Luminescence dating and palaeomagnetic age constraint on hominins from Sima de los Huesos, Atapuerca, Spain

Lee J. Arnold a,*, Martina Demuro a, Josep M. Parés a, Juan Luis Arsuaga b,c, Arantza Aranburu d, José María Bermúdez de Castro e, Eudald Carbonell e

a Centro Nacional de Investigación sobre la Evolución Humana, CENIEH, Paseo Sierra de Atapuerca s/n, 09002 Burgos, Spain
b Centro Mixto Universidad Complutense-Instituto de Salud Carlos III de Evolución y Comportamiento Humanos, Monforte de Lemos s/n, 28040 Madrid, Spain
c Departamento de Paleontología, Facultad de Ciencias Geológicas, Universidad Complutense de Madrid, 28040 Madrid, Spain
d Departamento de Paleontología, Facultad de Ciencias Geológicas, Universidad del País Vasco, Ap. 644, 48980 Bilbao, Spain
e Instituto de Paleontología Humana y Evolución Social, Área de Prehistoria, U. Rovira i Virgili, Plaça Imperial Tarracon, 1, E-43005 Tarragona, Spain

ABSTRACT

Establishing a reliable chronology on the extensive hominin remains at Sima de los Huesos is critical for an improved understanding of the complex evolutionary histories and phylogenetic relationships of the European Middle Pleistocene hominin record. In this study, we use a combination of 'extended-range' luminescence dating techniques and palaeomagnetism to provide new age constraint on sedimentary infills that are unambiguously associated with the Sima fossil assemblage. Post-infrared-infrared stimulated luminescence (pIR-IR) dating of K-feldspars and thermally transferred optically stimulated luminescence (TT-OSL) dating of individual quartz grains provide weighted mean ages of 433 ± 15 ka (thousands of years) and 416 ± 19 ka, respectively, for allochthonous sedimentary horizons overlying the hominin-bearing clay breccia. The six replicate luminescence ages obtained for this deposit are reproducible and provide a combined minimum age estimate of 427 ± 12 ka for the underlying hominin fossils. Palaeomagnetic directions for the luminescence dated sediment horizon and underlying fissile clays display exclusively normal polarities. These findings are consistent with the luminescence dating results and confirm that the hominin fossil horizon accumulated during the Brunhes Chron, i.e., within the last 780 ka. The new bracketing age constraint for the Sima hominins is in broad agreement with radiometrically dated Homo heidelbergensis fossil sites, such as Mauer and Arago, and suggests that the split of the H. neanderthalensis and H. sapiens lineages took place during the early Middle Pleistocene. More widespread numerical dating of key Early and Middle Pleistocene fossil sites across Europe is needed to test and refine competing models of hominin evolution. The new luminescence chronologies presented in this study demonstrate the versatility of TT-OSL and pIR-IR techniques and the potential role they could play in helping to refine evolutionary histories over Middle Pleistocene timescales.

Keywords: Geochronology Middle Pleistocene Brunhes Western Europe Homo heidelbergensis Homo neanderthalensis

Introduction

The endokarstic system of the Sierra de Atapuerca, north-central Spain, comprises an extensive network of cavities, passages and galleries spanning more than 4.7 km (Fig. 1). The 50 or so sediment-filled cavities explored to date have yielded an unparalleled archive of Early to Late Pleistocene hominin remains, lithic tools and associated faunal assemblages (e.g., Arsuaga et al., 1993, 1997a, 1999; Carbonell et al., 1995; Carbonell, 2008; Bermúdez de Castro et al., 1997, 1999; Cuenca-Bescós et al., 2010; Rodríguez et al., 2011; Ollé et al., 2013). One of the most prolific palaeoanthropological sites discovered at Atapuerca is Sima de los Huesos or the 'Pit of Bones' (herein abbreviated to 'Sima'), a small phreatic chamber in the lowest level of the Cueva Mayor-Cueva del Silo karst system, located >500 m from the cave entrance of Cueva Mayor and ~35 m below the present-day surface (e.g., Arsuaga et al., 1997a; Ortega et al., 2013) (Figs. 1 and 2a and b). The site is renowned for its unique accumulation of hominin fossils (Arsuaga et al., 1993, 1997a,b), which totals more than 6500 remains from at least 28 individuals (Bermúdez de Castro et al., 2004) and represents ~80% of the global Middle Pleistocene fossil record for the genus Homo. This fossil collection grows in number with every excavation season, providing new opportunities to assess intra-
inter-population variability in the Middle Pleistocene hominin record (e.g., Arsuaga et al., 1997b,c; Lorenzo et al., 1998; Martinón-Torres et al., 2006, 2012). The Sima hominin assemblage is also of importance for understanding the origins of the Neanderthal lineage since the fossils display some Neanderthal traits and have been cited as evidence for a Homo heidelbergensis–Homo neanderthalensis continuum in Europe (e.g., Arsuaga et al., 1993, 1997a,b; Rosas, 2001; Bermúdez de Castro et al., 2004; Hublin, 2009; Martinón-Torres et al., 2012). Debate continues, however, over the phylogenetic relationships of Early to Middle Pleistocene hominin populations, and the nature and timing of evolutionary divergences in the Eurasian hominin lineage (e.g., Arsuaga et al., 1997b; Endicott et al., 2010; Dennell et al., 2011; Stringer, 2012; Bermúdez de Castro and Martinón-Torres, 2013). Establishing unequivocal age constraint on the Sima fossils is therefore critical for understanding the palaeoanthropological history and palaeoclimatic context of this site and has broader implications for evaluating competing models of Middle Pleistocene human evolution.

Reliable dating of unheated sedimentary material within Middle Pleistocene karstic cave systems remains a challenge for many geochronological techniques. Two approaches that offer good potential for establishing age control in such settings are palaeomagnetic reversal dating and luminescence dating techniques. Palaeomagnetic dating has been extensively applied to the endokarstic infill sequences at Atapuerca (e.g., Parés and Pérez-González, 1995, 1999; Parés et al., 2000, 2006, 2010, 2013) and has been integral in establishing a Lower Pleistocene age for the Homo antecessor fossils at Sima del Elefante and Gran Dolina (Carbonell et al., 1995; Carbonell, 2008). Luminescence dating techniques, particularly single-grain optically stimulated luminescence (OSL) dating of quartz (e.g., Murray and Roberts, 1997; Duller, 2008), are increasingly being used to date Late Pleistocene archaeological deposits in cave and rock shelter settings (e.g., Bowler et al., 2003; Bouzouggar et al., 2007; Petraglia et al., 2007; Jacobs et al., 2008; Armitage et al., 2011). However, conventional quartz OSL methods are limited in their application to Middle Pleistocene sites because the main signal used for dating (the so-called 'fast component') becomes saturated with respect to radiation dose over longer timescales. Dose saturation effects typically impose a practical upper age limit of ~100–200 ka (thousands of years) for quartz OSL, depending on sample-specific luminescence properties and the specific activities of radionuclides in the sedimentary matrix. Nevertheless, the last decade has seen the emergence of several 'extended-range' luminescence dating techniques that make use of alternative luminescence signals with significantly higher dose saturation limits (e.g., Fattahi and Stokes, 2000; Singarayer et al., 2000; Huot et al., 2006; Wang et al., 2006a,b; Jain et al., 2007; Thomsen et al., 2008).

Two such extended-range luminescence dating techniques offer particularly good potential to circumvent conventional OSL restrictions at Middle Pleistocene sites such as Sima de los Huesos: namely, thermally transferred OSL dating of quartz (TT-OSL; Wang et al., 2006a,b) and post-infrared-infrared stimulated luminescence dating of K-feldspars (pIR-IR; Thomsen et al., 2008; Buylaert et al., 2009). Thermally transferred OSL is the light-sensitive signal observed following initial depletion of the main OSL signal and application of a preheat treatment to induce a transfer of charge into the conventional quartz dating trap (e.g., Duller and Wintle,
Previously published results of the simulation of the site showing the luminescence dating and palaeomagnetic sampling localities. (b) west north-west to east south-east longitudinal transect across the Sima de los Huesos. (c) Generalised stratigraphic sequence showing the K-feldspar (pIR-IR, pIR-IR$_{23}$) and quartz (single-grain TT-OSL) luminescence ages obtained at SRB and SRM. The inferred geomagnetic polarity zones (black = normal; white = reversed) are based on the new results presented in this study and the previously published results of Parés et al. (2000), AU and LU — allostratigraphic units and lithostratigraphic units.

Figure 2. Overview of the Sima de los Huesos site and sampling exposures. (a) Plan view of the site showing the luminescence dating and palaeomagnetic sampling localities. (b) west north-west to east south-east longitudinal transect across the Sima chamber, Sima ramp and Sima top sections. (c) Generalised stratigraphic sequence showing the K-feldspar (pIR-IR$_{23}$) and quartz (single-grain TT-OSL) luminescence ages obtained at SRB and SRM. The inferred geomagnetic polarity zones (black = normal; white = reversed) are based on the new results presented in this study and the previously published results of Parés et al. (2000), AU and LU — allostratigraphic units and lithostratigraphic units.

The pIR-IR dating method has been developed out of attempts to circumvent the problem of anomalous fading in feldspars (i.e., unexpected signal loss relative to predicted electron trap lifetimes), which causes apparent age underestimation (Spooner, 1992, 1994; Huntley and Lamothe, 2001). The pIR-IR approach involves measuring an elevated temperature infrared stimulated luminescence (IRSL) signal immediately after performing a lower temperature IRSL bleach (e.g., Thomsen et al., 2008), with the aim of accessing distant electron-hole pairs that have low probabilities of tunnelling recombination during burial. Both TT-OSL and pIR-IR have been used to derive accurate ages over Middle Pleistocene timescales across a range of depositional contexts (e.g., Wang et al., 2006a; Sun et al., 2010; Jacobs et al., 2011; Kang et al., 2011; Thiel et al., 2011; Buylaert et al., 2012; Li and Li, 2012; Lowick et al., 2012; Roberts, 2012; Vasliniu et al., 2012; Arnold et al., 2013; Rink et al., 2013). There are also a growing number of reliable TT-OSL and pIR-IR dating applications in potentially complex archaeological cave or rock shelter settings (Gliganic et al., 2012a, b; Pickering et al., 2013; Rink et al., 2013; Sun et al., 2013).

Here we present new luminescence dating and palaeomagnetic results for sedimentary infill deposits that are directly related to the hominin fossils at Sima de los Huesos. The main aim of the study is to provide a more reliable, bracketing chronological framework for the hominin assemblage using a multifaceted dating approach. For the luminescence dating analyses, we have opted to use both TT-OSL and pIR-IR techniques in tandem. This dual approach provides us with two semi-independent chronological datasets for different mineralogical components of each sample, thereby enabling internal checks on dating consistency and affording added robustness in the final ages. The resultant chronologies are discussed with respect to methodological considerations and other age constraints established at the site. We also examine the broader compatibility of our ages with other well-dated ancestral Neanderthal sites across Europe and with genetic divergence estimates for the Middle Pleistocene hominin record.

Site description and sample details

Details of the Sima de los Huesos stratigraphy have been published previously in Arsuaga et al. (1997a) and Bischoff et al. (1997, 2003, 2007). A revision of the sedimentary record will be presented in a forthcoming manuscript (Arsuaga et al., submitted), as part of an accompanying study to the present chronological research. The latest sedimentological interpretations, which focus on a more quantitative analysis of spatio-temporal relationships among different horizons and exposures, enable identification of twelve distinct lithostratigraphic units that can be grouped into five allostratigraphic members, each delimited by erosive contacts. Fig. 2c summarises the generalised stratigraphic column of the site according to the new classification scheme. In brief, the sedimentary sequence consists of re-deposited basal Miocene marls and well-sorted fluvial sands and silts, overlain unconformably by sterile red clays (lower red clays), red clay-silts containing hominin and carnivore bones (upper red clays) and clay-silt breccias dominated by bear bones and lithic clasts (caf6 con leche unit) (Fig. 2c). Continuous speleothems and broken, isolated fragments of flowstones are found above the basal marls and fluvial sands and silts (LU-2, LU-4), within the upper red clays (LU-2, LU-4) and intercalated with the clay-silt breccia horizons (LU-8). The whole sequence is sealed by an expansive flowstone (Colada or LU-10) and capped by a clay breccia unit containing distinct bat guano horizons.

