Chronological and post-depositional insights from single-grain IRSL dating of a Palaeolithic sequence at Stelida, Naxos (Greece)

Ninon Taffin [a,*], Christelle Lahaye [a], Daniel A. Contreras [b], Justin A. Holcomb [c,d], Danica D. Mihailović [e,f], Panagiotis Karkanas [d], Guillaume Guérin [g], Demetris Athanasoulis [h], Tristan Carter [i,*]

[a] Archéosciences Bordeaux, UMR 6034–CNRS/Université Bordeaux Montaigne, 33600 Pessac Cedex, France
[b] Department of Anthropology, University of Florida, Gainesville, FL 32611, USA
[c] Kansas Geological Survey, University of Kansas, Lawrence, KS 66047-3724, USA
[d] Malcolm H. Wiener Laboratory of Archaeological Science, American School of Classical Studies at Athens, Athens 10676, Greece
[e] Institute of Archaeology, Serbian Academy for Science and Arts, 11000 Belgrade, Serbia
[f] Department of Archaeology, University of Belgrade, 11000 Belgrade, Serbia
[g] Géosciences Rennes, UMR 6118, CNRS, Univ. Rennes, 35000 Rennes, France
[h] Ephorate of Antiquities, Greek Ministry of Culture, Athens 10555, Greece
[i] Department of Anthropology / School of Geography and Earth Sciences, McMaster University, Hamilton L8S 4L9, Canada

ARTICLE INFO

ABSTRACT

Since 2015, the *Stelida Naxos Archaeological Project* (SNAP) has excavated a prehistoric site on what today is the northwest coast of Naxos, the largest island of the Cycladic archipelago in the southern Aegean Sea (Greece). Survey and excavations at the site have produced artefacts spanning the Lower Palaeolithic through the Meso- lithic periods based on their techno-typological attributes. These discoveries suggest that exploitation of Stelida began as early as the Middle Pleistocene, challenging the long-standing model that the Cyclades were not inhabited until the Early Holocene. Due to the site's likely temporal depth and the lack of preserved organics, luminescence dating is the most appropriate method to scientifically date this activity. However, luminescence dating in this context is complicated by the site's complex hillslope formation processes. Experiments upon the Stelida sediments have demonstrated a lack of luminescence sensitivity of quartz at the site. To evaluate the potential for post-depositional mixing of previously acquired dates yielded from a stratigraphic sequence first published in 2019, as well as to evaluate the efficacy of multiple luminescence dating models, we measured and compared different infrared stimulated luminescence [IRSL] measurements on K-feldspars, with IR_{50} and $pIRIR_{290}$ multi-grain and $pIRIR_{290}$ single-grain signals. The single-grain results confirm the multi-grain results and provide additional and more precise information on the site's depositional and post-depositional events. The results of each approach demonstrate that feldspars were well-bleached, suggesting that in hillslope settings where quartz grains prove difficult to date, IR_{50} and $pIRIR_{290}$ multi-grain, and $pIRIR_{290}$ single-grain signals of feldspars can be used to achieve reliable results. Finally, when considered alongside field and laboratory ob- servations of site stratigraphy, these results suggest that colluvial and aeolian (windblown) deposits at Stelida retain a degree of stratigraphic integrity characterized by minimal post-depositional alteration following their most recent deposition. These support previous estimates of the deposition at the site, the new earliest deter- mination being 233 – 217 thousand years ago [ka], compared to the date of 198.4 ± 14.5 ka published in 2019. These dates represent the earliest – indirect – evidence for open sea crossings in the northern hemisphere, though it remains uncertain as to which species of the genus *Homo* was responsible for such maritime activity. These results also have implications for the preservation potential of similar deposits across the hillslope, as well as deposits preserved in similar geomorphic settings in Mediterranean landscapes.

* Corresponding authors.
E-mail addresses: ninon.taffin@u-bordeaux-montaigne.fr (N. Taffin), stringy@mcmaster.ca (T. Carter).

1. Introduction

Until relatively recently, the earliest evidence for the human occupation in the Cycladic islands of the southern Aegean was limited to the Early Holocene, as evidenced by a ca. 10,500-year-old Mesolithic village on the island of Kythnos (Sampson et al., 2010), despite hominin activity on the neighboring continental masses of Anatolia and Greece extending back to the Early and Middle Pleistocene respectively (Dinçer, 2016; Tourloukis and Harvati, 2018). This late settlement of the Cyclades seemed to fit a pan-Mediterranean (Cherry, 1981) and larger global pattern of island colonization as a Late Pleistocene–Holocene phenomenon associated with *Homo sapiens,* whereby seafaring has been considered an index of behavioral modernity (Gamble, 2013). Over the past decade, there have been several challenges to this late model of insular Aegean colonization, with new evidence potentially extending island occupation back to the Middle Pleistocene, if not earlier (Runnels et al., 2014). Unfortunately, there are only a few such cases supported by scientific dating (Galanidou et al., 2016; Strasser et al., 2011), with Pleistocene sea-level reconstructions suggesting that many of these now insular sites were joined to neighboring continents during large parts of the Palaeolithic (Lykousis, 2009; Tourloukis and Karkanas, 2012; Sakellariou and Galanidou, 2017).

This paper focuses on one of the key sites in this debate: Stelida on what today is the northwest coast of the Cycladic Island of Naxos (Figs. 1-2), where excavations since 2015 have produced deep stratigraphic sequences whose associated material culture suggests activity stretching back to the Lower – Middle Palaeolithic (Carter et al., 2014,

2017). Previous infrared stimulated luminescence [IRSL] multi-grain dates of one completed excavation sounding – trench DG-A/001 – demonstrated that its sequence of hillslope (colluvial) deposits contained stone tools that had been deposited as early as the Middle Pleistocene, ~200,000 years ago (Carter et al., 2019). However, these colluvial deposits pose challenges for luminescence dating due to the possibility of post-depositional mixing of sediments (Fuchs and Lang, 2009). To evaluate the possibility of such mixing we have re-evaluated the previously dated samples, this time using single-grain measurements. Comparing these results with the original multi-grain measurements using Central Age (Galbraith et al., 1999), Average Dose (Guérin et al., 2017), and Finite Mixture (Roberts et al., 2000) models of grain equivalent doses [CAM, ADM, FMM], we conclude that the new study's results (1) broadly support previous age estimations determined by multi-grain analysis, (2) suggest that the single-grain results account for some minor effects of incomplete bleaching, and (3) provide a means of evaluating the risk that the multi-grain results were affected by post-depositional mixing. Taken together, the results demonstrate that although one date is slightly younger than first reported, post-depositional mixing has not compromised the samples. As such, they argue for the reliability of luminescence dating of these colluvial deposits, refining the original chronology and making it more robust. These data have implications for future research seeking to apply luminescence dating techniques in colluvial settings throughout the Mediterranean and beyond. Moreover, these dates, in concert with new palaeogeographic reconstructions of the Pleistocene Aegean (Ferentinos et al., 2023), suggest that the earliest deposits from Stelida now provide
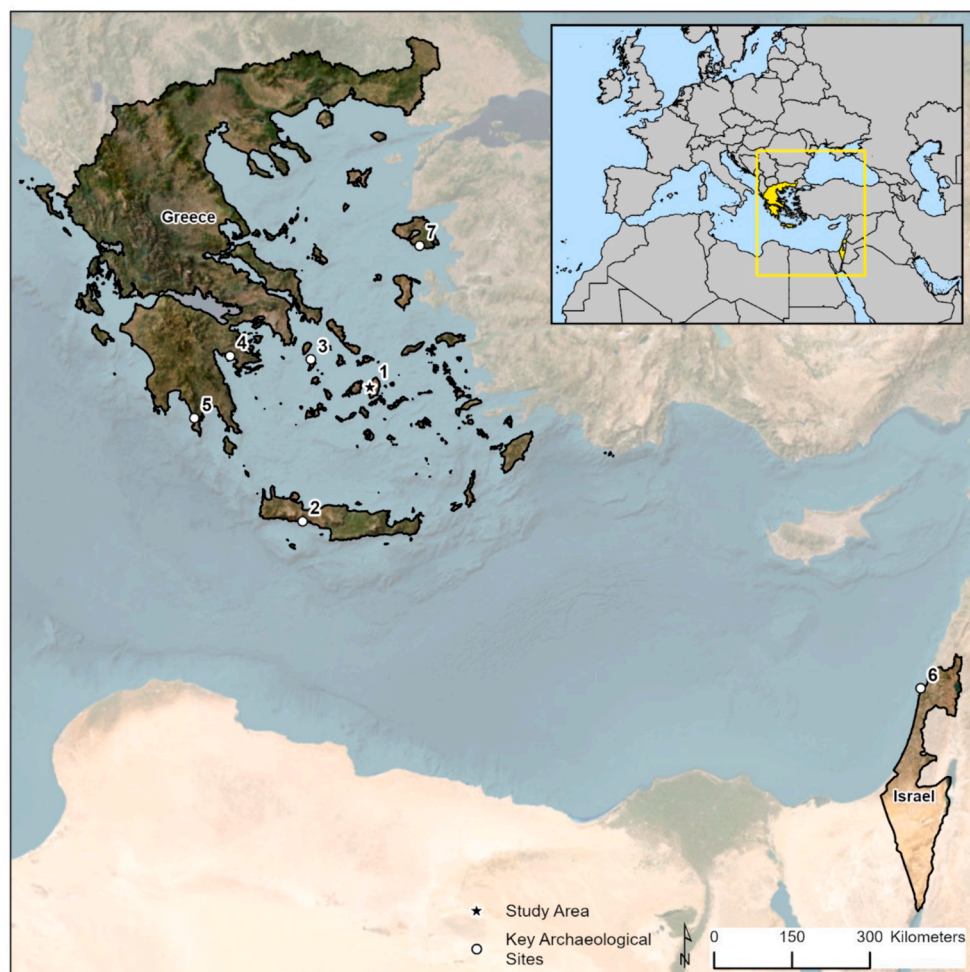

Fig. 1. Location of study area and key archaeological sites referenced in text: (1) Stelida, Naxos, (2) Plakias, Crete, (3) Maroulas, Kythnos, (4) Franchthi Cave, Argolid, (5) Apidima Cave, Mani, (6) Misliya Cave, Israel, (7) Rodafnidia, Lesvos (map by K. Andrzejewski).
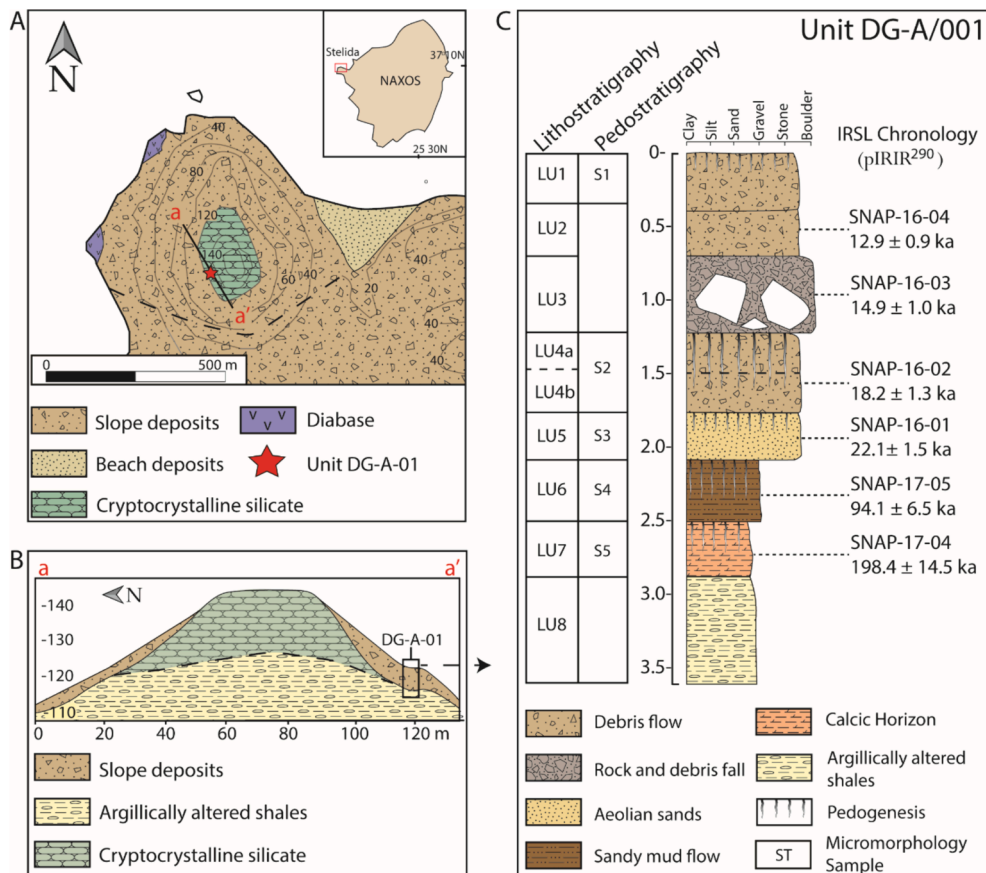
Fig. 2. Geoarchaeological framework, stratigraphic interpretation of the Stelida hillslope and excavation unit DG-A/001, plus the original luminescence dates which are now superseded by this study. (A) Generalized plan view of key geomorphic units observed on Stelida hillslope and location of Unit DG-A/001. (B) Generalized profile of cross-section a-a′ illustrating the upper half of the Stelida hillslope. (C) Stratigraphic profile, geoarchaeological interpretation, and geochronology of unit DG-A/001 with dates expressed as 68% confidence intervals (reproduced from Carter et al., 2019).

the earliest – indirect – evidence for open sea crossings in the northern hemisphere (see Gaffney 2021), though other work published since our first chronology paper have served to muddy the waters as to which human species this involved (Harvati et al., 2019).

2. Materials and methods

2.1. Archaeological remains found at the Stelida site

The double-peaked (151 masl) hill of Stelida is today located on the northwest coast of Naxos, the largest island of the Cycladic archipelago in the central Aegean Basin, Greece (Figs. 1-2). The hill consists of an uplifted outcrop of sediments silicified by hydrothermal alteration overlying Miocene shales (Skarpelis et al., 2017). Those silicified sediments ("chert") comprise the raw material that was exploited in prehistory for toolmaking, as evidenced by the hundreds of thousands of flaked artefacts that litter the site (Carter et al., 2016, 2017). The archaeological site was discovered in 1981 (Séfériadès, 1983), with the *Stelida Naxos Archaeological Project* (SNAP) initiated in 2013 to detail the site's archaeology, geology, and chronology. The techno-typological characteristics of the artefacts recovered from the 2013–14 survey suggested that the chert source had been exploited from the Lower Palaeolithic, through the Middle and Upper Palaeolithic and into the Mesolithic (Carter et al., 2014, 2016).

