLF1 OSL date for the Mousterian level, for example, has a  $\pm 2100$  year error and in calibrated space (as shown in the figure below) the result at 95.4% spans 39,500—48,500 cal BP.





Colle Rotondo's OSL dates are 14-18,000 BP so much too late to be of use in this debate. The dates fall within the Epigravettian (Villa et al. 2018: OSL dating performed on three samples collected at the base (CR3) and at the top (CR2, CR1) of the layer incorporating the artifacts yielded unexpected young ages ranging from  $18,180 \pm 950$  for CR3,  $15,870 \pm 1,100$  for CR 2 and  $14,640 \pm 960$  for CR1. At the time indicated by the OSL dates there are only Epigravettian assemblages in Italy. The lithic assemblage, unaffected by eolian transportation, is likely to have remained close to the exposed surface until around 14 ka, and then rapidly buried: a fact that fully justifies the OSL dates yielded by the sediment immediately below the artifacts (CR 3 sample, 18 ka) and immediately above them (CR2 and CR 1 samples, 15–14 ka). Thus the quoted OSL ages do not refer to the depositional event but to a later sun exposure: specifically the last time of exposure before rapid burial of the archaeological layer. The assemblage remains undated but the similarity to the La Fabbrica assemblage suggests that they may be broadly contemporaneous). For this reason we have mentioned in the updated text these sites, and thank the reviewer, but we have not included the results in the analysis.

The Grotta Reali, which is more extensively dated, also has some issues that are concerning. One issue is that the site does not have any post-Mousterian archaeology that can be dated. This means that there is a lack of constraint above the latest dated samples which means that the precision is lacking. The presence of the CI or well-dated Protoaurignacian levels would greatly benefit this, but of course it does not exist currently. For this reason the Reali site, though useful, is less useful in precisely dating this key phase. Nevertheless, we have included it, with caveats, in the updated analysis.

Furthermore, the very recent paper by Nicolas Teyssandier reviewing the Middle-to-Upper Paleolithic transition in Europe, and the Neanderthals/Homo sapiens technocomplexes among which the Uluzzian should be integrated in the discussion of the data too: Us and Them: How to Reconcile Archaeological and Biological Data at the Middle-to-Upper Palaeolithic Transition in Europe? *Journal of Paleolithic Archaeology* (2024) 7:1, <https://doi.org/10.1007/s41982-023-00166-8>

We had not seen this paper prior to submission. We note, however, that Teyssandier surprisingly fails to reference two of the key recent papers which present an alternative view on the authorship of the Uluzzian with specific reference to Grotta del Cavallo:

Moroni, A., Ronchitelli, A., Simona, A., Aureli, D., Bailey, S.E., Boscato, P., Boschini, F., Capocchi, G., Crezzini, J., Douka, K. and Marciani, G., 2018. Grotta del Cavallo (Apulia–Southern Italy). The uluzzian in the mirror. *Journal of Anthropological Sciences*, 96, pp.125-160.

Ronchitelli, A., Benazzi, S., Boscato, P., Douka, K. and Moroni, A., 2014. Comments on "Human-climate interaction during the Early Upper Paleolithic: Testing the hypothesis of an

adaptive shift between the Proto-Aurignacian and the Early Aurignacian” by William E. Banks, Francesco d’Errico, João Zilhão. *J. Hum. Evol.*, 73(107), p.e111.

To us, the review is useful, but in terms of the Italian sequence, somewhat selective (or lazy?) in terms of presenting the wider picture and significance particularly of the links between the Uluzzian and *H. sapiens* which we feel is compelling. Nonetheless, we quoted Teyssandier 2024. (SEE LINE 82).

Finally, a few typos in the Tables of the Supplementary materials:

Line 1169, Table 1, first row: *Ilex aquifolium* in Italic font

Line 1164: Tables for Italian paper / what does it mean?

Thank you, we have changed these.

Line 1183: inside the Table, sample Cvl 8, *Dentalium* sp. Italic font

In Arrighi et al 2020 an identification by a malacologist of the previously classified as “*Dentalia*” from Cavallo was carried out. Actually, the correct genus of these shells is *Antalis* and not *Dentalium*. Therefore, we corrected everywhere *Dentalium* to *Antalis* (Arrighi, S., Bortolini, E., Tassoni, L., Benocci, A., Manganelli, G., Spagnolo, V., Foresi, L.M., Bambini, A.M., Lugli, F., Badino, F., Aureli, D., Boschini, F., Figus, C., Marciani, G., Oxilia, G., Silvestrini, S., Cipriani, A., Romandini, M., Peresani, M., Ronchitelli, A., Moroni, A., Benazzi, S., 2020. Backdating systematic shell ornament making in Europe to 45,000 years ago. *Archaeological and Anthropological Sciences*, 12:59 <https://doi.org/10.1007/s12520-019-00985-3>.)

Line 1202: Table 4 in the table Sampler of charcoal instead of Sample

Done

Line 1203: *Oscurusciuto* not *Oscurusciuto*

Thank you for this, we have changed these. And thank you for the review!

Reviewer #2 (Remarks to the Author):

This paper presents a large, robust set of dates and chronological models that will be essential to future research on Neanderthal and *Homo sapiens* occupations of Europe. This is the right venue for top-level research, and I recommend this paper be published after revisions. The data and models and conclusions are very strong, but the organization of the text stands to be improved. As is, a broad readership will have difficulty following the argument. Here I list some suggestions for improvement.

