



The Southeast Asian prehistoric house: a correlation between archaeology and linguistics

C. F. W. Higham¹ · M. J. Alves²

Received: 15 August 2024 / Revised: 14 April 2025 / Accepted: 22 April 2025 / Published online: 29 April 2025
© The Author(s) 2025

Abstract

Austroasiatic languages are spoken from Southeast Asia to central India, and their divergences evidence a considerable passage of time since speakers of the proto language spread south from the lower Yangtze region. Proto-Austroasiatic etyma for the house and its component parts have raised the possibility of their validation through matching archaeological data. It is found that there is a correlation between linguistic, genetic and archaeological evidence that involved the expansion of Neolithic rice farmers who brought with them a cultural package of agriculture, pottery, domesticated animals and housing structures, and who integrated with the indigenous hunter gatherer communities they encountered.

Keywords Austroasiatic languages · The house · Neolithic expansion

1 Introduction

The conjunction of linguistics, archaeology, and genetics has contributed to resolving major events in the human past, including the expansion of farmers, dispersals that “constitute collectively the most important process in Holocene human history” (Diamond and Bellwood 2003). In the case of Indo-European origins, Heggarty et al. (2023) have identified and dated two epicentres for the origin and expansion of proto Indo-European speakers, south of the Caucasus, followed by a movement into the Pontic-Caspian steppes. This paper involves a similar synthesis from the three disciplines that consider the Austroasiatic (AA) family of languages that are distributed from India to Việt Nam, extending north into China. Academic interest in AA began in the mid nineteenth century (Mason 1854; Sidwell 2021), and 13 or possibly 14 major branches are recognized, comprising over 150 languages (Sidwell 2021; Eberhard et al. 2024).

Vietic and Khmeric are the national languages of Việt Nam and Cambodia. Mon, once spoken widely in Thailand and Myanmar, is now restricted to small enclaves. All minority languages are confined in distribution, due to later incursions by the speakers of now dominant languages. Katuic and Bahnaric are spoken in central to southern Việt Nam and bordering parts of Laos, and moving in a westerly direction, there are the Khmuic, Palaungic, and Khasic branches. Distant outliers are found in the Munda languages of northeast and central India, Aslian in Malaysia and Nicobarese on the Nicobar islands (Fig. 1). This language family has been hypothesized to have origins in southern China, but a dispersal locus in northern Việt Nam (Sidwell 2021). Higham (2002) has reviewed cognate AA words for rice and aspects of its cultivation, concluding that the distribution of AA languages today is consistent with the fifth millennium BP southerly diaspora of rice farmers from the Yangtze River region into Southeast Asia. He concluded that the linguistic evidence from AA cognates for rice supported the archaeological evidence for the expansion of rice farmers needed further testing, and one means of doing so was through archaeogenetics. This has been achieved: the incoming farmers were genetically related to ancestors in the lower Yangtze region (McColl et al. 2018; Lipson et al. 2018).

Alves (2023) has now provided an opportunity to explore this issue further by assembling recently re-evaluated proto-AA (pAA) etyma for domestic structures. These have been vetted by Sidwell and Alves (2023) to determine those which