Ultimately, the project was established with the aim of contributing to a major debate on the alleged seagoing capabilities of archaic humans in the Aegean, a hypothesis that countered the received wisdom that insular activity was (a) began in the Late Pleistocene (e.g., Perlès, 1987), and (b) was exclusively associated with *Homo sapiens* (Cherry, 1981;

Gamble, 2013). This debate had been initiated by an American-Greek team surveying in the Plakias area (Fig. 1) of southwest Crete (Strasser et al, 2010), which had recovered surface material of apparent Lower and Middle Palaeolithic date (see also Kopaka and Matzanas, 2009; Mortensen 2008), eroding from deposits subsequently shown to be at least 100,000 – 130,000 years old (Strasser et al., 2011). Given that Crete had been insular throughout the Pleistocene (Lykousis, 2009), and that *Homo sapiens* was believed to have only arrived in the Aegean region around 40,000 years ago (Douka et al., 2011), it was argued that the early visitation of Crete involved seafaring pre-*sapiens* populations (Runnels et al., 2014; Strasser et al., 2010). A larger review of claimed early insular material from elsewhere in the Aegean further contributed to the hypothesis that early hominin dispersals may not have been restricted to terrestrial routes as previously believed (Runnels et al., 2014). These papers prompted much debate, with many not fully convinced by the claims being made, due variously to concerns that (i) this was surface material, rather than from a excavation, (ii) the Plakias quartz 'tools' were natural, not cultural, (iii) the material need not be Acheulean (Lower Palaeolithic), but instead might be a North African Middle Stone Age type associated with early *Homo sapiens*, (iv) the distances involved – if indeed pre-*sapiens* populations were involved – are very modest ('seagoing', not 'seafaring') and do not challenge the significantly greater technological and cognitive capabilities of *Homo sapiens* that facilitated global colonization, and (v) such a small body of data is very unlikely to be representative of early intentional seagoing (Ammerman, 2014; Broodbank, 2014; Galanidou, 2014; Leppard, 2014, 2015; Phoca-Cosmetatou and Rabett, 2014, *inter alia*).

Given the potential implications of Palaeolithic insular activity for our understanding of pre-*sapiens* hominin capabilities (Leppard and

Runnels, 2017), it was argued that a more robust evidential basis was required to develop the debate, to be achieved by scientifically dating well-stratified, artefact-bearing deposits (Cherry and Leppard, 2018: 188). It was for this reason that SNAP segued in 2015 from its initial pedestrian survey mode to excavation, with >40 soundings dug over the next five years, the trenches established on both flanks, and from hilltop to modern coastal plain (Carter et al., 2017; Carter and Athanasoulis, 2021). One issue faced by the project was that most archaeological deposits were in secondary (downslope) context, due to the relatively steep hillslope, and long-term effects of erosional processes. While these processes have resulted in some areas being denuded of soil, in other instances chert outcrops have served as sediment traps, collecting meters of colluvial and aeolian deposits behind them, as with the case of trench DG-A/001, the focus of this paper's chronometric analysis.

Given the likely time depth-involved, and the absence of organic remains, luminescence dating was the obvious scientific dating technique to produce a chronology of the stratified deposits at Stelida. This process began in 2015 working with the University of Washington (Feathers et al., 2017), and since 2017 has been based at the Université Bordeaux Montaigne. In this study we detail how the luminescence results now provide an additional line of evidence with which to evaluate the stratigraphic integrity of the site's deposits, and an opportunity to further evaluate luminescence dating techniques in colluvial settings (Fuchs and Lang, 2009).

The luminescence-based chronologies are established on the most recent exposure of sediment grains from dated strata to sunlight. In Stelida, these depositional ages provide both *terminus ante quem* (*TAQ*) dates for when the stone tools recovered from the sampled lithostratigraphic deposits were buried (Carter et al., 2019) and chronological information about site formation processes. The *depositional* ages are particularly important at Stelida, as the excavated sediments are predominantly colluvial, but also include aeolian deposits. The colluvial deposits aggregate material both from the excavation location and up-slope, while aeolian deposits derive from sources further afield, likely during periods of sea-level lowstands when the continental shelf was exposed. Because sediment available for colluvial transport included cultural lithic material, that material was incorporated into the colluvial deposits. The luminescence dates thus date the time at which the artefacts came to rest and be buried, rather than their date of manufacture.

The first deep stratigraphic sequence at Stelida to be fully excavated to natural was trench DG-A/001, a *sondage* established on a debris cone at the base of a low cliff of outcropping chert on Stelida's uppermost western flanks (Fig. 2). Excavation of this 2 m^2 unit exposed 3.8 m of deposits, a sequence that was dug as 30 contexts representing eight lithostratigraphic units (LUs). These LUs are the product of colluvial deposition punctuated by distinct periods of aeolian deposition and subsequent pedogenesis that produced four buried paleosols (S2 – S5). Underlying this sequence of cultural deposits is a basal natural stratum of weathered saprolite (LU8), which was exposed during the deposition of the oldest artifact-bearing debris flow (LU7). Following the deposition of LU7, a period of stability facilitated a well-developed calcareous colluvial soil (S5). This soil was buried by a sandy mud flow, creating an erosional unconformity at the boundary between these two units. Following deposition of LU6, the hillslope witnessed another period of landscape stability as indicated by well-developed soil within LU6 (S4). Unconformably overlying LU6 was a deposit of moderately sorted fine to medium aeolian sand, marking a shift in depositional regime. Following deposition, a third period of stability resulted in a moderately developed soil (S3). This deposit was capped by two colluvial events (debris flows) (LU4a/4b). Following the deposition of these colluvial units, a fourth soil developed (S2) during a brief period of stability. These deposits were buried by a mass movement boulder-filled dry fall (LU3), which was in turn buried by continued colluvial deposition (LU2) and (LU1). A final (modern) soil developed at the surface of LU1, marking the most recent period of landscape development on the Stelida hillslope (S1).

Lithics were abundant in all but the deepest LU (LU8), with ~12,000

artefacts recovered (excluding heavy residue), >9000 of which derived from sealed and dated Pleistocene strata (Fig. 3) (Carter et al., 2019). The flaked stone traditions represented in the assemblages from these LUs appeared to span the Lower Palaeolithic through Mesolithic, based on detailed comparisons with excavated material from elsewhere in Greece, the Balkans, and the Eastern Mediterranean more generally (e. g., Darlas, 2007; Galanidou et al., 2016; Gopher and Barkai, 2014; Kaczanowksa et al., 2010; Kuhn et al., 1996; Perlès, 1987 *inter alia*).

2.2. Previous numerical dating and open questions

The first major publication of the DG-A/001 stratigraphy, lithic material, and chronology (Carter et al., 2019) involved multi-grain infrared stimulated luminescence [IRSL] dates from six samples that spanned most of the sequence (LU2 – LU7) (Fig. 2). These IRSL dates demonstrated that the artefact-bearing deposits extended into the Middle Pleistocene, the earliest dated to at least (*TAQ*) 198.4 ± 14.5 ka (Carter et al., 2019). The results of the multi-grain measurements from Carter et al., 2019 are reproduced in Fig. 2.

However, in colluvial settings where sufficient bleaching must be checked for, and where post-depositional disturbances can alter age estimates, it is necessary to compare different dating techniques (Murray et al., 2012). In this paper we refine previous chronometric work through the evaluation of single-grain pIRIR$_{290}$ measurements and the application of different statistical models (Central Age, Average Dose and Finite Mixture), and consider the implications of the IRSL-dating results for our understanding of site formation processes. The presence of poorly bleached grains will be more identifiable with single-grain than with multi-grain analyses, which is useful in the context of such a site as Stelida with its complex colluvial sediments. Specifically, if the grains have been poorly bleached, the apparent equivalent dose may be different for multi-grain and single-grain and will provide single-grain overdispersion estimates that permit a more in-depth analysis.

2.3. Sampling strategy and sample preparation

The ages presented here refer to six samples taken in the summers of 2016 and 2017: SNAP16 01, SNAP16 02, SNAP16 03, SNAP16 04, SNAP17 04, and SNAP17 05 (Fig. 2). These are the same samples as those detailed in Carter et al. (2019; Table 3). In addition, four modern analogue samples were taken from the surface, for residual dose measurements: SNAP20 S1, SNAP20 S2, SNAP20 S3, and SNAP20 S4.

All six samples were etched but the modern analogues were not. This choice facilitated some observations, notably underlined by Duller (1992), about anisotropic removal of the surface on K-feldspars (Porat et al., 2015). Although in theory the use of Hydrofluoric Acid [HF] should be able to attack the grains on the surface, thus eliminating the contribution of alpha radiation, it has been observed that this does not take place homogeneously on the grain surface (Duval et al., 2018). Indeed, etching occurs in a heterogeneous way on the grains, whereby it is judged more judicious to stop this type of treatment and to consider the alpha contribution. Measurements were carried out on SNAP17 04 with and without etching to check whether the results are similar by recording the alpha values for measurements carried out on the part without etching. The study indicated that the ages were comparable to 1σ, so it was decided not to proceed with etching on the modern analogues (there is no age calculation for the modern analogues, only the equivalent dose).

2.4. Single-grain post IR-IRSL

2.4.1. Experimental protocols

Measurements were carried out only on potassium-feldspar enriched fractions, since the OSL measurements on quartz did not give any suitable luminescence signal (Carter et al., 2019). IRSL measurements on potassic feldspars were made using the SAR (single-aliquot-

Fig. 3. Select artifacts from LU5 to LU7. Flakes unless otherwise noted. a, scraper; b, backed flake; c, bladelet; d, piercer; e, piercer on blade-like flake; f, piercer; g, combined tool (burin and scraper on chunk); h, nosed scraper; i, combined tool (inverse scraper/denticulate/notch); j, denticulate (LU5); k, flake; l, denticulated blade-like flake (LU7); m, piercer; n, denticulate; o, denticulate; p, piercer; q, combined tool (linear retouch/denticulate); r, scraper; s, convergent denticulate (Tayac point); t, blade; u, scraper; v, denticulate; w, linear retouch; x, tranchet; and y, blade-like flake (LU6). Photographed by J. Lau and modified and page set by N. Thompson (reproduced from Carter et al., 2019).

regenerative-dose) protocol (Murray and Wintle, 2000) either with a stimulation temperature at 50 °C (IR$_{50}$) (Auclair et al., 2003) or with one at 50 °C and one at 290 °C (pIRIR$_{290}$) (Thiel et al., 2011). IR$_{50}$ measurements were made on multi-grain aliquots, when relevant, and pIRIR$_{290}$ measurements were made on both multi-grain and single-grain aliquots. Multi-grain aliquots (MG) or single grains (SG) used to determine the equivalent doses (D$_e$) were selected after applying the commonly used selection criteria (recycling ratio less than 10 % and recuperation ratio less than 5%). For the single-grain measurements presented in this study, the dose response curve of each grain was fitted with a saturating exponential of the form $a\,(1\text{-}exp(D/D_0)$, where a is the asymptotic limit for the normalized signal, D is the dose and D_0 is the curvature parameter.

We decided to perform measurements with a pIRIR$_{290}$ SAR protocol for two reasons: to avoid fading measurements (Thiel et al., 2011; Kars et al., 2012) and in view of the high equivalent doses for some samples after estimating these D$_e$ by IR$_{50}$ measurements. Indeed, elevated temperature pIRIR measurements involve recombination of more distant electron hole-pairs (Jain et al., 2015). Thus, these signals are less prone to fading (e.g., Thomsen et al., 2008). In addition, the accuracy of pIRIR$_{290}$ has been demonstrated on many samples extending back to the Middle Pleistocene (Buylaert et al., 2012).

To go beyond the multi-grains results, which correspond to signals summed over a dozen of grains (aliquot diameter = 1 mm) (Duller,

2008), single-grain measurements were carried out on each sample. With multi-grain measurements, in the case of poor bleaching, there may be an overestimation of the age since it is impossible to isolate the poorly bleached grains. We then speak of apparent age since we have an age estimate which may not reflect the reality of the distribution of grain ages. However, the single-grain method makes it possible to process the results statistically by isolating equivalent dose (D$_e$) populations.

For these single-grain post-IRSL protocol, a Risø TL/OSL DA 20 reader (Bøtter-Jensen et al., 2003, 2010) was used with a stimulation by IR laser (870 nm; Vishay TSFF5210) with a power density of stimulation of 300 mW/cm^2, detected by an EMI 9635 QA photomultiplier tube (PMT) with combination of optical filters (BG-39 in combination with BG3). For dose estimation, the signal is integrated over the first 0.06 s of stimulation; the background, over the last 0.62 s of stimulation. Laboratory irradiations were made using one internal calibrated ^{90}Sr/^{90}Y beta sources, delivering a dose rate of 0.093 ± 0.002 Gy/s and 0.091 ± 0.002 Gy/s at the times of measurements. pIRIR$_{290}$ measurements were performed with a preheat at 320 °C for 60 s for both natural and regenerated signals. A final IR stimulation bleach at 325 °C for 200 s was also applied at the end of each SAR cycle, to minimize recuperation signals.

2.4.2. Residual doses measured with the pIRIR$_{290}$

Elevated temperature pIRIR signals are known for presenting

residual doses, even for samples for null age (Buylaert et al., 2011; Murray et al., 2012. Kars et al., 2014). Thus, dating using pIRIR$_{290}$ signal involves estimating the remaining dose of each aliquot after light exposure and before burial. For this purpose, we implemented two approaches.

First, measurements were carried out on the six samples from trench DG-A/001, based on aliquots bleached in a solar simulator (Hönle SOL500) and giving an evaluation of the remaining dose after a given time (bleaching time ranging from 15 min to 48 h). For the *a priori* youngest samples from the top of the stratigraphic section and for the samples from the bottom of the same trench, a residual dose of ~ 6–7 Gy was previously measured (Carter et al., 2019). This represents a maximum of around 30% of the smallest equivalent dose obtained and less than 1% for the highest equivalent dose at stake in this study. This value, compared to the equivalent dose obtained, does not have a significant impact for most high dose samples, which means that taking them into account does not significantly affect the ages (see also the discussion in Kars et al., 2014).

Another approach to estimate the potential residual dose is to measure the D_e of modern analogues to the samples (Murray and Olley, 2002; Jain et al., 2004; Porat et al., 2010; Buylaert et al., 2011; Murray et al., 2012). For that purpose, the equivalent doses of two samples collected from a depth of ca. 4 cm (divided into 2 samples each, one from the 2 cm closer to the surface and the other between 2 and 4 cm deep from the surface) of the site, were measured. Results are presented in Table 1. These are the upper estimate of the possible residual values, since the measured D_e may be the sum of both a residual value and the D_e accumulated since sediment burial; we assume the latter event to be quite recent, without precise knowledge other than their current depth.. We obtained similar residual doses compared to our bleaching experiments, between ~ 4 and 6 Gy but with important overdispersion (as high as 32% for SNAP20 S2, which is quite significant for multigrain aliquots that tend to average D_e values).