Abstract. The abstract should not be an introduction, but summarize the conclusions and results, including the temporal range for the three lithic industries in three conclusions. I would emphasize that models are consistent and reliable across three independent types of chronological markers (with almost no outliers): 14C, OSL, and regional tephra. This is a remarkable result for a chronology paper. Also, the models produce much narrower error ranges than before.

We have rewritten the abstract along these lines (SEE LINES 47-61). We also added a reference in the conclusion regarding the synchronicity of the 14C and OSL (+ tephra) results which we agree are very encouragingly similar (SEE LINES 747-749). Thank you for this.

As a non-regional specialist, I find the first paragraph quite confusing, with a soup of names of periods and industries, regional debates about definitions, details about excavations at a single site, transitional lithics, a history of ideas, and comments on the latest studies. I think some of this is relevant, but in the discussion and/or Supplementary Material (SM).

We have rewritten the introduction to make this more straightforward and understandable for the wider readership. We also moved sections to the Supplementary Materials to provide more details there regarding the different industries and so on.

The opening of paper needs to be written for a much more general audience, with a clearer statement of what we know now and the research questions and motivations. I think it's important to clarify that for this paper, the data cannot speak to the species that made each lithic industry, but they can provide temporal limits for the industries. In this paper, if I've understood correctly, we are dealing with dates associated with three industries: Mousterian, Uluzzian, and Protoaurignacian.

1. Please define these three lithic industries and differences between them, with photographs. Are there any other material differences, for example, are shell ornaments only found in association with Uluzzian lithics?

2. I would highlight this fundamental point: "Stratigraphically, [Uluzzian lithics] are always found above Mousterian [lithics], and a sedimentological unconformity is present in most cases between the two" (line 84). What is stratigraphic situation for Protoaurignacian lithics? On the map, please include the undated sites that have this stratigraphic setup – this will make your regional generalizations much stronger, and you can hypothesize that chronologies are broadly similar at undated sites.

We added other (undated) sites to the location map. (SEE LINES 225-231)

3. It seems to me that the consistent stratigraphic setup is what motivates this study – it suggests there is a regional pattern we need to track in time. Hence, the study's goal is to make a reliable, quantitative estimate for the final regional use of Mousterian lithics, the first use of Uluzzian lithics, and the first use of Protoaurignacian lithics. And make comparisons between northern and southern Italy. Another possible goal: define the gap between the three industries, which could be quantified with Interval query (across all site-specific boundaries). This is mentioned (line 474, 644), but I think it needs to be part of the study's goals, since it could tell us if the makers of the three industries possibly interacted (as noted in abstract).

Yes this is correct. Concerning points 1), 2) and 3), we broadly modified the introduction following point by point the reviewer's suggestions (SEE LINES 85-104). We introduced a figure illustrating the the different assemblages (Figure 1, caption LINES 174-182) and we provided a more extended description of the Italian late Mousterian, Uluzzian, Protoaurignacian and Early Aurignacian (SEE LINES 119-270), and in the SI as well. Additionally:

- In recent weeks a detailed and updated study on the Aurignacian layers of Castelcivita (rsa', gic and ars) was completed, which evidenced that the assemblages from gic and ars are to be classified as Early Aurignacian instead of Protoaurignacian (Falcucci, A., Arrighi, S. Spagnolo, V., Rossini, M., Higgins, O., Muttillio, B., Martini, I., Crezzini, J., Boschin, F., Ronchitelli, A., Moroni, A., A pre-Campanian Ignimbrite techno-cultural shift in the Aurignacian sequence of Grotta di

Castelcivita, southern Italy). Therefore we corrected their attribution in the paper (SEE LINES 188, 235, 317-319, 446).

- In the article we prefer to use the term “technocomplexes” or the more generic “cultural entities” instead of lithic industries as we refer to all the elements of the material culture available, including, besides lithics, bone tools, faunal remains, subsistence strategies, features, space management and, when present, items connected to symbolic behaviour like colouring substances, personal ornaments and art.

4. I recommend emphasizing that we can now build a much more reliable chronology than ever before, not just because of new dates, but because:

- Significant IntCal20 updates to this age range. This is lost in the conclusion, but to me is a motivation to do this the study now.
- State-of-the-art sample prep for both 14C and OSL and reduced AMS error ranges can make significant improvements over existing chronologies. These are key contributions from the list of authors.
- Models have strong agreement indices even when cross-referenced with three independent time markers: 14C, TL, and tephra layers.

Yes, we agree, this is a good idea. We have rewritten the Methods introductory section to describe these developments in more detail along these exact lines. (SEE LINES 273-285)

Alignments with the NGRIP record (Figs 3, 10) and genetic studies appear later, but are not in the introduction. They are interesting, but I would cut them, since the paper's data do not directly address these issues.

We have reduced the discussion of the genetic studies significantly, and linked this with the possibility that the Aurignacian and Uluzzian might have been produced by two different human populations (SEE LINES 727-741). We hope the reviewer agrees with this. We think it is important to at least refer to the important new genetic data that has appeared recently, so that the reader is aware of the possibilities in interpreting the data we have produced in this paper.

The intro and conclusion mention Uluzzian lithics in Greece, but these are not mapped or discussed. I recommend a map of these sites, which would give the paper much more regional impact. Why don't you include sites from other parts of Europe? Are there no similar lithics outside of Italy and Greece? Limiting the study area and comparisons to modern national boundaries places reflects a modern bias on data patterns created long before these countries existed.

Thank you for this suggestion. We have modified the map in this light (Figure 2) and also put in a new location map in the Supplementary (Fig. S6) showing all of the mapped occurrences of the Uluzzian.

