

Homo floresiensis and *Homo luzonensis* are not temporally exceptional relative to *Homo erectus*

DAVID L. ROBERTS,¹  IVAN JARIĆ,^{2,3,4}  STEPHEN J. LYCETT,⁵ DYLAN FLICKER⁶  and ALASTAIR KEY^{6*} 

¹Durrell Institute of Conservation and Ecology, School of Anthropology and Conservation, University of Kent, Canterbury, Kent, CT2 7NR, UK

²Biology Centre of the Czech Academy of Sciences, Institute of Hydrobiology, České Budějovice, Czech Republic

³University of South Bohemia, Faculty of Science, Department of Ecosystem Biology, České Budějovice, Czech Republic

⁴Université Paris-Saclay, CNRS, AgroParisTech, Ecologie Systématique Evolution – IDEEV, 12 rue 128, 91190, Gif-sur-Yvette, France

⁵Department of Anthropology, University at Buffalo, SUNY, Amherst, NY, 14261, USA

⁶Department of Archaeology, University of Cambridge, Downing Street, Cambridge, CB2 3DZ, UK

Received 26 March 2022; Revised 30 November 2022; Accepted 5 December 2022

ABSTRACT: The presence of *Homo floresiensis* and *Homo luzonensis* in southeast Asia 90,000 to 60,000 years ago is considered surprising by many, and has been used to support their designation as unique species and the islands they were discovered on as refugia. Here, we statistically test the null hypothesis that *H. floresiensis* and *H. luzonensis* represent temporally uninterrupted occurrences relative to *Homo erectus*. We do this using the ‘surprise test’ for the exceptionality of a new record. Results demonstrate that *H. floresiensis* and *H. luzonensis* are not temporally distinct relative to *H. erectus*. Their late persistence should, therefore, not be considered surprising, they cannot reliably be inferred to be outside of *H. erectus*’ temporal range, and – temporally – the islands of Luzon and Flores are not supported as refugia. Similarly, late *H. erectus* at Ngandong, Java, is not demonstrated to be temporally distinct relative to earlier, principally mainland-Asian, *H. erectus*. Further, we demonstrate that substantial numbers of fossil discoveries would be needed before *H. floresiensis* and *H. luzonensis* are outside of *H. erectus*’ expected temporal range. If *H. floresiensis* and *H. luzonensis* are descended from *H. erectus* populations, our results point toward either geographic processes of allopatric speciation or behavioural processes leading to a sympatric speciation event. © 2023 The Authors *Journal of Quaternary Science* Published by John Wiley & Sons Ltd.

KEYWORDS: extreme order statistics; Flores; Gumbel domain of attraction; hominin; Luzon; Philippines; southeast Asia; species association

Introduction

Homo floresiensis and *H. luzonensis* represent exceptional fossil discoveries (Brown *et al.*, 2004; Détroit *et al.*, 2019). The discovery of both species was highly unexpected, they represent rare examples of southeast Asian fossil hominin occurrences, and both are anatomically unique. Indeed, on an anatomical basis, there are strong arguments for their classification as separate species relative to the *H. erectus* populations from which they potentially evolved (Jungers *et al.*, 2009; Aiello, 2010; Baab *et al.*, 2013; Détroit *et al.*, 2019; Zanolli *et al.*, 2022).

Distinguishing *H. floresiensis* and *H. luzonensis* from *H. erectus* in behavioural terms is more difficult. Both *H. floresiensis* and *H. erectus* are associated with expedient stone tool technologies, presumably for the butchery of animals and the processing of plant materials (Corvinus, 2004; Moore *et al.*, 2009; Brumm *et al.*, 2010, 2016; Dennell, 2018). Early *H. luzonensis* may also display similar behaviours (Ingicco *et al.*, 2020). Other behavioural evidence in Asia is sparse for all species, albeit with a couple of notable exceptions for *H. erectus* (Joordens *et al.*, 2015; Dennell, 2016), although such instances could easily be a result of varying search effort. Yet, anatomical differences between these three species would have impacted their respective energy budgets, manual capabilities,

locomotor efficiencies, cognitive processes and dietary emphases (Steudel-Numbers, 2006; Lucas *et al.*, 2008; Navarrete *et al.*, 2011; Zollikofer and de Leon, 2013; Key and Lycett, 2018). Thus, even in the absence of clear archaeological evidence, distinctions in at least some aspects of their behaviour were likely.

One difference between these species that has only been considered at a very broad level is their temporal presence. Indeed, despite the late occurrence of *H. floresiensis* and *H. luzonensis* (90–60 ka) being widely known and recognised for its importance (Aiello, 2010; Baab, 2016; Sutikna *et al.*, 2018; Détroit *et al.*, 2019; Tocheri, 2019), no work has investigated how this relates to the temporal occupancy of *H. erectus* in Asia and questions of population separation. These questions are even more pertinent after Sutikna *et al.*’s (2016) recently revised dating of *H. floresiensis* and Rizal *et al.*’s (2020) discovery that *H. erectus* persisted to c. 112.5 ka at Ngandong, Java. As such, the current state of the debate is that the two smaller hominin species are thought to have persisted for more than 40,000 years longer than *H. erectus*, but little is known about the implications of this knowledge for how the species are related, or whether this inference is reliable. Several fundamental questions remain unanswered about the temporal character of these three species, including: Given the fragmented nature of the Asian fossil record, is the late persistence of *H. floresiensis* and *H. luzonensis* considered unexpected or surprising? Can *H. floresiensis* and *H. luzonensis* be considered temporally

*Correspondence Alastair Key, as above.
Email: ak2389@cam.ac.uk

separate from *H. erectus* in the late Pleistocene? Does the late persistence of *H. floresiensis* and *H. luzonensis* support the notion that the islands of Flores and Luzon (respectively) acted as refugia for these populations?