2.4.3. Dose recovery tests and selection of grains based on curvature parameter D_0

To test the ability of our SAR protocol to accurately measure a known dose, SG pIRIR$_{290}$ measurements were made on the six samples after giving a known dose, trying to retrieve it (*Dose recovery ratio* [DRT]). Three SG discs were first bleached in the same conditions as for the bleaching tests mentioned above (solar simulator (Hönle SOL500) for 48 h). The grains were then given a dose corresponding to the multigrain equivalent doses measured on multi-grain aliquots (Carter et al., 2019), with a given dose of 1140 Gy for SNAP17 04. Grains were selected following the curvature parameter of their dose response curve (also called D_0 by Thomsen et al., 2016). Thus, all samples were sorted according to their D_0 value, i.e., we strictly constrained the doses with a D_0 criterion for which we no longer find any grains with low D_0. The *Dose recovery ratio* gave a good result equal to 1 at 1σ (knowing the range of acceptability from 0.9 to 1.1), except SNAP17 04 but the ratio was made on only 10 grains (this small number of grains is mainly due to a concern for the conservation of grains during the measurement of the three discs and are not coming from saturation problems). Furthermore, single-grain over-dispersal dose recovery ratio (OD DRT) values, without subtracting the residual dose, give an overdispersion between 7

and 20% (Table 2).

Grains were sorted by their D_0 value, to investigate – and resolve – issues arising from early saturation of the luminescence signals. Indeed, low D_0 grains inevitably bias dose distributions when the fraction of saturated grains is not negligible (Singh et al., 2017; Guérin et al., 2015; Thomsen et al., 2016). Previous work on the application of the D_0 criterion has been carried out mainly on quartz, because feldspars have a higher D_0 than quartz (e.g., Thomsen et al., 2016). However, in our case, we have samples with high doses – even for K-feldspar pIRIR$_{290}$.

Saturated grains are defined by either a natural signal lying above the dose response curve, or by the impossibility to estimate an uncertainty of their equivalent dose with the Analyst software (Duller, 2015). Such grains are problematic because statistical models such as the CAM, ADM, FMM, etc. only include sets of D_e values and their uncertainties. In addition, low D_0 grains saturate early – biasing dose distributions. Following Singh et al. (2017), Guérin et al. (2015), and Thomsen et al. (2016), we selected grains with an increasing minimum value. Thus, we only consider grains whose D_0 value is higher than a certain threshold and we increase the minimum accepted D_0 value for each sample. As expected, this procedure only affected the De estimates of the higher dose samples, especially sample SNAP17 04.

Fig. 4 shows the effect of increasing the minimum D_0 value on the fraction of saturated grains and on the central dose of this sample, estimated with the ADM. About one third of the total of accepted grains (before D_0-based grain selection) appear in saturation. This fraction remains stable until 300Gy, after what it drops to zero quickly. In parallel, the ADM dose slightly increases (by 13%) to reach a plateau. Interestingly, this plateau is consistent with – and much closer to – the multi-grains dose estimates of Carter et al. (2019).

For all other samples, the fraction of saturated grains was smaller than ~67 % and D_0-based grain selection did not affect ADM doses.

2.4.4. The effect of test dose size

Regarding the choice of the value of the test doses, it is common to test different test dose sizes (Murray & Wintle, 2000; Qin and Zou, 2012; Li et al., 2014; Yi et al. 2016; Colarossi et al., 2018; Lamothe et al., 2018). However, it has been shown that choosing a small test dose (compared to the palaeodose) may lead to significant thermal transfer, while too large a test dose can lead to an underestimation of the palaeodose (Murray & Wintle, 2000; Murray et al., 2021). Consequently, two dose measurements – one with a large test dose (75–80% of the natural dose) and another with a smaller test dose (25–30% of the natural dose) – were performed on the 2017 single-grain samples (Table 6). The measured doses are statistically indistinguishable at 2σ, although the doses measured with a 75–80% test dose are slightly smaller (Murray & Wintle, 2000).

Table 1

Equivalent doses measured for modern analogues. The multi-grain (MG) doses have been calculated with the ADM model, as well as their σ_d (Guérin et al., 2017). '*n*' denotes the number of measured aliquots.

Modern analogues	Depth (cm)	n	D_e (Gy) MG ADM	σ_d
SNAP20 S1	Surface − 2	5	5.5 ± 0.4	0.13 ± 0.04
SNAP20 S2	2-4	4	6.3 ± 1.2	0.32 ± 0.15
SNAP20 S3	Surface − 2	5	4.1 ± 0.2	0.13 ± 0.05
SNAP20 S4	2–4	5	4.2 ± 0.2	0.12 ± 0.03

Table 2

Results of dose recovery tests performed with the pIRIR$_{290}$ measurements. There were no OD results for SNAP17 04 due to the low number of grains.

Samples	LU	n	DRT (mean measured/given dose ratio)	OD DRT (%) (corresponds to σ_m in ADM)
SNAP16 01	LU5	123	1.21 ± 0.15	9 ± 1
SNAP16 02	LU4a	85	1.22 ± 0.14	8 ± 1
SNAP16 03	LU3	96	1.08 ± 0.13	10 ± 1
SNAP16 04	LU2	127	1.19 ± 0.15	10 ± 1
SNAP17 04	LU7	10	1.93 ± 0.58	–
SNAP17 05	LU6	55	0.95 ± 0.13	18 ± 2

Table 3

G-values of snap16 01 and snap17 05.

g-values SNAP16 01 (%/decade)	g-values SNAP17 05 (%/decade)
0.37 ± 0.53	0.17 ± 1.09

2.4.5. pIRIR$_{290}$ fading tests

Even if the pIRIR$_{290}$ signal seems to be insignificantly affected by anomalous fading (Lamothe et al., 2003; Jain and Ankjærgaard, 2011; Buylaert et al., 2008; Thiel et al., 2011; Kars et al., 2012), fading measurements were performed on two samples – SNAP16 01 and SNAP17 05 – on 500 and 300 grains respectively, to verify the absence of significant anomalous fading in the specific case of Stelida's samples. It should be noted that these grains were measured on single-grain disc. As can be seen in Table 3, the g-values (Auclair et al., 2003) are very low and are consistent with zero, within 1σ. Consequently, no fading correction has been applied to the pIRIR$_{290}$ D$_e$ estimates.

3. Results

3.1. Comparison and evaluation of central dose models (CAM, ADM)

For all samples, out of 500 measured grains, at least 100 grains gave a suitable signal, except for sample SNAP16-02 (Table 6). D$_e$ results are presented in Table 6, as determined with the Central Age Model (CAM; Galbraith et al., 1999, 2012) and the Average Dose Model (ADM; Guérin et al., 2017).

Initially, the CAM model was employed, as had been done for the multi-grain measurements (Carter et al., 2019). However, since the CAM systematically leads to age underestimates (especially in single-grain since overdispersion values are greater than for multi-grain measurements (Guérin et al., 2017)), the ADM model was applied. The ADM estimates the central dose not as a median value (as does the CAM) but as an average value, which is consistent with the average dose rate determined (Guérin et al., 2017; Heydari et al., 2018). Thus, the use of these two models makes it possible to compare the results in SG with those in MG. To run the ADM, the intrinsic overdispersion (σ$_m$) was determined for each sample based on the dose recovery tests described

Table 4

Doses values for each sample after FMM (Galbraith et al., 1999).

Components	Dose values of SNAP16 01	Dose values of SNAP16 02	Dose values of SNAP16 03	Dose values of SNAP16 04	Dose values of SNAP17 04	Dose values of SNAP17 05
80–100 %	–	–	–	–	–	412 ± 13 (83 %)
60–80 %	–	96 ± 4 (74 %)	–	–	947 ± 31 (63 %)	–
40–60 %	–	–	–	–	–	–
20–40 %	120 ± 9 (37 %) 100 ± 19 (31 %) 84 ± 7 (24 %)	–	70 ± 2 (39 %) 51 ± 2 (26 %) 100 ± 5 (23 %)	–	706 ± 23 (32 %)	–
0–20 %	177 ± 12 (8 %)	74 ± 6 (16 %) 115 ± 25 (7 %) 180 ± 20 (2 %)	154 ± 9 (9 %) 37 ± 9 (3 %)	–	341 ± 15 (4 %) 104 ± 9 (<1%)	309 ± 90 (8 %) 940 ± 72 (5 %) 177 ± 18 (3 %)

Table 5

Equivalent dose results obtained with different models (arithmetic mean, Central Age Model, Average Dose Model) in multi-grains.

			pIRIR$_{290}$ multi-grains results (Carter et al., 2019)				
Samples	LU	n	De (Gy)(arithmetic mean value)	De (Gy) CAM	OD (%) CAM	De (Gy) ADM	σ$_d$ ADM
SNAP16 01	LU5	10	122 ± 5	122 ± 2	1 ± 2	122 ± 2	0.01 ± 0.01
SNAP16 02	LU4a	10	106 ± 7	105 ± 3	6 ± 2	105 ± 3	0.06 ± 0.02
SNAP16 03	LU3	10	77 ± 2	77 ± 1	2 ± 2	77 ± 1	0.02 ± 0.01
SNAP16 04	LU2	10	68 ± 2	68 ± 1	1 ± 2	68 ± 1	0.01 ± 0.01
SNAP17 04	LU7	10	1019 ± 58	1014 ± 22	0	1014 ± 22	0.00 ± 0.01
SNAP17 05	LU6	9	479 ± 45	476 ± 16	7 ± 3	478 ± 16	0.07 ± 0.02

Table 6

The different expression of dose equivalent results in single-grain.

			pIRIR$_{290}$ single-grain results								
Sample	LU	n	De (Gy) Arithmetic mean	De (Gy) CAM	OD (%) CAM	De (Gy) CAM Test dose 75–80 % De	De (Gy) ADM	ADM (σ$_d$)	De BayLum (68 %)	n	De (Gy) after D$_0$ selection (ADM)	De (Gy) after D$_0$ selection (CAM)
SNAP16 01	LU5	170	112 ± 2	107 ± 2	21 ± 1	–	109 ± 2	0.18 ± 0.0 2	109–113	144	108 ± 2	106 ± 2
SNAP16 02	LU4a	46	97 ± 3	95 ± 3	17 ± 2	–	96 ± 3	0.13 ± 0.04	93–99	46	96 ± 3	95 ± 3
SNAP16 03	LU3	134	80 ± 3	73 ± 2	35 ± 2	–	78 ± 3	0.34 ± 0.0 2	78–83	107	77 ± 3	72 ± 3
SNAP16 04	LU2	150	63 ± 2	58 ± 2	38 ± 2	–	62 ± 2	0.36 ± 0.02	62–66	50	57 ± 3	55 ± 3
SNAP17 04	LU7	248	911 ± 28	825 ± 21	35 ± 2	777 ± 13	872 ± 16	0.33 ± 0.05	1099–1162	27	988 ± 48	953 ± 66
SNAP17 05	LU6	283	435 ± 10	406 ± 7	29 ± 1	376 ± 10	416 ± 8	0.22 ± 0.00	410–424	180	414 ± 12	402 ± 9

above.

It should be noted here that the ADM and CAM equivalent dose values give essentially the same results for multi-grain measurements (Table 5). This is not surprising since the OD of multi-grain D_e distributions is very close to 0 and never exceeds 10%.

Table 5 presents the different results obtained from the multi-grain pIRIR$_{290}$ De distributions. Here, two analytical choices are particularly relevant: (i) whether the CAM or the ADM is used; and (ii) whether grains are selected based on their D_0 value.

About (i), let us first note that OD values – the treatment of which makes the difference between CAM and ADM estimates – range from 17 ± 2 to 38 ± 2%. Such values are commonly encountered for well-bleached distributions of quartz D_e values (Guérin et al., 2015). In addition, according to Heydari and Guérin (2018), for such OD values one would expect ~5–10% discrepancy between CAM and ADM estimates. This is exactly what is observed for our samples: the three samples displaying the greatest OD values display ADM dose values 6–7% greater than those calculated with the CAM.

3.2. Sources of overdispersion and use of FMM

The overdispersion in single-grain equivalent dose distributions is a result of a combination of two series of factors, some extrinsic (dose rate dispersion, post-depositional mixing of sediment, partial/incomplete bleaching) and the others intrinsic (inherent luminescence characteristics of grains, instrument reproducibility uncertainties), following the terminology of Thomsen et al. (2005). The latter source can be determined through dose recovery tests, in which all grains receive the same dose. In an ideal scenario where the sediments are homogeneous, were fully bleached prior to burial, and have not undergone any post-

depositional mixing, only intrinsic factors are expected to contribute to the natural overdispersion (Thomsen et al., 2005, 2012; Guérin et al., 2017; Smedley et al., 2020; Murray et al., 2021).

For all 2016 samples, σ_m values range between 0.07 and 0.11 while σ_m of the 2017 samples are greater (11 ± 7% and 18 ± 2%) (Table 2; see the OD from the DRTs). SNAP17 04 and SNAP17 05 samples have a greater contribution of intrinsic dispersion to the overdispersion than the other samples, while they have similar natural OD values compared to other samples (SNAP16 04, SNAP16 03; see Table 6). This is probably due to the luminescence characteristics of the grains and perhaps related to the fact that these samples are characterized by higher doses (see Thomsen et al., 2012, for such an effect on quartz OSL). Nevertheless, the extrinsic contribution remains more important than intrinsic factors, except for SNAP17 04, pointing out that there is still a greater contribution to this dispersion from a potentially heterogeneous environment or from a potential poor bleaching of the samples, or from post-depositional mixing.