Excavations have been conducted in three interconnected areas of the site: (i) Sima ramp (SR): a 9-m long linear chamber with a surface inclination of ~32°; (ii) Sima top section (ST): a shorter, ~1.5 m horizontal exposure found at the top of the ramp and...
immediately below the 13 m vertical entry shaft; and (iii) Sima chamber (SH): a small, 32 m² terminal chamber at the base of the ramp, from which the vast majority of hominin and faunal remains have been recovered (Fig. 2a and b). Hominin bones from the same individuals are found dispersed throughout different sectors of the Sima chamber and ramp (e.g., Arsuaga et al., 1997a; Bischoff et al., 2007). The spatial distribution of these bones provides confidence that the chronology established on individual exposures of
sedimentary infill are directly applicable to the main fossil accumulations found in the SH chamber.

Four luminescence dating samples were collected from two excavation exposures at Sima de los Huesos (Figs. 2 and 3). One sample was taken from exposure SRB, a ~1.5 m deep excavation trench located directly adjacent to the main Sima chamber at the base of the inclined Sima ramp. The remaining three samples came from exposure SRM, a similarly sized trench located ~5 m to the southeast of SRB and midway along the inclined Sima ramp. Luminescence dating samples were collected from cleaned exposures at SRB and SRM using opaque PVC tubes and were wrapped in light-proof bags for transportation and storage. Bulk sediment samples were also collected from the surrounding few centimetres of each sample tube for dose rate determination and water content analysis.

Samples SH12-1A, SH12-2A and SH12-3A were collected from the dark yellowish brown café con leche clay breccia horizon, which contains abundant bear bones at SRB and SRM. Isolated hominin bones are also present in the café con leche deposits at SRB but these are thought to have been reworked at the time of deposition from the underlying hominin-bearing red clay breccia (upper red clays). Samples SH12-1A to SH12-3A provide minimum age estimates for the main archaeological horizon found below the café con leche horizon. Luminescence dating samples were not collected directly from the fossiliferous red clays because these units contain relatively low yields of suitably sized silicate minerals and are potentially derived from material that infiltrated slowly through fissures in the bedrock, meaning that the grains may not have been exposed to external daylight immediately prior to deposition. Sample SH12-1A was collected from café con leche deposits, found 33 cm below the ~5 cm-thick capping flowstone (Colada) at SRB and 13 cm above the contact with the underlying hominin-bearing red clay breccia. Sample SH12-2A was taken from the lower café con leche deposits at SRM, 43 cm below the ~10 cm-thick capping flowstone and 5 cm above the underlying fossiliferous red clay deposits. Sample SH12-3A was collected 40 cm to the right and 30 cm above SH12-2A, on the eastern face of SRM and 15 cm below the capping flowstone. This sample was taken from the upper part of the café con leche mud breccia, which contains a distinct horizon of greenish white, carbonate-rich, allochthonous Miocene marl pods (Bischoff et al., 1997) (Fig. 3b). Sample SH12-4A was collected 106 cm to the left of sample SH12-2 from a 40 cm-thick layer of dark yellowish-brown clay overlying the capping flowstone along the northern wall of SRM. This unit is interbedded with several ~5 cm-thick, dark brownish-black bat guano horizons and contains modern rodent fossils (Guerra-Bescós et al., 1997).

This unit shares similar sedimentological properties (colour, texture, structure) to the café con leche deposits but is devoid of bear bones. Sample SH12-4A was taken from the base of this capping clay breccia, 5 cm above the contact with the flowstone and 5 cm below a prominent guano horizon. Uranium series and ^{14}C ages obtained on the immediately underlying flowstone (Bischoff et al., 1997) indicate that the SH12-4A sediments were deposited sometime after 18–68 ka.

The poorly sorted, massive nature of the café con leche accumulations, together with the presence of well-rounded pebbles and conglomerate fragments sourced from nearby surficial Miocene deposits, indicates emplacement of these allochthonous units by a series of high-energy, debris flow event(s) originating from the cave exterior (Arsuaga et al., 1997a). Importantly, the café con leche deposits display some shared mineralogical and sedimentological properties with the sub-aerial surface deposits adorning the surrounding hillslopes (e.g., Bischoff et al., 1997; Arsuaga et al., submitted), revealing a likely aeolian origin for the silt-sized silicate grains being dated in this study before they were washed into the cave. From a luminescence dating perspective, high-energy transportation within the cave system is advantageous because it significantly reduces the chances that the café con leche sediments suffered protracted residency times or punctuated re-transportation episodes within higher parts of the endokarst complex prior to their final deposition. It seems improbable that there would have been sufficient daylight exposure to reset (bleach) the previously accumulated luminescence signals of all sediment grains during these high-energy debris flow events. However, the aeolian source sediments from which these mudflow deposits were derived are likely to have received prolonged and direct sunlight exposure prior to entrainment and transportation into the cave. Providing that the source sediments had accumulated relatively recently before being washed into the cave (i.e., compared with the expected Middle Pleistocene ages of the Sima archaeological deposits), it seems reasonable to assume that the café con leche sediments could have been fully bleached at deposition, or else retained only very minor residual luminescence signals. Indeed, Berger et al. (2008) demonstrated that the present-day surface sediments surrounding the Atapuerca endokarst system have very low residual IRSL signals of 0.8–19.3 Gy (average multi-grain aliquot equivalent dose (D_e) = 3.2 ± 1.3 Gy; n = 6). If similar surface sediments were washed into the caves in the past, they would have only yielded residual ages of 0.3–6.0 ka, which would be effectively inconsequential over the Middle Pleistocene burial timescales considered here. To directly test the assumption that the natural pIR-IR and TT-OSL signals of the café con leche Sima samples could have been adequately bleached prior to their arrival in the endokarst system, we collected a modern sediment sample from the hillside located directly outside the entrance to Cueva Mayor (El Portalon). While the geological provenance of the quartz and feldspar grains in this modern sample may not necessarily match that of the prehistoric Sima deposits, the sample does represent a useful analogue for the geomorphic and transportation history of the café breccia sediments being dated in this study.

A total of 105 oriented palaeomagnetic samples were collected from sedimentary clay horizons and flowstones at eight localities within the main Sima chamber and along the inclined Sima ramp (Table 1, Figs. 2 and 3). Thirty-three sediment samples were collected from the hominin-bearing upper red clay horizon at exposures SH(T17) and SH(R17), located along the northern and southern walls of the main Sima chamber, respectively, and at exposure SRA, a 1-m trench located at the head of the Sima ramp. The underlying sterile lower red clays were sampled at an adjacent exposure along the northern wall of the main chamber (SH(T16)), and an additional 18 samples were collected from gradational zones spanning the base of the upper red clays and top of the lower red clays in excavation squares SH(R11) and SH(R16). The overlying café con leche clay breccia deposits were sampled at excavation trench ST-K (Sima ramp Tenerife or excavation square SH(K22)), which represents a laterally continuous exposure of SRB along the southwest margin of the inclined Sima ramp. Twenty-four samples were collected from two flowstones preserved at exposure SRA: speleothem LU-10, which caps the entire sedimentary sequence, and speleothem LU-8, found overlying in situ café con leche deposits. Eleven samples were also collected from a large flowstone at the adjacent exposure of ST, which comprises three generations of speleothem denoted ST-a, ST-b and ST-c. Speleothems ST-b and ST-c correspond to the lower and upper parts of speleothem SRA-3 reported in Bischoff et al. (2003, 2007). In these earlier studies, speleothem SRA-3 was interpreted as being correlated to flowstone LU-8 of SRA, which clearly overlies the fossiliferous layers. However, with subsequent excavations and more detailed stratigraphic and petrographic analysis, it is now clear that both speleothem ST-a and the lower part of SRA-3 (i.e., ST-b) correspond to LU-2 at
Tabk
Palaeomagnetk
Luminescence dating procedures
recording the azimuth and dip. Ancillary sediment samples at each sampling locality for prevent water dose signal treatments. Samples were stored in a laboratory cooler to prevent water loss and other post-sampling alterations. Core samples were sliced into standard 2.1 cm-long specimens prior to measurement.

**Experimental approaches**

**Luminescence dating procedures**

In the laboratory, quartz and K-feldspar grains of 90–125 μm and 125–180 μm diameter were extracted from the un-illuminated centres of the PVC tube samples under safe (dim red) light conditions and prepared for burial dose estimation using standard procedures (Aitken, 1998). Separated fractions of quartz and K-feldspar grains were etched with 48% hydrofluoric (HF) acid for 40 min and 10% HF acid for 10 min, respectively, to remove their alpha-irradiated external layers. Thermally transferred OSL and pIR-IR measurements were made using the same experimental apparatus described by Arnold et al. (2013). Ultraviolet OSL emissions from quartz were detected using an EMI 9235QA photomultiplier tube fitted with 7.5 mm-thick Hoya U-340 filters. Potassium feldspar IRSL emissions were detected in the blue region of the spectrum using a combination of 3 mm-thick BG39 and 4 mm-thick Corning 7-59 filters. Samples were irradiated with a mounted 90Sr/90Y beta-source that had been calibrated to administer known doses to multi-grain aliquots and single-grain discs. For single-grain measurements, spatial variations in beta-dose rates across the disc plane were taken into account by undertaking hole-specific calibrations using gamma-irradiated quartz.

Laboratory bleaching studies have revealed that natural K-feldspar pIR-IR signals are reset relatively rapidly by simulated daylight, albeit an order of magnitude slower than for conventional quartz OSL signals, i.e., initial signal intensities are depleted by over 90% within several tens of seconds to several tens of minutes using a SOL2 solar simulator (depending on choice of preheat and pIR-IR stimulation temperature, e.g., Thomsen et al., 2008; Buylaert et al., 2012). In the present study, we have performed pIR-IR stimulations at 225 °C following a preheat of 250 °C for 60 s (i.e., a pIR-IR250 procedure). Modern samples from a range of well-bleached aeolian deposits typically display relatively low residual doses on the order of <5 Gy for pIR-IR250 procedures (e.g., Buylaert et al., 2011), and up

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Palaeomagnetic data summary for the Sima de los Huesos samples.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Palaeomagnetism Site reference Sedimentary unit</td>
<td>Sample depth (m)</td>
</tr>
<tr>
<td>SH1 SH(R11)</td>
<td>Lower/upper red clays</td>
</tr>
<tr>
<td>SH2 SH(T-17)</td>
<td>Upper red clays</td>
</tr>
<tr>
<td>SH3 SH(R16)</td>
<td>Lower/upper red clays</td>
</tr>
<tr>
<td>SH4 SH(R17)</td>
<td>Upper red clays</td>
</tr>
<tr>
<td>SH5 SH(T16)</td>
<td>Lower red clays</td>
</tr>
<tr>
<td>SH6 SRT(K22)</td>
<td>Café con leche clays</td>
</tr>
<tr>
<td>SRA1 SRA</td>
<td>LU-10 and LU-8 flowstones</td>
</tr>
<tr>
<td>SRA1 SRA</td>
<td>LU-2 (including ST-a, b) and LU-8 (ST-c) flowstones</td>
</tr>
</tbody>
</table>

Dec: Declination, Inc: Inclination, a95: Radius of 95% confidence; k: fisher parameter, n/N: number of used/measured specimens.

exposure SRA, and therefore underlie the red clay horizons and the sand and silt horizons in the revised lithostratigraphic scheme (Fig. 2). Only the upper portion of SRA-3 (i.e., ST-c) can be correlated with LU-8 at exposure SRA, as detailed in the new stratigraphic scheme of Arsuaga et al. (submitted). In addition to collecting the main samples for determining magnetic polarity, we also took ancillary sediment samples at each sampling locality for rock-magnetism measurements.

