The Middle to Upper Paleolithic transition: dating, stratigraphy, and isochronous markers

S.P.E. Blockley*, C. Bronk Ramsey, T.F.G. Higham

Research Laboratory for Archaeology, University of Oxford, Dyson Perrins, South Parks Road, Oxford OX1 3QY, United Kingdom

Abstract

Accurate and precise dating is vital to our understanding of the Middle to Upper Paleolithic transition. There are, however, a number of uncertainties in the chronologies currently available for this period. We attempt to examine these uncertainties by utilizing a number of recent developments in the field. These include: the precise dating of the Campanian Ignimbrite (CI) tephra by 40Ar/39Ar; the tracing of this tephra to a number of deposits that are radiocarbon dated; the publication of revised radiocarbon calibration data for the period, showing a much better convergence with other available data than during the recent IntCal comparison; and a layer-counted ice-core chronology extending beyond 40,000 cal BP. Our data comparisons suggest that a reasonable overall convergence between calibrated radiocarbon ages and calendar dates is possible using the new curves. Additionally, we suggest that charcoal-based radiocarbon ages, as well as bone-based radiocarbon determinations, require cautious interpretation in this period. Potentially, these issues extend far beyond the sites in this study and should be of serious concern to archaeologists studying the Middle to Upper Paleolithic. We conclude by outlining a strategy for moving the science forward by a closer integration of archaeology, chronology, and stratigraphy.

© 2008 Elsevier Ltd. All rights reserved.

Introduction

Understanding the transition from the Middle to Upper Paleolithic in Europe and the Near East is a process that is heavily dependent on reliable chronological information. In most cases, this means radiocarbon dating. In the time period from ~30,000–40,000 calendar years BP, however, there are significant dating uncertainties (Pettitt et al., 2003). Thus, our understanding of regional patterns and comparisons to other data, such as paleoclimate information, are also uncertain. These problems have been reviewed at various levels of detail in the literature (e.g., Bird et al., 1999; Lowe and Walker, 2000; Santos et al., 2001; Goodwin et al., 2004; van der Plicht et al., 2004; Higham et al., 2006a,b) and can be broken down into the following areas: 1) the ability to successfully remove contamination from individual radiocarbon dates; 2) the problems of taphonomic processes within sites, distorting the stratigraphical relationship between dates (so-called sedimentological time averaging); 3) the reliability of the conversion from radiocarbon to calendar time, particularly in relation to dramatic shifts in the rate of radiocarbon production in the atmosphere; and 4) the rarity of other reliable dating information with which to evaluate radiocarbon chronologies. A robust chronology of Paleolithic archaeology requires us to disentangle the competing influences of these problems.

By comparing radiocarbon dating with other methods, it is apparent that developing reliable chronologies during the Middle to Upper Paleolithic transition is difficult. While some cross dating studies have shown general agreement between radiocarbon dating and other methods, it is within fairly broad, millennial scale uncertainties. At Geißenklösterle, for example, a very careful 14C dating program on bone material, compared to a thermoluminescence and electron spin resonance study for the same Early Aurignacian levels, agreed within ~2,000 years, the discrepancy being put down to calibration issues (Richter et al., 2000). Other similar studies have shown greater discrepancies between different methods. At Devil’s Lair in Australia, for example, a traditional radiocarbon age of 38,240 ± 1250 14C BP compares to other ages for the same levels of ~42–46 ka 14C BP (see O’Connell and Allen, 2004). While clearly there is potential to get reasonable agreement between 14C ages and other dating methods, there is still uncertainty. Where high precision comparisons are required, such as the potential impact of Heinrich events on human populations, then we need to be able to test these uncertainties with better than millennial precision.