✉ C. F. W. Higham
charles.higham@otago.ac.nz

M. J. Alves
mark.alves@montgomerycollege.edu

¹ Archaeology Programme, University of Otago, Dunedin, New Zealand

² Montgomery College, Maryland, USA

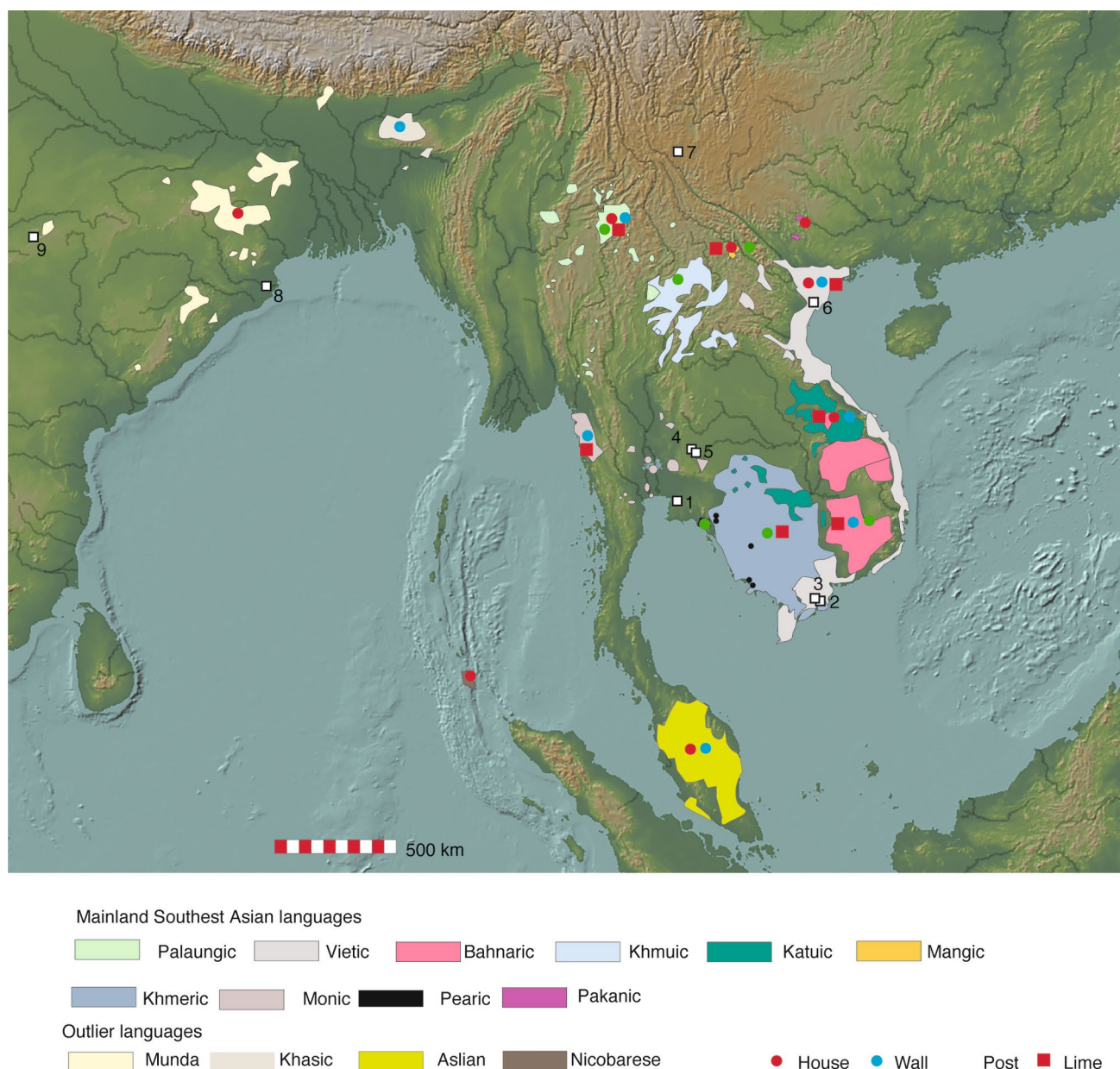


Fig. 1 Map showing the location of the main branches of Austroasiatic languages, the presence of key cognates for the house and sites mentioned in the text. 1. Khok Phanom Di, 2. Rạch Núi, 3. Bưng Bắc, 4. Non Ban Jak, Non Muang

Kao, 5. Noen U-Loke, Ban Non Wat, 6. Đông Sơn, 7. Shizhaishan, 8. Mahanadi Delta, 9. Satpura Hills, 10. Đền Đò, 11. Xóm Rền. (Figure by C.F.W. Higham, courtesy GeoMapApp (www.geomapp.org), CC by Ryan et al. (2009))

can be validly considered to connect to an original pAA period. Later lexical innovations that could have spread through borrowing have been excluded from this study. This approach relies on attestations between the core group of MSEA languages, and the remote, external branches located in India, the Nicobar Islands and peninsular Malaysia, because the isolation of the latter branches rules out later borrowings.