Even if assemblages resembling the Uluzzian were discovered in India and Ceylon, their actual degree of identity has never been directly verified. To date the only Uluzzian sites formally recognized are in Italy and Greece; thus, we prefer to mention only these.

The authors seem to be writing for regional specialists, which makes word choice confusing. It seems the authors are mixing terms for cultures, people, layers, phase, and occupations, when in fact we are talking about dates associated with three lithic industries: Mousterian, Uluzzian, and Protoaurignacian.

As said above, we are not talking exclusively about dates associated with lithic industries. These dates have been obtained from samples collected in specific layers which were occupied by different human groups that had their own productions, tool-kits, use of the living space, subsistence strategies, symbolic manifestations (sometimes also according to the function of site) and lived in different environments.

We tried to simplify the terminology in order to make the text more understandable to a general audience and hope we have succeeded.

For example:

Uluzzian and Protoaurignacian archaeology (line 440), populations (529), Protoaurignacian phase (331), Mousterian layer, Uluzzian occupation (476), the Uluzzian, the Mousterian, Uluzzian technocomplex (41), the identity of the Uluzzian (495) / Ulizzians (645).

We have modified these throughout the text to make it more consistent.

I suggest the authors be more careful and consistent in referring to lithic industries, layers with these industries, dates associated with the lithics, and people who made and used these lithics. Mixing terms for cultures and artifacts recalls a bad culture history habit and can be quite confusing for non-archaeologists. For example: "... a reduction in Neanderthal populations prior to the spread of Uluzzians in the region (line 644). What you mean, I think, is that there a potential occupation hiatus between Hominids using Mousterian lithics and those using Uluzzian lithics.

As specified above, are dealing not only with lithic industries but with defining a behaviour basing on the totality of material culture made available by the archaeological record. This justifies the use of a more general terminology. Using only the lithic evidence would be reductive and not sufficient to highlight differences and similarities among cultural entities and human groups.

In Uluzzo Bay the fact that the Late Mousterian was made by Neanderthals is proven by the retrieval of Neanderthal human remains at Cavallo in layer F.

We changed "Ulizzians" to "Uluzzian". The term Ulizzians now no longer appears.

For the purposes of this paper, you seem to be following the following simplification: Mousterian lithics = Neanderthals and Uluzzian or Protoaurignacian lithics = Homo sapiens. This is fine, and would help guide the reader, but state this generalization more clearly in the introduction, since a general reader will become quickly confused. The discussion is the place to mention possible problems with this simplification.

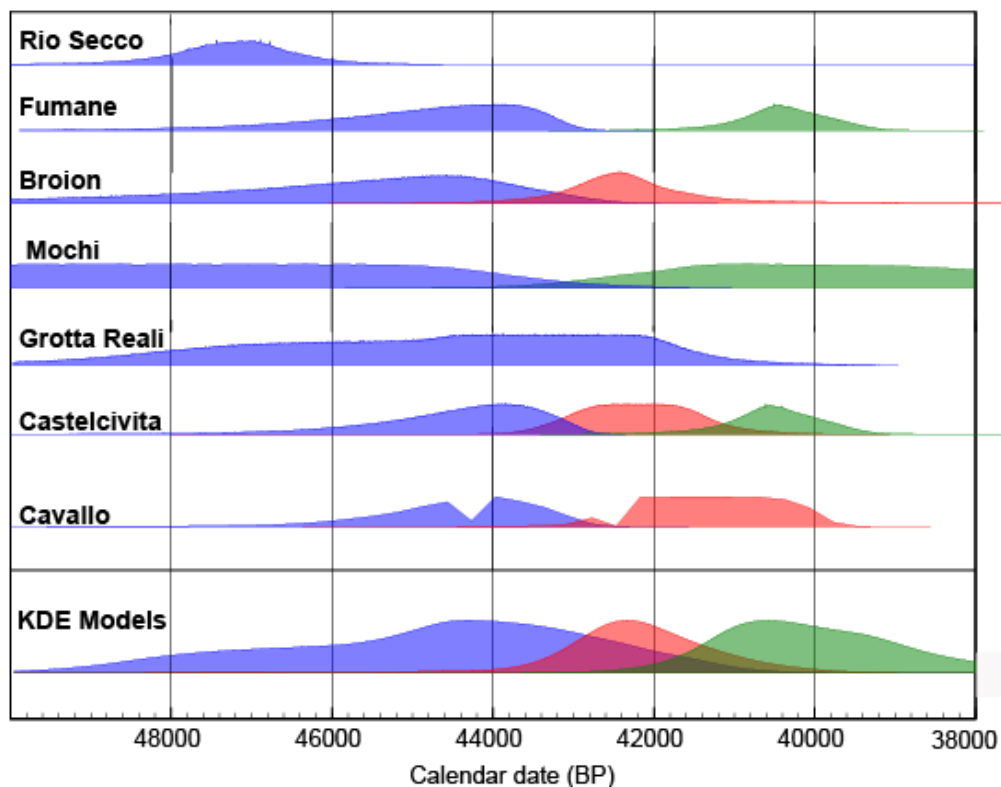
Done. We stated this generalization in the introduction (SEE LINES 74-84).

Figures

Figs 7–10. It would be clearer to see all of this in a single figure. I am imagining a single line for each site, with stacked HPDs for Mousterian end, Uluzzian start, and Protoaurignacian start (and if they fit, KDEs as well). This would make it clear which sites have which data. The same figure could include the summary HPDs based on the phases-of-boundaries approach. Northern Italy at the top; Southern at the bottom. This single figure could 1) summarize all the important results of the paper, and 2) make it visually clear how similar (or different) boundaries are at each site and 3) make it clear how many sites have data that inform regional conclusions, for example, the final use of Mousterian lithics is more robust

since more sites have similar boundaries; Uluzzian and Protoaurignacian regional boundary trends are only based on three sites.