A potential method of addressing these questions can be found in Solow and Smith's (2005) frequentist test for assessing the temporal exceptionality of new occurrences relative to a prior known sample. In effect, the method asks the question: how surprising is a new record? With 'surprising' in this instance being defined relative to α (i.e. if $p < \alpha$, a new occurrence can be considered surprising). This method has been previously used to test the exceptionality of early cases of COVID-19 (Roberts *et al.*, 2021), the persistence of the dodo on an inshore island off the coast of Mauritius (Roberts, 2013), exceptionality in athletic records (Solow & Smith, 2005), and the rediscovery of the polecat in Scotland (Solow *et al.*, 2006). However, in the context of the present study, the most relevant example is Solow and Smith's (2005) investigation of Palaeolithic art from Chauvet Cave. At the time it was the oldest elaborate cave art in the world (Valladas *et al.*, 2001) and had been dated to be around 5300 years older than the next oldest dated occurrence in southern Europe. On this basis alone – that is, its temporal presence relative to other cave art sites – it was considered a major discovery (Valladas *et al.*, 2001; Sadier *et al.*, 2012). Solow and Smith's (2005) analysis revealed that the surprise over the age of the artwork in Chauvet Cave was unwarranted. Here, we examine the exceptionality of the temporal relationship between *H. erectus* and its supposedly more recent relatives, *H. floresiensis* and *H. luzonensis*.

Methods

In a statistical test for the exceptionality of a new record, interest lies in testing the null hypothesis that the new record was generated by the same process that created the previous records (Solow and Smith, 2005). In the present scenario, this means testing the null hypothesis that the fossil occurrence for *H. floresiensis* at Liang Bua was generated by the same process responsible for the *H. erectus* records elsewhere in Asia (i.e. that they stem from the same overall *H. erectus* population). In effect, this means testing whether the population responsible for the *H. floresiensis* and *H. erectus* fossils can – at least in terms of their temporal associations – be considered the same. The same argument applies to *H. luzonensis* at Callao Cave and *H. erectus* fossil occurrences.

It was necessary to identify the youngest *H. erectus* fossil occurrences in Asia, along with the dates associated with *H. floresiensis* and *H. luzonensis* (Table 1). Due to the long temporal presence of *H. erectus* in the region (Swisher *et al.*, 1994; Morwood *et al.*, 2016; Matsu'ura *et al.*, 2020) we restricted our search effort to occurrences dated to less than one million years. This had the added benefit of keeping k within ranges where the test is known to have good explanatory power (Solow and Smith, 2005). Dating evidence from sites characterised only by artefacts could theoretically have been used; however, in the absence of strong species associations for stone tool and butchery sites in the region this would introduce analytical confound. The only exception was the marked *Pseudodon* shell from Trinil, which was included in a single modelled scenario due to its – albeit not conclusive – association with the *H. erectus* type specimens from the same location (Joordens *et al.*, 2015) and the presence of few other reliably dated *H. erectus* occurrences in southeast Asia. Several

other scenarios were investigated due to some fossil occurrences having weaker species associations or certainty in their dating. Fossil occurrences with 'highly ambiguous' or 'ambiguous but likely not *H. erectus*' associations were excluded. If there were no secure dates associated with fossils then these occurrences were also excluded. Sites could be included more than once so long as dates represented distinct fossil occurrences with date ranges that do not overlap.

Temporal data for *H. floresiensis* and *H. luzonensis* came from the most recent dating of their respective fossil sites (Sutikna *et al.*, 2016; Déroît *et al.*, 2019). Two dates were used for both *H. floresiensis* and *H. luzonensis*, meaning that two versions of the analyses were performed (see below). The first was the mean age of the dates associated with each species' fossils. For *H. luzonensis* this was the mean of the dates returned for the third metatarsal and maxillary tooth, as published by Grün *et al.* (2014) and Déroît *et al.* (2019). For *H. floresiensis*, this was the mean of the LB1, LB2 and LB6 ulnae, alongside the dated sediment samples from the fossil layer (notably, we do not include the dated sediment associated with the stone tools) (Sutikna *et al.*, 2016, 2018). To provide a scenario that investigated the minimum possible temporal gap with *H. erectus*, we also used the oldest dated individual sample for each species. In the case of *H. floresiensis* this was the 89 ka age for the fossil sediment, while for *H. luzonensis* it was the 66.7 ka metatarsal. The 700 ka *H. floresiensis* fossils from Mata Menge were not included in the present study as interest focuses on the late persistence of the species (van den Bergh *et al.*, 2016). Temporal overlap between early (700 ka) *H. floresiensis* and *H. erectus* is well evidenced (Fig. 1).

In total, five *H. erectus* occurrence scenarios were investigated here. The first four were tested for their temporal exceptionality relative to both of the *H. floresiensis* and *H. luzonensis* dates. The *H. erectus* scenarios are:

- 1) All *H. erectus* fossil occurrences that are secure in their species identification and their published dates can be considered reliable.
- 2) All *H. erectus* fossil occurrences that are secure in their species identification and their published dates can be considered reliable, with the exception of Sangiran and Ngandong on Java. This scenario tested occurrence exceptionality between *H. floresiensis*/*H. luzonensis* and mainland Asia *H. erectus* only.
- 3) All *H. erectus* fossil occurrences that are secure in their species identification and their published dates can be considered reliable. In addition, fossils with debated but likely *H. erectus* status and dates that are reasonable but not secure are also included.
- 4) Scenario 3 fossil occurrences with the addition of the dated *Pseudodon* shell from Trinil.

The fifth *H. erectus* scenario did not test against *H. floresiensis* or *H. luzonensis*.

- 5) Scenario 3 fossil occurrences. Here, instead of *H. floresiensis* or *H. luzonensis*, the latest dated *H. erectus* occurrence at Ngandong (Java) is examined for its temporal exceptionality relative to all other *H. erectus* sites.

The sites included in these model scenarios are presented in Table 1 and Fig. 1.