In this context, it is important to identify whether or not the etyma for the domestic buildings found in pAA were

present in the Neolithic settlements of the donor region in the lakeland of the middle course of the Yangtze River, and on the low-lying terrain bordering Hangzhou Bay, where Hemudu has been intensively analysed (Sun 2013). Houses there were raised on wooden posts and beams. Woven reed matting was recovered, and loom components, along with spindle whorls and rope, provide evidence of weaving. There were foundations for fenced areas, with corresponding lexical artefacts in AA languages referring to fences and to the act of making fenced enclosures. Rectangular houses raised

on wooden posts and fencing have also been identified at Tianluoshan, dating between 6500 and 7000 BP (Fig. S11; ZPICA et al. 2008, Li 2009), and matting was recovered from Jingtoushan a millennium earlier still (ZPICA et al. 2022). While such early evidence predates pAA, it likely contributed to cultural elements represented by pAA etyma. Rice was abundant in these sites, as were the remains of domestic pigs and dogs, components of the diet of those who emigrated into Southeast Asia.

As Alves has made clear, “For a historical-linguistic study that involves a cultural domain, archaeohistorical information is necessary to check the validity of such reconstructions and to consider the sociocultural context of such words.” (Alves 2023:3). The question is, can reconstructed proto-language words be identified and documented archaeologically? In his article, Alves concentrates on the evidence from the few documented early Neolithic occupation sites in Mainland Southeast Asia. However, similarities in material culture and the analysis of isotopes in human teeth from Neolithic to Iron Age sites in the region of the upper Mun River indicate that there was continuity of settlement with no instance of inward migration (King et al. 2015). This investigation will, therefore, incorporate all prehistoric evidence for domestic buildings while at the same time, stressing that regardless of later periods, the vetted pAA etyma provide strong support for their dating back to the Neolithic AA dispersal, now placed in the 5th millennium BP, because the lexical reconstructions reviewed below could not have been spread to the Munda languages in India or the languages of the Nicobar Islands in later periods.

2 Proto austroasiatic etyma for household-residential structures

In this review, we adopt Sidwell and Alves’s (2023) approach of distinguishing between AA branches in Mainland Southeast Asia (MSEA) and the external ones, namely, Khasic, Munda, Aslian and Nicobarese. Lexical attestations in the latter demonstrate the time-depth of these words: thus, the words in this study must have been brought with pAA speakers in a very early period. There is no historical phonological means of determining ages of the AA branches, but the dramatic typological restructuring of Munda and Nicobarese indicate likely substantial time-depth. Rau and Sidwell (2019) have combined linguistic, archaeological and genomic evidence to support a hypothesis that Munda speakers made a maritime crossing over the Bay of Bengal from Southeast Asia, to make landfall in the region of the Mahanadi Delta before expanding inland over time as far west as the Satpura Hills of Maharashtra. They place this migration ca. 4000–3500 BP. Archaeologically, this is entirely tenable. There is little doubt that pAA speakers were migrating down

the coast of Việt Nam in the 5th millennium BP, and the Neolithic settlement Khok Phanom Di, located on the estuary of the Bang Pakong River in Central Thailand, is a likely further extension of this movement. This site is now dated from ca. 4300 BP, and one key finding from male skeletal remains is the development of musculature compatible with managing watercraft (Tayles 1999). Moreover, the Nicobar Islands could only have been colonised by sea, meaning the Nicobaric branch is concrete evidence of long-distance travel over water.

The Southeast Asian origins of the Munda are also supported by the evidence from DNA. The Y chromosome haplogroup O2a is found in both Munda and Southeast Asian AA speakers. This Southeast Asian component in Munda speakers supports a dispersal event from Southeast Asia followed by admixture with South Asians (Chaubey et al. 2011). Tätte et al. (2019) have analysed Munda and comparative DNA, again finding that the former have significant Southeast Asian ancestry, with a time depth since admixture of about 4000 years. This is virtually identical with the archaeological evidence. In the absence of significant maritime contact after the initial migration into eastern India, any words shared between the MSEA and the external branches must evidence early cognacy.