These are really good suggestions, and we agree that it is a better option. We explored two possibilities, first using boundaries (end Mousterian, Start Uluzzian/Protoaurignacian) as the reviewer suggests, alongside the KDE plots at the bottom (SEE Figure 10, LINES 689-690). These were a little bit hard to interpret, so we opted for the Date ranges from OxCal as well as the KDE plots at the bottom like this:



We have also added discussion of the results (using `Order`) to explain more about the significance of the various comparisons (SEE LINES 622 and 649). This has really improved the discussion we think.

Fig 7. Clarify that the inset is based on a Phase of the ending Boundaries. Is this HPD a Date query from within this phase? A KDE is preferable (though will be nearly the same result, probably). I agree with your approach of grouping these boundaries in a phase, but I think a broader readership needs the clarification that this collapses intra-regional variability, for example, Mousterian lithics were last used at Rio Secco well before their final use at other sites.

Yes, the HPD in Fig 7 is an end boundary from a single phase within which the priors are embedded, each of which is an end boundary from the various site Bayesian models. We have changed this to add comment to the meaning behind it and commented on the Rio Secco case specifically. (SEE LINES 583-585). Thanks again for this suggestion.

Fig 8. Comparisons of starting boundaries for the Uluzzian. The authors suggest there is no statistically significant difference between these boundaries based on the command Difference. However, if we queried them with Order, we would see a strong likelihood that Uluzzian lithics appeared at Castelcivita before Cavallo (as noted in line 609), both sites in the south. While the HPDs do overlap, three boundaries are not a strong basis for regionally synchronous changes. Perhaps we don't need to force them into a phase? I would say three boundaries are too few to group, as in Figure 10, which has three Protoaurignacian starting boundaries that are not grouped. It seems quite plausible that these industries first appeared at different times; there may not be a regional trend here (at this smaller temporal scale).

These are not in a Phase, they are just plotted and the Difference calculated. We agree that it would be better to use Order here though, and we have modified to text to discuss this (as mentioned above). In the SI we also included Order statistics for all of the dated sites and levels. Thank you for suggesting this.

Fig 9. Here I have the same question as the inset in Fig 7. All three boundaries seem to be the main conclusions of the paper. I would find it very useful to see a KDE summary of the dates from each lithic industry next to the boundaries, as in Bronk Ramsey (2017:Fig 3). With these KDEs, you could speak to the spans that people used these lithics, not only the boundaries between them.

Agree, it is a good idea. We have modified this in the text (LINES 661-667) and changed the figure to include a detailed KDE analysis.

Figs 11–12. Clarify if these maps are based on the results from Figs 7 and 8 (which seems to be the case). Also, clarify that these circles show ending and starting boundaries are probabilities, not the durations of when these lithic industries were in use. I would suggest caution with these figures, which have a tendency to be misread. In this case, the largest circles that appear on the most maps are in fact the least precise boundaries, which is often confused for duration or significance. I suggest the authors consider a more intuitive figure: circles representing KDEs for each lithic industry. In this case, change circle sizes at each site would better show declining Mousterian dates and rising Uluzzian and Protoaurignacian dates, perhaps in a single figure.

We have remade these figures along the lines you suggest. The KDE model data is now included for each of the main technocomplexes on the map, rather than start/end boundaries as previously used.

Details to clarify in the Supplementary Material:

- Anything that does not directly deal with the research questions should be moved to the supplementary material, for example: model sensitivity to Firenze dates (line 379), dates with caveats (paragraph starting on line 383), and model variants. If these are relevant to the paper's questions, clarify how.

Yes we agree, and have moved this.

- Clarify that in Europe, OSL and 14C have been robustly correlated. TL and OSL remain less reliable in regions with younger geology, including the Andes (Marsh et al. 2021). This paper's results provide further confirmation of this correlation.

Agree

- Please clarify the basis for the Delta\_R on shell dates ( $4\pm 189$  years). Are there paired dates?

This value is generated from the Calib website Marine Reservoir Database.

- Add Cavallo's unpublished field notes here (line 474) – this is too fine a detail for the main text. The discussion should focus on inter-site comparisons.

We removed this and put it into the Supplementary Information.

- The authors should not assume readers know about site details, as this is not a regional journal, for example, line 325: "As is widely known, the Protoaurignacian layers are capped by a series of flowstones...". This seems like a detail for the Supplementary Material.

Agreed, we modified this text.

In addition:

- To avoid confusion and make the reader more familiar with the terms used in the text we added in the SI a paragraph with the description of the excavation methodology followed at Cavallo, Cala, Castelcivita and Oscurusciuto, in with the different terms are defined.
- For the same reasons we added in the SI a more extended description of the Italian Late Mousterian, Uluzzian, Protoaurignacian and Early Aurignacian technocomplexes
- We have eliminated the Italian term "taglio" everywhere, replacing it with its counterpart "spit".
- in the discussion we have given more space to the hypotheses on the Uluzzian/Aurignacian relationship

## References

Bronk Ramsey, C., 2017. Methods for Summarizing Radiocarbon Datasets. Radiocarbon 59, 1809–1833. <https://doi.org/10.1017/RDC.2017.108>

Marsh, E.J., Korpisaari, A., Mundt, S.P., Gasco, A., Durán, V., 2021. Radiocarbon vs. luminescence dating of archaeological ceramics in the southern Andes: a review of paired dates, bayesian models, and a pilot study. Radiocarbon 63, 1471–1501. <https://doi.org/10.1017/RDC.2021.82>

Thank you for this reference, we have included it in the paper!