With radial plots (Galbraith et al., 1999) and abanico plots (Dietze et al., 2016) related to the D_e distributions without D_0 sorting (since there is no significant difference in dose with D_0 sorting, and for some samples too few grains remain as for SNAP17 04), we observe for each SNAP16 sample a dose distribution with a large kurtosis, skewed peak around a central value (Figs. 5-8). On the other hand, we can observe that the dose distribution observed on the SNAP17 05 plots shows presence of a positive gaussian kurtosis peak, surrounded by some smaller peaks corresponding to lower and higher doses and, for SNAP17 04, one positive gaussian kurtosis peak with one smaller peak corresponding to a lower dose (Figs. 9-10). The plots also allow us to see that we do not have a large mixture of grains since they are composed mainly of a single major component. Indeed, there is no conservation of the
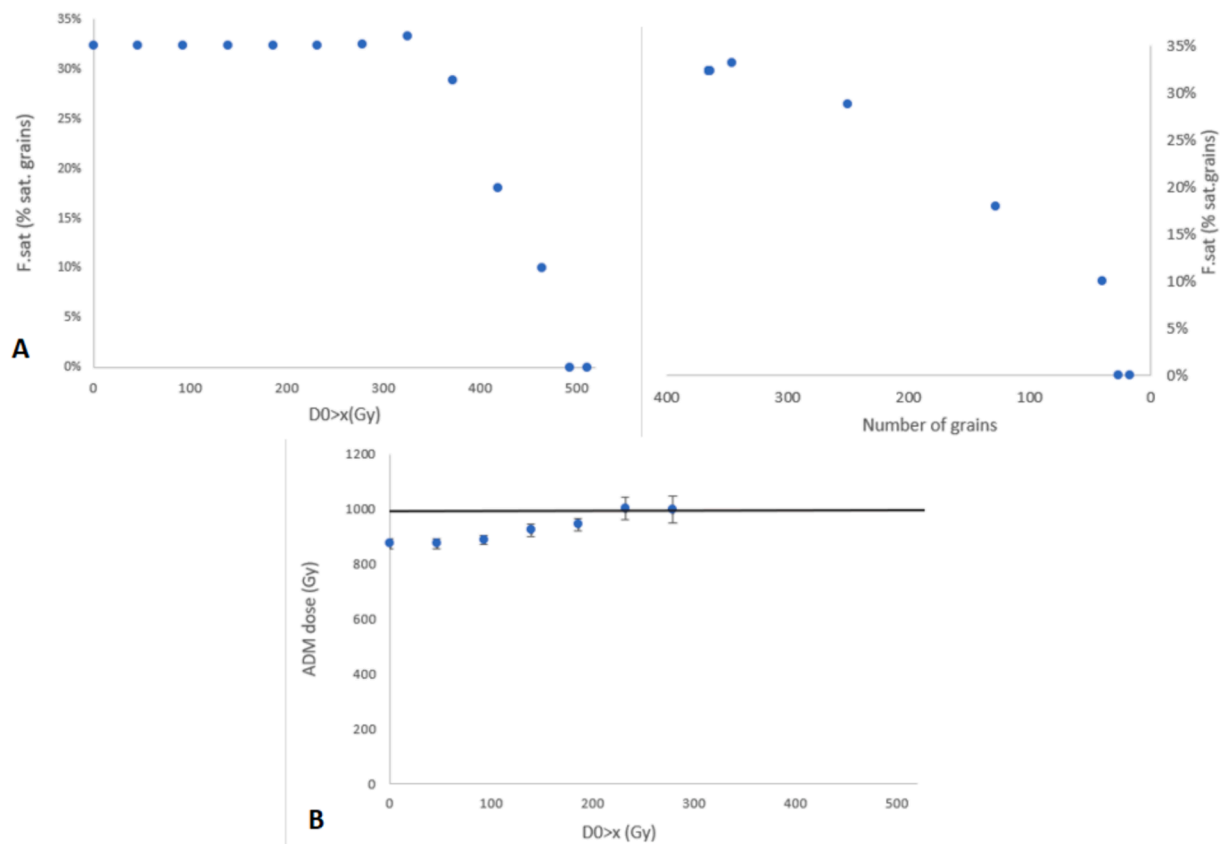

Fig. 4. SNAP17 04 – Effect of D_0-based grain selection for the single-grain pIRIR$_{290}$ signals measured from sample SNAP17 04. The decision to present the results of this sample is based on the significant contribution of sorting by D_0 on the results for single grain. A: Fraction of saturated grains as a function of the minimum accepted D_0 value (with a comparison of the number of remaining grains as a function of saturation). B: effect of D_0 on ADM SG D_e; the solid line indicates the multi-grain estimate: 1000 ± 24 Gy (Carter et al., 2019).
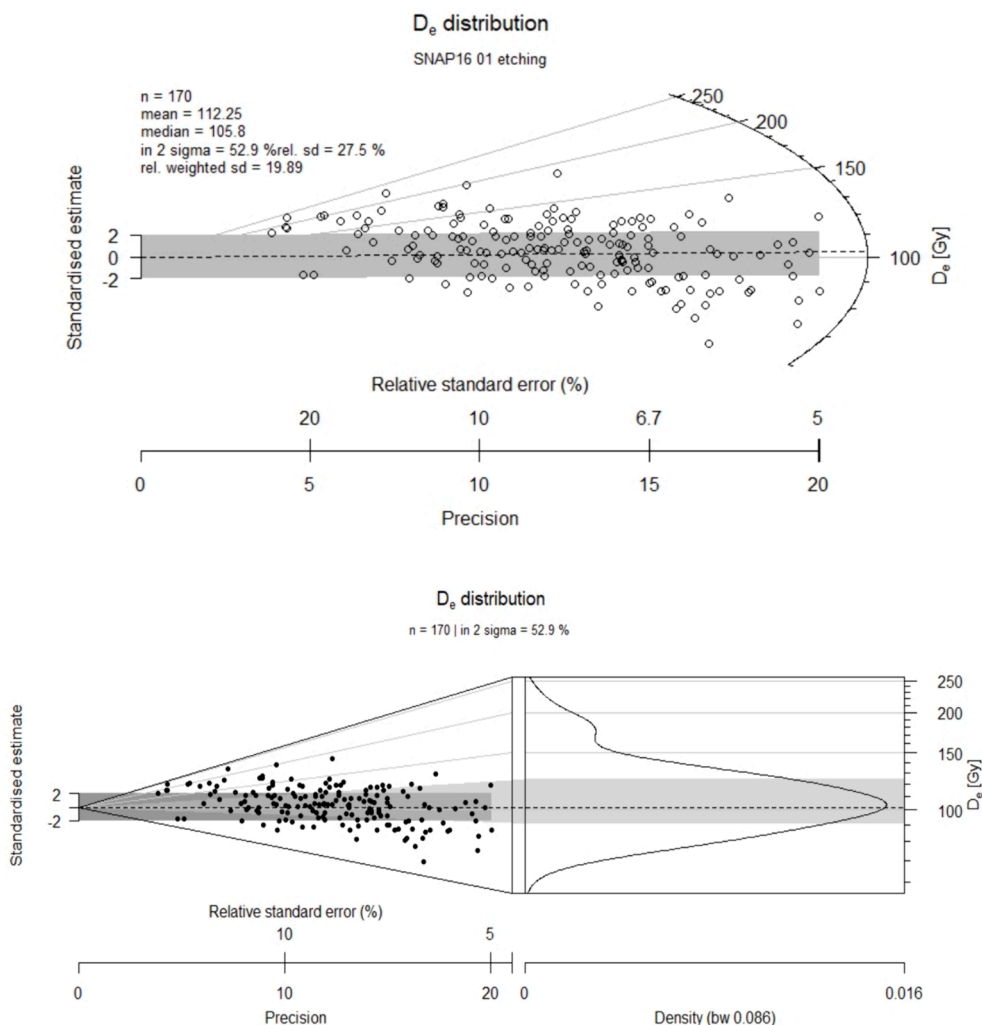
Fig. 5. Plots (kernel and abanico) of SNAP16 01 showing the distribution of the measured De.

doses corresponding to several major geological events but perhaps some rare poorly bleached grains and far lower doses (which correspond to between 3 and 4 % for all samples).

Regarding these plots and the difference in overdispersion between D_e and DRT (Nian et al., 2012), it appears important to check if there are potentially several discrete dose components that could influence the result of the central dose value estimated with the CAM or ADM. Thus, we tested the implementation of the FMM (Roberts et al., 2000). Several numbers of components have been tested (from 2 up to 8 components) and the σ_b value was chosen equal to the intrinsic OD of each sample. The sigma-b value represents the minimum expected overdispersion in the data should the sample be well-bleached (Thomsen et al., 2005; Cunningham & Wallinga, 2012). We observe that for some samples, we are facing a major component and minor components for the highest dose samples (SNAP17). Indeed, both have major components corresponding to the SG ADM dose (Table 4) and one or more minority components corresponding to less than 10% of the grains. In addition, for sample SNAP17 04, there are minority dose components with values higher than the central dose. Regarding all SNAP16 samples (except SNAP16 04), we can observe that apparent D_e values may be fitted by three components with similar grain fractions (between 30 and 40%) (Table 4). For SNAP16 01, there are also minority components with doses lower than the apparent dose.

To summarize, observation of the plots and FMM results indicate that the observed scattering of some sample is due to insufficient bleaching of a few (very few) poorly bleached grains (less than 10% of the measured

grains) and the presence of rare grains with lower D_e values probably corresponds to bioturbation (Smedley et al., 2015). These have no significant influence on other components on the apparent De determined, notably for the highest of them. Indeed, the FMM doses of the dominant components are consistent with the ADM doses and with the multi-grain doses reported by Carter et al. (2019).

3.3. Dose rates

Dose rates were calculated based on *in situ* dosimeter measurements for the gamma and cosmic contributions (following the approach of Kreutzer et al., 2017) and on radioelement contents from the sediment samples, for beta and alpha components. The radioelement contents were measured by high-resolution, low background gamma spectrometry BEGe (Guibert and Schvoerer, 1991) and converted to dose rates using the factors from Guérin et al. (2011). Concerning the dose rate from K, U, Th, no real significant disequilibrium in the ^{238}U series was observed (Table 7), even if we can observe a slight excess of U(^{226}Ra) compared to U(^{238}U) (terminology according to Guibert *et al.*, 1994) in the sediment samples. We chose the U(^{226}Ra) value for the calculation of the annual dose rate because the radioelements of this part of the U-series (including ^{214}Pb, ^{214}Bi and ^{226}Ra) represent around 80% of the contribution to the dose-rate (Aitken, 1985). For the internal dose rate of potassium feldspar grains, the internal beta dose rate was calculated assuming a ^{40}K-content of 12.5 ± 0.5% (Huntley and Baril, 1997) and grain size dependent self-dose factors were also considered (Guérin
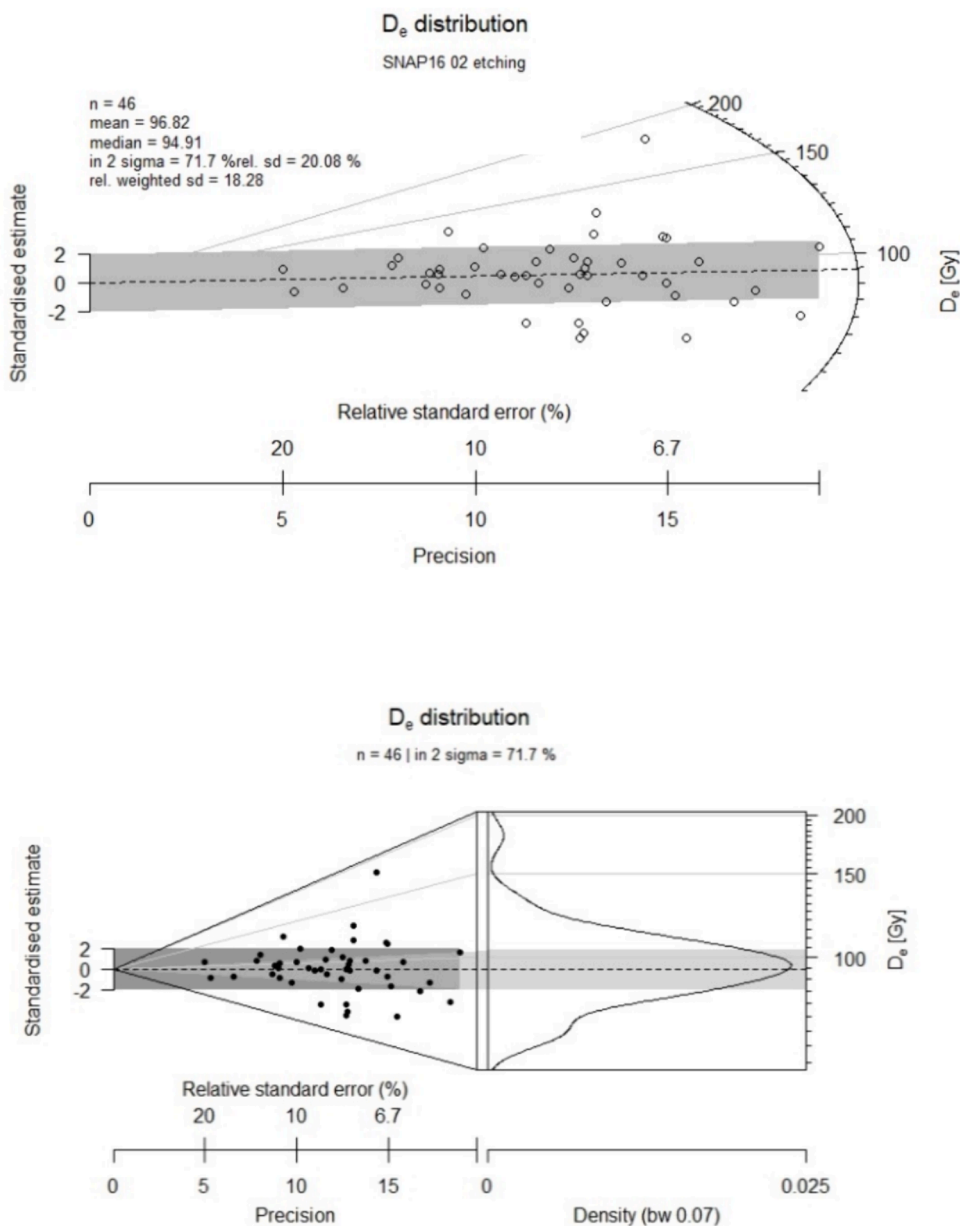
Fig. 6. Plots (kernel and abanico) of SNAP16 02 showing the distribution of the measured De.

et al., 2012). For the alpha component, an a-value of 0.08 ± 0.02 was used (Rees-Jones, 1995). Concerning water content, it has been estimated at $8 \pm 6\%$ during burial for all samples, based on measurements at the time of sampling.

3.4. BayLum ages

In our previous paper, we used BayLum, a software implementing a Bayesian model (Combes et al., 2015; Combes and Philippe, 2017; Philippe et al., 2017; Guérin et al., 2021), to calculate final ages. Using this model, one may constrain ages from the same trench using stratigraphic constraints and covariance in OSL ages arising from systematic errors (Carter et al., 2019). BayLum ages were considered as the most representative of sediment deposition at the site; in general, BayLum has been shown to provide more accurate, precise, and robust results than conventional approaches (Heydari and Guérin, 2018; Heydari et al., 2020, 2021; Guérin et al., 2022; Chevrier et al., 2020). In this study, BayLum was also used, without sorting by D_0 in view of the low number

of grains preserved in some samples (SNAP17 04; Table 6), to see if the results are still close to what we can see with multi-grain results. The same systematic error parameters as previously used (Carter et al., 2019) were conserved for water content, internal dose rate, radioelement concentration and D_e measurements.