Here we use recently available chronological and stratigraphic information to evaluate the magnitude of the current problem, as it applies to the chronology of the Middle to Upper Paleolithic transition in Europe. In particular, we consider: 1) revised calibration...
data from the Cariaco varve record (Hughen et al., 2007), which shows much better agreement with other available calibration data than had previous comparisons; 2) new layer counted Greenland ice-core data to greater than 40,000 cal BP (Svensson et al., 2007); 3) the precise age of the Campanian Ignimbrite (CI) tephra, generated by multiple $^{40}$Ar/$^{39}$Ar dating of 15 different exposures from proximal deposits (de Vivo et al., 2001); and 4) radiocarbon dates of multiple samples in stratigraphic association with this tephra (Deino et al., 1994; Ton-That et al., 2001; Sinitsyn, 2003; Anikovich et al., 2007). The potential now exists to compare these data and to attempt to quantify the scale of the problem. Because of the significant uncertainties of some of the dates, we additionally utilize a formal Bayesian model to frame this comparison. We then consider the implications of the model output for the timing and environmental context of the Middle to Upper Paleolithic transition. We use this study to develop an improved strategy for future chronological research in this area.

**Chronostratigraphic issues in Middle and Upper Paleolithic archaeology**

**Radiocarbon quality assurance and asymptotic radiocarbon dating**

A number of important sites covering the Middle to Upper Paleolithic transition exhibit similar radiocarbon ages through more than one section of a sequence. One potential cause of these phenomena is the asymptotic radiocarbon effect. This effect was first noted by Australian researchers examining long Quaternary records that spanned most of the last glacial and interglacial periods. The sections returned radiocarbon ages of 30–40,000 radiocarbon years above the instrument background, even at the interglacial base of the sequence (Chappell et al., 1996). The site in question was dated by several other radiometric methods, including K-Ar dating of interspersed lavas and luminescence ages on sediments, all of which reported an expected increase in age with depth. However, at the base of the sequence the radiocarbon ages yielded results with a defined above background measurement and error, but which were in fact incorrect by around 90,000 years. This result prompted the development of much more rigorous pre-treatment strategies for older charcoal samples, particularly the ABOX (a wet oxidation technique with stepped combustion of samples; Bird et al., 1999), at some radiocarbon laboratories. This method has proved successful at returning reliable ages for old sites when tested against other methods (Santos et al., 2001; Turney et al., 2001).

The other significant advance in the pre-treatment of samples for archaeological dating relates to the improvements made in the dating of bone using ultrafiltration. These methods seem to have captured the archaeological imagination much more than the advances made on charcoal samples. For the appropriate sites, ultrafiltration will, we believe, significantly improve our ability to examine both site chronologies and the integrity of site stratigraphy. The use of an additional ultrafiltration step in the gelatinization of archaeological bone has resulted in substantial improvements in the removal of contaminants (Bronk Ramsey et al., 2004; Higham et al., 2006a,b). For Paleolithic age bone, the difference in age between samples analyzed with and without ultrafiltration is often very large. In almost all cases the ultrafiltrated gelatin determinations are older, sometimes by several thousand years. In addition, the regular collection of analytical data enables an assessment of the quality of bone preservation and the collagen extracted. Additional improvements revolve around increased levels of measurement precision and reduced backgrounds in AMS dating. Together these developments have resulted in an extension of the radiocarbon range to ca.50–55 ka (Bronk Ramsey et al., 2004; Higham et al., 2006a,b).

As yet, relatively few radiocarbon dates on sites from the Middle to Upper Paleolithic transition are analyzed under these exacting criteria. One key aim of this paper is to demonstrate, at a number of sites, that the asymptotic radiocarbon effect is a real phenomenon affecting the archaeology of this period in Europe. We further demonstrate that radiocarbon dating problems reported in the literature cannot be explained by site specific taphonomic processes of sedimentary time averaging.