## Reviewer #3 (Remarks to the Author):

The authors of this manuscript present results of an impressive project to explore whether or not there was overlap between Neanderthals (associated with the Mousterian tool industry) and modern humans (associated with Uluzzian tool industry) in four cave sites in southern Italy. They then compare the timing of the Uluzzian tool industry to that of another modern human tool industry (Protoaurignacian) using spatial-temporal mapping to interpret the results. The project shows the power of careful radiocarbon and optical stimulated luminescence (OSL) dating combined with stratigraphy in Bayesian models. The radiocarbon dating uses the most advanced pretreatment techniques as well as using ZooMS to identify the species of bone fragments.



However, I feel that the interpretation of the model results regarding the Mousterian ending prior to the beginning of the Uluzzian may be slightly overstretched. At Grotta del Cavallo there is a clear end to the Mousterian below a tephra layer and apparently some time elapsed before for the start of the Uluzzian but there appears to be potential overlap between the end of the Mousterian and the start of the Uluzzian in the Grotta di Castelcivita age model (Figure 3). The individual model boundaries for the start of the Uluzzian at these two sites are used as priors in a phase model along with that from Riparo Brion (Figure 8) to produce a single HPD (Figure 9). With only three sites, does the later start and tight HPD of the Cavolo Uluzzian disproportionately shift the combined HPD to age younger age? If so, that may be forcing the start of the Uluzzian to be later than the end of the Mousterian rather than overlapping with it. I think the authors should acknowledge that this conclusion is somewhat uncertain.

It's a good point but it is important to remember that our interpretation is based not only on the chronology but also on the stratigraphic evidence that undoubtedly testifies to the overlying of the Uluzzian above the Mousterian and, in addition, to a gap between the two occupations (Mousterian and Uluzzian) recorded in spit 18. The fact that the start of the Uluzzian at Castelcivita is later than the end of the Mousterian is, therefore, undeniable, simply on the basis of the stratigraphy. However, it is also true that the EIII layer is not as well dated as we would like (as described in the paper) and more work is needed (currently underway), so we leave an open mind on whether the Uluzzian at Cavallo might be slightly earlier than it appears. We have reanalysed the end date for the Mousterian based on other dated sites previously published and now present this in Figure 8 (SEE LINES 569-570). We also undertook a new statistical analysis using OxCal's `Order` command (SEE LINES 623-630) which allows us to assign a probability estimate to determine whether or not one PDF is earlier than another (see Table S8). What this shows is that the probability that Mousterian end boundaries are earlier than Uluzzian end boundaries on a case-by-case basis shows that this is 70-70% likely, the only exception being Grotta Reali, which, as we say in the paper, is not robust as a single PDF end boundary because it is not constrained by any dated archaeology that is post-Mousterian (SEE LINES 573-574). We therefore do not rely on this in the summary model, preferring instead to use a Date derived PDF. So we think that this evidence shows that the Mousterian is demonstrably earlier than the Uluzzian on the basis of the model builds, and also the stratigraphy. The Order stats show that while there might be a slight later shift in the PDFs from the three analysed sites, this is not significant statistically.

Also, in the conclusions the authors state: 'We also see evidence that there may have been reduction in Neanderthal populations prior to the spread of Uluzzians in the region'. As far as I can tell this evidence is not presented in the paper.

With this we intend a reduction at a regional scale in the sense that when modern humans reached the sites at issue, these and the surrounding territories had been definitively abandoned by Neanderthals who may have moved elsewhere. At Grotta della Cala, for instance, after the arrival of the Uluzzian, Neanderthals were not present in the area anymore. However, there is a site, Grotta dei Vallicelli (which is still under study), where a very late Neanderthal occupation is recorded possibly coeval to that of the Uluzzian at Cala. This cave is located in the inner part of Cilento on Monte Cervati, about 100 km far from Cala. On the basis of statistical analysis of the dated PDFs, and the stratigraphy, we think this evidences a general absence of Neanderthals prior to the Uluzzian spread.

Otherwise, I think the paper is a significant contribution to the field and should be published

with minor revisions.

Thank you for your comments on the paper.

Specific comments:

Line 137 and elsewhere with regards to Castelvita: What does Taglio refer to? It is confusing since it doesn't seem to correspond to stratigraphic units and this is not shown on Fig S2.

Taglio is an Italian word which equates to the English word "spit", and refers to a designated stratigraphic unit of depth used when excavating a sediment sequence. (SEE LINES 458). We modified this in the text to define it better and changed all occurrences.

Line 192: 'Shell carbonates' Are the shell carbonates all marine shells?

Yes correct.

Line 360: 'We calculated and applied a new local marine reservoir correction (or Delta\_R) for the shell dates in the model of  $4 \pm 189$  years, along with the Marine20 calibration' In the supplementary info it is stated that Delta\_R has been updated to Marine20 but the authors doesn't state where the original reservoir age data is from. Is the data from the [calib.org/marine](http://calib.org/marine) website or is it based on previously published marine/terrestrial pairs?

Yes that is correct, Calib. We have updated this in the text (SEE LINES 477-479).

Line 496-498: 'There appears to be a slight delay in the Cavallo site by comparison, as hinted at above, although it is not statistically significant compared with Castelvita.' How was the significance tested?

Thank you for this observation. We have now added the statistical backing for this assertion which was made using the Order function in OxCal (SEE LINES 623-629). We also added the table of all of the stats in the Supplementary Methods (Table S7).