Clay horizons were sampled for palaeomagnetic dating using standard 8 cm³ cubic boxes. The sedimentary facies at SH, SRT and SRA were all soft enough to allow the sampling boxes to be gently pushed into the exposures and extracted using a ceramic knife after recording the azimuth and dip. Flowstone samples were collected at SRA and ST using a portable electric drill equipped with a 1-inch non-magnetic drill bit. Following collection and transportation, all palaeomagnetic samples were stored in a laboratory cooler to prevent water loss and other post-sampling alterations. Core samples were sliced into standard 2.1 cm-long specimens prior to measurement.

### Table 2

**Single-aliquot regenerative-dose (SAR) procedures used for D<sub>n</sub> determination.**

<table>
<thead>
<tr>
<th>Step</th>
<th>Treatment</th>
<th>Signal</th>
<th>Step</th>
<th>Treatment</th>
<th>Signal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Dose (natural or laboratory)</td>
<td></td>
<td>1</td>
<td>Dose (natural or laboratory)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Preheat 1 (250 °C for 60 s)</td>
<td></td>
<td>2</td>
<td>Preheat 1 (250 °C for 10 s)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>IR stimulation (50 °C for 200 s)</td>
<td></td>
<td>3</td>
<td>Single-grain OSL stimulation (125 °C for 3 s)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>pIR-IR stimulation (225 °C for 200 s)</td>
<td></td>
<td>4</td>
<td>Preheat 2 (250 °C for 10 s)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Test dose (100 Gy)</td>
<td></td>
<td>5</td>
<td>Single-grain TT-OSL stimulation (125 °C for 3 s)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Preheat 2 (250 °C for 60 s)</td>
<td></td>
<td>6</td>
<td>OSL stimulation (280 °C for 400 s)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>IR stimulation (50 °C for 200 s)</td>
<td></td>
<td>7</td>
<td>Test dose (250 Gy)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>pIR-IR stimulation (225 °C for 200 s)</td>
<td></td>
<td>8</td>
<td>Preheat 3 (220 °C for 10 s)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Repeat measurement cycle for different sized regenerative doses</td>
<td></td>
<td>9</td>
<td>Single-grain OSL stimulation (125 °C for 3 s)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Repeat measurement cycle for different sized regenerative doses</td>
<td></td>
<td>10</td>
<td>Preheat 4 (250 °C for 10 s)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Repeat measurement cycle for different sized regenerative doses</td>
<td></td>
<td>10</td>
<td>Single-grain TT-OSL stimulation (125 °C for 3 s)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Repeat measurement cycle for different sized regenerative doses</td>
<td></td>
<td>12</td>
<td>OSL stimulation (290 °C for 400 s)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Repeat measurement cycle for different sized regenerative doses</td>
<td></td>
<td>13</td>
<td>Repeat measurement cycle for different sized regenerative doses</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Repeat measurement cycle for different sized regenerative doses</td>
<td></td>
<td>14</td>
<td>Repeat measurement cycle for different sized regenerative doses</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Repeat measurement cycle for different sized regenerative doses</td>
<td></td>
<td>15</td>
<td>Repeat measurement cycle for different sized regenerative doses</td>
<td></td>
</tr>
</tbody>
</table>

Each SAR measurement cycle was repeated for the natural-dose, three to four different sized regenerative doses, a 0 Gy regenerative-dose (to measure signal recuperation) and a replicate of the first regenerative-dose cycle (to assess the suitability of the test-dose sensitivity correction). In the case of the single-grain TT-OSL SAR procedure, the OSL IR depletion ratio of Duller (2003) was measured separately and used to check for the presence of feldspar contaminants. L<sub>n</sub> = regenerative-dose signal response; L<sub>x</sub> = natural-dose signal response; T<sub>n</sub> = test-dose signal response for a laboratory dose cycle; T<sub>x</sub> = test-dose signal response for the natural-dose cycle.
Table 3
Environmental dose rate data for the Sima de los Huesos luminescence dating samples.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Sample depth (m)</th>
<th>Grain size (µm)</th>
<th>Water content (%)</th>
<th>Radionuclide specific activities (Bq/kg)</th>
<th>Environmental dose rate (Gy/ka)</th>
<th>Total dose rate (quartz)</th>
<th>Total dose rate (K-feldspar)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>238U</td>
<td>226Ra</td>
<td>210Pb</td>
<td>228Ra</td>
</tr>
<tr>
<td>Sima ramp middle (SRM)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SH12-4A</td>
<td>0.35</td>
<td>90–125</td>
<td>28 ± 6</td>
<td>44.0 ± 6.8</td>
<td>31.6 ± 3.8</td>
<td>31.1 ± 2.6</td>
<td>38.9 ± 3.3</td>
</tr>
<tr>
<td>SH12-3A</td>
<td>0.53</td>
<td>90–125</td>
<td>22 ± 4</td>
<td>36.0 ± 4.8</td>
<td>36.6 ± 2.7</td>
<td>40.1 ± 0.8</td>
<td>36.1 ± 3.0</td>
</tr>
<tr>
<td>SH12-2A</td>
<td>0.83</td>
<td>90–125</td>
<td>26 ± 5</td>
<td>32.5 ± 4.4</td>
<td>30.2 ± 2.1</td>
<td>30.0 ± 0.5</td>
<td>37.8 ± 3.2</td>
</tr>
<tr>
<td>Sima ramp bottom (SRB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SH12-1A</td>
<td>0.33</td>
<td>90–125</td>
<td>20 ± 4</td>
<td>30.6 ± 4.9</td>
<td>27.4 ± 2.0</td>
<td>27.5 ± 2.3</td>
<td>26.3 ± 2.2</td>
</tr>
</tbody>
</table>

a Field water content, expressed as % of dry mass of mineral fraction, with an assigned relative uncertainty of ±20%.
b Measurements made on dried and powdered samples by high-resolution gamma-ray spectrometry.
c Gamma dose rates were calculated from in situ measurements made at each sample position with a NaI(Tl) detector, using the 'energy windows' approach detailed in Arnold et al. (2012a). Beta dose rates were calculated using a Riso GM-25-5 low-level beta counter (Butter-Jensen and Mejdahl, 1988), after making allowance for beta dose attenuation due to grain size effects and HF etching (Brennan, 2003).
d Radionuclide concentrations and specific activities have been converted to dose rates using the conversion factors given in Guérin et al. (2011) and Readhead (2002), making allowance for beta-dose attenuation (Mejdahl, 1979; Brennan, 2003).
e Cosmic-ray dose rates were calculated using the approach of Prescott and Hutton (1994) and assigned a relative uncertainty of ±10%.
f Assumed internal alpha + beta dose rate for quartz with an assigned relative uncertainty of ±30%, based on intrinsic 238U and 232Th contents published by Mejdahl (1987), Bowler et al. (2003), Jacobs et al. (2006) and Pawley et al. (2008), and an ε-value of 0.04 ± 0.01 (Rees-Jones, 1995; Rees-Jones and Tite, 1997).
g Assumed internal feldspar dose rate. Internal 40K and 40Ar concentrations were estimated as being 12.5 ± 0.5% (Huntley and Bari, 1997) and 400 ± 100 ppm (Huntley and Hancock, 2001), respectively, yielding an internal beta dose rate of 0.43 ± 0.03 Gy/ka for the 90–125 µm K-feldspar grains. An additional internal alpha + beta dose rate of 0.06 ± 0.03 Gy/ka has been calculated using assumed intrinsic 238U and 232Th concentrations of 0.15 ± 0.03 ppm and 0.35 ± 0.07 ppm, respectively (following Mejdahl, 1987; Huntley and Bari, 1997; Huntley and Clague, 1985; Huntley and Lian, 1999; Alappat et al., 2010), and an ε-value of 0.09 ± 0.03 (following Rees-Jones, 1995; Lang and Wagner, 1997; Banerjee et al., 2001; Lang et al., 2003; Berger et al., 2008; Feathers et al., 2012).
h Mean ± total uncertainty (68% confidence interval), calculated as the quadratic sum of the random and systematic uncertainties.
to several tens of Gy for pIR-IR stimulations performed at 290 °C following a preheat of 320 °C for 60 s (i.e., a pIR-IR$_{290}$ procedure) (Buylaert et al., 2012). Together, these results suggest that unbleachable/thermally transferred residual pIR-IR signals are unlikely to yield noticeable age inaccuracies for most Middle Pleistocene age samples and that the use of multi-grain aliquot dating procedures is nominally suitable in the present depositional context.

Unlike pIR-IR signals, natural quartz TT-OSL signals have been shown to bleach very slowly (Duller and Wintle, 2012), taking on the order of several weeks to be depleted by >80–90% in solar simulator experiments (e.g., Tsukamoto et al., 2008; Porat et al., 2009) and several months to be significantly depleted by natural daylight (e.g., Jacobs et al., 2011; Arnold et al., 2013). Suitably low $D_\alpha$ residuals of a few tens of Gy have, however, been recorded in well-bleached modern samples (Duller and Wintle, 2012). The slowly bleaching nature of the TT-OSL signal dictates that there is greater potential for insufficient signal resetting (and concomitant age overestimation) in most depositional settings when using multi-grain aliquot TT-OSL approaches. For this reason, we have cautiously chosen to use a single-grain TT-OSL dating procedure in the present study, which ensures that any partially bleached grain populations can be identified more reliably and treated with an appropriate statistical model in the final age calculation. The potential for applying single-grain TT-OSL dating procedures to Middle Pleistocene deposits from the Sierra de Atapuerca is greatly enhanced by the inherently bright OSL signal intensities of quartz from the upper Duero Basin (Arnold et al., 2012), and this approach has been recently used to reliably date nearby Middle Pleistocene fluvial terrace deposits in the Pico valley (Arnold et al., 2013). Single-grain TT-OSL ages obtained in the study of Arnold et al. (2013) were in good agreement with an associated single-grain OSL ‘supergrain’ age, as well as with a series of quartz ESR ages (Moreno et al., 2012), $T_l$ ages and palaeomagnetic polarity results (Benito-Calvo et al., 2008) for bracketing fluvial terraces.

Quartz and K-feldspar $D_\alpha$ values were determined using pIR-IR and TT-OSL versions of the single-aliquot regenerative-dose (SAR) procedure (Murray and Wintle, 2000) shown in Table 2. In these SAR protocols, sensitivity changes are corrected for by measuring the TT-OSL/pIR-IR signal response to a test-dose ($T_x$) given after the natural ($L_\alpha$) or regenerated ($L_\beta$) TT-OSL/pIR-IR signals have been measured. Sensitivity-corrected luminescence responses are then determined by dividing the natural and regenerated TT-OSL/pIR-IR intensities by their respective test-dose signal intensities ($L_{\alpha,\beta}/T_x$, respectively). Further details of the SAR procedures, including dose-recovery validation tests and quality assurance criteria used to eliminate unreliable grains and aliquots, are provided in the Supplementary Online Material (SOM).

Environmental dose rates have been calculated using a combination of field gamma-ray spectrometry and low-level beta counting (Table 3). High-resolution gamma spectrometry was additionally performed on each of the luminescence dating samples to assess the present-day state of secular (dis)equilibrium in the $^{238}$U and $^{232}$Th decay series. Full details of the procedures used to calculate the environmental dose rates for the quartz and K-feldspar fractions are provided in the SOM.