Solow and Smith's surprise method

Let $t_1 > t_2 > \dots > t_k$ be the k most recent specimens of *H. erectus* ordered from the most recent to the earliest specimen. The method of Solow and Smith (2005) assumes that these

Table 1. *Homo erectus* sites and dates forming the main species occurrence range in each of the five investigated fossil scenarios (i.e. without *H. floresiensis* and *H. luzonensis*, and in the case of Scenario 5, without the Ngandong *H. erectus* occurrence). Note that central values are used here and are derived from the published date ranges available in Supplementary Information 1.

Ranked occurrence		<i>Homo erectus</i> fossil scenarios				
		Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
1	Site	Ngandong, Java	Chaoxian, China	Ngandong, Java	Ngandong, Java ¹	Chaoxian, China
	Date BP	112,500	335,000	112,500	112,500	335,000
2	Site	Chaoxian, China	Hexian (Longtan), China	Chaoxian, China	Chaoxian, China ²	Hexian (Longtan), China
	Date BP	335,000	412,000	335,000	335,000	412,000
3	Site	Hexian (Longtan), China	Zhoukoudian, China	Hexian (Longtan), China	Hexian (Longtan), China ³	Zhoukoudian, China
	Date BP	412,000	450,000	412,000	412,000	450,000
4	Site	Zhoukoudian, China	Tangshan (Hulu), China	Zhoukoudian, China	Zhoukoudian, China ⁴	Tangshan (Hulu), China
	Date BP	450,000	577,000	450,000	450,000	577,000
5	Site	Tangshan (Hulu), China	Yiyuan, China	Tangshan (Hulu), China	Trinil, Java ⁵	Yiyuan, China
	Date BP	577,000	640,000	577,000	540,000	640,000
6	Site	Yiyuan, China	Zhoukoudian, China	Yiyuan, China	Tangshan (Hulu), China ⁶	Chenjiawo, China
	Date BP	640,000	770,000	640,000	577,000	650,000
7	Site	Zhoukoudian, China	Yunxian & Nankin, China	Chenjiawo, China	Yiyuan, China ⁷	Bailongdon, China
	Date BP	770,000	936,000	650,000	640,000	760,000
8	Site	Sangiran, Java		Bailongdon, China	Chenjiawo, China ⁸	Zhoukoudian, China
	Date BP	900,000		760,000	650,000	770,000
9	Site	Yunxian & Nankin, China		Zhoukoudian, China	Bailongdon, China ⁹	Sangiran, Java
	Date BP	936,000		770,000	760,000	900,000
10	Site			Sangiran, Java	Zhoukoudian, China ¹⁰	Yunxian & Nankin, China
	Date BP			900,000	770,000	936,000
11	Site			Yunxian & Nankin, China	Sangiran, Java ¹¹	
	Date BP			936,000	900,000	
12	Site				Yunxian & Nankin, China ¹²	
	Date BP				936,000	

¹Rizal *et al.* (2020); ²Shen *et al.* (2010); ³Grün *et al.* (1998); ⁴Shen *et al.* (2001); ⁵Joordens *et al.* (2015); ⁶Zhao *et al.* (2001); ⁷Guo *et al.* (2019); ⁸Zhisheng and Kun (2017); ⁹Liu *et al.* (2015); ¹⁰Shen *et al.* (2009); ¹¹Matsu'ura *et al.* (2020); ¹²de Lumley *et al.* (2008).

represent the k largest values of a larger collection of values generated from a distribution from the Gumbel domain of attraction. The Gumbel distribution represents a type of generalised extreme value distribution, used to model maximums and minimums of various distributions. Suppose that a more recent specimen is dated at time y , interest centres on assessing the exceptionality of this more recent specimen. Under the null hypothesis that the new case was generated by the same process as the earlier cases, Solow and Smith (2005) showed that the quantity,

$$S_k = \frac{y - t_1}{(y - t_1) + \sum_{j=1}^{k-1} (j+1)(t_j - t_{j+1})}, \quad (1)$$

has a β distribution with parameters 1 and $k-1$ so that the p -value corresponding to an observed value S_k is:

$$P = (1 - S_k)^{k-1}. \quad (2)$$

Simulations conducted by Solow and Smith (2005) assessed the adequacy of the asymptotic approximation to the p -value for finite n , and the effect of k on the power of the test. Simulations indicated adequate performance of the method for k of 10 (Solow and Smith 2005), which also was the average k value applied in the present study.

Main analysis: testing the known fossil record

First we applied the Solow and Smith (2005) method to central age estimates for all records (Table 1). However, since most of the *H. erectus* records are represented by age ranges, we also applied a resampling approach in order to address the uncertainty of the age estimates. Dates of all *H. erectus* specimens were randomly drawn from a normal distribution, using the central estimate as the mean value, and half of the difference between the central estimate and range bounds as the standard deviation. Such randomly generated datasets were assessed with the Solow and Smith (2005) method, and the whole procedure was repeated 10,000 times, with results expressed as mean values across all iterations. We performed the central age and resampling approaches for both sets of dates for *H. floresiensis* and *H. luzonensis*.

Secondary analysis: how many *H. erectus* fossil discoveries are needed to guarantee that *H. floresiensis* and *H. luzonensis* are temporally distinct?

Given the relatively sparse Middle-to-Late Pleistocene hominin fossil record in Asia, and the spacing of known occurrences, we anticipated that one or more scenarios may not return a