Figure 1 or on Figures 11-12: Please label etc. on the figure as these locations are discussed with regards to the spatial-temporal discussion and in the conclusions.

Figures 11-12: It would make more sense to reverse the order of these so that the oldest plots were first. Also in captions, please explain what the size of the circle means.

Thank you for the comment. We have changed these and addressed the circle size too (SEE LINE 702).

## Reviewers' Comments:

### Reviewer #1:

#### Remarks to the Author:

In my opinion the Authors accepted to review the critical points or argued convincingly on others. Therefore, I am satisfied with the new version of the paper that is thus recommended for publication.

### Reviewer #2:

#### Remarks to the Author:

This revision is a major overhaul. This is a much improved presentation of the very strong data and methods I saw in the first version. The authors have rewritten significant sections and it now has the format and content I would expect for this journal. I think the text is much more accessible to a broad readership and they have clarified and organized the definition of the lithic technocomplexes associated with the radiocarbon dates. They have conscientiously addressed the points I brought before, and I see that they found some suggestions useful. In this version, I see a few minor points that need correcting.

#### Marine $\Delta R$ and shell dates from Cavallo

This methodological point needs clarification in the Supplementary Materials. In the code, I see two  $\Delta R$ s:  $4 \pm 189$  (Cavallo) and  $58 \pm 85$  (Mochi model). In the marine  $\Delta R$  database, I cannot reproduce these values – instead, I am getting large negative  $\Delta R$ s. The authors need to follow Reimer's recommendation and cite the values they are using from the database, in addition to the database. Even so, nearly all of these data points are from modern shells with known ages. It is quite unlikely the  $\Delta R$  was the same 30,000 years ago – we simply have no reliable information on this. I would recommend letting the  $\Delta R$  float freely for these dates, and the model will define it as best it can based on the other parameters (with other reliable parameters, this can be a good way to define a premodern  $\Delta R$ ).

```
Delta_R("Undefined Local Marine",U(-500,500));
```

This will widen their error ranges, I think an honest reflection of the lack of information we have. Hence these shell dates are very imprecise and this should be reflected in the models.

OxA/OxA-X-21072 and 19256 are not in code. Why not?

In Table 2, define \*\* as duplicate samples.

Only 19242 is calibrated with a marine  $\Delta R$  of  $4 \pm 189$ .

Other shell dates are calibrated with IntCal20, but shouldn't use Marine20 and a  $\Delta R$  correction: 19254, 19255, 20631, 19257, 19258, 2280-16.

The text mentions OxA-19257 as an outlier run on shell (as discussed in Douka et al. 2014, Journal

of Human Evolution). The model code, however, calibrates this date with IntCal (in three models), which should explain why it pops out as an outlier. Calibrating it with a Delta\_R will improve the model – I recommend a free floating one, not a modern one from the database.

Composite models (Uluzzian, etc.)

For these models, the authors use the original 14C dates. Instead, it would be much better to use the priors from the site-specific models. Otherwise the composite models lose the improvements made in the site-specific models. This can be done with saved priors or cross-referenced parameters (if run together). Since the site-specific models reduce error ranges, I think this has the potential to significantly improve the overall results, namely by making them more precise.

I think this applies to all parameters in the composite models – it first came to my attention with the combined dates:

2280-16 is correctly combined with 19256 in the site model, but this combination is excluded from the general models (Cavallo Uluzzian, Uluzzian).

Same question for charcoal dates 40125 and 40126.

Same question for 2733-13 and 2772-7 (bones dates from Castelcivita).

Also, Douka's PhD thesis is missing from the bibliography.

Intervals between double boundaries?

It would be interesting to add an Interval query between double boundaries, to estimate the lapses between deposition.

In the composite models in the discussion, this same approach could give a quantitative estimate for how long might have passed between the last Neanderthal occupation and first H. sapiens occupation, for example, by querying the Difference between the starting and ending boundaries at specific sites. The text mentions these differences generally, but a quantitative estimate (with error ranges) would be preferable.

Color choices

The stacked KDEs are a drastic improvement from the previous version of the paper. This is mostly an artistic preference, but I find the color choices a little confusing. In three figures, the colors refer to boundaries. In two other figures the same colors refer to KDE time slices. What my eye is used to (personally) is green=start/go, red=end/stop, and blue/grey=KDE. Personally, I would reserve red and green for boundaries and use three different colors for Mousterian, Uluzzian, and Protoaurignacian. But the most important thing is consistency, and perhaps using more colors to avoid the confusion between boundaries and KDEs, which are very different model results and need to be read differently.

Fig 8: all curves are ending boundaries (grayish blue and gray)

Fig 5: neon green curves are for marine shells.

In Figs 4 and 9: blue, red and green (or neon green) refers to boundaries: end Mousterian, start Uluzzian, and start Protoaurignacian.

In Figs 10 and 11, KDEs are blue, red, and dark green are Mousterian, Uluzzian, and Protoaurignacian.

(though in Fig 11 the blue looks purple to me).