Palaeomagnetic dating procedures

Palaeomagnetic analyses were carried out at the Palaeomagnetism Laboratory of the Centro Nacional de Investigación sobre la Evolución Humana (CENIEH), Burgos, Spain. Measurements of the Natural Remanent Magnetisation, and its decay during demagnetisation, were carried out on a 2G Enterprises 755-4K Superconducting Rock Magnetometer, housed in a low-magnetic environment and created using a set of large Helmholtz coils. The magnetometer has a built-in degaussser for alternating field (AF) demagnetisation up to 170 mT. Thermal demagnetisation apparatus includes a TD-48SC oven (ASC Scientific). Hysteresis and isothermal remanent magnetisation (IRM) curves were obtained with a Princeton Measurements Corporation Vibrating Sample

![Figure 4](image-url)
<table>
<thead>
<tr>
<th>Sample</th>
<th>Sample depth (m)</th>
<th>Grain size (µm)</th>
<th>Total dose rate (Gy/ka)</th>
<th>D_s type and resolution</th>
<th>Accepted/measured D_s values</th>
<th>Fading rate (R_{1.25\text{ yrs}}) (%/decade)</th>
<th>Weighted skewness</th>
<th>Critical skewness (95% CI)</th>
<th>CAM results</th>
<th>MAM results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CAM D_s (Gy)</td>
<td>MAM-3 D_s (Gy)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CAM age (ka)</td>
<td>MAM-4 D_s (Gy)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modern analogue sample(^d)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SH12-3A</td>
<td>0.05</td>
<td>90–125</td>
<td>—</td>
<td>MG (350 grains)</td>
<td>6/6</td>
<td>—</td>
<td>—0.24</td>
<td>±0.00</td>
<td>24 ± 7</td>
<td>7.3 ± 0.8</td>
</tr>
<tr>
<td>Sinam ramp middle (SRM)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SH12-3A</td>
<td>0.35</td>
<td>90–125</td>
<td>2.26 ± 0.14</td>
<td>MG (350 grains)</td>
<td>11/12</td>
<td>1.38 ± 0.11</td>
<td>—0.02</td>
<td>±1.48</td>
<td>4 ± 1</td>
<td>943 ± 15</td>
</tr>
<tr>
<td>SH12-3A</td>
<td>0.53</td>
<td>90–125</td>
<td>2.26 ± 0.12</td>
<td>MG (350 grains)</td>
<td>12/12</td>
<td>1.19 ± 0.06</td>
<td>0.33</td>
<td>±1.41</td>
<td>4 ± 1</td>
<td>971 ± 14</td>
</tr>
<tr>
<td>SH12-2A</td>
<td>0.83</td>
<td>90–125</td>
<td>2.20 ± 0.12</td>
<td>MG (350 grains)</td>
<td>11/12</td>
<td>1.62 ± 0.06</td>
<td>—0.08</td>
<td>±1.48</td>
<td>5 ± 1</td>
<td>940 ± 17</td>
</tr>
<tr>
<td>Sinam ramp bottom (SRB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SH12-1A</td>
<td>0.33</td>
<td>90–125</td>
<td>2.09 ± 0.10</td>
<td>MG (350 grains)</td>
<td>12/12</td>
<td>1.71 ± 0.10</td>
<td>0.90</td>
<td>±1.41</td>
<td>7 ± 2</td>
<td>923 ± 21</td>
</tr>
</tbody>
</table>

\(^a\) D_s type: MG (350 grains) = multi-grain aliquots containing ~350 grains each.

\(^b\) Number of \(D_s\) measurements that passed the SAR rejection criteria/total number of multi-grain aliquots analysed.

\(^c\) Weighted skewness scores have been calculated on log-transformed \(D_s\) values using Eqs. (7)–(8) of Arnold and Roberts (2009).

\(^d\) Critical skewness scores have been calculated using Eq. (16) of Bailey and Arnold (2006). \(D_s\) distributions are considered to be significantly skewed if the weighted skewness value is greater than the corresponding critical skewness value.

\(^e\) CAM = central age model, MAM-3 = 3-parameter minimum age model, MAM-4 = 4-parameter minimum age model (Galbraith et al., 1999).

\(^f\) Total uncertainty includes a systematic component of ±2% associated with laboratory beta-source calibration.

\(^g\) Final pIR-IR ages (shown in bold) have been calculated using the CAM \(D_s\) for these multi-grain aliquot measurements (see main text for details).

\(^h\) The modern analogue sample was collected from surface sediments for the purpose of assessing residual \(D_s\) values in this sedimentary context. Since it was not practical to measure the in situ gamma dose rate for this sample we have not calculated a final age. The fading rate \(R_{1.25\text{ yrs}}\) was also not measured for this sample.
Table 5
Summary statistics for the TT-OSL $D_a$ measurements, age model results and final ages of the Sima de los Huesos samples.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Sample depth (m)</th>
<th>Grain size ($\mu$m)</th>
<th>Total dose rate (Gy/ka)</th>
<th>$D_a$ type and resolution</th>
<th>Accepted/ measured $D_a$ values</th>
<th>Weighted skewness</th>
<th>Critical skewness (95% CI)</th>
<th>CAM results</th>
<th>FMM results</th>
<th>FMM age (ka)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Modern analogue sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SH12-5A</td>
<td>0.05</td>
<td>125–180</td>
<td>–</td>
<td>MG (1100 grains)</td>
<td>6/6</td>
<td>−0.60 ± 2.00</td>
<td>11 ± 8</td>
<td>7.3 ± 1.2</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><strong>Sima ramp middle (SRM)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SH12-4A</td>
<td>0.35</td>
<td>90–125</td>
<td>1.80 ± 0.13</td>
<td>SG (18 grains)</td>
<td>39/1500</td>
<td>−0.12 ± 0.78</td>
<td>22 ± 5</td>
<td>767 ± 41</td>
<td>426 ± 39</td>
<td>426 ± 39</td>
</tr>
<tr>
<td>SH12-3A</td>
<td>0.53</td>
<td>90–125</td>
<td>1.80 ± 0.11</td>
<td>SG (18 grains)</td>
<td>70/2300</td>
<td>−0.81 ± 0.59</td>
<td>42 ± 5</td>
<td>713 ± 42</td>
<td>396 ± 35</td>
<td>396 ± 35</td>
</tr>
<tr>
<td>SH12-2A</td>
<td>0.83</td>
<td>90–125</td>
<td>1.74 ± 0.12</td>
<td>SG (18 grains)</td>
<td>76/1800</td>
<td>−0.08 ± 0.56</td>
<td>21 ± 4</td>
<td>728 ± 27</td>
<td>419 ± 34</td>
<td>419 ± 34</td>
</tr>
<tr>
<td><strong>Sima ramp bottom (SRB)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SH12-1A</td>
<td>0.33</td>
<td>90–125</td>
<td>1.63 ± 0.10</td>
<td>SG (18 grains)</td>
<td>54/1400</td>
<td>−0.26 ± 0.67</td>
<td>19 ± 5</td>
<td>701 ± 31</td>
<td>429 ± 32</td>
<td>429 ± 32</td>
</tr>
</tbody>
</table>

† $D_a$ type: MG (1100 grains) = multi-grain aliquots containing ~1100 grains each; SG (18 grains) = single-grain measurements containing ~18 grains per grain-hole position.

‡ Number of $D_a$ measurements that passed the SAR rejection criteria/total number of grains or multi-grain aliquots analysed.

§ Weighted skewness scores have been calculated on log-transformed $D_a$ values using Eqs. (7)–(8) of Arnold and Roberts (2009).

‖ Critical skewness scores have been calculated using Eq. (16) of Bailey and Arnold (2006). $D_a$ distributions are considered to be significantly skewed if the weighted skewness value is greater than the corresponding critical skewness value.

¶ CAM = central age model (Galbraith et al., 1999), FMM = finite mixture model (Galbraith and Green, 1990).

± Mean ± total uncertainty (68% confidence interval), calculated as the quadratic sum of the random and systematic uncertainties.

‡‡ Total uncertainty includes a systematic component of ±2% associated with laboratory beta-source calibration.

∞ Final ages (shown in bold) have been calculated using the CAM $D_a$ for these single-grain TT-OSL measurements (see main text for details).

†† The FMM was fitted by varying the common overdispersion parameter ($\sigma_0$) between 5 and 30% and incrementally increasing the specified number of $K_a$ components. The FMM parameter values shown here were obtained from the optimum FMM fit (i.e., the fit with the lowest BIC score; Arnold and Roberts, 2009). Using this approach, the $D_a$ distributions of SH12-1A, SH12-2A and SH12-4A are shown to contain a single dose component $K_1$, whereas sample SH12-3A contains two discrete dose populations ($K_1$ and $K_2$). The FMM has not been applied to the multi-grain $D_a$ dataset of SH12-5A because it is only suitable for use with single-grain $D_a$ datasets (Arnold and Roberts, 2009).

‡‡‡ The modern analogue sample was collected from surface sediments for the purpose of assessing residual $D_a$ values. As it was not practical to measure the in situ gamma dose rate for this sample we have not calculated a final age.
Magnetometer (Micromag 3900 VSM). Specimens were stored in a μ-metal shielded cylindrical can during analysis to prevent acquisition of viscous magnetisation.

In addition to acquiring the primary magnetisation, it is well known that rocky materials can acquire secondary magnetisations as a result of natural processes such as weathering, heating for an extended time and diagenesis. Moreover, materials can be artificially magnetised by proximity to strong magnetic fields during sampling, drilling, transportation, and storage. One of the major tasks in any palaeomagnetic study is to resolve the Natural Remanent Magnetisation, i.e., the sum of all magnetisations carried by a rock, into its constituent primary and secondary components. For this purpose, both thermal and alternating field demagnetisation were used to isolate the Characteristic Remanent Magnetisation (ChRM) component. The ChRM directions were calculated for all specimens using Principal Component Analysis (Kirschvink, 1980), guided by visual inspection of orthogonal demagnetisation plots (Zijderveld, 1967). Mean directions and associated statistical parameters were estimated using Fisher's (1953) method (dispersion k and confidence angle α95).

**Luminescence dating results and analyses**

**Luminescence characteristics**

Representative pIR-IR25/TT-OSL decay curves and doresponse curves for grains/ aliquots that passed the SAR quality assurance criteria are shown in Fig. 4. The pIR-IR25 Tₐ signals typically decay to <5% of initial intensities within the first 100-150 s of stimulation and reach stable background levels within the last 50 s of stimulation (e.g., Fig. 4a). The associated pIR-IR25 multi-grain aliquot dose-response curves are generally well-represented by either a single saturating exponential function or a saturating exponential plus linear function. The accepted quartz grains typically display Tₐ (200 Gy) TT-OSL signals of 100-500 cts/0.17 s (e.g., Fig. 4b). However, all samples contain a small number of relatively bright accepted grains (e.g., Fig. 4c) displaying Tₐ TT-OSL signal intensities >1000 counts/0.17 s (Fig. 4b). Single-grain TT-OSL dose-response curves display more intra-sample variability than their multi-grain aliquot pIR-IR25 counterparts but are generally characterised by continued signal growth at very high (10³ G) doses and are well-represented by a single saturating exponential function (e.g., Fig. 4b-c).

**Modern analogue residual doses**

Tables 4 and 5 summarise the Dₑ values and age estimates obtained for the Sima samples. Also included in these tables are the weighted mean Dₑ results obtained for the modern sample (SH12-5A) collected outside the cave mouth. The average pIR-IR25 and TT-OSL residual doses for this sample are 7.3 ± 0.8 Gy and 7.3 ± 1.2 Gy, respectively. These natural Dₑ residuals are comparable with those obtained for modern aeolian samples in previous TT-OSL studies (e.g., Wang et al., 2006a; Porat et al., 2009; Stevens et al., 2009) and previous pIR-IR25 studies (e.g., Thomsen et al., 2008; Buylaert et al., 2009). They are also consistent with the range of residual IRSL doses measured for a wider range of present-day surface sediments in the surroundings of the Sierra de Atapuerca (Berger et al., 2008). The residual Dₑ values of SH12-5A are largely inconsequential in comparison with the empirical Dₑ values measured for the Sima samples; equating to <1% of the weighted mean quartz Dₑ values and <0.7% of the weighted mean K-feldspar Dₑ estimates for samples SH12-1A to SH12-4A. As such, we have not considered an additional residual dose subtraction in the age estimates presented in Tables 4 and 5. These results provide good support for the potentially well-bleached nature of the Sima source sediments prior to their transportation into the cave system. The residual Dₑ values would correspond to age offsets of only ~3–4 ka if the Sima sample had been sourced from similar surface deposits in the past. Any such systematic biases would be well within the existing 1σ uncertainties of our final ages.