We now consider the pAA reconstructions for terms for the house and related structural elements and consider relevant archaeological studies that provide both supporting evidence to date these etyma to the time of the MSEA Neolithic expansion and lexical evidence of these early practices in pAA culture (See SI Table 1, which contains a complete list of the reconstructed Austroasiatic words (from Sidwell 2024) referred to in this study). While the pAA lexical reconstructions are of a general nature, together, they make up a cultural domain, so all the available items must be considered. While some elements, such as ‘wall’ and ‘post’, seem to be obvious elements, lexical gaps in the reconstructed system would weaken cultural reconstructions. That they are available permits the association of this proto-language lexical system with a past cultural package.

3 HOUSE (‘house’ **ja:ʔ* and ‘home, clan territory’ **du:ŋ*)¹

Central to these etyma is the word for a house. Cognates for **ja:ʔ* ‘house’ are identified in five branches of AA in mainland Southeast Asia, and three further afield (Fig. 1).

¹ The asterisk is a standard notation for historical linguistic reconstructions. Asterisks mark word forms that have been reconstructed based on recurring phonological patterns evident in various languages and/or textual evidence. An asterisk distinguishes a reconstruction from their modern attestations.

The second etymon, *đu:ŋ ‘house, clan territory’, is present in Katuic, Khmuic, Monic and Pearic within MSEA, and in Aslian outside that region. While most branches have only one or the other etymon, both are found in Katuic and possibly Aslian, suggesting that both could have been in pAA, albeit with slightly different semantic scopes, as *đuŋ/*đu:ŋ has an additional sense of ‘clan territory.’ Whether this cultural feature of that etymon can be connected to early pAA culture would need to be considered through additional ethnographic investigation.

The archaeological investigation of houses in Southeast Asia has until recently been dominated by the notion that they were constructed on posts raised above ground level (Carter 2023). This has been founded in part on the illustrations of such houses on Đông Sơn Iron Age drums in Việt Nam, and house models from the Dian chiefdoms of Yunnan (Higham 1996 figs. 4.29–32; 5.18). One reason for this practice projected back into prehistory is that raised houses obviated the risks of flooding from monsoon rains. It is certainly the case that post holes are a common characteristic of prehistoric settlements, and most excavations have covered too small an area to identify possible building plans,

Early Neolithic sites, however, have provided convincing evidence for houses built at ground level. The Phùng Nguyên culture sites of the lower reaches of the Red River in northern Việt Nam were initially occupied in the late third millennium BCE, probably by means of coastal expansion from the north. At Xóm Rền, clay floors have been identified (Han 2009). Đền Dồi is a Neolithic settlement dated to the early 2nd millennium BCE that was constructed in a coastal mangrove environment in Central Việt Nam (Nguyễn et al. 2024). A succession of living floors there were constructed of compacted oyster shells on raised platforms, although the evidence does not unequivocally reveal whether the houses themselves were built at ground level, or raised on posts. Excavations at Rạch Núi, a coastal Neolithic site in southern Việt Nam, have identified similar rectangular lime-plastered floors (see discussion on ‘lime’ below) that were consistently replaced over a period of about 150 years (Oxenham et al. 2015). It is beyond reasonable doubt that they represent domestic structures built at ground level. Fragments of wattle and daub along with the posts that flank these floors indicate the structure of the walls (see discussion on ‘wall’ below).

At Khok Phanom Di in Central Thailand, the rectangular room of a house has been revealed, through the floor of which graves were cut (Fig. 2). A floor and wall foundations containing post holes lay over a series of earlier floors that were initially interpreted as a raised platform (Higham and Bannanurag 1990). It seems highly likely that construction techniques in early Neolithic coastal settlements varied with the availability of local building materials, with lime