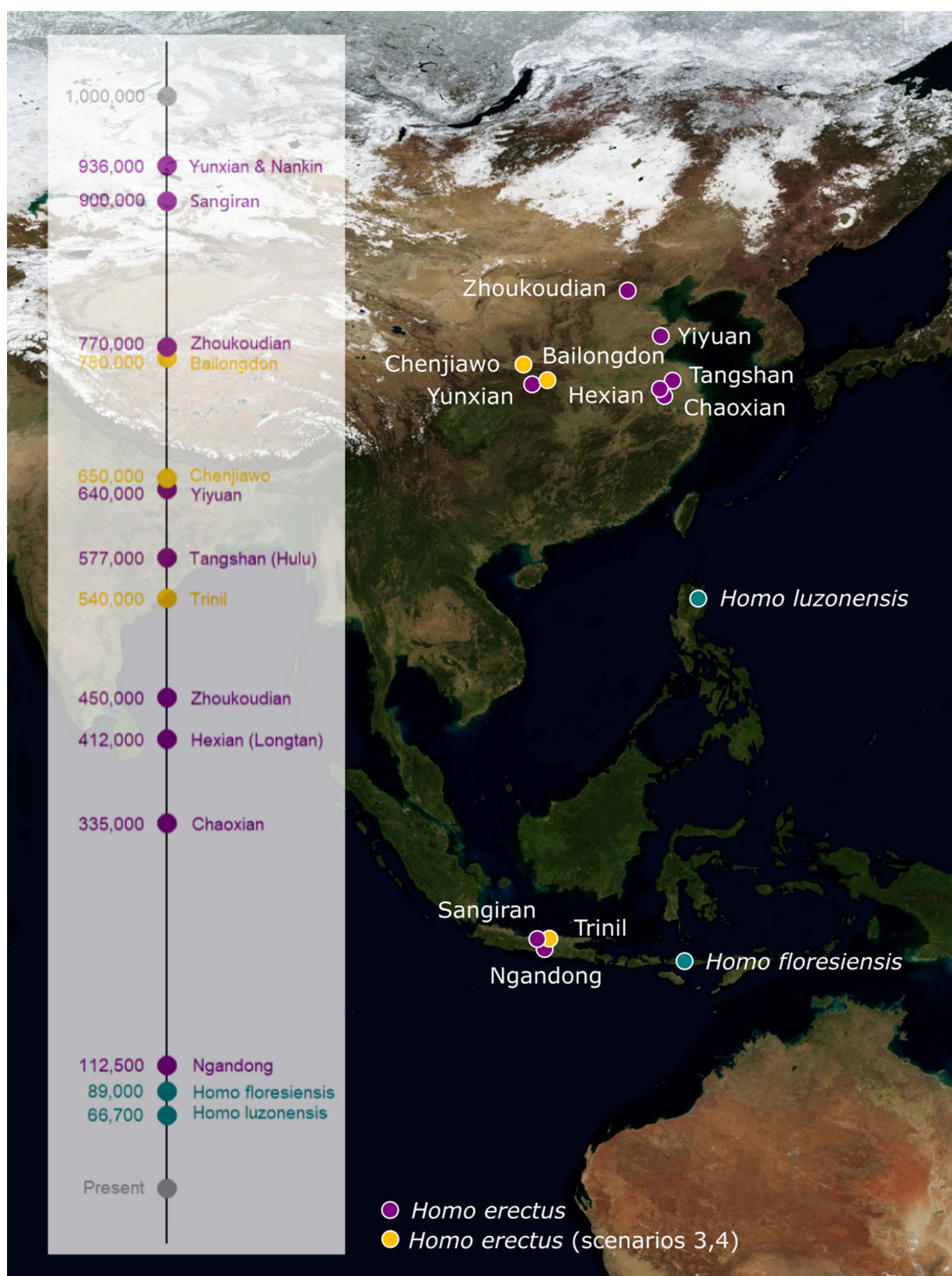


Figure 1. The location and age of all fossil and archaeological sites used in the present analysis. Note that these only represent occurrences dating to less than one million years ago. *Homo floresiensis* and *Homo luzonensis* dates represent the oldest dated individual samples for each species. [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

significant *p*-value. To help understand how a less porous fossil record – that is, a fossil record where increased effort had previously been put into discovering hominin fossils – could have influenced the main analysis' results, we performed a second set of analyses using additional, randomly discovered *H. erectus* fossils. That is, we created new randomly dated

H. erectus fossil occurrences for all five scenarios using a uniform distribution, and investigated how many of these 'discoveries' were needed before a significant result was guaranteed.

We made no changes to the temporal range of *H. erectus*, so the newly created fossil discoveries were bounded between

112.5 ka and 936 ka for Scenarios 1, 3 and 4, and between 335 ka and 936 ka for Scenarios 2 and 5 (Table 1). For each scenario, we randomly increased the number of additional records (k) by 1 and then re-ran the exceptionality test, with each new record being sampled from within their respective bounded range. We did this in a stepwise process, such that the test was re-run after one additional 'discovery' was made and we modelled up to 1000 new fossils being discovered. We only used central estimate data and did not run the resampling process outlined in the main analysis.

To avoid the effect of increased k and to focus only on the effect of increased density in *H. erectus* records, only the 10 most recent records were used in each re-run of the test. This meant that when additional fossil 'discoveries' were older than the existing tenth youngest record, the new record was not used in the analysis but the number of total records required still increased by 1. We performed 1000 iterations of this process for each k for each scenario, and identified the median resulting p -value among all iterations. From this, we identified which value of k would make the median p -value < 0.05 , or in other words which value of k resulted in $> 50\%$ of iterations displaying $p < 0.05$.

There are two limitations to these analyses that need to be highlighted. First, the newly added records were sampled from a uniform distribution, while the Solow and Smith (2005) method assumes that the record is characterised by a Gumbel distribution. Second, there is an inherent assumption that all new records fall within the current *H. erectus* temporal range, and not between the youngest *H. erectus* record and the tested 'outlier' record. Real-world fossil discoveries are unlikely to precisely match these two modelled fossil-discovery scenarios. All analyses were undertaken in R version 4.0.3 using code written for the present analyses (Supplementary Information 2).

Results

All modelled scenarios in the main analysis returned a p -value greater than α (i.e. 0.05) (Table 2), indicating that the late persistence of *H. floresiensis* and *H. luzonensis* is not surprising or unexpected relative to known *H. erectus* fossil occurrences. Scenario 2, which investigated *H. floresiensis* and *H. luzonensis* against mainland Asian *H. erectus* fossils was the closest to significance ($p = 0.584$ – 0.626 ; Table 2). Nor can the late Javan *H. erectus* occurrence at Ngandong be considered surprising relative to earlier, principally mainland Asian, *H. erectus* occurrences ($p = 0.577$ and 0.573 ; Table 2). In all scenarios, mean estimate and resampling versions of the tests were closely aligned in their results.