Reviewer #3:

Remarks to the Author:

The authors have revised the paper to answer all of my questions/comments, except that the issue of the Delta\_R values used for calibration of shell samples in the models is still confusing. The authors state that the Delta\_R values used were from the calib.org database, however I cannot reproduce the value of  $4 \pm 189$  14C yrs stated in line 478. Were all database values within a certain distance from the archaeological sites averaged, only the suspension feeders or what exactly? The citation 'Reimer and Reimer, 2001; 2002 and the calib.org website' is also confusing. I suspect that the 2002 citation should be Reimer and McCormac 2002 (which is in the references but not otherwise cited in the text), but it doesn't make sense that only data from that study would be utilised since there is now much more data available. In any case, using only the Reimer & McCormac 2002 values or a subset of these from the database I still do not get a weighted average of  $4 \pm 189$  14C yrs. Also, in the supplemental material the KDE model for Mochi Protoaurignacian line 1720, the Delta\_R used is  $58 \pm 85$ . This appears to be the average value for the Mediterranean from Reimer & McCormac 2002, but those measurements have not been recalculated with Marine20.

Also please note that at Line 479: 'along with the Marine20 calibration (Reimer et al., 2020)' the citation should be Heaton et al. 2020 (which is in References)

Tyrrhenian Sea is misspelled on Figure 2.

## RESPONSE TO REVIEWERS

Reviewer #1 (Remarks to the Author):

In my opinion the Authors accepted to review the critical points or argued convincingly on others. Therefore, I am satisfied with the new version of the paper that is thus recommended for publication.

Thank you for your reviews and comments. We appreciate your time and efforts to improve the paper.

Reviewer #2 (Remarks to the Author):

This revision is a major overhaul. This is a much improved presentation of the very strong data and methods I saw in the first version. The authors have rewritten significant sections and it now has the format and content I would expect for this journal. I think the text is much more accessible to a broad readership and they have clarified and organized the definition of the lithic technocomplexes associated with the radiocarbon dates. They have conscientiously addressed the points I brought before, and I see that they found some suggestions useful. In this version, I see a few minor points that need correcting.

Thank you again for your comments on the previous manuscript, and this revision.

Marine  $\Delta R$  and shell dates from Cavallo

This methodological point needs clarification in the Supplementary Materials. In the code, I see two  $\Delta R$ s:  $4\pm 189$  (Cavallo) and  $58\pm 85$  (Mochi model). In the marine  $\Delta R$  database, I cannot reproduce these values – instead, I am getting large negative  $\Delta R$ s. The authors need to follow Reimer's recommendation and cite the values they are using from the database, in addition to the database. Even so, nearly all of these data points are from modern shells with known ages. It is quite unlikely the  $\Delta R$  was the same 30,000 years ago – we simply have no reliable information on this. I would recommend letting the  $\Delta R$  float freely for these dates, and the model will define it as best it can based on the other parameters (with other reliable parameters, this can be a good way to define a premodern  $\Delta R$ ).

Delta\_R("Undefined Local Marine",U(-500,500));

This will widen their error ranges, I think an honest reflection of the lack of information we have. Hence these shell dates are very imprecise and this should be reflected in the models.

This is a good suggestion and we have followed your advice here. Fortunately, it did not have a hugely significant effect on the precision, and the posterior values are virtually unchanged, but we agree it is a better approach with fewer assumptions so we have changed both the Mochi and the Cavallo models to allow this floating DeltaR value to operate and updated all of the models and summary datasets following.

OxA/OxA-X-21072 and 19256 are not in code. Why not?

OxA-19256 is in the code, but we left out OxA-21072 because it is a substantial outlier. We made reference to this in the text.

In Table 2, define \*\* as duplicate samples.

Done, thank you for identifying this.

Only 19242 is calibrated with a marine  $\Delta R$  of  $4 \pm 189$ .

Other shell dates are calibrated with IntCal20, but shouldn't use Marine20 and a  $\Delta R$  correction: 19254, 19255, 20631, 19257, 19258, 2280-16.

We modified these in the models (as outlined above) and confirm they are now correct.

The text mentions OxA-19257 as an outlier run on shell (as discussed in Douka et al. 2014, Journal of Human Evolution). The model code, however, calibrates this date with IntCal (in three models), which should explain why it pops out as an outlier. Calibrating it with a Delta\_R will improve the model – I recommend a free floating one, not a modern one from the database.

Thank you for this observation.

Composite models (Uluzzian, etc.)

For these models, the authors use the original 14C dates. Instead, it would be much better to use the priors from the site-specific models. Otherwise the composite models lose the improvements made in the site-specific models. This can be done with saved priors or cross-referenced parameters (if run together). Since the site-specific models reduce error ranges, I think this has the potential to significantly improve the overall results, namely by making them more precise.

We think the reviewer might be confused over this issue. All of the composite models are indeed based on the priors derived from the individual site models. The only time individual dates are referenced outside the individual site models mentioned are in the data for the KDE models.

I think this applies to all parameters in the composite models – it first came to my attention with the combined dates:

2280-16 is correctly combined with 19256 in the site model, but this combination is excluded from the general models (Cavallo Uluzzian, Uluzzian).

Same question for charcoal dates 40125 and 40126.

Same question for 2733-13 and 2772-7 (bones dates from Castelcivita).

Yes we left these as single data points in the KDE models. Although we could R\_Combine these paired samples we left them in the KDEs and independent results.

Also, Douka's PhD thesis is missing from the bibliography.

This has been added.

Intervals between double boundaries?

It would be interesting to add an Interval query between double boundaries, to estimate the lapses between deposition.

Yes this is a good suggestion.

In the composite models in the discussion, this same approach could give a quantitative estimate for how long might have passed between the last Neanderthal occupation and first H. sapiens occupation, for example, by querying the Difference between the starting and ending boundaries at specific sites. The text mentions these differences generally, but a quantitative estimate (with error ranges) would be preferable.