**Radiocarbon calibration**

A major area of concern in dealing with radiocarbon chronologies is the availability of reliable calibration data to convert radiocarbon time into calendar time. Currently, we are in the unfortunate position of having no agreed upon calibration curve for the period in question (van der Plicht et al., 2004). The data available when the current international curve was established suggested that no consensus calibration existed for dates older than ~26,000 cal BP. This is an issue that is particularly problematic for archaeologists. While we readily accept that a calibration curve has not yet been agreed upon and ratified, therefore making calibration not yet possible, it is potentially useful to compare radiocarbon dates to a range of proposed calibration data in specific circumstances. One recently published set of calibration data makes such a comparison both interesting and timely. The revision of the Cariaco calibration data set (Hughen et al., 2007) improves the agreement between two of the main archives that make up the IntCal curve (the Fairbanks et al. (2005) coral data and the Cariaco data). The new calibration data of Hughen et al. (2007) for Cariaco has added an additional 187 determinations on foraminifera, increasing the $^{14}$C data set for this curve by two-thirds. Additionally, the generation of a calendar chronology for the new Cariaco curve has been revised and now the Cariaco grayscale has been compared to the radiometrically-dated Hulu Cave speleothem. This revised Cariaco archive, when compared to the Fairbanks data set, shows significantly improved agreement between the two records, particularly for the period of interest in this special issue, ~40,000 cal BP, where the two curves agree within errors.

For the period of interest here, there is now a sufficiently robust body of calibration data to make a comparison with radiocarbon dates potentially useful. This is because one of the central assumptions behind the generation of the Cariaco curve is that climatic change at the global or hemispherical level is synchronous. This may not always be the case (Lowe et al., 2001; Blockley et al., 2006). However, for the period around 40,000 cal BP, the newly layer-counted NGRIP ice core (Rasmussen et al., 2006; Svensson et al., 2007) is now annually counted back to beyond 40,000 cal BP and has fully quantified uncertainties. Despite disagreements in some periods between Hulu Cave and NGRIP, both related to the timing and nature of climate change, there is close correspondence between the two independently dated climate archives for the period 30,000 to 42,000 cal BP. We feel, therefore, that for the time period in question there is sufficiently robust calibration data to compare with different radiocarbon dates and $^{40}$Ar/$^{39}$Ar dated records.

**Known age tie-points**

One other recent development that is of interest here is the location of the Campanian Ignimbrite (CI) tephra in archaeological deposits very far from the source volcano (Pyle et al., 2006). A recent extensive re-dating of this tephra (de Vivo et al., 2001) by $^{40}$Ar/$^{39}$Ar methods, using both single-crystal total fusion analyses (SCF) and bulk laser incremental heating (LIH) of 15 different exposures from proximal deposits, gives an age of unusually high precision for this event, 39,280 ± 110 cal BP. This age is in good
agreement with other published ages of this tephra and is by far the most extensive study of the age of this ash. We accept here Pyle and colleagues’ (2006) recommendation that this be the accepted geological age of the eruption. For the sake of completeness, however, we will also include other radiometric ages for this tephra that have been fully published (Ton-That et al., 2001). One date that we no longer use for this tephra (Deino et al., 1994) was never fully published and we follow the protocol of Pyle et al. (2006) and view it as being superseded by more recent analyses.

One other tie point that is of value here is the Laschamp geomagnetic excursion. This is known to slightly precede the Campanian Ignimbrite eruption stratigraphically and has been radio-metrically dated in lava flows to 40,400 ± 1100 BP (Ar/39Ar (Guillou et al., 2004).