The number of new *H. erectus* fossil discoveries (k) required for *H. floresiensis*, *H. luzonensis* or *H. erectus* at Ngandong to be considered temporally surprising was substantial in all instances (Table 3). Scenario 2 displayed the lowest required number, ranging between 33 and 40 new fossils, depending on the dates used for *H. floresiensis* or *H. luzonensis*. Scenario 5 was similar, with an additional 46 *H. erectus* fossil discoveries being required before the occurrence at Ngandong is temporally distinct from earlier *H. erectus*. Scenarios 1, 3 and 4 required considerably greater numbers of new fossil discoveries, ranging from 304 to 700.

Discussion and conclusions

Three questions about the temporal character of late Pleistocene fossil hominin occurrences in Asia were posed in the introduction. Each has been addressed using Solow and

Table 2. Significance values for all modelled scenarios in the main analysis using Solow and Smith's (2005) surprise test for the exceptionality of a record relative to other dated occurrences. Central estimates – results based on central age estimates of all fossil records; resampling – mean values based on random sampling of record dates within age ranges.

Minimum possible temporal gap with <i>H. erectus</i>				
Scenario	<i>H. luzonensis</i> (66.7 ka)		<i>H. floresiensis</i> (89 ka)	
	Central estimates	Resampling	Central estimates	Resampling
1	0.915	0.914	0.955	0.955
2	0.601	0.584	0.626	0.609
3	0.905	0.904	0.950	0.949
4	0.904	0.902	0.949	0.949
5	Central estimates		Resampling	
	0.577		0.573	
Mean age of the dates associated with <i>H. floresiensis</i> and <i>H. luzonensis</i> fossils				
Scenario	<i>H. luzonensis</i> (58.4 ka)		<i>H. floresiensis</i> (71.4 ka)	
	Central estimates	Resampling	Central estimates	Resampling
1	0.901	0.900	0.924	0.923
2	0.592	0.574	0.606	0.589
3	0.889	0.888	0.914	0.913
4	0.888	0.886	0.913	0.912
5	Central estimates		Resampling	
	0.577		0.573	

Table 3. The number of additional *H. erectus* fossil discoveries (k) required before *H. floresiensis*, *H. luzonensis* and *H. erectus* at Ngandong can be considered to be temporally surprising.

Scenario	<i>H. luzonensis</i> (66.7 ka)	<i>H. floresiensis</i> (89 ka)	<i>H. luzonensis</i> (58 ka)	<i>H. floresiensis</i> (71 ka)
1	362	695	304	403
2	34	40	33	36
3	363	700	304	409
4	363	689	304	403
5	<i>H. erectus</i> at Ngandong			
	46			

Smith's (2005) test for assessing the temporal exceptionality of new discoveries relative to a prior known sample of dated occurrences. The results are unanimous in demonstrating that the late persistence of *H. floresiensis* and *H. luzonensis* should not be considered surprising or unexpected relative to the fossil record for *H. erectus* (Table 2). In other words, *H. floresiensis* and *H. luzonensis* are not outside the expected temporal range for *H. erectus*. The discovery of these small hominin species 90–60 ka does, therefore, fit within the wider temporal distribution established by *H. erectus* fossils in Asia.

Our results further suggest that – in terms of their temporal presence – these *H. floresiensis* and *H. luzonensis* occurrences cannot currently be considered sufficiently distinct from *H. erectus* for them to belong to separate populations. That is, if one were to discard anatomical arguments for their distinction (which we do not), we would be unable to refute the null hypothesis that the *H. floresiensis* and *H. luzonensis* fossils belong to the same temporal unit of fossil occurrences, and therefore

species, as *H. erectus*. Indeed, one way to define distinct populations or species of closely related organisms, including hominins (Du and Alemseged, 2019; Du *et al.*, 2020; Bobe and Wood, 2021), is via the temporal character of 'sightings'. In part, this is because genetic mutations accumulate over time; the greater the temporal distinction between populations, the greater the genetic distinction. If distinct clusters of 'sightings' (in this case fossil occurrences) can be temporally distinguished using statistical methods, it lends support to the notion that an earlier population belongs to a distinct species, may have gone extinct, may be heavily reduced in number, or have left a given region, and a wholly separate and/or genetically distinct population is evidenced at a later date. In the present context, if a significant temporal gap had been identified between the two smaller hominin species and *H. erectus*, this would lend support to their designation as temporally discrete populations. In turn, lending additional support to their designation as separate species. Our results therefore indicate that temporal evidence should not be used in addition to anatomical evidence when determining species-level differences between *H. erectus* and either *H. floresiensis* or *H. luzonensis*.

Our results indicate that all three populations may have co-occurred for a portion of time in the late Pleistocene of southeast Asia. This reduces the chance of the two smaller hominins having evolved by evolutionary processes involving temporal isolation, potentially hinting at a degree of interaction with *H. erectus* that continued until 90–60 ka. In the case of *H. floresiensis*, the potential for interaction with *H. erectus* may have been present for upwards of 650,000 years (van den Bergh *et al.*, 2016). Accordingly, our results (as informed by current fossil evidence) point toward either geographic isolation (allopatric speciation) or behavioural factors driving population divergence leading to sympatric speciation of *H. floresiensis* and *H. luzonensis* from *H. erectus*. Alternatively, our results would also be consistent with scenarios proposing that *H. floresiensis* and *H. luzonensis* did not descend from *H. erectus* but from some, as yet, unidentified fossil species (Jungers *et al.*, 2009; Baab, 2016; Détroit *et al.*, 2019; Tocheri, 2019).