**pIR-IR25 Dₑ results**

The multi-grain aliquot K-feldspar Dₑ distributions (Fig. 5) each cover a fairly limited range of doses (relative Dₑ range = 0.16–0.32); they are consistent with a single dose population centred on the weighted mean Dₑ value (as indicated by the large proportion of aliquots lying within the grey bands in Fig. 5) and they are not considered to be significantly skewed at the 95% C.I. when assessed using a log weighted skewness test (Arnold and Roberts, 2009, 2011). The Dₑ distributions all display very low overdispersion values of 4–7% (Table 4), which suggests that any dose dispersion originating from sources such as partial bleaching are relatively unimportant in relation to the size of the existing Dₑ measurement uncertainties. While it is difficult to rule out multi-grain averaging effects in any multi-grain aliquot study (e.g., Demuro et al., 2008, 2013; Arnold and Roberts, 2009; Arnold et al., 2013), the apparently fully bleached nature of these samples is consistent with the multi-grain aliquot results obtained for the modern analogue. Moreover, the effects of any undetectable, averaged-out residual doses should be relatively minor in this instance given the very large doses absorbed by the sediments during their prolonged burial periods. In light of these multi-grain aliquot Dₑ characteristics, we have estimated the final burial doses from each sample's weighted mean Dₑ estimate, calculated using the central age model (CAM) of Galbraith et al. (1999). For comparative purposes, Table 4 also shows the Dₑ values and corresponding ages obtained using the 3- and 4-parameter minimum age models (MAM-3, MAM-4) of Galbraith et al. (1999). The MAM-3 and MAM-4 pIR-IR25 ages are consistent with their CAM counterparts at 1σ for all four Sima samples, indicating that our final chronologies are insensitive to age model choice or assumptions concerning bleaching history.

**pIR-IR25 fading tests**

To assess the long-term (athermal) stability of the pIR-IR25 signal, we undertook anomalous fading measurements on three of the aliquots used to derive Dₑ values for each sample. Laboratory fading rates were measured using the pIR-IR25 SAR protocol shown in Table 2a, following the procedure suggested by Auclair et al. (2003). After Dₑ measurements were completed, aliquots were subjected to a series of nine repeated SAR Lₐ/Tₑ measurements that included different storage periods after each regenerative-dose preheat step (i.e., between steps 2 and 3 in Table 2a). The storage times included in these repeat Lₐ/Tₑ cycles varied between ‘prompt’ measurements of 0.17 h to maximum delays of 24 h. The total delay times used to calculate fading rates comprise half of the Lₑ irradiation time, the intervening preheating and thermal ramping times, carousel rotation times between measurements, and the storage periods. Empirical fading rates were quantified in terms of the g-value (Aitken, 1985), which expresses the percentage decrease in sensitivity-corrected pIR-IR signal per decade of storage time since irradiation. The g-values were determined from the repeated Lₐ/Tₑ values using Eq. (4) of Huntley and Lamoureux (2001), and were subsequently normalised to a measurement delay time of two days (Tₑ2days) to enable direct comparisons with previously published values.

The pIR-IR25 laboratory fading rates obtained for the Sima samples are all very low (Table 4). Weighted mean g-values for each
sample ranged between $1.19 \pm 0.07\%$ and $1.71 \pm 0.10\%/\text{decade}$, and the combined weighted average $g$-values for all individual aliquots ($n = 12$) was $1.47 \pm 0.07\%/\text{decade}$. By comparison, weighted mean $g$-values obtained using the conventional IRSL signal (measured as part of the pIR-IR$_{225}$ SAR protocol; IR$_{50/225}$) were much higher, ranging between $2.75 \pm 0.26\%$ and $3.52 \pm 0.14\%/\text{decade}$ per sample. These results are consistent with previous observations (e.g., Thomsen et al., 2008; Buylaert et al., 2009), which suggest that pIR-IR signals are inherently more stable than conventional IRSL signals. The pIR-IR$_{225}$ $g$-values are also in good agreement with previously published pIR-IR$_{225}$ fading rates, which typically vary between $1-2\%/\text{decade}$ (e.g., Thomsen et al., 2008; Buylaert et al., 2009, 2011; Kars et al., 2012; Lowick et al., 2012; Roberts, 2012; Vasiliniuc et al., 2012).

A number of studies have demonstrated good agreement between pIR-IR$_{225}$ ages that have not been corrected for fading and associated age control for samples that display $g$-values of $1-2\%/\text{decade}$ (e.g., Gliganic et al., 2012b; Kars et al., 2012; Roberts, 2012). In several cases, uncorrected pIR-IR$_{225}$ ages appear equally, if not more, reliable in comparison to uncorrected pIR-IR$_{290}$ ages for known-age samples (Buylaert et al., 2011; Lowick et al., 2012; Roberts, 2012; Vasiliniuc et al., 2012). It is worth noting that replicate pIR-IR$_{225}$ and pIR-IR$_{290}$ fading rates reported so far in the literature for the same samples are statistically indistinguishable from each other at $2\sigma$ (e.g., Buylaert et al., 2011; Lowick et al., 2012; Roberts, 2012; Vasiliniuc et al., 2012). Additionally, there is evidence to suggest that natural pIR-IR$_{290}$ (and pIR-IR$_{225}$) signals can reach field-saturation in some samples and are therefore seemingly unaffected by anomalous fading (e.g., Buylaert et al., 2012; Thiel et al., 2011; Roberts, 2012; Vasiliniuc et al., 2012). These collective observations have led to suggestions that the very low $g$-values measured for pIR-IR$_{225}$ signals are likely to be an artefact of laboratory measurement procedures (e.g., they could be caused by unreliable sensitivity correction following storage) and that they do not necessarily provide reliable estimates of anomalous fading (e.g., Buylaert et al., 2012; Roberts, 2012). If this is the case, fading corrections based on potentially problematic $g$-value assessments may actually introduce additional age uncertainties. Additional support for this interpretation comes from $g$-value measurements performed on quartz, which does not typically suffer from anomalous fading. Empirical $g$-values measured for various quartz samples by Thiel et al. (2011) and Buylaert et al. (2012) ranged between $0.8$--$1.9\%/\text{decade}$ at $2\sigma$, suggesting that pIR-IR fading rates may be overestimated by a similar order of magnitude.

**Figure 5.** Multi-grain aliquot K-feldspar pIR-IR$_{225}$ $D_e$ distributions for the four SH12 samples, shown as radial plots. The $D_e$ value for each aliquot is read by drawing a line from the origin of the y-axis ("Standardised Estimate"), through the data point of interest, to the radial axis (plotted on a log scale) on the right-hand side; the point of intersection is the $D_e$ (in Gy). The measurement error on this $D_e$ is obtained by extending a line vertically to the x-axis, where the point of intersection is the relative standard error (shown as a percentage of the $D_e$ value) and its reciprocal ("Precision"). In radial plots, the most precise estimates fall furthest to the right, and the least precise estimates fall furthest to the left. Here, the grey bands are centred on the weighted mean $D_e$ values used to calculate the pIR-IR$_{225}$ ages, which we estimated using the central age model of Galbraith et al. (1999).
The low pIR-IR$_{225}$ fading rates obtained in the present study (Table 4) suggest that the signal being used for dating is potentially as stable as that reported elsewhere for pIR-IR$_{290}$ and quartz OSL procedures. Taking into consideration this factor, as well as published evidence suggesting that such low empirical fading rates may be unreliable or artefacts of laboratory procedures (e.g., Buylaert et al., 2012; Lowick et al., 2012; Roberts, 2012; Vasiliniuc et al., 2012), we have not applied a fading correction to the final pIR-IR$_{225}$ ages of the Sima samples.

**TT-OSL $D_e$ results**

Samples SH12-1A, SH12-2A and SH12-4A share similar, homogeneous single-grain TT-OSL $D_e$ distributions (Fig. 6a,b,d) characterised by moderate dose dispersion (relative $D_e$ range = 1.3–14), single dose populations that are not significantly skewed according to the criterion outlined by Arnold and Roberts (2011), $D_e$ scatter that is well-represented by the weighted mean band at 2$\sigma$, and low overdispersion values of 20–23% (Table 5). The individual overdispersion estimates are all consistent with a value of <20% at 1$\sigma$ and <15% at 2$\sigma$, and are in agreement with published single-grain OSL overdispersion datasets for ideal, well-bleached sedimentary samples that have not been affected by post-depositional mixing (see summary in Arnold and Roberts, 2009). The single-grain TT-OSL radial plots do not display distinct ‘leading-edges’ of low $D_e$ values or elongated, asymmetric ‘tails’ of higher $D_e$ values, which are characteristic of partially bleached single-grain $D_e$ distributions (e.g., Olley et al., 1999, 2004; Arnold et al., 2007, 2009). Any residual doses arising from insufficient bleaching therefore appear to have been relatively minor compared with the large burial doses accumulated by these deposits. The absence of multiple dose components in the single-grain $D_e$ distributions (Table 5) also suggests that sediment mixing (e.g., Arnold et al., 2013) and beta-dose heterogeneity in the natural burial environment (e.g., Jacobs et al., 2008) are unimportant. Finally, the single-grain TT-OSL $D_e$ distributions are consistent with the corresponding multi-grain aliquot pIR-IR$_{225}$ $D_e$ characteristics for these three samples, lending further support to the interpretation that these samples were fully bleached prior to deposition and remained in situ thereafter. We have therefore used the CAM to derive representative burial dose estimates and final ages. Multi-grain (~1100 grain) aliquot TT-OSL measurements made on sample SH12-2A yielded a weighted mean natural $D_e$ value of 795 ± 97 Gy ($n = 5$) and a corresponding age of 458 ± 65 ka. Both the single-grain and multi-grain aliquot replicate

Figure 6. Single-grain quartz TT-OSL $D_e$ distributions for the four SH12 samples, shown as radial plots. For samples SH12-1A, SH12-2A and SH12-4A, the grey bands are centred on the weighted mean $D_e$ values used to calculate the TT-OSL ages, which we estimated using the central age model of Galbraith et al. (1999). For sample SH12-3A, the weighted mean burial dose estimate of the dominant FMM component (i.e., that containing the highest proportion of individual $D_e$ values) is shown as a dark grey shaded band on the radial plot. The additional dose component identified by the optimum FMM fit is shown as a light grey shaded band on the radial plot. The percentage of grains associated with each fitted FMM component is also shown on the radial plots. The weighted mean $D_e$ value used to calculate the TT-OSL age of this sample is shown as a dashed band.
ages for this sample are consistent at 1σ (Table 5), providing added confidence in the final burial dose estimates.

The single-grain TT-OSL $D_e$ distribution for sample SH12-3A is distinctly different from its neighbouring samples (Fig. 6c). This $D_e$ dataset displays a large amount of scatter (relative $D_e$ range = 1.7) and a significant proportion of individual $D_e$ values fall outside the 2σ standardised estimate of the CAM burial dose. The $D_e$ distribution also shows statistically significant negative skewness at 95% CI. and has a higher overdispersion value of 42%. Application of the finite mixture model (FMM) (Galbraith and Green, 1990) using the implementation procedure outlined in the supplementary information of Arnold et al. (2013) confirmed the presence of multiple dose populations in this dataset. The results of the optimum FMM fit (obtained using an overdispersion parameter of 20%, which is consistent with empirical values obtained for the well-bleached, unmixed $D_e$ datasets of samples SH12-1A, 2A and 4A) are summarised in Table 5 and reveals two discrete burial dose populations: a major dose population (containing 85% of $D_e$ values) centred on 844 ± 43 Gy, and a smaller component (containing 15% of $D_e$ values) with a mean dose of 319 ± 46 Gy (Fig. 6c). Various extrinsic (field-related) and/or intrinsic (experimentally derived) sources of $D_e$ scatter could explain the presence of these two discrete dose populations, and these factors are discussed in further detail in the SOM. Taking into consideration the stratigraphic context, sedimentary history and $D_e$ characteristics of SH12-3A (and bracketing samples), we would suggest that spatial variations in the beta dose rate experienced by individual grains could explain the high dose overdispersion of this sample.