Comparing radiocarbon dates, calibration data, and known-age tie points

Simple comparison

The stratigraphical relationships between the known-age isochron tie points discussed above, a number of associated charcoal-based radiocarbon determinations, and proposed calibration data can now be assessed. We test the relationship between the different data sets to examine the overall reliability of constructing radiocarbon-based chronologies for this period. We have attempted to de-couple site stratigraphical issues, such as re-working, from problems of sample contamination and the asymptotic radiocarbon effect. To achieve this, we include radiocarbon ages from different sites in known stratigraphical relationships with the Campanian Ignimbrite (CI). We use radiocarbon dates from immediately below the CI layers at Kostenki 1, 12, and 14 (Sinitsyn, 2003; Anikovich et al., 2007), as well as dates on wood that are from within the proximal ash deposits of the CI (Deino et al., 1994). All of the archaeological dates are on charcoal and use the standard acid-base-acid technique. We also use a date from marine foraminifera (using a standard reservoir correction with a regional AR; Reimer et al., 2004) found in association with the CI in a Mediterranean marine core (Ton-That et al., 2001). Although reworking of material is possible in all cases, the sealing of the Kostenki material by a defined volcanic tuff minimizes this problem, and the use of dates from wood within the CI seems an unlikely context for reworking of younger material. By employing this strategy we de-couple reworking issues from radiocarbon sample integrity and test whether there is a reliable corpus of radiocarbon ages for this eruption.

We initially compared these data to the radiometric ages for the CI and the new calibration data of Hughen et al. (2007). We did not use the Fairbanks data set here. There is sufficiently close agreement between the two calibration data sets in this time period, and the revised Cariaco data has many more analyses. We have not compared any of our data to the NOTCAL04 curve, as this is a composite curve that does not constitute calibration data and was never intended to be used as such (van der Plicht et al., 2004). The results of this first comparison (Fig. 1) suggest that there is still significant disagreement and scatter in the radiocarbon determinations, although within broad uncertainties there is agreement between the data. The significant uncertainties in many of the available radiocarbon dates mean, however, that initial interpretation of Fig. 1 has to be done with caution. All of the dates relate to material that should agree with each other broadly, and all are slightly older than the age of the CI eruption. Taking the 39,280 ± 110 cal BP age as the best available estimate of the CI eruptions, it is clear, however, that the radiocarbon dates fall into two separate groups. The majority of dates are younger than, or just slightly overlapping with, the age of the CI and do not convincingly fit their stratigraphical relationship to the ash (i.e., to be from material that is slightly older). The dates from Kostenki are stratigraphically below the tephra and should be older than the ash age. One date, from wood embedded in the ash, also appears to be too young. Only dates OxA-15482 from Kostenki, one date from wood in the ash, and the marine radiocarbon determination appear to be slightly older than the ash, although there are very large errors on the marine age. Given that several sites are involved and the contextual security of the samples, which were embedded in or sealed by an ash layer, this seems to suggest that reworking is not the only issue here. Instead, it is more likely that this discrepancy relates either to a small amount of contaminating carbon in the samples, as implied in other studies of old radiocarbon dates (see above), or calibration problems.

In order to test this issue, we attempted this same exercise using the old Cariaco 2004 data set (Hughen et al., 2004). The results were very similar (Fig. 2), suggesting that differences between dates are not the result of the calibration data set used for comparison. Due to the significant uncertainties on some of the dates, we next used a formal model to constrain and test the data.

Formal Bayesian modeling

In order to more rigorously test the simple, visual analyses described above, a formal model was constructed using a Bayesian framework. Bayesian analysis is now well-established in archaeological and paleoenvironmental chronometry and seems an ideal tool for this task (see Buck et al., 1991; Bronk Ramsey, 2000, 2001, 2008; Blockley et al., 2004, 2007). With this model we hoped to mitigate the significant errors in some of the dates and see if we can more reliably state that a) there was disagreement between dates, effectively two groups, when compared to the calibration data, and b) only one of the two groups (dates ranging around 35,000 14C BP) were consistent with the known-age tie points and their stratigraphical relationships.