Fig. 11 Unlike BayLum's multi-grain results, the use of covariance matrices failed to converge the MCMC chains (Carter et al., 2019). Unfortunately, unlike when used for multi-grain datasets (Carter et al., 2019), BayLum did not converge when calculating single-grain ages with stratigraphic constraints and a theta matrix to account for systematic errors. Nevertheless, the results suggest that BayLum manages the presence of saturated grains better, reducing their influence on the results, particularly regarding sample SNAP17 04 (Table 6). Indeed, the values are comparable to, and consistent with, those in multi-grain, except for sample SNAP17 05 which is closer to the single-grain result after D_0 sorting. Thus, BayLum results obtained from single-grain pIRIR$_{290}$ confirm the previous results of Carter et al. (2019). Finally, the BayLum results obtained with single grain are consistently more

D_e distribution

SNAP16 03 etching

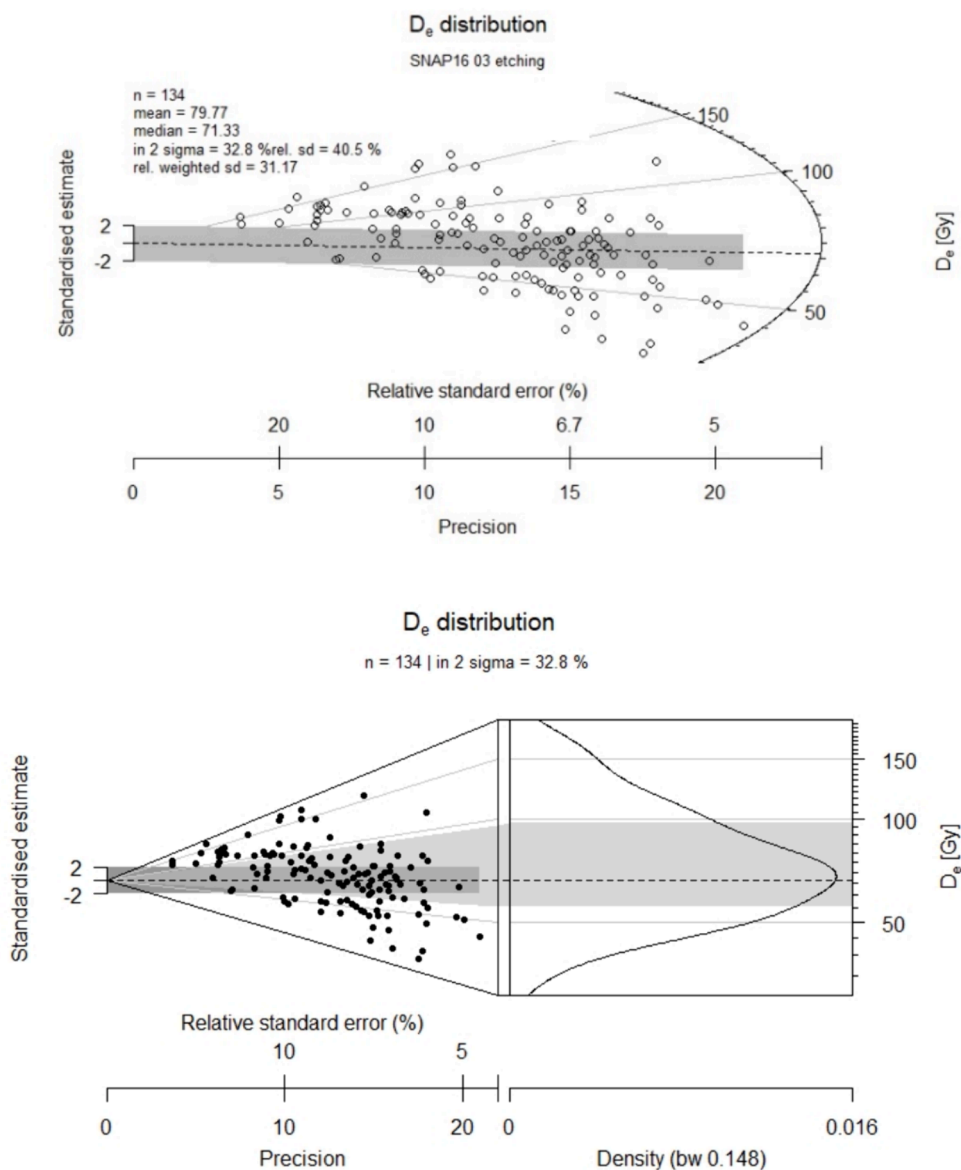## D_e distribution

n = 134 | in 2 sigma = 32.8 %

Fig. 7. Plots (kernel and abanico) of SNAP16 03 showing the distribution of the measured De.

precise than multi-grain, average dose model [ADM] ages.

4. Discussion

4.1. Final SG ages and comparison with results from Carter et al., 2019

This study quantified and compared different infrared stimulated luminescence [IRSL] measurements on K-feldspars, namely IR_{50} and $pIRIR_{290}$ multi-grain measurements and $pIRIR_{290}$ single-grain measurements from trench DG-A/001, a sequence that was first published in 2019. The results of the new residual dose measurements show that the same order of magnitude is measured with bleaching under laboratory conditions in a solar simulator and with modern analogues. This observation indicates that the laboratory measurements are close to the "real" conditions. The residual doses represent only a small part of the equivalent doses, especially for the older samples, and it is difficult to know exactly which dose should be subtracted. Furthermore, if we consider that we subtract about 7 Gy (maximum value) from the youngest samples, which are SNAP16 04 and 03, this only changes the age by about a thousand years (which is 8% of the age for the youngest age, that of sample SNAP16 04), and does not fundamentally change the

interpretation of these deposits from geological and archaeological perspectives. We therefore decided not to subtract any residual dose from the equivalent dose obtained for each sample, as other authors have already argued (Kars et al., 2014).

All samples measured in IR_{50} and $pIRIR_{290}$ in multigrain are similar at the 1σ or 2σ level for three samples (SNAP16 02,03,04 and SNAP17 05) but are statistically different for SNAP16 01. However, the data are close (17 ± 1 ka for IR_{50} and 23 ± 1 ka for MG $pIRIR_{290}$) and furthermore, it should be noted that the $pIRIR_{290}$ single-grain and the new IR_{50} multi-grain results (from the 2019 publication) are similar at 2σ (Table 8). Multi-grain and single-grain ADM $pIRIR_{290}$ ages are consistent with each other. The single-grain samples generally have a lower equivalent dose, and therefore a slightly younger age than the multi-grain measurements. However, when we keep only grains with a large D_0, we observe that the values become closer to those in multi-grain for sample SNAP17-04, which represents the oldest age, the new single-grain analysis now extending this date slightly further into the Middle Pleistocene (Table 8). Except for one sample, SNAP17 05, the new BayLum SG ages without sorting by D_0 presented in this article are similar (with smaller age intervals) to the previous age estimates (Table 8). However, the results are very close for SNAP17 05.
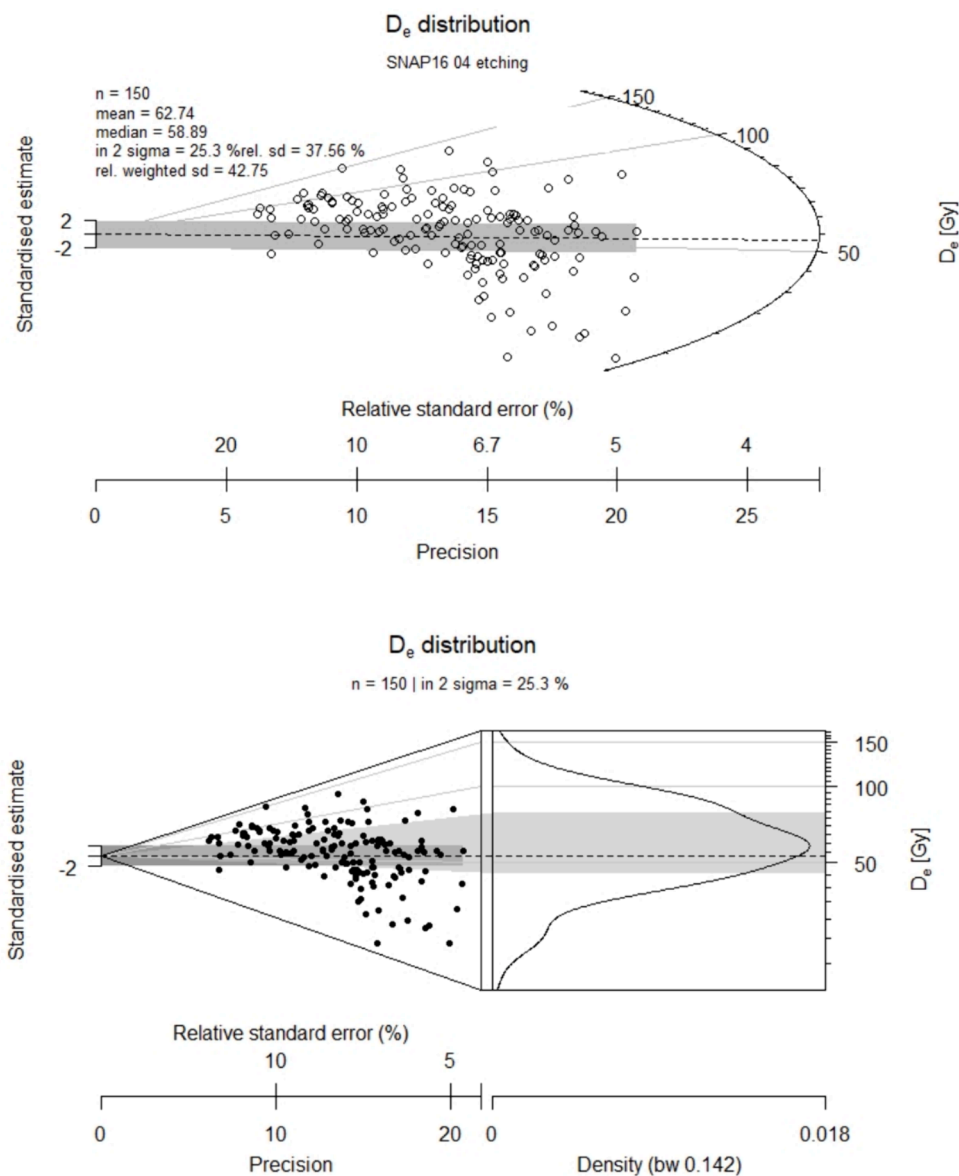
Fig. 8. Plots (kernel and abanico) of SNAP16 04 showing the distribution of the measured De.

These new more precise and more reliable measurements confirm by single-grain experiments what were only hypotheses in the first measurements in multi-grain IRSL: that the deposition of each archaeological layer is the result of a single main depositional event that may confidently be dated by IRSL signals. More generally, regarding over-dispersion, the low over-dispersion value of multi-grain distributions led Carter et al. (2019) to assume that the grains were well bleached at deposition. The single-grain measurements suggest that the colluvial packages at Stelida, once deposited, retained their stratigraphic integrity. Single-grain measurements also provide more information about overdispersion than could be seen in multi-grain. The results indicate that the intrinsic dispersion value of the samples only accounts for about 4–20% of the dispersion. However, the fact that for almost all samples (including the oldest ones) the multi-grain and single-grain equivalent doses are equal within uncertainties, underlines that the results are moderately impacted by grain dispersion. In sum, the equivalent dose results presented here suggest that no major modification is necessary to the current luminescence chronological framework of Stelida, except that the results in SG are much more accurate than those in multi-grain, since the results of the new single-grain feldspar measurements agree with the former age estimates (Carter et al., 2019) when we extract

saturated grains by selecting only high D_0 grains. The consistency between multi-grain and single-grain results suggests that aeolian and colluvial deposits at Stelida have not been significantly affected by post-depositional processes since their most recent deposition and that most of the grains were well-bleached prior to deposition.

4.2. Implications

Based on the abanico plots, it appears that few grains with a higher and/or lower dose than the apparent dose exist, corresponding presumably to few partial bleached grains and bioturbation (for small doses) (Smedley et al., 2015). Thus, the most likely explanation for the apparent dispersion in samples − like SNAP17 05 − is the presence of a very small number of grains with different doses from the main component, which may have very high different dose values corresponding to partial bleaching. However, the OD values of D_e below 50% and without asymmetry of De distributions (Trauerstein et al., 2014) and the OD values of DRT (Brown et al., 2015) suggest that the samples are mostly composed of well bleached grains, not affected by major post-depositional mixing. Furthermore, FMM results seem to identify components that do correspond only to statistical artifacts.
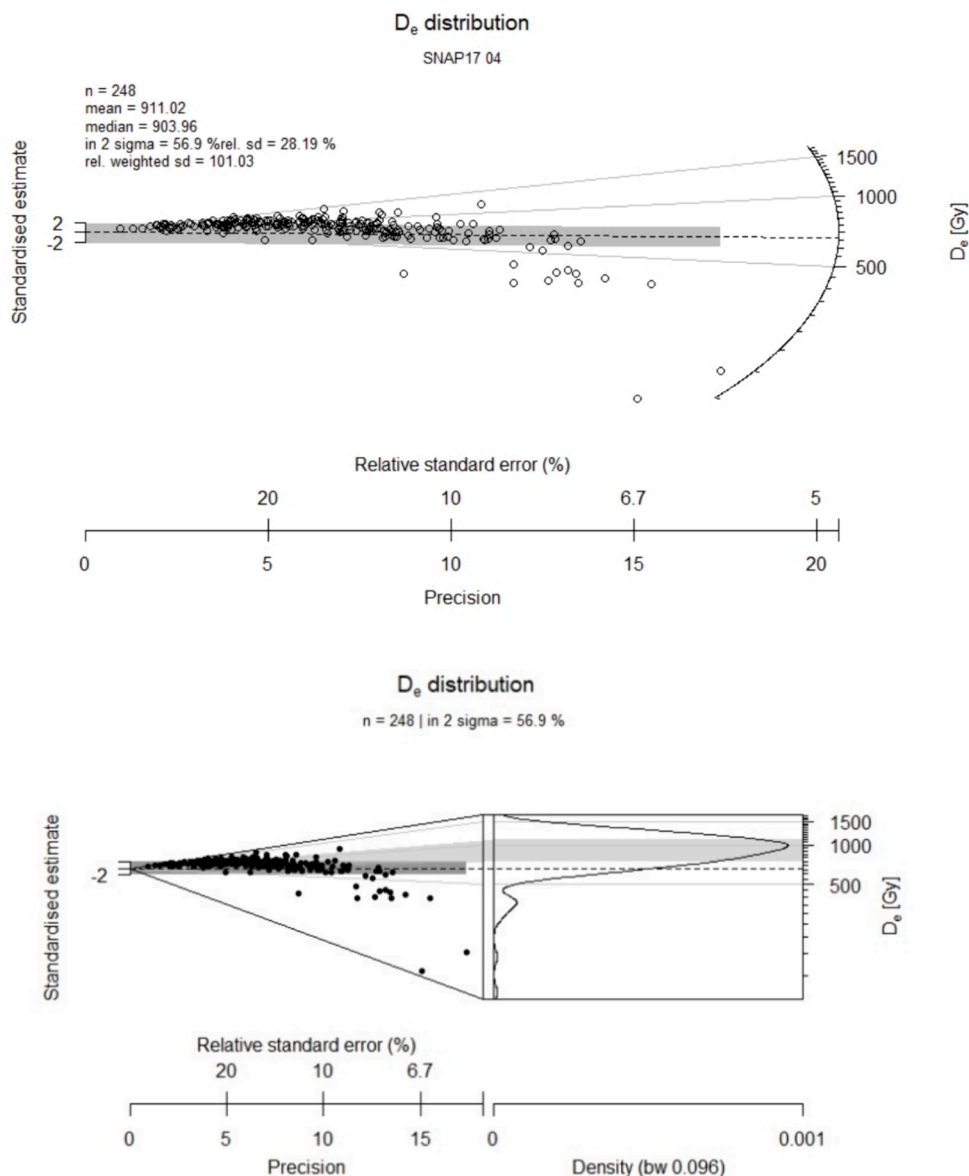
Fig. 9. Plots (kernel and abanico) of SNAP17 04 showing the distribution of the measured De.