Assuming that the two discrete dose components of SH12-3A can be explained by small-scale differences in beta dose rates, it follows that the bulk (sample-average) dose rate will not be representative of that experienced by either dose component during burial. There are differing views on how to derive appropriate age estimates from such complex $D_e$ datasets, namely whether to calculate beta-adjusted dose rates for separate dose components (e.g., Jacobs et al., 2008) or whether to use the sample-averaged beta dose rate in conjunction with a weighted mean $D_e$ estimate for the combined $D_e$ dataset (e.g., Guérin et al., 2013). We have applied these two approaches to the SH12-3A $D_e$ dataset and both yield consistent TT-OSL age estimates. Specifically, the CAM age derived using the approach of Guérin et al. (2013) (396 ± 35 ka) overlaps at 1σ with the FMM age obtained for the main dose component using either the sample-average dose rate (468 ± 39 ka; Table 5) or the beta-adjusted dose rate calculation of Jacobs et al. (2008) (445 ± 38 ka). Deriving component-specific environmental dose rates for the two FMM dose populations remains an open-ended problem in the present context, since there are too many unconstrained parameters to retrospectively reconstruct accurate in situ dose rate geometries for the measured grains. In the absence of a more definitive solution, therefore, we have pragmatically used the CAM $D_e$ and sample-averaged beta dose rate to derive the final age estimate for this sample. The overall consistency of the results obtained using the approaches of Jacobs et al. (2008) and Guérin et al. (2013) provides us with confidence that our final age model selection is reliable within empirical uncertainties.

Palaeomagnetic results and analyses

Most of the clay samples taken from the café con leche and fossiliferous red clay units showed stable behaviour during demagnetisation, leading to an accepted-to-measured specimen ratio of 0.8 (Table 1). Both alternating field (AF) and thermal demagnetisation produced similar results, hence the former was chosen for the majority of the samples collected. Maximum AF peak fields of about 40 mT or temperatures of around 500 °C were high enough to isolate the ChRM directions in the clay samples (Fig. 7a and b). These observations suggest that magnetite is the main carrier of stable remanence in the sediments, consistent with previous results for similar sediments from the Trinchera de Atapuerca sites (e.g., Parés and Pérez-González, 1995, 1999; Parés et al., 2010), and from other cave deposits (e.g., Bosák and Pruner, 2011). The hysteresis curves are also indicative of a low coercivity magnetic phase, and the associated $M_r/M_s$ and $H_{c1}/H_{c2}$ ratios reveal magnetite in the pseudo-single domain (Fig. 8a). The ChRM directions for the café con leche clays, hominin-bearing upper red clays and sterile lower red clays are north directed and have positive inclination, producing Virtual Geomagnetic Pole positions that cluster around the north geomagnetic pole. The normal polarities for these sediments are consistent with the luminescence dating results obtained on the café con leche horizons and fossil evidence for a Middle Pleistocene age for the red clay deposits (Bischoff et al., 1997; Cuenca-Bescós et al., 1997; García et al., 1997; García and Arsuaga, 2011). The palaeomagnetic results therefore confirm that the hominin fossil horizon accumulated during the Brunhes Chron, and is younger than 780 ka.

The ST and SRA flowstone samples were also generally well-behaved upon thermal or alternating field demagnetisation. Typically, a soft component was removed by 10 mT or 150 °C, above which the demagnetisation decayed linearly towards the origin in the Zijderveld diagrams (Fig. 7c). Both thermal and alternating field demagnetisation produced similar ChRM directions. In all cases, the ChRM directions are north directed and have positive inclinations. The palaeomagnetic directions recorded by the LU-10, LU-8 and LU-2 flowstones are all statistically indistinguishable from the expected reference direction for a normal polarity pole during the Pleistocene, which suggests that these speleothems have not been tilted (horizontally or vertically) since their formation and most likely formed in their present locations and orientations. Maximum unblocking temperatures (590 °C), median destructive field, and IRM acquisition curves (Fig. 8b) suggest the presence of magnetite as the main ferromagnetic mineral responsible for the remanence. These trends are consistent with previous results obtained for other flowstones from neighbouring Atapuerca archaeological sites (e.g., Parés and Pérez-González, 1995, 1999; Parés et al., 2010), as well as with speleothem rock-magnetism studies undertaken elsewhere (Latham and Ford, 1993; Perkins and Maher, 1993). These observations provide us with confidence that the magnetisation of the sediments and flowstones is primary rather than secondary, and reflects the polarity of the geomagnetic field at the time of deposition. As such, the palaeomagnetic samples are considered suitable for geochronological purposes. In the case of speleothems LU-10 and LU-8, which overlie the luminescence dated café con leche horizon, the normal polarity results can be confidently interpreted as corresponding to the Brunhes Chron. Ascertaining a corresponding geomagnetic chron for the underlying speleothem (LU-2) at ST is less straightforward. The palaeomagnetic directions for speleothem LU-2 are consistent with formation during the Brunhes Chron, but we cannot rule out the possibility that they may pertain to a subchron (e.g., Jaramillo, Cobb Mountain) or a shorter-lived normal polarity event (e.g., Santa Rosa, Kamikatsu) within the Matuyama Chron.

Discussion

Revised minimum age constraint for the Sima hominin fossils

Both the TT-OSL and pIR-IR$_{25}$ chronological datasets are internally consistent and statistically indistinguishable at 1σ (Tables 4 and 5). Replicate TT-OSL and pIR-IR$_{25}$ ages obtained for individual samples are also in agreement at 1σ (Fig. 2c), lending support to
the reliability of our extended-range luminescence dating approaches at this site. The lack of a distinguishable age-depth profile for the three café con leche samples reveals a relatively rapid accumulation by one or more depositional episodes spanning a period that is shorter than our empirical error ranges. The internal consistency of the two chronological datasets permits calculation of a weighted mean TT-OSL age of 416 ± 19 ka and a weighted mean pIR-IR255 age of 433 ± 15 ka for the café con leche unit. The pooled weighted mean luminescence age for this horizon (n = 6) is therefore 427 ± 12 ka. These results indicate sediment deposition during early Marine Isotope Stage (MIS) 11 or the terminal stages of late MIS 12 (Lisiecki and Raymo, 2005), and provide corresponding minimum age estimates for the underlying red clay archaeological horizon.

There has been some debate in the literature over the thermal stability of the TT-OSL signal and, in particular, whether the lifetime of the TT-OSL trap is sufficient enough to enable reliable age determination over Middle to Early Pleistocene timescales in different thermal environments (e.g., Adamiec et al., 2010; Jacobs et al., 2011; Duller and Wintle, 2012; Thiel et al., 2012). Adamiec et al. (2010) calculated an electron-retention lifetime for the TT-OSL source trap of ~4.5 Ma at a storage temperature of 10 °C. This predicted lifetime is three orders of magnitude lower than for the source trap giving rise to the conventional (fast component) OSL signal and implies that TT-OSL ages may underestimate true burial ages by ~10% for a 1 Ma sample experiencing a long-term burial temperature of 10 °C. However, the uncertainties associated with laboratory derived estimates of trap lifetimes (which are typically not quoted) are likely to be particularly broad owing to the difficulties of accurately and precisely measuring kinetic parameters with standard luminescence measurements (Aitken, 1985; Duller and Wintle, 2012). It is also unclear whether the limited number of laboratory lifetime measurements made so far adequately reflects possible variations in TT-OSL thermal stabilities between samples from different geological provenances, and — in the case of single-grain TT-OSL measurements — whether they are suitably representative of inter-grain variations within a given sample. Indeed, accurate TT-OSL ages have been reported for a range of Early and Middle Pleistocene samples that either have associated independent age control (e.g., Wang et al., 2006a; Sun et al., 2010; Arnold et al., 2013; Pickering et al., 2013) or have comparative conventional luminescence chronologies (e.g., Jacobs et al., 2011; Arnold et al., 2013). These empirical results reveal that, for some localities at least, TT-OSL appear to produce reliable
age estimates over timescales that would otherwise not be anticipated from existing lifetime calculations.

Regardless of discussions about the long-term stability of the TT-OSL signal, the potential implications for our single-grain TT-OSL ages appear to be relatively unimportant. Closed cave chambers such as Sima de los Huesos represent ideal thermal environments for Middle Pleistocene TT-OSL applications because the temperature (and moisture) conditions are naturally regulated and remain effectively constant over diurnal, seasonal and annual timescales. Present-day temperatures inside the deeper chambers of Cueva Mayor are characterised by a stable temperature of \( \sim 10 \, ^\circ\text{C} \) (Martin-Chivelet et al., 2011). If we assume that the laboratory lifetime measurements of Adamiec et al. (2010) are directly applicable to our samples, the expected TT-OSL age underestimation would only amount to \( \sim 4.7\% \) according to the fractional loss equation of Aitken (1985:59). Moreover, this value would likely represent an uppermost estimate of long-term TT-OSL signal loss over the sample burial periods because average cave temperatures across the northern Iberian Peninsula are thought to have been lower during past glacial stages (e.g., Moreno et al., 2010). The difference between the measured weighted mean TT-OSL age for the café con leche sediments (416 ± 19 ka) and the revised age obtained using a signal loss correction factor of 4.7% (436 ± 21 ka) is statistically insignificant at \( \alpha \). Given the similarity of these results, we can be reasonably confident that our existing single-grain TT-OSL ages are not masking any significant systematic biases arising from thermal instability.

In deriving a final TT-OSL age for the café con leche sediments we would favour the original, rather than corrected, weighted mean TT-OSL age of 416 ± 19 ka because there are insufficient constraints on the precision of the published lifetime estimates and on their direct applicability to our single-grain TT-OSL results. It is notable that the replicate pIR-IR225 ages for all three of the café con leche samples are consistent with both the original TT-OSL ages and the revised TT-OSL ages obtained using an additional thermal stability correction. All things being equal, these results would support our earlier contention that an empirical fading correction does not appear to be necessary for the pIR-IR225 signals of these samples. Obtaining replicate ages using two independent luminescence signals therefore provides a useful additional check on dating reliability in this study.

The replicate single-grain TT-OSL ages and pIR-IR225 ages for the capping clay breccia (sample SH12-4A) are indistinguishable, and are also in agreement with the luminescence ages obtained for the underlying café con leche deposits. Taken at face value, the correspondence between the café con leche and the capping clay breccia ages might appear contradictory given that the latter sediments overlie a prominent flowstone that has been previously dated to \( \sim 18-68 \, \text{ka} \) (Bischoff et al., 1997). However, the older than-expected ages for SH12-4A are consistent with the common sedimentological properties of these two yellowish brown clay horizons and the latest lithostratigraphic interpretations, which confirm a shared origin for the two deposits. As detailed in Arsuaga et al. (submitted), the capping clay breccia is not thought to represent a new infilling event of allochthonous sediments sourced directly from outside the cave. Instead, the matrix originates from internal reworking of café con leche sediments that had previously accumulated as chimney 'plugs' during the original high-energy debris flow event (i.e., LU-7). Sometime after the formation of the colada flowstone at \( \sim 18-68 \, \text{ka} \), these café con leche chimney accumulations fell (by gravity) onto the floor of the Sima ramp, where they were consolidated to form the matrix of the capping clay breccia. The age equivalence obtained in the present study for the main café con leche horizon and the capping clay breccia would appear to confirm these sedimentological interpretations. The agreement between the replicate TT-OSL and pIR-IR225 ages for SH12-4A is also consistent with the suggestion that the capping clay breccia was re-transported and re-deposited within a closed cave system without being subjected to any daylight exposure. If this was not the case, the two luminescence dating signals would likely display markedly different residual \( \Delta \theta \) values on account of their distinctly different bleaching sensitivities. In effect, sample SH12-4A represents a homogeneously 'unbleached' sample; i.e., the burial event recorded by its natural luminescence signals pertains to the penultimate, rather than the most recent, erosion—transportation—deposition cycle. Sample SH12-4A serves as a useful example of how an understanding of site accumulation processes can be critical for accurately interpreting luminescence dating datasets in complex cave environments.