There are two possible explanations for the lack of a clear temporal separation between these species in the late Pleistocene, which have implications for eventually determining which of these alternative evolutionary scenarios is correct. Potentially, the current Asian hominin fossil record accurately depicts the nature of each population's temporal range during the Pleistocene. That is, all three species were contemporaneous at the end of the Pleistocene, potentially being limited to their respective islands of Java, Flores and Luzon. Results from Scenario 5 further suggest that the late *H. erectus* fossil occurrence at Ngandong (Java) cannot currently be considered as a temporally distinct population relative to earlier, principally mainland Asian, *H. erectus*. Potentially indicating the concurrent presence of *H. erectus* in mainland Asia, and the Malay Archipelago during the Late Middle Pleistocene.

Alternatively, the Asian hominin fossil record may currently be too porous for us to have a reliable understanding of the temporal character of *H. erectus*. This may mean that although our present analyses present the most plausible scenario based on current evidence, this may change as additional discoveries are made. Indeed, as more fossils are discovered and dated we can expect gaps in the *H. erectus* fossil record to be filled in. If the gap to the next *H. erectus* occurrence prior to Ngandong is reduced substantially, it would increase the likelihood of the 20–50 ka gap to *H. floresiensis* and *H. luzonensis* being significant. For example, if four new sites aged between Ngandong and 130 ka were discovered, results would start to approach significance (for example, Scenario 1 returns $p=0.117$ and 0.072 for *H. floresiensis* and *H. luzonensis*, respectively, when using mean ages, central estimates and four

randomly sampled new sites between 112.5 and 130 ka). Should the sites of Sambungmacan and Ngawi ever be reliably dated to this period (they were excluded from this analysis due to the insecurity of their ages (Yokoyama *et al.*, 2008; Joordens *et al.*, 2015)), then such a scenario may be realised. In turn, this would provide more robust grounds to suggest that the two diminutive species persisted to a date that was unexpected relative to the temporal range of *H. erectus*.

To help understand whether additional fossil discoveries may overturn our current inferences (assuming no changes to the present temporal range for *H. erectus*), we modelled the discovery of randomly dated *H. erectus* fossil occurrences within existing temporal boundaries. We did this in a stepwise process, such that the surprise test was re-run after one additional fossil 'discovery' was made. This was repeated until a significant p -value was returned and *H. floresiensis*, *H. luzonensis*, or *H. erectus* at Ngandong, could be stated to be temporally distinct from the main *H. erectus* occurrences. We ran this process 1000 times for each k (record) within each scenario, and identified which value of k would make the median p -value <0.05 . Using this technique, another 304–700 *H. erectus* fossil discoveries were needed before *H. floresiensis* and *H. luzonensis* (respectively) could be considered to be temporally distinct from *H. erectus*' current temporal range. The exceptionally high numbers required reinforces our above inference and suggests that, given current discovery rates, these three species may never display clear evidence in support of their temporal distinction. The latter point being strengthened by future *H. erectus* discoveries potentially being dated to <112.5 ka (which we did not model).

The temporal discontinuity between mainland (or principally mainland) Asian *H. erectus* and *H. floresiensis*, *H. luzonensis* and *H. erectus* at Ngandong is greater (Scenarios 2 and 5), with only 33 to 46 new fossil discoveries required. Nonetheless, we consider this to be a high number – relative to current discovery rates – and that at present the fossil record suggests the late persistence of all three species in the Malay Archipelago to be within the expected temporal range of *H. erectus* in mainland Asia. Our fossil-discovery simulations emphasise two key points. First, the Asian hominin fossil record is currently highly porous, with a substantial number of discoveries required before *H. erectus* fossil sites – without additional modelling – can be considered to display a reliably defined 'origin' or 'extinction' date. Secondly, these simulations demonstrate how unreliable it currently is to suggest that *H. floresiensis* and *H. luzonensis* are temporally distinct relative to *H. erectus*. Although this does not mean that more recent (younger) *H. floresiensis* and *H. luzonensis* fossils will not be discovered in the future, and if the present *H. erectus* range does not also change, such younger finds could yet demonstrate the former to be outside of the expected temporal range of the latter (i.e. the relative gap between *H. erectus* and the new, younger *H. floresiensis* or *H. luzonensis* would increase).

To be clear, here we do not demonstrate that *H. erectus* potentially survived to 90–60 ka; other techniques would be required to suggest this (Roberts and Solow, 2003; Du *et al.*, 2020; Key *et al.*, 2021). What the surprise test has done is demonstrate that, given the current fossil record, it would not be surprising if *H. erectus* fossils are observed at these dates in the future. We infer this to mean that currently, *H. floresiensis* and *H. luzonensis* cannot reliably be stated to have persisted later than *H. erectus*. Faith *et al.* (2021) discuss the implications this could have on our understanding of their behavioural and biological evolution at length. Our results reinforce Bobe and Wood's (2021) recommendation that we should move away from using the earliest of latest known fossils as a fixed boundary for understanding the emergence, extinction, presence and absence of hominin populations.

Finally, the late occurrence of *H. floresiensis* and *H. luzonensis* should not be used to support the hypothesis that the islands of Flores and Luzon (respectively) acted as refugia for these species. It does not mean that Flores and Luzon were not refugia (from ecological change in other areas of Asia, or other hominin species), but simply, that temporal data do not currently support this inference. This conclusion follows the same logic as above, insofar as the *H. floresiensis* and *H. luzonensis* fossil occurrences are feasibly generated by the same processes as those that led to other fossil hominin occurrences in Java and China. Geographic, topographical and ecological factors – including the presence of the Wallace Line – may still support the suggestion that the islands (and wider region) acted as refugia (Louys and Turner, 2012; Stewart and Stringer, 2012; Dennell, 2017; Zachwieja *et al.*, 2020).

To our knowledge, this study presents the first application of Solow and Smith's (2005) temporal exceptionality test within palaeoanthropology. In a discipline where fossil discoveries can be sensationalised as the earliest and latest of their kind, the surprise test provides a route for assessing whether any temporally based fanfare is warranted. More importantly, the method provides a way to assess population structure in species and genera with temporally fragmented fossil discoveries. The surprise test therefore joins an increasing number of statistical methods that can improve our understanding of when and where different species were present using existing hominin fossil occurrences (Du *et al.*, 2020; Bobe and Wood, 2021; Deviese *et al.*, 2021; Key *et al.*, 2021; Djakovic *et al.*, 2022).