For sample SNAP 17 04, which corresponds from technological considerations to the oldest archaeological artefacts unearthed at Stelida, we consider that the ages calculated with BayLum are the most reliable; in multi-grain we obtain, at the 68% confidence interval, an age interval between 220,000 and 189,000 years; based on single-grain measurements, this 68% interval becomes 233,000-217,000 ka (209,000 – 241,000 ka at the 95 % confidence interval). These data, more accurate and more precise than multi-grain results, therefore, support previous assertions that Stelida was visited during the Middle Pleistocene, but compatible only with Marine Isotope Stage 7 [MIS 7], evidence that part-formed the basis of an argument that the central Aegean represented a key throughfare for early humans during the initial peopling of southeastern Europe (Carter et al., 2019).

These results also have implications for geomorphologically-informed site prospection models seeking to find Palaeolithic sites in the Central Aegean (Holcomb et al. 2020). While hillslopes are typically avoided by archaeologists because they represent environmental settings characterized by erosion, our work at Stelida suggests that boulders can cause localized sediment traps preserving deposits up to 3 m in depth. These settings have the potential to preserve soils spanning the

Middle Pleistocene to late Holocene. At Stelida, such processes led to the preservation of well-developed argillic horizons associated with the Last Glacial Maximum (LGM) and the last Interglacial (MIS 5), as well as well-developed calcic horizons associated with MIS 7. Holcomb et al. (2020) suggested that coastal alluvial fans and internally drained basins represent key geomorphic settings conducive to Paleolithic site preservation. To this, we can add hillslopes found along coastal settings, especially those like Stelida which are associated with the Naxos-Paros detachment fault along the western coast of Naxos, eastern coast of Paros, and north-east coast of Mykonos (Skarpelis et al. 2017).

5. Conclusion

In this study, previously measured multi-grain IRSL values were compared against new single-grain measurements to increase the robustness of the site chronology and provide another line of evidence with respect to site taphonomy. This reanalysis of the DG-A/001 chronological sequence involved making a new kind of residual dose measurement and employing different statistical models (CAM, ADM, FMM) and plots (radial plot and abanico plot) on single-grain equivalent dose
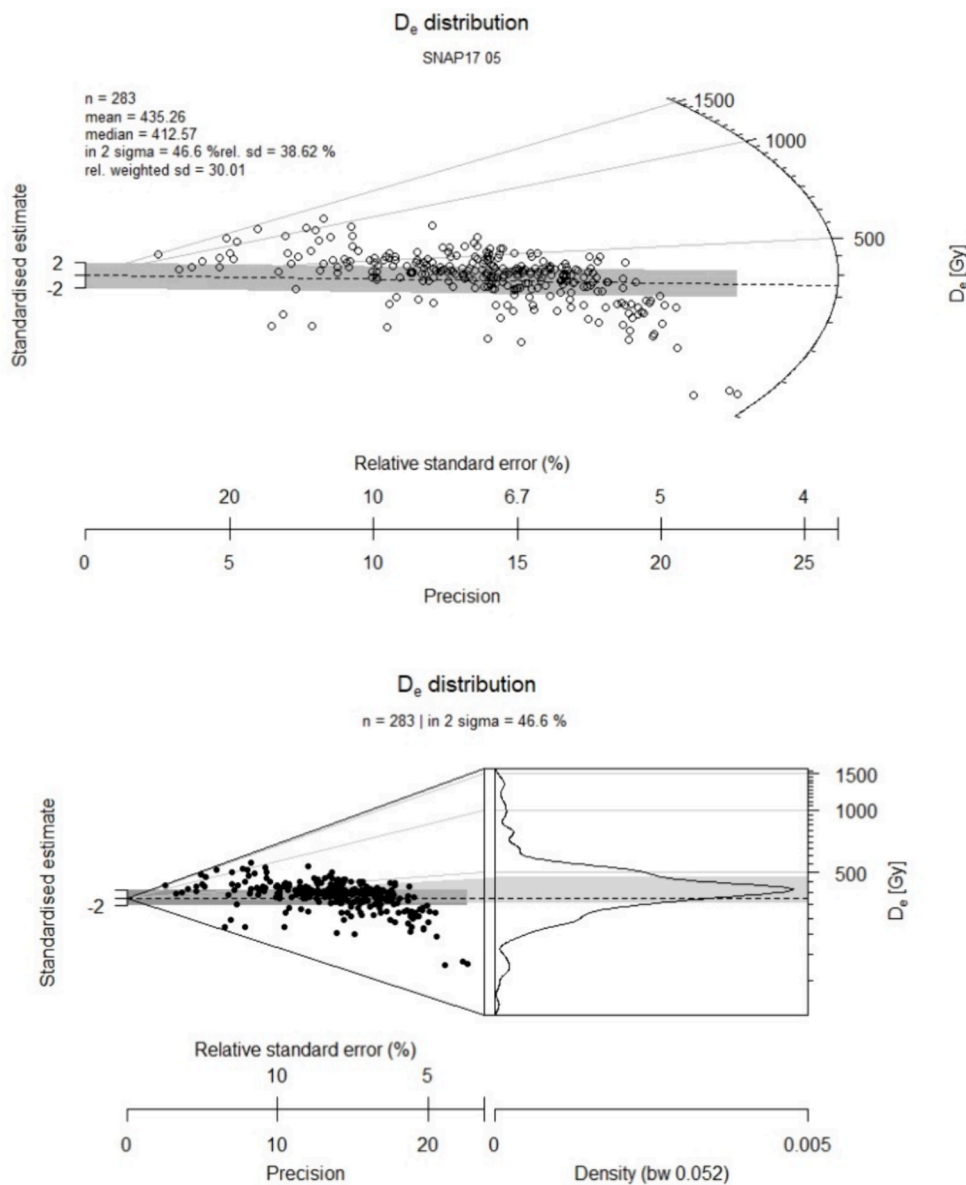
Fig. 10. Plots (kernel and abanico) of SNAP17 05 showing the distribution of the measured De.

Table 7

Radioelements contents determined by high-resolution low background gamma spectrometry and dose rate values (including *in situ* dosimeter values). The difference in internal dose rate value between SNAP16 and SNAP17 is due to the difference in particle size after final sieving after HF treatment: 180–250 μm for one and 200–250 μm for the other.

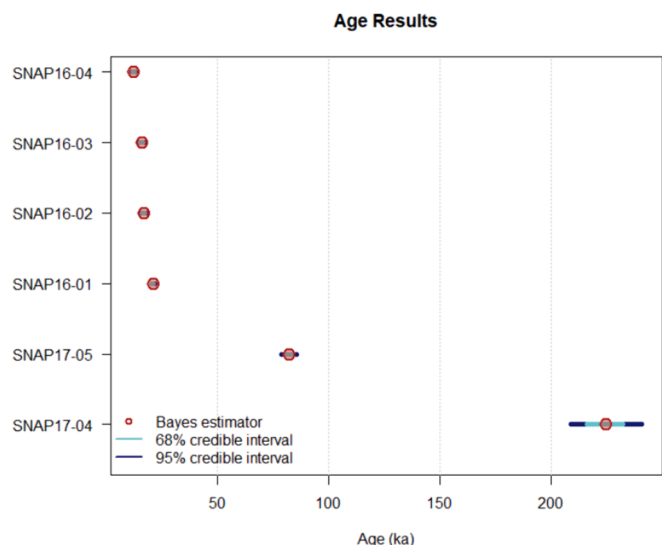| Samples	Trench	US level	K (%)	U(^{238}U) (ppm)	U(^{226}Ra) (ppm)	Th (ppm)	Internal Dose rate (Gy.ka^{-1})	External dose rate (alpha + beta + gamma + cosm) (Gy.ka^{-1})	Envir. Dose rate (gamma + cosmic) (Gy.ka^{-1})	Total dose rate (Gy. ka^{-1})
SNAP16 01	DG-A/ 001	LU5	2.62	3.25 ± 0.12	4.01 ± 0.03	22.12 ± 0.16	0.76 ± 0.03	4.54 ± 0.04	1.80 ± 0.02	5.30 ± 0.05
SNAP16 02	DG-A/ 001	LU4	2.30	4.54 ± 0.17	5.24 ± 0.04	28.28 ± 0.20	0.76 ± 0.03	4.96 ± 0.10	2.13 ± 0.10	5.71 ± 0.11
SNAP16 03	DG-A/ 001	LU3	2.60	4.09 ± 0.16	4.90 ± 0.04	25.46 ± 0.18	0.76 ± 0.03	4.27 ± 0.09	1.35 ± 0.08	5.03 ± 0.09
SNAP16 04	DG-A/ 001	LU2	2.56	4.39 ± 0.15	4.64 ± 0.04	26.58 ± 0.18	0.76 ± 0.03	4.51 ± 0.05	1.63 ± 0.04	5.27 ± 0.06
SNAP17 04	DG-A/ 001	LU7	2.72	2.63 ± 0.14	2.77 ± 0.03	22.62 ± 0.17	0.79 ± 0.03	4.24 ± 0.11	1.72 ± 0.11	5.03 ± 0.11
SNAP17 05	DG-A/ 001	LU6	2.52	3.57 ± 0.16	4.58 ± 0.04	25.29 ± 0.19	0.79 ± 0.03	4.28 ± 0.03	1.64 ± 0.03	5.07 ± 0.06

Age Results

Fig. 11. Single-grain post-IR IRSL chronology obtained with BayLum for the sequence of Stelida.

distributions. These results, obtained either with the ADM after selecting only high D_0 grains or with BayLum, generally confirm and improve the initial multi-grain results (Tables 5 and 6). Despite the potential complications inherent to colluvial settings, we observe through the single-grain results (and their statistical treatment) that each dated deposit is well dated by K-feldspar $pIRIR_{290}$ signals. Indeed, no significant bleaching problems were observed. No indication of significant post-depositional disturbance could be observed.

This study demonstrates that multi-grain $pIRIR_{290}$, multi-grain IR_{50}, and single-grain $pIRIR_{290}$ ages from trench DG-A/001 are similar and overlap at 2σ uncertainties. The chronologies produced by these different IRSL protocols agree with one another in terms of obtained ages and demonstrate the chronostratigraphic coherence of the DG-A/001 deposits. These results demonstrate that despite a complex taphonomy, colluvial sediments in semi-arid environments like Aegean hillslopes can still be well dated by potassic feldspars through IRSL, yielding both age estimates and post-depositional information.

Henceforth, in reporting the dates for the DG-A/001 sequence, we therefore consider it is more accurate and more precise to use the single-grain results than the multi-grain results. The precision of the results could be considered higher with the single-grain results, because the analysis is carried out grain by grain, considering each equivalent dose and more accurately characterizing overdispersion than multi-grain analyses. In our case, the BayLum ages based on single-grain analyses are more precise than BayLum ages based on multi-grain analyses, as we can especially see for SNAP17 05 (Table 8, Fig. 11); they also confirm the accuracy of the initial (Carter et al., 2019) ages. Furthermore, Bay-Lum SG ages seem to manage grains with a small D_0 as we can see for SNAP17 04. It should be noted, that although the age ranges are different for SNAP17 04, both intervals are consistent with an attribution to Marine Isotope Stage 7.

Consequently, based on the single-grain BayLum results, the *TAQ* dates for the cultural lithics indicate that the site was first visited at least 233,000–217,000 years ago, an age range which is broadly consistent with the chrono-cultural interpretation of the oldest stone tools. We also note the presence of chronological gaps (i.e., allostratigraphy – a discontinuous sequence) that correspond to erosional events leading to unconformities (Fig. 12). Indeed, three erosional unconformities were identified in the field, marked by sharp undulating lower boundaries between LU3 and LU4a, LU5 and LU6, and LU6 and LU7. Future excavations at Stelida will evaluate whether the deposits absent from DG-A/001 are represented elsewhere across the hillslope. Understanding the nature of deposition and erosion at Stelida is crucial for reconstructing past environmental change, and how past foragers interacted with the hillslope since the Middle Pleistocene. The next stage is to expand this work by detailing the integrated results of IRSL-dating, macro- and micro-stratigraphic analyses, surface dating techniques (e.g., Gliganic et al., 2021), and artefact studies from the other Pleistocene *sondages* around the site to produce a deep-time chronological framework and landscape history for Stelida.

In the first paper to present luminescence dating results from these deposits (Carter et al., 2019) we argued that the results of the luminescence dating showed Stelida to be the oldest cultural site of the Central Aegean Basin; this point still stands. The major frustration at the time was that because all the dates were minimum ages (*TAQ*, because they derived from secondary deposits), it was impossible to connect those episodes of early prehistoric activity with the then accepted palaeogeographic reconstructions of the Pleistocene Aegean (Lykousis, 2009). With it thus not possible to prove positively that pre-*sapiens* had accessed Stelida by boat, we took the more cautious approach and suggested that early visitations to the chert source occurred during those periods of glacial lowstands (e.g., during Marine Isotope Stage 8 [300 – 250 ka]), at which point terrestrial crossing likely existed joining the nearby Eurasian continents via what today is Naxos (Carter et al., 2019: Fig. 1). Things have subsequently changed, with more recent Ice Age sea-level reconstructions indicating that Naxos was insular throughout the Pleistocene (Ferentinos et al., 2023). As such, we can now argue that the exploitation of Stelida at least between 233 and 217 hundred thousand years ago (as determined by this study) necessarily involved open water crossings. Stelida thus now represents the earliest indirect evidence for seagoing in the northern hemisphere (see Gaffney, 2021), albeit at a date conceivably much later than those documented in South-East Island Asia, with the island of Flores accessed by boat – potentially by *Homo erectus* – at least one million years ago (Brumm et al. 2011). While these updated results are hugely significant, we are now in a poorer position than in 2019 with regards to knowing who might have been responsible for these early sea crossings to Stelida. Up until a week or so prior to the publication of our DG-A/001 dates, the received wisdom was that *Homo sapiens* only arrived in the Aegean around 40,000 years ago (Douka et al., 2011), whereby any earlier cultural activity necessarily involved the precursor populations, namely Neanderthals, or *Homo heidelbergensis* (Harvati et al., 2009; Tourloukis and Harvati, 2018). This model has now been upended by the republication of two

Table 8

DG-A/001 chronology by Lithostratigraphic Unit [LU]. SG ages sorted by D_0 are shown here for comparison purposes only, due to the small number of grains.