The model makes the following stratigraphical assumptions:

1. All of the radiocarbon dates relate to material that is stratigraphically below the CI tephra or to material taking in carbon before the time of the eruption.
2. Following from this, all of the radiocarbon ages should slightly predate the age of the eruption with a decreasing but finite probability of being from a time significantly before the eruption.
3. No date should be younger than the age of the eruption.
4. The radiocarbon ages should equate to the Laschamp excursion, due to the stratigraphical and chronological relationship between the Laschamp and CI.
5. The most conservative and best estimate of the age of the CI is a statistical combination within the model of the two reported 40Ar/39Ar ages (41,100 ± 2100 BP and 39,282 ± 110 BP).

We built these assumptions into a Bayesian framework (Appendix 1) using the calibration package OxCal 4.0.1. For details of the mathematical underpinnings of Bayesian analyses and use of this software see Bronk Ramsey (2008). We used an exponential prior for the radiocarbon determinations, due to the assumption in clause 2 of the prior model, and separated the different phases using boundaries. The two groups of dates (from oldest) in the model are: a) a phase of dated events measured on material most likely to be just prior to the age of the eruptions; and b) the eruption event itself, with associated direct dating information that can be combined. We compared this model to the calibration data of Hughen et al. (2007). We used the agreement indices of the model and the dates overall to assess the reliability of the individual ages in the model and the calibration data.
Results

Initially, the model would not run and returned unacceptably low agreement indices for many dates and the model overall (less than 60%; see Bronk Ramsey, 2008). The dates with the least likelihood of matching the model were those that appeared to be too young in the simple observation, despite the large uncertainty ranges involved. These dates have a significantly low probability of fitting the model and the other data. Since the model is based on simple stratigraphical observation, we regarded these dates as being suspect. In order for the model to run successfully, these dates were excluded from further analyses. This left only one of the two dates from the wood embedded in the proximal ash deposit, the foraminifera radiocarbon age, and OxA-15482 from Kostenki. Re-running the model and comparing it to the Hughen et al. (2007) calibration data resulted in a statistically-acceptable, model-data comparison. The results suggest that for this one time-slice, there are reasonable calibration data with which to compare (Fig. 3). They also suggest that a radiocarbon age of around 35,000 14C BP is an appropriate radiocarbon age for the CI. This assumes, of course, that the revised Cariaco curve (Hughen et al., 2007), and by extension the Fairbanks curve (Fairbanks et al., 2005), are good calibration data. This assumption, however, seems valid at present. Perhaps the most important point is that there is strong evidence that even slight differences in pre-treatment and preservation can radically influence radiocarbon dates of this age. We believe that the differences highlighted by our modeling are both significant and cannot be explained away by taphonomic issues.

Finally, we compare the model output to the new NGRIP ice core record for this period (Svensson et al., 2007). The new Greenland record is layer counted beyond 40,000 cal BP, and despite repeating the pattern of stadial/interstadial cycles found in previous ice core records, the chronology is much improved. This record has shifted the timing of some of the interstadial events as they are recorded in Greenland and agrees well with the Hulu Cave speleothem record for the 30,000 to 40,000 time period. We have compared our model age for the CI, based on the combined 40Ar/39Ar ages, and the timing from our model of the end of the archaeological signature immediately below this tephra in stratigraphic sequence (using the end boundary of this phase of the model; see Appendix 1; Fig. 4).

The levels immediately below the CI tephra are, at Kostenki particularly, interpreted as transitional Upper Paleolithic, with the
Aurignacian proper beginning within the CI layer (Anikovich et al., 2007). Although there is still archaeological debate as to the nature of these pre-Aurignacian industries, the CI has been taken as an onset marker of the Aurignacian proper in some studies. Zilhão and D’Errico (1999) have suggested that there is no evidence for Aurignacian industries in northern Spain and France prior to 36,500 radiocarbon years. d’Errico and Sanchez-Goni (2003) suggested that the spread of the Aurignacian might coincide with the short and weak amplitude warming phase of Greenland Interstadial 9, one of the shortest interstadials of marine oxygen isotope stage 3. Our model for sites containing the CI suggests that the transitional Upper Paleolithic in these sites is consistent with Interstadial 9, as recorded in NGRIP. However, the CI and the associated onset of the Aurignacian now appear not to correlate as well with GIS9. This suggests that the spread of the Upper Paleolithic, but not the Aurignacian proper, is associated with interstadials 9 and 10 in the sites in this study.