![Figure 8](image-url)
Contextualisation and implications of the new chronologies

The bracketing ages obtained in this study using luminescence dating and palaeomagnetism constrain the hominin-bearing sediments at Sima to between 427 ± 12 and ~780 ka. To place these new ages into a broader chronological context, we have summarised in Table 6 the reliable numerical and correlative ages published thus far for Sima de los Huesos. Two chronological techniques have previously been applied to the sedimentary facies lying below the capping colada flowstone, namely: (i) U-series dating of the ST-b (lower SRA-3) and ST-a (upper SRA-3) speleothems at exposure ST (Fig. 2) (Bischoff et al., 2003, 2007); and (ii) palaeomagnetic dating of the fossiliferous red clays, sterile red clays and basal sands and silts at SH (Area A), SH (Area B) and SRB (Parés et al., 2000). Generally speaking, the macro- and micro-faunal records of the café con leche and upper red clay breccia horizons do not permit well-constrained, unequivocal age assignment beyond that of the broader early or mid-Middle Pleistocene (Cuenca-Bescós et al., 1997; García et al., 1997, García and Arsuaga, 2011); hence we have not included biochronological indicators in this age comparison.

The published U-series ages for the SRA-3 speleothem were originally presented as minimum age constraint on the underlying archaeological horizon. However, as already mentioned, recent excavations have revealed that this previous interpretation was incorrect and that the lower part of SRA-3 (ST-b) represents a significantly older generation of flowstone (LU-2) that actually underlies the lower red clays and the basal sands and silts (Arsuaga et al., submitted). Uranium series ages of 153 ± 5 ka to 281±23 ka were obtained on the upper part of SRA-3 using thermal-ionisation mass-spectrometry (TIMS) (Bischoff et al., 2003) (Table 6). The ages for this speleothem, which corresponds to flowstone LU-8 at SRA and overlies the café con leche horizon, are in good agreement with (i.e., younger than) the luminescence ages obtained in this study. A non-finite age of >350 ka was originally obtained for the lower part of SRA-3 speleothem using TIMS (Bischoff et al., 2003). This age estimate was subsequently revised upwards to a minimum age of

<table>
<thead>
<tr>
<th>Deposit</th>
<th>Lithostratigraphic unit</th>
<th>Section</th>
<th>Dating method</th>
<th>Age (ka)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Re-worked café con leche</td>
<td>LU-11</td>
<td>SRM</td>
<td>Single-grain TT-OSL</td>
<td>426 ± 39</td>
<td>This study</td>
</tr>
<tr>
<td>Clay breccia (bat guano)</td>
<td>SRM</td>
<td>Multi-grain pTR-IRSL</td>
<td>417 ± 27</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>Capping speleothem (Colada)</td>
<td>LU-10</td>
<td>SRM</td>
<td>Conventional 14C</td>
<td>17.8 ± 0.1</td>
<td>Bischoff et al., 1997</td>
</tr>
<tr>
<td>Colada — embedded fox bones</td>
<td>LU-10</td>
<td>SRM</td>
<td>Palaeomagnetism</td>
<td>&lt;780 (n = 20)</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Conventional 14C (collagen)</td>
<td>25.1 ± 0.2</td>
<td>Bischoff et al., 1997</td>
</tr>
<tr>
<td>CVP flowstone</td>
<td>LU-10</td>
<td>SRB (R16)</td>
<td>Conventional 14C</td>
<td>68 ± 6 (n = 6)</td>
<td>Bischoff et al., 1997</td>
</tr>
<tr>
<td>Speleothem interbedded in café con leche horizon</td>
<td>LU-8</td>
<td>SRA</td>
<td>Palaeomagnetism</td>
<td>&lt;780 (n = 3)</td>
<td>Bischoff et al., 1997</td>
</tr>
<tr>
<td>Speleothem ST-c (upper portion of SRA-3)</td>
<td>LU-8</td>
<td>ST</td>
<td>Palaeomagnetism</td>
<td>153 ± 3</td>
<td>Bischoff et al., 2003</td>
</tr>
<tr>
<td>Café con leche clay breccia</td>
<td>LU-8</td>
<td>ST</td>
<td>Palaeomagnetism</td>
<td>&lt;780 (n = 1)</td>
<td>This study</td>
</tr>
<tr>
<td>Fossiliferous upper red clays</td>
<td>LU-6</td>
<td>SRT(K22)</td>
<td>Palaeomagnetism</td>
<td>&lt;780 (n = 7)</td>
<td>This study</td>
</tr>
<tr>
<td>Gradational upper-lower red clays</td>
<td>SH(R11 &amp; 16)</td>
<td>SRA (Area B)</td>
<td>Palaeomagnetism</td>
<td>&lt;780 (n = 12)</td>
<td>Parés et al., 2000</td>
</tr>
<tr>
<td>Sterile lower red clays</td>
<td>SH(T16)</td>
<td>SRA</td>
<td>Palaeomagnetism</td>
<td>&lt;780 (n = 6)</td>
<td>Parés et al., 2000</td>
</tr>
<tr>
<td>Sterile sands and silts</td>
<td>SH(Area A)</td>
<td>Palaeomagnetism</td>
<td>&gt;780 (N = 5)</td>
<td>Parés et al., 2000</td>
<td></td>
</tr>
<tr>
<td>Speleothem ST-b (lower portion of SRA-3)</td>
<td>LU-2</td>
<td>SRB, SRM</td>
<td>Multi-grain pTR-IRSL</td>
<td>430 ± 25</td>
<td>Bischoff et al., 2003</td>
</tr>
<tr>
<td></td>
<td>LU-2</td>
<td>SRB, SRM</td>
<td>Single-grain TT-OSL</td>
<td>441 ± 25</td>
<td>Bischoff et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SRM</td>
<td>Palaeomagnetism</td>
<td>&lt;780 (n = 4)</td>
<td>Parés et al., 2000</td>
</tr>
<tr>
<td></td>
<td>LU-2</td>
<td>ST</td>
<td>Palaeomagnetism</td>
<td>&gt;350 (n = 9)</td>
<td>Bischoff et al., 2007</td>
</tr>
<tr>
<td>Speleothem ST-a</td>
<td>LU-2</td>
<td>ST</td>
<td>Palaeomagnetism</td>
<td>&gt;780 (n = 6)</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>LU-2</td>
<td>ST</td>
<td>Palaeomagnetism</td>
<td>&gt;780 (n = 2)</td>
<td>This study</td>
</tr>
</tbody>
</table>

The ages are presented in order of their relative stratigraphic positions in the latest lithostratigraphic scheme (Fig. 2c). Published ages that were not considered to be methodologically reliable by the original authors have been omitted from this synthesis, as have ages that were presented without associated error ranges, supporting methodological descriptions, or specific details of sampling positions.
>530 (600 \pm 60) ka following reanalysis of the same material using multi-collector inductively-coupled plasma mass-spectrometry (MC ICP-MS) (Bischoff et al., 2007) (Table 6). The latest U-series ages for the lower part of SRA-3, which range from >571 ka to >571 ka, are in stratigraphic agreement with the luminescence ages obtained for the overlying cave leche horizon. Interpretation of these U-series ages as a reliable maximum age for the hominin fossils is complicated by the close proximity of the $^{230}$Th/$^{234}$U ratios to internal isotopic equilibrium, as is apparent from the large or non-finite upper error bounds. Since the existing U-series ages for the lower part of SRA-3 are effectively minimum ages, they do not provide any additional constraint on whether the normal polarity interval reported here for the LU-2 speleothem corresponds to the Brunhes Chron or to a shorter-lived intra-Matuyama polarity subchron.

The palaeomagnetic results obtained in the present study for 70 sediment samples complement those published previously at Sima de los Huesos (Parés et al., 2000) and help to extend the spatial coverage of the original palaeomagnetic dataset to include the cave leche horizon and red clay horizons at several new exposures. Parés et al. (2000) reported a dominant normal polarity magnetisation for the upper red clays in Sima areas A and B, consistent with our results for this horizon. However, that study also reported traces of reversed magnetisation in some of the upper red clay samples, which suggested that additional magnetic acquisition could have taken place during one of several short-lived polarity subchrons within the Brunhes Chron (e.g., Channell et al., 2012). We do not see any traces of reversed polarity in our upper red clay samples, reinforcing the original interpretation of Parés et al. (2000) that the short duration of this reversal event (excursion or cryptochron), combined with the lower intensity of the geomagnetic field during the reversal phase, and prolonged diagenesis of the clays thereafter, are likely to have resulted in the preservation of reversed magnetisation in some samples but not in others. Moreover, the existence of localised areas of high porosity and concentrated water drainage in the Sima chamber would have also acted to restrict the zones displaying both normal and reversed magnetisation traces in the red clay horizon.

Four samples collected by Parés et al. (2000) towards the base of the non-fossiliferous lower red clays at exposure SRB produced stable, south and upwards (reversed polarity) magnetisation directions (Table 6). Palaeomagnetic results obtained in the present study from a higher stratigraphic position within the same unit (i.e., closer to the contact with the upper red clays) reveal exclusively normal polarities. There are several possible explanations for the apparent differences between these two sets of results: (i) spatial variations in the preserved dominant polarity related to locally variable drainage and porosity conditions in the Sima chamber, as detailed above; (ii) the basal part of the sterile lower red clays was deposited during the Matuyama Chron, i.e., before 780 ka, whereas the upper part of this horizon was deposited during the Brunhes Chron (with or without an intervening hiatus); or (iii) the sterile lower red clay horizon was deposited during the Middle Pleistocene and records an intra-Brunhes reverse polarity event towards its base. Parés et al. (2000) also reported that the underlying sand and silt horizon preserved in Area A of the Sima chamber display a dominant reversed magnetisation (Table 6). The most conservative interpretation is that these sands and silts were deposited during the Matuyama Chron >780 ka, since (i) the reversal is continuous and spans a significant vertical coverage of sediment accumulation (>30 cm); and (ii) this unit is sedimentologically similar to a large sand body preserved in the nearby Sala de los Ciclopes, which is also of Matuyama age (Parés et al., 2010). Nevertheless, direct numerical dating of the lower red clays and the basal sand and silt deposits is required before this assertion can be confirmed.

In sum, the combined palaeomagnetic results from both the Parés et al. (2000) study and the current work provide a statistically significant sample dataset (11 exposures, 135 specimens) and confirm that the hominin-bearing red clays were deposited during the Brunhes Chron. In the absence of additional radiometric age control, however, it remains difficult to assign a definitive geomagnetic polarity chron to the underlying sediment horizons (lower red clays, sands and silts, white marls) and flowstones (LU-2) (Fig. 2c).

The revised stratigraphic interpretations of the existing U-series ages (Bischoff et al., 2007) means that our new luminescence chronologies provide the first unequivocal minimum age control for the archaeological red clay horizon and the Sima hominin fossils. The stratigraphic relation between the luminescence ages and the archaeological horizon is well-constrained, since the samples were collected from spatially continuous sedimentary infills with a clear and direct association with the underlying archaeological clays. Our luminescence dating results for the cave leche horizon are both reproducible and precise, and are based on two consistent luminescence datasets from different types of mineral grains. The luminescence ages are also in good agreement with bracketing U-series ages for the upper part of SRA-3 (LU-8) and lower part of SRA-3 (LU-2). The robustness of the new ages for the archaeological horizon permits a more secure age assignment for the Sima hominin fossils, and enables more reliable comparisons with other pencontemporaneous ancestral Neanderthal remains from across Europe. In particular, the new minimum age of 427 ± 12 ka for the Sima fossils is younger than the 609 ± 40 ka age obtained for the H. heidelbergensis holotype at Mauer, Germany, using infrared radiofluorescence dating of bracketing sediments and combined ESR-U-series dating of herbivore teeth (Wagner et al., 2010). The difference between these two chronologies could be interpreted as being consistent with the more Neanderthal-derived morphologies (Arsuaga et al., 1991, 1993, 1997b) and dentition (Martínón-Torres et al., 2012) of the Sima fossils. However, our luminescence dating results provide minimum age constraint for the Sima fossils, and therefore do not necessarily preclude a broadly contemporaneous age for the two assemblages. The new age for the Sima fossils is also consistent with U-series ages of >350 ka obtained for speleothems overlying H. heidelbergensis cranial and mandibular remains at the cave site of Arago in France (Falguères et al., 2004). Nonetheless, it is worth noting that the Arago ages are limited by internal isotopic equilibrium and they do not directly bracket the archaeological horizon, hence representing non-finite (minimum) ages for the associated H. heidelbergensis fossils. Subsequent attempts to refine the Arago age control using combined ESR-U-series dating of herbivore teeth from the main archaeological horizon have unfortunately proved problematic due to post-depositional geochemical alterations of the fossil-bearing sediments (Han et al., 2010).