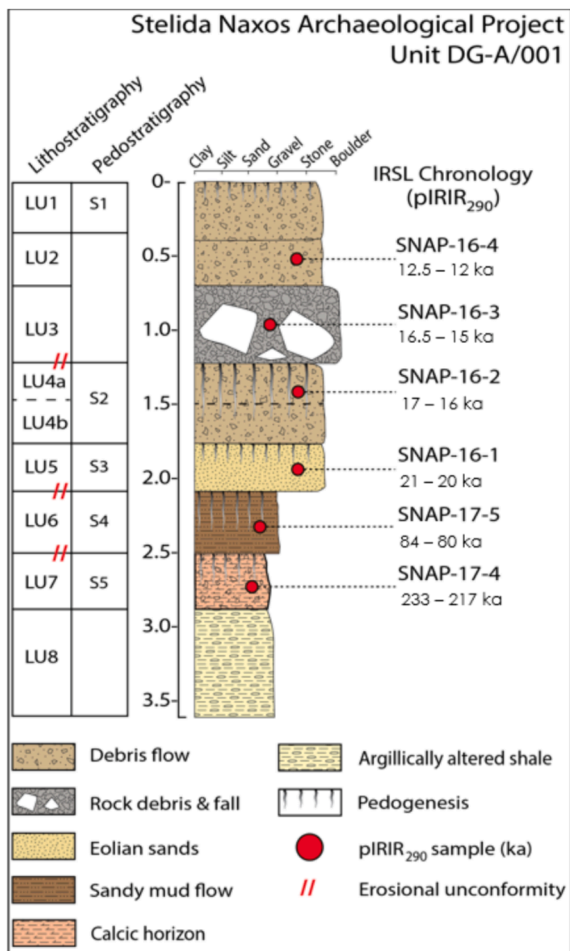| Echantillons DG-A/001	LU	IR_{50} MG ages DRC corrected (ka)	ADM		BayLum (68 %)	
			$pIRIR_{290}$ MG ages (ka)	$pIRIR_{290}$ SG (after D_0 selection) ages (ka)	$pIRIR_{290}$ MG ages (ka)	$pIRIR_{290}$ SG ages (ka)
SNAP16 04	LU2	11 ± 1	13 ± 1	12 ± 1	14–12	12.5 – 12
SNAP16 03	LU3	14 ± 1	16 ± 1	15 ± 1	16–14	16.5 – 15
SNAP16 02	LU4a	16 ± 1	19 ± 1	17 ± 1	20–17	17–16
SNAP16 01	LU5	17 ± 1	23 ± 1	21 ± 1	24–21	21–20
SNAP17 05	LU6	91 ± 13	94 ± 6	82 ± 6	100–86	84–80
SNAP17 04	LU7	–	198 ± 14	196 ± 14	220–189	233–217

Fig. 12. The new BayLum single-grain ages with the stratigraphy of DG-A/001 (modified from Carter et al., 2019); ages are indicated as 68% credibility intervals.

hominin skulls from the Apidima Cave in the southern Greek mainland (Fig. 1), the oldest of which – dated to over 210 ka – is *Homo sapiens*, the oldest example of this species in Eurasia (Harvati et al., 2019). This skull, along with a *Homo sapiens*' jawbone from the Misliya Cave (Israel) dated to at least 177 ka (Hershkovitz et al., 2018), indicates a Middle Pleistocene dispersal of early modern humans into the region, albeit possibly a failed/short-term one, prior to these populations being replaced by Neanderthals, as with the younger ($\geq$170 ka) skull from Apidima (Harvati et al., 2019). In sum, we cannot with any confidence assign the earliest cultural activity at Stelida – and by extent these Middle Pleistocene sea-crossings, to an exclusive *Homo sapiens* population, as the site's chronology *could* theoretically relate to *Homo sapiens*, or precursor populations, and with tool-traditions such as Levallois common to both early modern humans and Neanderthals (e.g., Hublin et al., 2017), we have to await new forms of evidence to resolve this question.

CRediT authorship contribution statement

Ninon Taffin: Writing – original draft, Resources, Methodology, Conceptualization. **Christelle Lahaye:** Writing – review & editing, Validation, Supervision, Resources, Data curation. **Daniel A. Contreras:** Writing – review & editing, Resources. **Justin A. Holcomb:** Writing – review & editing, Resources. **Danica D. Mihailović:** Resources. **Panagiotis Karkanas:** Resources. **Guillaume Guérin:** Writing – review & editing, Validation, Software, Data curation. **Demetris Athanasoulis:** Validation. **Tristan Carter:** Writing – review & editing, Visualization, Validation, Supervision, Resources, Investigation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Author contributions

T.C. and D.A. directed the excavation; IRSL sampling and analyses by N.T., C.L. and G.G..; geoarchaeological and stratigraphic analyses by J.A.H., D.A.C., and P.K.; lithic studies and related data analysis by D.D.M. and T.C; N.T., T.C., D.A.C, J.A.H., C.L. and G.G. wrote the main text, with T.C., J.A.H. and C.L. undertaking the post-review revisions.

References

Aitken, M.J., 1985. Thermoluminescence Dating. Academic Press, London.

Ammerman, A.J., 2014. Setting our sights on the distant horizon. Eurasian Prehistory 11, 203–236.

Auclair, M., Lamothe, M., Huot, S., 2003. Measurement of anomalous fading for feldspar IRSL using SAR. Radiat. Meas. 37 (4–5), 487–492.

Bøtter-Jensen, L., Thomsen, K.J., Jain, M., 2010. Review of optically stimulated luminescence (OSL) instrumental developments for retrospective dosimetry. Radiat. Meas. 45 (3–6), 253–257.

Bøtter-Jensen, L., Wintle, A.G., McKeever, S.W.S., 2003. Optically Stimulated Luminescence Dosimetry. Elsevier, Amsterdam.

Broodbank, C., 2014. So... what? Does the paradigm currently want to budge so much? JMA 27 (2), 267–272.

Brown, N.D., Rhodes, E.J., Antinao, J.L., McDonald, E.V., 2015. Single-grain post-IR IRSL signals of K-feldspars from alluvial fan deposits in Baja California Sur. Mexico. Quat. Int. 362, 132–138.

Brumm, A., Jensen, G.M., van den Bergh, G.D., Morwood, M.J., Kurniawan, I., Aziz, F., Storey, M., 2010. Hominins on Flores, Indonesia, by one million years ago. Nature 464 (7289), 748–752.

Buylaert, J.-P., Murray, A.S., Thomsen, K.J., Jain, M., 2009. Testing the potential of an elevated temperature IRSL signal from K-feldspar. Radiat. Meas. 44, 560–565.

Buylaert, J.-P., Jain, M., Murray, A.S., Thomsen, K.J., Thiel, C., Sohbati, R., 2012. A robust feldspar luminescence dating method for Middle and Late Pleistocene sediments. Boreas 41, 435–451.

Buylaert, J.P., Thiel, C., Murray, A.S., Vandenberghe, D.A., Yi, S., Lu, H., 2011. IRSL and post-IR IRSL residual doses recorded in modern dust samples from the Chinese Loess Plateau. Geochronometria 38, 432–440.

Carter, T., Athanasoulis, D., 2021. Stelida Naxos Archaeological Project. Mouseion: Journal of the Classical Association of Canada 18, 264–269.

Carter, T., Contreras, D.A., Doyle, S., Mihailović, D.D., Moutsiou, T., Skarpelis, N., 2014. The Stélida Naxos Archaeological Project: New data on the Mesolithic and Middle Palaeolithic Cyclades. Antiquity Project Gallery 88 (341). http://archive.antiquity.ac.uk/projgall/carter341.

Carter, T., Contreras, D.A., Doyle, S., Mihailović, D.D., Skarpelis, N., 2016. Early Holocene interaction in the Aegean islands: Mesolithic chert exploitation at Stélida (Naxos, Greece) in context. In: Ghilardi, M. (Ed.), Géoarchéologie Des Îles De Méditerranée. CNRS éditions, Paris, pp. 275–286.

Carter, T., Contreras, D.A., Holcomb, J., Mihailović, D.D., Karkanas, P., Guérin, G., Taffin, N., Athanasoulis, D., Lahaye, C., 2019. Earliest occupation of the Central Aegean (Naxos), Greece: Implications for hominin and *Homo sapiens*' behaviour and dispersals. Sci. Adv. 5 (10), eaax0097. https://advances.sciencemag.org/content/5/10/eaax0997.

Carter, T., Contreras, D.A., Holcomb, J. Mihailović, D.D., Skarpelis, N., Campeau, K., Moutsiou, T., Athanasoulis, D., 2017. The Stélida Naxos Archaeological Project: New studies of an early prehistoric chert quarry in the Cyclades. In Rupp, D.W., Tomlinson, J. (Eds.), From Maple to Olive: Proceedings of a Colloquium to Celebrate the 40th Anniversary of the Canadian Institute in Greece, 10-11 June 2016. Publications of the Canadian Institute in Greece 10, Athens, pp. 75-103.

Cherry, J.F., 1981. Pattern and process in the earliest colonization of the Mediterranean islands. PPS 47, 41–68.

Cherry, J.F., Leppard, T.P., 2018. Patterning and its causation in the pre-Neolithic colonization of the Mediterranean islands (Late Pleistocene to Early Holocene). J. i. Coast. Archaeol. 13, 191–205.

Chevrier, B., Lespez, L., Lebrun, B., Garnier, A., Tribolo, C., Rasse, M., Guérin, G., Mercier, N., Camara, A., Ndiaye, M., Huysecom, E., 2020. New data on settlement and environment at the Pleistocene/Holocene boundary in Sudano-Sahelian West Africa: Interdisciplinary investigation at Fatandi V. Eastern Senegal. PLOS ONE 15 (12), e0243129.

Colarossi, D., Duller, G.A.T., Roberts, H.M., 2018. Exploring the behaviour of luminescence signals from feldspars: implications for the single aliquot regenerative dose protocol. Radiat. Meas. 109, 35–44.

Combes, B., Philippe, A., Lanos, P., Mercier, N., Tribolo, C., Guérin, G., Guibert, P., Lahaye, C., 2015. A Bayesian central equivalent dose model for optically stimulated luminescence dating. Quat. Geochronol. 28, 62–70.

Combès, B., Philippe, A., 2017. Bayesian analysis of individual and systematic multiplicative errors for estimating ages with stratigraphic constraints in optically stimulated luminescence dating. Quat. Geochronol. 39, 24–34.

Cunningham, A.C., Wallinga, J., 2012. Realizing the potential of fluvial archives using robust OSL chronologies. Quat. Geochronol. 12, 98–106.

Darlas, A., 2007. Le Moustérien de Grèce a' la lumière des récentes recherches. L'anthropologie 111, 346–366.

Dietze, M., Kreutzer, S., Burow, C., Fuchs, M.C., Fischer, M., Schmidt, C., 2016. The abanico plot: visualising chronometric data with individual standard errors. Quat. Geochronol. 31, 12–18.

Dinçer, B., 2016. The Lower Paleolithic in Turkey: Anatolia and hominin dispersals out of Africa. In: Harvati, K., Roksandic, M. (Eds.), Paleoanthropology of the Balkans and Anatolia. Springer, New York, pp. 213–228.

Douka, K., Perlès, C., Valladas, H., Vanhaeren, M., Hedges, R.E.M., 2011. Franchthi Cave revisited: the age of the Aurignacian in south-eastern Europe. Antiquity 85 (330), 1131–1150.

Duller, G.A.T., 1992. Luminescence Chronology of Raised Marine Terraces, South-West North Island, New Zealand. University of Wales. Aberystwyth. Doctoral dissertation,.

Duller, G.A., 2008. Single-grain optical dating of Quaternary sediments: why aliquot size matters in luminescence dating. Boreas 37 (4), 589–612.

Duller, G.A.T., 2015. The Analyst software package for luminescence data: overview and recent improvements. Ancient TL 33, 35–42.

Duval, M., Guilarte, V., Campaña, I., Arnold, L., Miguens, L., Iglesias, J., González-Sierra, S., 2018. Quantifying hydrofluoric acid etching of quartz and feldspar coarse grains based on weight loss estimates: implication for ESR and luminescence dating studies. Ancient TL 36 (1), 1–14.

Feathers, J., Carter, T., Contreras, D.A., Campeau, K., 2017. Luminescence dating of a Palaeolithic site in the Aegean. Poster presentation, 82nd annual meeting of the Society for American Archaeology, Vancouver.

Ferentinos, G., Gkioni, M., Prevenios, M., Geraga, M., Papatheodorou, G., 2023. Archaic hominins maiden voyage in the Mediterranean Sea. Quat. Int. 646, 11–21.

Fuchs, M., Lang, A., 2009. Luminescence dating of hillslope deposits—a review. Geomorphology 109 (1–2), 17–26.

Gaffney, D., 2021. Pleistocene water crossings and adaptive flexibility within the *Homo* genus. J. Archaeol. Res. 29 (2), 255–326.

Galanidou, N., 2014. Archaic hominins on Crete: fact or fiction? JMA 27 (2), 260–267.

Galanidou, N., Athanassas, N.C., Cole, J., Iliopoulos, G., Katerinopoulos, A., Magganas, A., McNabb, J., 2016. The Acheulian site at Rodafnidia, Lisvori, on Lesbos, Greece: 2010–2012. In: Harvati, K., Roksandic, M. (Eds.), Paleoanthropology of the Balkans and Anatolia. Springer, New York, pp. 119–138.

Galbraith, R.F., Roberts, R.G., Laslett, G.M., Yoshida, H., Olley, J.M., 1999. Optical dating of single and multiples grains of quartz from Jinmium rock shelter, northern Australia: Part I, experimental design and statistical models. Archaeometry 41, 339–364.

Gamble, C., 2013. Settling the Earth: The Archaeology of Deep Human History. Cambridge University Press, Cambridge.

Gliganic, L.A., Meyer, M.C., May, J.H., Aldenderfer, M.S., Tropper, P., 2021. Direct dating of lithic surface artifacts using luminescence. Sci. Adv. 7 (23), eabb3424. https://doi.org/10.1126/sciadv.abb3424.

Gopher, A., Barkai, R., 2014. Middle Paleolithic open-air industrial areas in the Galilee, Israel: The challenging study of flint extraction and reduction complexes. Quat. Int. 331, 96–102.

Guerin, G., Christophe, C., Philippe, A., Murray, A.S., Thomsen, K.J., Tribolo, C., Urbanova, P., Jain, M., Guibert, P., Mercier, N., Kreutzer, S., Lahaye, C., 2017. Absorbed dose, equivalent dose, measured dose rates, and implications for OSL age estimates: Introducing the Average Dose Model. Quat. Geochronol. 41, 163–173.

Guérin, G., Mercier, N., Adamiec, G., 2011. Dose-rate conversion factors: Update. Ancient TL 29, 5–8.

Guérin, G., Discamps, E., Lahaye, C., Mercier, N., Guibert, P., Turq, A., Dibble, H.L., McPherron, S.P., Sandgathe, D., Goldberg, P., Jain, M., Thomsen, K., Patou-Mathis, M., Castel, J.-C., Soulier, M.-C., 2012. Multi-method (TL and OSL), multi-material (quartz and flint) dating of the Mousterian site of Roc de Marsal (Dordogne, France): Correlating Neanderthal occupations with the climatic variability of MIS 5–3. J. Archaeol. Sci. 39, 3071–3084.