**Discussion**

Our understanding of the later Paleolithic of Europe depends heavily on reliable chronologies. We believe that this study has enabled us to discuss the different factors affecting the reliability of radiocarbon dating. Even within the considerable dating uncertainties often involved in this time period, there are serious issues in radiocarbon dating to be addressed by the wider community. This supports the need for the detailed work now being done on improving radiocarbon dating of bone, but also reinforces the need to consider appropriate strategies for the dating of charcoal. Regardless of which calibration data are used, the differences between the radiocarbon ages compiled here remain. We chose a number of sites to mitigate the influence of site taphonomy, and suggest that the best explanation for the patterning in these data is differential removal of contamination due to the very low levels of original 14C in such old samples. This conclusion is in line with several detailed studies of the difficulties in analyzing very old material (Bird et al., 1999).

One important point here is that we are only aware of these problems because we have well-dated marker horizons in known stratigraphical relationships with the radiocarbon dates from which to build a model. As yet, there are few sites with such markers in European Paleolithic archaeology. It appears from this study that demographic patterning of the transition from Middle to Upper Paleolithic, that is based on a variety of radiocarbon dates unconstrained by well-dated marker horizons, may well be an artifact of dating uncertainties.

We make the following recommendations:

1. For every site, it would be advisable to develop a radiocarbon inventory (Lowe et al., 2001), in which different materials are
tested for their reliability as $^{14}$C chronometers. Careful consideration of the pre-treatment strategies used on the samples is a matter for archaeological consideration, as well as for radiocarbon laboratories.

2. Detailed reporting of site taphonomic issues is required, as is the consideration, where possible, of information such as the isotopic variation of shell carbonates to test for mixing (Goodwin et al., 2004).

3. Because full calibration is not yet possible, more than one calibration data set should be considered. We note that although for the time window on which we have focused (around 40,000 cal BP) there is good convergence between Carriaco 07 and the Fairbanks data, this relationship breaks down in some parts of the curves. We also note that at younger than 30,000 cal BP there is disagreement between the Hulu Cave record, upon which the calendar chronology for Carriaco 07 is based, and the new NGRIIP $^{18}$O data (Svensson et al., 2007). This suggests that the underlying assumption behind the construction of Carriaco 07, over the synchronicity of climate change (within errors), is not yet validated.

4. Well-dated marker horizons, particularly tephras, should be sought. Recent developments in the identification and extraction of tephras from host sediments (Turney et al., 1997; Blockley et al., 2005), allow distal tephras that are invisible to the naked eye (microtephra or cryptotephra) to be traced in records in which they were thought to be absent. This has radically extended the scope for tephrochronology (e.g., Wastegard et al., 2000; Davies et al., 2002). Eruptions that are far smaller than the CI can deposit distal ash thousands of kilometers from the host volcano (see Blockley et al., 2006). Often these deposits are found more than twice as far away from the host volcano as any previously discovered visible deposits. A long term project by members of the Oxford Research Laboratory for Archaeology, in collaboration with colleagues at Royal Holloway, University of London, and in several international centers, aims to test the usefulness of tephrochronology for Paleolithic archaeology. Not only can tephras provide independently dated marker horizons, but the tracing of tephras into both archaeological sites and local environmental settings can test and constrain presumed relationships between global environmental changes, such as interstadials, Heinrich events, and human responses. Distal ash can be deposited and conserved in a range of environments as long as there has been reasonably stable preservation. Publications from several groups are in progress, showing that archaeological sites can contain identifiable tephras. Most of Europe and the Mediterranean Basin are potentially open to microtephra study. Although many sites may not have the