To conclude, the present study demonstrates that temporal spacing in the current Asian fossil hominin record does not reliably indicate *H. luzonensis* or *H. floresiensis* persisted outside of the expected temporal range for *H. erectus*. In other words, Liang Bua and Callao Cave are the most recent non-modern human fossil occurrences in the region (Sutikna *et al.*, 2016; Détroit *et al.*, 2019), but given how fragmented the *H. erectus* fossil record is, it cannot reliably be used to infer temporal distinctions between species. Our results point toward either geographic processes of allopatric speciation or change in behavioural processes leading to a sympatric speciation event if *H. floresiensis* and *H. luzonensis* are descended directly from *H. erectus* populations. Alternatively, our results would also be consistent with scenarios suggesting an alternative fossil ancestry for *H. floresiensis* and *H. luzonensis*. Should additional fossils come to light in the future, then it may be possible to determine which of these scenarios is more likely, but our analysis shows this may require a substantial number of discoveries.

Supporting information

Additional supporting information can be found in the online version of this article.

References

- Aiello LC. 2010. Five years of *Homo floresiensis*. *American Journal of Physical Anthropology* **142**(1): 167–179.
- Baab KL. 2016. The place of *Homo floresiensis* in human evolution. *Journal of Anthropological Sciences* **94**: 5–18.
- Baab KL, McNulty KP, Harvati K. 2013. *Homo floresiensis* contextualized: A geometric morphometric comparative analysis of fossil and pathological human samples. *Plos One* **8**(7): e69119.
- Bobe R, Wood B. 2021. Estimating origination times from the early hominin fossil record. *Evolutionary Anthropology* **31**(2): 92–102.
- Brown P, Sutikna T, Morwood MJ, *et al.* 2004. A new small-bodied hominin from the Late Pleistocene of Flores, Indonesia. *Nature* **431**: 1055–1061.
- Brumm A, Moore MW, van den Bergh GD, Kurniawan I, Morwood MJ, Aziz F. 2010. Stone technology at the Middle Pleistocene site of Mata Menge, Flores, Indonesia. *Journal of Archaeological Science* **37**(3): 451–473.
- Brumm A, van den Bergh GD, Storey M, *et al.* 2016. Age and context of the oldest known hominin fossils from Flores. *Nature* **534**: 249–253.
- Corvinus G. 2004. *Homo erectus* in East and Southeast Asia, and the questions of the age of the species and its association with stone artifacts, with special attention to handaxe-like tools. *Quaternary International* **177**(1): 141–151.
- Dennell R. 2016. Life without the Movius Line: The structure of the East and Southeast Asian Early Palaeolithic. *Quaternary International* **400**: 14–22.
- Dennell R. 2017. Human colonization of Asia in the Late Pleistocene: the history of an invasive species. *Current Anthropology* **58**(s17): s383–s396.
- Dennell R. 2018. The Acheulean assemblage of Asia. In *The Emergence of the Acheulean in East Africa and Beyond*, Gallotti R, Mussi M (eds). Springer, Cham: 195–214.
- Détroit F, Mijares AS, Corny J, *et al.* 2019. A new species of *Homo* from the Late Pleistocene of the Philippines. *Nature* **568**: 181–186.
- Deviese T, Abrams G, Hajdinjak M, *et al.* 2021. Reevaluating the timing of Neanderthal disappearance in Northwest Europe, *PNAS* **118**(12) e2022466118.
- Djakovic I, Key A, Soressi M. 2022. Optimal linear estimation models predict 1400–2800 years of co-existence between Neandertals and *Homo sapiens* in western Europe. *Scientific Reports* **12**: 15000.
- Du A, Alemseged Z. 2019. Temporal evidence shows *Australopithecus sediba* is unlikely to be the ancestor of *Homo*. *Science Advances* **5**(5) eaav9038.
- Du A, Rowan J, Wang SC, Wood BA, Alemseged Z. 2020. Statistical estimates of hominin origination and extinction dates: A case study examining the *Australopithecus anamensis*–*afarensis* lineage. *Journal of Human Evolution* **138**: 102688.
- Grun R, Huang P-H, Huang W, *et al.* 1998. ESR and U-series analyses of teeth from the palaeoanthropological site of Hexian, Anhui Province, China. *Journal of Human Evolution* **34**(6): 555–564.
- Ingicco T, Reyes MC, de Vos J, *et al.* Taphonomy and chronosequence of the 709 ka Kalinga site formation (Luzon Island, Philippines). *Scientific Reports*, **10**: 11081.
- Joordens JCA, d'Errico F, Wesselingh FP, *et al.* 2015. *Homo erectus* at Trinil on Java used shells for tool production and engraving. *Nature* **518**: 228–231.
- Jungers WL, Larson SG, Harcourt-Smith W, Morwood MJ, Sutikna T, Awe RD, Djubiantono T. 2009. Descriptions of the lower limb skeleton of *Homo floresiensis*. *Journal of Human Evolution* **57**(5): 538–554.
- Key AJM, Lycett SJ. 2018. Investigating interrelationships between Lower Palaeolithic stone tool effectiveness and tool user biometric variation: implications for technological and evolutionary changes. *Archaeological and Anthropological Sciences* **10**: 989–1006.
- Key A, Roberts D, Jarić I. 2021. Reconstructing the full temporal range of archaeological phenomena from sparse data. *Journal of Archaeological Science* **135**: 105479.
- Liu X, Shen G, Tu H, Lu C, Granger DE. 2015. Initial 26Al/10Be burial dating of the hominin site Bailong Cave in Hubei Province, central China. *Quaternary International* **389**: 235–240.
- Louys J, Turner A. 2012. Environment, referred habitats and potential refugia for Pleistocene *Homo* in Southeast Asia. *Comptes Rendus Palevol* **11**(2–3): 203–311.
- Lucas PW, Constantino PJ, Wood BA. 2008. Inferences regarding the diet of extinct hominins: structural and functional trends in dental and mandibular morphology within the hominin clade. *Journal of Anatomy* **212**(4): 486–500.
- de Lumley H, Hkati B, Mestour B, *et al.* 2008. Étude du paléomagnétisme du site de l'homme de Yunxian. Le site de l'homme de Yunxian, Quyanhekou, Quingqu, Yunxian, province du Hubei. *CNRS Editions* **589**: 2008.