Guérin, G., Frouin, M., Talamo, S., Aldeias, V., Bruxelles, L., Chiotti, L., Dibble, H.L., Goldberg, P., Hublin, J.-J., Jain, M., Lahaye, C., Madelaine, S., Maureille, B., McPherron, S.P., Mercier, N., Murray, A.S., Sandgathe, D., Steele, T.E., Thomsen, K. J., Turq, A., 2015a. A multi-method luminescence dating of the Palaeolithic sequence of La Ferrassie based on new excavations adjacent to the La Ferrassie 1 and 2 skeletons. J. Archaeol. Sci. 58, 147–166.

Guérin, G., Combès, B., Lahaye, C., Thomsen, K.J., Tribolo, C., Urbanova, P., Guibert, P., Mercier, N., Valladas, H., 2015b. Testing the accuracy of a Bayesian central-dose model for single-grain OSL, using known-age samples. Radiat. Meas. 81, 62–70.

Guérin, G., Lahaye, C., Heydari, M., Autzen, M., Buylaert, J.P., Guibert, P., Jain, M., Kreutzer, S., Lebrun, B., Murray, A.S., Thomsen, K.J., Urbanova, P., Philippe, A., 2021. Towards an improvement of optically stimulated luminescence (OSL) age uncertainties: modelling OSL ages with systematic errors, stratigraphic constraints and radiocarbon ages using the R package BayLum. Geochronology 3 (1), 229–245.

Guibert, P., Schvoerer, M., 1991. TL dating: Low background gamma spectrometry as a tool for the determination of the annual dose. Int. J. Rad. Appl. Instrum. D 18, 231–238.

Harvati, K., Panagopoulou, E., Runnels, C., 2009. The Paleoanthropology of Greece. Evol. Anthropol. 18 (4), 131–143.

Harvati, K., Röding, C., Bosman, A.M., Karakostis, F.A., Grün, R., Stringer, C., Karkanas, P., Thompson, N.C., Koutoulidis, V., Moulopoulos, L.A., Gorgoulis, V.G., Kouloukoussa, M., 2019. Apidima Cave fossils provide earliest evidence of *Homo sapiens* in Eurasia. Nature 571 (7766), 500–504.

Hershkovitz, I., Weber, G.W., Quam, R., Duval, M., Grün, R., Kinsley, L., Weinstein-Evron, M., 2018. The earliest modern humans outside Africa. Science 359 (6374), 456–459.

Heydari, M., Guérin, G., 2018. OSL signal saturation and dose rate variability: Investigating the behaviour of different statistical models. Radiat. Meas. 120, 96–103.

Heydari, M., Guérin, G., Kreutzer, S., Jamet, G., Kharazian, M.A., Hashemi, M., Nasab, H. V., Berillon, G., 2020. Do Bayesian methods lead to more precise chronologies? 'BayLum'and a first OSL-based chronology for the Palaeolithic open-air site of Mirak (Iran). Quat. Geochronol. 59, 101082. https://doi.org/10.1016/j.quageo.2020.101082.

Heydari, M., Guérin, G., Zeidi, M., Conard, N.J., 2021. Bayesian luminescence dating at Ghār-e Boof, Iran, provides a new chronology for Middle and Upper Paleolithic in the southern Zagros. J. Hum. Evol. 151, 102926. https://doi.org/10.1016/j.jhevol.2020.102926.

Holcomb, J.A., Runnels, C., Wegmann, K.W., 2020. Deposit-centered archaeological survey and the search for the Aegean Palaeolithic: A geoarchaeological perspective. Quant. Intl. 550, 169–183.

Hublin, J.J., Ben-Ncer, A., Bailey, S.E., Freidline, S.E., Neubauer, S., Skinner, M.M., Gunz, P., 2017. New fossils from Jebel Irhoud, Morocco and the pan-African origin of Homo sapiens. Nature 546 (7657), 289–292.

Huntley, D.J., Baril, M.R., 1997. The K content of the K-feldspars being measured in optical dating or in thermoluminescence dating. Ancient TL 15, 11–13.

Jain, M., Ankjærgaard, C., 2011. Towards a non-fading signal in feldspar: Insight into charge transport and tunnelling from time-resolved optically stimulated luminescence. Radiat. Meas. 46, 292–309.

Jain, M., Murray, A.S., Botter-Jensen, L., 2004. Optically stimulated luminescence dating: How significant is incomplete light exposure in fluvial environments? [Datation par luminescence stimulée optiquement : quelle signification en cas de blanchiment incomplet des sédiments fluviatiles?]. Quaternaire 15, 143–157.

Jain, M., Sohbati, R., Guralnik, B., Murray, A.S., Kook, M., Lapp, T., Prasad, A.K., Thomsen, K.J., Buylaert, J.P., 2015. Kinetics of infrared stimulated luminescence from feldspars. Radiat. Meas. 81, 242–250.

Kaczanowksa, M., Kozlowski, J.K., Sobczyk, K., 2010. Upper Palaeolithic human occupations and material culture at Klissoura Cave 1. Eurasian Prehist. 7, 133–285.

Kars, R.H., Busschers, F.S., Wallinga, J., 2012. Validating post IR-IRSL dating on K-feldspars through comparison with quartz OSL ages. Quat. Geochronol. 12, 74–86.

Kars, R.H., Reimann, T., Ankjærgaard, C., Wallinga, J., 2014. Bleaching of the post-IR IRSL signal: New insights for feldspar luminescence dating. Boreas 43, 780–791.

Kopaka, K., Matzanas, C., 2009. Palaeolithic industries from the island of Gavdos, near neighbour to Crete in Greece. Antiquity Project Gallery 83 (321). http://www.antiquity.ac.uk/antiquityNew/projgall/kopaka321/.

Kreutzer, S., Martin, L., Guérin, G., Tribolo, C., Selva, P., Mercier, N., 2018. Environmental dose rate determination using a passive dosimeter: techniques and workflow for α-Al2O3: C chips. Geochronometria 45 (1), 56–67.

Kuhn, S.L., Arsebük, G., Howell, F.C., 1996. The Middle Pleistocene lithic assemblage from Yarimburgaz cave, Turkey. Paléorient 22, 31–49.

Lamothe, M., Auclair, M., Hamzaoui, C., Huot, S., 2003. Towards a prediction of long-term anomalous fading of feldspar IRSL. Radiat. Meas. 37, 493–498.

Lamothe, M., Brisson, L.F., Hardy, F., 2018. Dose recovery performance in double IRSL/pIRIR SAR protocols. Radiat. Meas. 120, 120–123.

Leppard, T.P., 2015. Passive dispersal versus strategic dispersal in island colonization by hominins. Current Anthropology 56 (4), 590–595.

Leppard, T.P., Runnels, C., 2017. Maritime hominin dispersals in the Pleistocene: advancing the debate. Antiquity 91 (356), 510–519.

Li, B., Jacobs, Z., Roberts, R., Li, S.-H., 2014. Review and assessment of the potential of post-IR IRSL dating methods to circumvent the problem of anomalous fading in feldspar luminescence. Geochron. 41, 178–201.

Lykousis, V., 2009. Sea-level changes and shelf break prograding sequences during the last 400 ka in the Aegean margins: subsidence rates and palaeogeographic implications. Cont. Shelf Res. 29 (16), 2037–2044.

Mortensen, P., 2008. Lower to Middle Palaeolithic artefacts from Loutró on the south coast of Crete. Antiquity Project Gallery 82 (317). http://www.antiquity.ac.uk/projgall/mortensen/.

Murray, A.S., Andrew, S., Wintle, A.G., 2000. Luminescence dating of quartz using an improved single-aliquot regenerative-dose protocol. Radiat. Meas. 32, 57–73.

Murray, A.S., Andrew, S., Olley, J.M., 2002. Precision and accuracy in the optically stimulated luminescence dating of sedimentary quartz: a status review. Geochronometria 21, 1–16.

Murray, A.S., Arnold, L.J., Buylaert, J.P., Guérin, G., Qin, J., Singhvi, A.K., Smedley, R., Thomsen, K.J., 2021. Optically stimulated luminescence dating using quartz. Nat. Rev. Methods Primers 1 (1), 1–31.

Murray, A.S., Thomsen, K.J., Masuda, N., Buylaert, J.P., Jain, M., 2012. Identifying well-bleached quartz using the different bleaching rates of quartz and feldspar luminescence signals. Radiat. Meas. 47 (9), 688–695.

Nian, X., Bailey, R.M., Zhou, L., 2012. Investigations of the post-IR IRSL protocol applied to single K-feldspar grains from fluvial sediment samples. Radiat. Meas. 47 (9), 703–709.

Perlès, C., 1987. Les Industries Lithiques Taillées de Franchthi (Argolide, Grèce) I. Indiana University Press, Bloomington, Présentation Générale et Industries Paléolithiques.

Philippe, A., Guérin, G., Kreutzer, S., 2019. BayLum-An R package for Bayesian analysis of OSL ages: An introduction. Quat. Geochronol. 49, 16–24.

Phoca-Cosmetatou, N., Rabett, R.J., 2014. Reflections on Pleistocene Island Occupation. JMA 27 (2), 255–259.

Porat, N., Faerstein, G., Medialdea, A., Murray, A.S., 2015. Re-examination of common extraction and purification methods of quartz and feldspar for luminescence dating. Ancient TL 33, 22–30.

Porat, N., Naomi, 2010. Abandonment ages of alluvial landforms in the hyperarid Negev determined by luminescence dating. J. Arid Environ. 74, 861–869.

Qin, J.T., Zhou, L.P., 2012. Effects of thermally transferred signals in the post-IR IRSL SAR protocol. Radiat. Meas. 47, 710–715.

Rees-Jones, J., 1995. Optical dating of young sediments using fine-grain quartz. Ancient TL 13 (2), 9–14.

Roberts, R.G., Galbraith, R.F., Yoshida, H., Laslett, G.M., Olley, J.M., 2000. Distinguishing dose populations in sediment mixtures: a test of single-grain optical dating procedures using mixtures of laboratory-dosed quartz. Radiat. Meas. 32, 459–465.

Runnels, C., DiGregorio, C., Wegmann, K.W., Gallen, S.F., Strasser, T.F., Panagopoulou, E., 2014. Lower Palaeolithic artifacts from Plakias, Crete: implications for hominin dispersals. Eurasian Prehistory 11 (1–2), 129–152.

Sakellariou, D., Galanidou, N., 2017. Aegean Pleistocene landscapes above and below sea-level: Palaeogeographic reconstruction and hominin dispersals. In: Bailey, G.N., Harff, J., Sakellariou, D. (Eds.), Under the Sea: Archaeology and Palaeolandscapes of the Continental Shelf. Springer, New York, pp. 335–359.

Sampson, A., Kaczanowska, M., Kozłowski, J.K., 2010. The Prehistory of the Island of Kythnos (Cyclades, Greece) and the Mesolithic Settlement at Maroulas. Polish Academy of Arts and Sciences, Kraków.

Séfériadès, M., 1983. Un centre industriel préhistorique dans les Cyclades: Les ateliers de débitage du silex à Stélida (Naxos). In: Rougement, G. (Ed.), Les Cyclades: Matériaux Pour Une Étude De Géographie Historique. CNRS, Lyon, pp. 67–73.

Singh, A., Thomsen, K.J., Sinha, R., Buylaert, J.P., Carter, A., Mark, D.F., Mason, P.J., Densmore, A.L., Murray, A.S., Jain, M., Paul, D., Gupta, S., 2017. Counter-intuitive influence of Himalayan river morphodynamics on Indus Civilisation urban settlements. Nat. Commun. 8 (1), 1–14.

Skarpelis, N., Carter, T., Contreras, D.A., Mihailović, D.D., 2017. Characterization of the siliceous rocks at Stélida, an early prehistoric lithic quarry (Northwest Naxos, Greece), by petrography and geochemistry: A first step towards chert sourcing. J. Archaeol. Sci. Rep. 12, 819–833.

Smedley, R.K., Duller, G.A.T., Roberts, H.M., 2015. Bleaching of the post-IR IRSL signal from individual grains of K-feldspar: Implications for single-grain dating. Radiat. Meas. 79, 33–42.

Smedley, R.K., Duller, G.A.T., Rufer, D., Utley, J.E.P., 2020. Empirical assessment of beta dose heterogeneity in sediments: implications for luminescence dating. Quat. Geochronol. 56, 101052.

Strasser, T.F., Panagopoulou, E., Runnels, C.N., Murray, P.M., Thompson, N., Karkanas, P., McCoy, F.W., Wegmann, K.W., 2010. Stone Age seafaring in the Mediterranean: evidence from the Plakias region for Lower Palaeolithic and Mesolithic habitation of Crete. Hesperia 79 (2), 145–190.

Strasser, T.F., Runnels, C., Wegmann, K., Panagopoulou, E., Mccoy, F., Digregorio, C., Karkanas, P., Thompson, N., 2011. Dating Palaeolithic sites in southwestern Crete. Greece. J. Quat. Sci. 26 (5), 553–560.

Thiel, C., Buylaert, J.-P., Murray, A., Terhorst, B., Hofer, I., Tsukamoto, S., Frechen, M., 2011. Luminescence dating of the Stratzing loess profile (Austria)—Testing the potential of an elevated temperature post-IR IRSL protocol. Quat. Int. 234, 23–31.

Thomsen, K.J., Murray, A.S., Bøtter-Jensen, L., 2005. Sources of variability in OSL dose measurements using single grains of quartz. Radiat. Meas. 39 (1), 47–61.

Thomsen, K.J., Murray, A.S., Jain, M., Bøtter-Jensen, L., 2008. Laboratory fading rates of various luminescence signals from feldspar-rich sediment extracts. Radiat. Meas. 43, 1474–1486.

Thomsen, K.J., Murray, A., Jain, M., 2012. The dose dependency of the over-dispersion of quartz OSL single grain dose distributions. Radiat. Meas. 47 (9), 732–739.

Thomsen, K.J., Murray, A.S., Buylaert, J.P., Jain, M., Hansen, J.H., Aubry, T., 2016. Testing single-grain quartz OSL methods using sediment samples with independent age control from the Bordes-Fitte rockshelter (Roches d'Abilly site, Central France). Quat. Geochronol. 31, 77–96.

Tourloukis, V., Harvati, K., 2018. The Palaeolithic record of Greece: a synthesis of the evidence and a research agenda for the future. Quat. Int. 466, 48–65.

Tourloukis, V., Karkanas, P., 2012. The Middle Pleistocene archaeological record of Greece and the role of the Aegean in hominin dispersals: new data and interpretations. Quat. Sci. Rev. 43, 1–15.

Trauerstein, M., Lowick, S.E., Preusser, F., Schlunegger, F., 2014. Small aliquot and single grain IRSL and post-IR IRSL dating of fluvial and alluvial sediments from the Pativilca valley. Peru. Quat. Geochronol. 22, 163–174.

Wintle, A.G., Murray, A.S., 2006. A review of quartz optically stimulated luminescence characteristics and their relevance in single-aliquot regeneration dating protocols. Radiat. Meas. 41, 369–391.

Yi, S., Buylaert, J.-P., Murray, A.S., Lu, H., Thiel, C., Zeng, L., 2016. A detailed post-IR IRSL dating study of the Niuyangzigou loess site in northeastern China. Boreas 45, 644–657.