## Color choices

Thanks for your comments on this. We have in the past (in our 2014 Nature article on Neanderthal disappearance, as well as the 2019 article in the same journal on Denisova Cave) used the same colour scheme for Neanderthal dates and genomes (blue), and a green for Uluzzian in the former case, with red for Chatelperronian. We feel as though some continuity is good, so we thought let's continue with these schema. We think as long as the key describes the colours used and what they refer to, and are consistent throughout, it's ok.

The rather ugly neon green in OxCal is used automatically to denote shell/marine dates.

Thank you again for your constructive comments on the paper so far, we feel it has really helped to improve it. It's been a nice experience whoever you are!

The stacked KDEs are drastic improvement from the previous version of the paper. This is mostly an artistic preference, but I find the color choices a little confusing. In three figures, the colors refer to boundaries. In two other figures the same colors refer to KDE time slices. What my eye is used to (personally) is green=start/go, red=end/stop, and blue/grey=KDE. Personally, I would reserve red and green for boundaries and use three different colors for Mousterian, Uluzzian, and Protoaurignacian. But the most important thing is consistency, and perhaps using more colors to avoid the confusion between boundaries and KDEs, which are very different model results and need to read differently.

See above.

Fig 8: all curves are ending boundaries (grayish blue and gray)

Fig 5: neon green curves are for marine shells.

In Figs 4 and 9: blue, red and green (or neon green) refers to boundaries: end Mousterian, start Uluzzian, and start Protoaurignacian.

In Figs 10 and 11, KDEs are blue, red, and dark green are Mousterian, Uluzzian, and Protoaurignacian.

(though in Fig 11 the blue looks purple to me).

## Reviewer #3 (Remarks to the Author):

The authors have revised the paper to answer all of my questions/comments, except that the issue of the Delta\_R values used for calibration of shell samples in the models is still confusing. The authors state that the Delta\_R values used were from the [calib.org](http://calib.org) database, however I cannot reproduce the value of  $4 \pm 189$  14C yrs stated in line 478. Were all database values within a certain distance from the archaeological sites averaged, only the suspension feeders or what exactly? The citation 'Reimer and Reimer, 2001; 2002 and the [calib.org](http://calib.org) website' is also confusing. I suspect that the 2002 citation should be Reimer and McCormac 2002 (which is in the references but not otherwise cited in the text), but it doesn't make sense that only data from that study would be utilised since there is now much more data available. In any case, using only the Reimer & McCormac 2002 values or a subset of these from the database I still do not get a weighted average of  $4 \pm 189$  14C yrs. Also, in the supplemental material the KDE model for Mochi Protoaurignacian line 1720, the Delta\_R used is  $58 \pm 85$ . This appears to be the average value for the Mediterranean from Reimer & McCormac 2002, but those measurements have not been recalculated with Marine20.

Thank you. See above for a comment on this. We have gone with the advice of Reviewer 2 and used a floating DeltaR value for both the Cavallo and Mochi (comparative) models. So this avoids the confusion here, which results from using two different values, one from the published



Frouin et al. paper for Mochi and the other from this paper for Cavallo. Using a floating value at this antiquity is, we think, a better option.

Also please note that at Line 479: 'along with the Marine20 calibration (Reimer et al., 2020)' the citation should be Heaton et al. 2020 (which is in References)

Thank you for noticing that! We have now corrected this.

Tyrrhenian Sea is misspelled on Figure 2.

Thank you, this has been corrected.

Reviewers' Comments:

Reviewer #2:

Remarks to the Author:

The authors have diligently addressed all of my concerns. I realize the floating  $\Delta R$  has a minimal impact on the final results, but this strengthens the methods behind a substantial dataset. The colors choices are just preferences. Excellent paper. It was really worth looking for ways to polish it. I strongly recommend publication.

Reviewer #3:

Remarks to the Author:

The authors have implemented the useful suggestion (made by another reviewer) to use an undefined  $\Delta R$  for the models given that the available  $\Delta R$  data is from the last 100 years or so and not necessarily valid 30,000 years ago. I agree this is the best approach.

There is still a very minor issue of the citation of Reimer and Reimer 2002 in line 480. No such reference exists. I suspect that the 2002 citation should be Reimer and McCormac 2002 (which is in the references but not otherwise cited in the text). It still doesn't make sense to me that only data from that one paper would be utilised since there is now much more data available but since the Reimer and McCormac data are not actually used, it is not crucial. I would suggest either citing Reimer and McCormac here or getting rid of the 2002 reference.

Otherwise the revised version is fine.

## Responses to Reviewers

### Reviewer #2 (Remarks to the Author):

The authors have diligently addressed all of my concerns. I realize the floating  $\Delta R$  has a minimal impact on the final results, but this strengthens the methods behind a substantial dataset. The colors choices are just preferences. Excellent paper. It was really worth looking for ways to polish it. I strongly recommend publication.

Thank you for your comments!

### Reviewer #3 (Remarks to the Author):

The authors have implemented the useful suggestion (made by another reviewer) to use an undefined  $\Delta R$  for the models given that the available  $\Delta R$  data is from the last 100 years or so and not necessarily valid 30,000 years ago. I agree this is the best approach.

There is still a very minor issue of the citation of Reimer and Reimer 2002 in line 480. No such reference exists. I suspect that the 2002 citation should be Reimer and McCormac 2002 (which is in the references but not otherwise cited in the text). It still doesn't make sense to me that only data from that one paper would be utilised since there is now much more data available but since the Reimer and McCormac data are not actually used, it is not crucial. I would suggest either citing Reimer and McCormac here or getting rid of the 2002 reference.

Otherwise the revised version is fine.

Thank you. I deleted to Reimer and Reimer reference.