The luminescence ages for the Sima fossils are slightly older than the $^{40}$Ar/$^{39}$Ar age of 353 ± 4 ka obtained on a volcanoclastic layer found immediately above the Ceprano cranium (Nomade et al., 2011), which has been variously classified as a morphologically primitive example of H. heidelbergensis (e.g., Mounier et al., 2011) or a separate lineage that co-existed with H. heidelbergensis during the Middle Pleistocene (e.g., Manzi et al., 2001; Mallegnì et al., 2003). Adding to the taxonomically diverse mosaic of Middle Pleistocene European hominin records is a recently dated mandible from Mala Balanica Cave, southern Serbia, which appears to lack derived Neanderthal features (Rakočević et al., 2011; Roksandić et al., 2011, Rink et al., 2013). Minimum ages of between 397 and 525 ka have been obtained for this hominin mandible using luminescence (pIR-IR$^{25}$) dating of overlying sediments and combined ESR-U-series dating of teeth found above the
hominin remains. The stratigraphic relationship between these dated samples and the associated mandible is comparable with that of our luminescence dating samples and the underlying Sima fossils. The minimum age estimate for the Balanica mandible overlaps with the new minimum ages for the Sima remains, indicating the potential co-existence of two distinct hominin lineages within different regions of Europe during the mid-Middle Pleistocene. We stress, however, that the Sima hominin morphology has been described for all skeletal parts, whereas the Balanica partial mandible is an incomplete fossil and should not be assigned the same taxonomic/phylogenetic value. Middle Pleistocene populations could have exhibited different combinations of primitive traits (for the modern human/Neanderthal clade) and different types of Neanderthal specialisations. The possibility remains, therefore, that the Balanica individual could have displayed Neanderthal autopomorphies in skeletal regions that were not preserved in the incomplete fossil record.

It is also interesting to compare the new ages for the Sima assemblage with the preserved archaeological history of other related sites within the Sierra de Atapuerca endokarstic complex. Of particular interest is the Middle Pleistocene lithostratigraphic sequence at Gran Dolina (e.g., Carbonell et al., 1995; Bermúdez de Castro et al., 1997), which comprises five sedimentary units (TD7-11) associated with age control from combined ESR / U-series dating of ungluate teeth (Falguères et al., 1999) and IRS and thermoluminescence (TL) dating of polymineral sediments (Berger et al., 2008). Though lacking in hominin remains, TD10 preserves an extensive record of human activity beginning at ~430 ± 59 ka (Berger et al., 2008) and continuing until ~337 ± 29 ka (Falguères et al., 1999) or 240 ± 44 ka (Berger et al., 2008). This important phase of hominin occupation at Gran Dolina is preceded by an apparent hiatus in the archaeological record from at least 780 ka (TD7) to 480 ± 130 ka (TD9) (Berger et al., 2008), despite a semi-continuous sedimentary sequence being recorded during this time period. Our luminescence ages for the café con leche horizon at Sima de los Huesos are consistent with the combined ESR/U-series and TL ages obtained on the lower-middle horizons of TD10, suggesting that the two sedimentary fills were deposited broadly synchronously. The coeval nature of these two Middle Pleistocene deposits, and that of the fossiliferous red clay horizon at Sima de los Huesos, is also supported by their shared faunal and lithic records (e.g., Cuenca-Bescós et al., 2010; Rodríguez et al., 2011; Olé et al., 2013). The similarities in the preserved lithostratigraphic record at Gran Dolina and apparent hiatus in the Middle Pleistocene archaeological record prior to ~480 ka might suggest that the Sima hominin remains and the café con leche deposits were deposited in close temporal association, or even quasi-contemporaneously. In this sense, our minimum age estimates for the Sima fossils may approximate the actual depositional ages of the hominin remains. In the absence of more direct evidence, however, we maintain that the luminescence dating results should be considered as minimum ages for the Sima fossils. Further radiometric dating of the upper sedimentary layers at the nearby site of Sima de la Elefante is currently in progress and might help shed further light on potentially correlative occupation and sedimentary histories of the Atapuerca endokarst during the mid-Middle Pleistocene. Direct luminescence dating of the fossil-bearing red clays at Sima remains a challenge from a methodological perspective. However, successful isolation of suitably bleached silicate grains from this unit would also enable us to tighten the current maximum age limits on the Sima fossils.

The previously published minimum age estimate of >530 ka for Sima de los Huesos (Bischoff et al., 2007) received considerable attention vis-à-vis its perceived incompatibility with morphological and genetic evidence for hominin evolutionary divergences during the Middle Pleistocene (e.g., Endicott et al., 2010; Stringer, 2012). The revised stratigraphic interpretations of this U-series age presented earlier (and further detailed in Arsuaga et al., 2010) effectively negates any direct comparisons with existing genetic divergence estimates, since these U-series results are now known to provide a non-finite maximum age estimate on the overlying fossils. However, it is worth considering whether our new luminescence ages fit with the latest phylogenetic and de novo predictions of hominin evolution. Molecular phylogenetic reconstructions of a population split between H. neanderthalensis and H. sapiens based on nuclear DNA sequencing were previously placed at between 270 and 440 ka (Green et al., 2010). However, a recent revision of the nuclear mutation rate based on updated generation lengths for apes and improved generational mutation rates for modern humans (Langergraber et al., 2012) suggests that previous H. neanderthalensis–H. sapiens divergence estimates should be pushed back to somewhere on the order of 400–600 ka (Sclay and Dubin, 2012) or even 423–781 ka (Langergraber et al., 2012). Phylogenetic estimates of the most recent common ancestor (MRCA) for H. neanderthalensis and H. sapiens lineages using complete mitochondrial DNA sequencing (Endicott et al., 2010) range from 315 to 506 ka and 338 to 538 ka assuming prior Homo-Pan calibration splits at 6–7 Ma and 6.5–7.5 Ma, respectively. The new minimum ages for the Sima hominins, which were ancestral to Neandertals and display clear Neanderthal synapomorphies, are compatible with (i.e., potentially post-date) the updated population divergence and MRCA estimates of sometime between 315 to 781 ka for the H. neanderthalensis and H. sapiens lineages. Considerable debate remains, however, over the suitability of existing taxonomic classifications, the nature of phylogenetic relationships and the spatio-temporal evolutionary histories of the Middle Pleistocene hominin record (e.g., Arsuaga et al., 1997b; Hublin, 2009; Endicott et al., 2010; Dennell et al., 2011; Stringer, 2012; Bermúdez de Castro and Martín-Torres, 2013).

It is evident from the literature that relatively few of the Middle Pleistocene hominin remains across Europe have reliable numerical age control based on modern and widely accepted dating techniques. Out of the Middle Pleistocene sites that have been variously reported to contain H. heidelbergensis or ancestral Neanderthal assemblages, the list of undisputed, radiometrically dated localities is perhaps only limited to Mauer (Wagner et al., 2010), Sima de los Huesos (this study), Arago (Falguères et al., 2004) and Ceprano (Nomade et al., 2011). From a geochronological perspective, appropriate caution must be exercised when making temporal comparisons with other Middle Pleistocene sites that have been dated using relative or correlational techniques (e.g., Roberts et al., 1994; Stringer and Hublin, 1999), or that have been dated using methodologies with potentially low reliability by present-day standards. Accurate interpretation of published ages for Middle Pleistocene hominin sites also requires due consideration of associated contextual information and chronological caveats, particularly for results that provide terminus ante quem (e.g., Falguères et al., 2004) or terminus post quem age constraints on associated fossils, and results that are quoted as non-finite or even tentative ages due to methodological restrictions (e.g., Bischoff et al., 2003; Han et al., 2010). Overlooking this essential quality control information is likely to confuse rather than clarify the complex relationships in the Middle Pleistocene fossil record.

The paucity of wider numerical age control for key Middle Pleistocene fossil sites largely reflects the historical difficulties of dating this troublesome time period using traditional geochronological techniques. Recent advances in techniques such as uranium-lead dating (U–Pb) (e.g., Woodhead et al., 2006), cosmogenic burial dating (e.g., Carbonell, 2008; Balco and Shuster, 2009) and extended-range luminescence dating (e.g., Buylaert et al., 2012;
Duller and Wintle, 2012) offer a new set of geochronological tools for redressing previous methodological limitations. The development of more widespread and reliable age control is critical for unravelling the evolutionary histories and phylogenetic relationships of the European Middle Pleistocene record. The new chronologies presented in this study for Sima de los Huesos demonstrate the versatility of TT-OSL and pIR-IR dating methods and highlight the potential role they could play in this endeavour.

Conclusions

To address uncertainties in existing chronological control at Sima de los Huesos, we have obtained new minimum age estimates for the hominin remains using two semi-independent extended-range luminescence dating techniques. From this latest geochronological study, we can constrain the age of the café con leche clay breccia horizon to early MIS 11 or terminal MIS 12 and we are able to provide a combined minimum age estimate of $427 \pm 12$ ka for the underlying red clay breccia and its associated hominin remains. Our luminescence ages are internally consistent and support previous sedimentological interpretations that the café con leche horizon was deposited during a relatively short-lived, high-energy flooding episode. The very small natural residual doses (~7 Gy) measured for a modern analogue provide confidence that these sediments were adequately bleached by prolonged sunlight exposure prior to deposition inside the cave system. Our novel single-grain TT-OSL approach has also provided additional insight into dating suitability at this potentially complex cave site, by enabling us to explicitly assess the extent of signal bleaching in the absence of multi-grain averaging influences. The consistency and reproducibility of our single-grain TT-OSL ages support the recent findings of Arnold et al. (2013) and demonstrate the broader potential of this approach for dating Middle Pleistocene deposits that contain suitably bright quartz grains.

New palaeomagnetic analyses undertaken at exposures within the main Sima chamber and along the Sima ramp reveal that the café con leche and upper red clay breccia were deposited during the Brunhes Chron, confirming the previous results of Parés et al. (2000). These correlative dating results provide a maximum age constraint of ~780 ka for the associated hominin fossils at Sima de los Huesos. Our new ages provide an important piece of the puzzle for reconstructing the complex history of the Sima hominin assemblage, and are compatible with the latest genetic predictions for a H. neanderthalensis–H. sapiens divergence sometime before ~315 ka and no earlier than ~781 ka. However, there remains a clear need for more widespread and improved numerical chronology at other penecontemporaneous sites to better resolve the evolutionary history and phylogenetic relationships of the European Middle Pleistocene hominin fossil record.

Acknowledgements

Financial support for this research was provided by grant CGL2010-16821 awarded to J.M.P., L.A. and M.D. from the Ministerio de Economía y Competitividad de España (Plan Nacional de I + D + i 2008–2011), by project GL2012-38434-C03 awarded to J.L.A. and A.A., and by a Marie Curie International Reintegration Grant (PIRG08-GA-2010-276810) awarded to M.D. within the 7th European Community Framework Programme. L.A. and M.D. thank Carlos Pérez Garrido for his assistance with preparing and measuring the luminescence dating samples. We also thank the multidisciplinary research team of the Sierra de Atapuerca (EIA) for scientific and logistic support. Site access and permission to collect samples at Atapuerca was granted by the Junta de Castilla y León. L.A. and M.D. are grateful to Sébastien Huot (Université du Québec à Montréal) for providing an Excel package to calculate K-feldspar fading rates. C. Mac Niocaill and an anonymous reviewer are thanked for their constructive comments on this paper.

Appendix A. Supplementary material

Supplementary material related to this article can be found at http://dx.doi.org/10.1016/j.jhevol.2013.12.001.

References

grains from the Palaeolithic site of Bonneval (Eure-et-Loir, France). Quatern. Geochronol. 5, 342–347.