- Matsu'ura S, Kondo M, Danhara T, *et al.* 2020. Age control of the first appearance datum for Javanese *Homo erectus* in the Sangiran area. *Science* **367**(6474): 210–214.
- Moore MW, Sutikna T, Jatmiko W, Morwood MJ, Brumm A. 2009. Continuities in stone flaking technology at Liang Bua, Flores, Indonesia. *Journal of Human Evolution* **57**(5): 503–526.
- Morwood MJ, O'Sullivan P, Susanto EE, Aziz F. 2016. Revised age for Mojokerto 1, an early *Homo erectus* cranium from East Java, Indonesia. *Australian Archaeology* **57**(1): 1–4.
- Navarrete A, van Schaik CP, Isler K. 2011. Energetics and the evolution of human brain size. *Nature* **480**: 91–93.
- Rizal Y, Westaway KE, Zaim Y, *et al.* 2020. Last appearance of *Homo erectus* at Ngandong, Java, 117,000–108,000 years ago. *Nature* **577**: 381–385.
- Roberts DL, Solow AR. 2003. When did the Dodo become extinct? *Nature* **426**: 245.
- Roberts DL. 2013. Refuge-effect hypothesis and the demise of the dodo. *Conservation Biology* **27**: 1478–1480.
- Roberts DL, Rossman JS, Jarić I. 2021. Dating first cases of COVID-19. *PLoS Pathogens* **17**(6): e1009620.
- Sadier B, Delannoy J-J, Benedetti L, *et al.* 2012. Further constraints on the Chauvet cave artwork elaboration. *PNAS* **109**(21): 8002–8006.
- Shen G, Ku T-L, Cheng H, Edwards RL, Tuan Z, Wang Q. 2001. High-precision U-series dating of Locality 1 at Zhoukoudian, China. *Journal of Human Evolution* **41**(6): 679–688.
- Shen G, Gao X, Gao B, Granger DE. 2009. Age of the Zhoukoudian *Homo erectus* determined with ²⁶Al/¹⁰Be burial dating. *Nature* **458**: 198–200.
- Shen G, Fang Y, Bischoff JL, Feng Y-X, Zhao J-X. 2010. Mass spectrometric U-series dating of the Chaoxian hominin site at Yinshan, eastern China. *Quaternary International* **211**(1–2): 24–28.
- Solow AR, Kitchener AJ, Roberts DL, Birks JDS. 2006. Rediscovery of the Scottish polecat, *Mustela putorius*: Survival or reintroduction? *Biological Conservation* **128**: 574–575.
- Solow AR, Smith WK. 2005. How surprising is a new record? *The American Statistician* **59**: 153–155.
- Stewart JR, Stringer CB. 2012. Human evolution out of Africa: the role of refugia and climate change. *Science* **335**: 1317–1321.
- Steudel-Numbers KL. 2006. Energetics in *Homo erectus* and other early hominins: The consequences of increased lower-limb length. *Journal of Human Evolution* **51**(5): 445–453.
- Sutikna T, Tocheri MW, Faith JT, *et al.* 2018. The spatio-temporal distribution of archaeological and faunal finds at Liang Bua (Flores, Indonesia) in light of the revised chronology for *Homo floresiensis*. *Journal of Human Evolution* **124**: 52–74.
- Sutikna T, Tocheri MW, Morwood MJ, *et al.* 2016. Revised stratigraphy and chronology for *Homo floresiensis* at Liang Bua in Indonesia. *Nature* **532**: 366–369.
- Swisher CC, Curtis GH, Jacob T, Getty AG, Suprijo A, Widiasmoro S. 1994. Age of the earliest known hominids in Java, Indonesia. *Science* **263**(5150): 1118–1121.
- Tocheri MW. 2019. Previously unknown human species found in Asia raises questions about early hominin dispersals from Africa. *Nature* **568**: 176–178.
- Valladas H, Clottes J, Geneste J-M, *et al.* 2001. Evolution of prehistoric cave art. *Nature* **413**: 479.
- Van den Bergh G, Kaifu Y, Kurniawan I, *et al.* 2016. *Homo floresiensis*-like fossils from the early Middle Pleistocene of Flores. *Nature* **534**: 245–248.
- Yokoyama Y, Falgueres C, Semah F, Jacob T, Grun R. 2008. Gamma-ray spectrometric dating of late *Homo erectus* skulls from Ngandong and Sambungmacan, Central Java, Indonesia. *Journal of Human Evolution* **55**(2): 274–277.
- Zachwieja AJ, Bacon A-M, Nguyen TMH, *et al.* 2020. Understanding Late Pleistocene human land preference using ecological niche models in an Australasian test case. *Quaternary International* **563**: 13–28.
- Zanolli C, Kaifu Y, Pan L, *et al.* 2022. Further analyses of the structural organization of *Homo luzonensis* teeth: Evolutionary implications. *Journal of Human Evolution* **163**: 103124.
- Zhao J-X, Collerson KD, Xu H-K. 2001. Thermal ionization mass spectrometry U-series dating of a hominid site near Nanjing, China. *Geology* **29**(1): 27–30.
- Zhisheng A, Kun HC. 2017. New magnetostratigraphic dates of Lantian *Homo erectus*. *Quaternary Research* **32**(2): 213–221.
- Zollikofer CPE, de Leon MSP. 2013. Pandora's growing box: Inferring the evolution and development of hominin brains from endocasts. *Evolutionary Anthropology* **22**(1): 20–33.