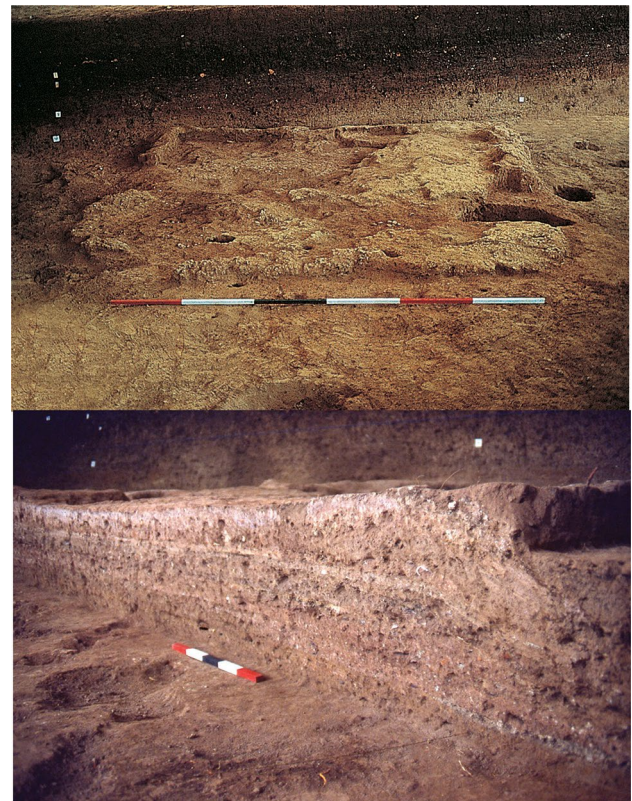


Fig. 2 The room of a house at Khok Phanom Di, ca. 4000 BP, and the section through the base showing layers of earlier floors

plastered floors at Rạch Núi, clay at Khok Phanom Di and shells at Đền Dồi.

Ban Non Wat is a moated site in Northeast Thailand with a cultural sequence spanning about 2500 years from initial Neolithic settlement. Excavations there have traced hard floor surfaces dating to ca 3700 BP for the Neolithic and continuing into the Iron Age occupation, that contain higher levels of potassium and phosphorus than in other contexts, and these have been seen as evidence for occupation debris. Phytoliths suggest that the floors were working surfaces for processing rice, while the diatoms probably come from the clay that was brought onto the site to lay the floors. Unfortunately, the areas excavated were too constrained to trace any floor plans that might represent houses (Kanthilatha et al. 2014, 2017). For clear evidence for such residences, one has to turn to the much larger areas exposed at nearby Non Ban Jak, another Neolithic to Iron Age moated settlement nearby. Although considerably later than the initial settlement of Southeast Asia by rice farmers, the consistent evidence for cultural continuity until the end of the Iron Age, seen in the analysis of ceramic typology and isotopes, highlights the relevance of the later periods to this enquiry (Higham 2024; King et al. 2013). At Non Ban Jak the foundations

for ground-level houses dating ca. 200–800 CE have been revealed (Higham 2017, Higham and Kijngam 2020).

One area excavated contained the remains of a building that had been destroyed in a conflagration (Fig. 3). This house, constructed with the initial Iron Age settlement, comprised clay wall foundations, within which there were thick deposits of burnt material including charred structural timber and rice grains. In the southeastern corner of the building, there was an incinerated clay floor, on which lay a stove, two hearth stones and a large quantity of carbonised rice grains. The most southerly part of this building revealed burnt and blackened wooden foundations. Structural daub incorporated the profiles of wattle supports, and this is interpreted as part of a collapsed wall. The clay floor lips up against the edge of the walls. Further north again, there are parallel ridges, seen as impressions of bamboo joists. (Fig. SI2).

This house was replaced by a second on a slightly different orientation (Fig. 3). It had clay and laterite outer walls, incorporating post holes that probably supported the walls. Within, there were thin subdivisions of laterite dividing the structure into rectangular areas that may have been the foundations for internal divisions that partitioned the



Fig. 3 At basal Non Ban Jak, a burnt house adjacent to a ceramic kiln was replaced by a second burnt house. A kitchen with rice scattered over the floor is seen at the rear right of this image

structure. Particularly in the northern part of this building, the narrow partitions were marked by post holes which may have supported furniture or have acted as subdivisions. The interior contained burnt clay and laterite floors. These contained concentrations of carbonised rice and broken ceramic vessels. The structure measured 3.40 by 6.40 m, although interior partition walls proceeded south and beyond the excavation square.

These houses set a pattern that continued until the site was abandoned, as further house foundations accumulated. One of the best-preserved residential areas at this site incorporated a lane between two houses. The rooms are well preserved and it is possible to enter them and explore (Fig. 4). One contained a plastered clay floor containing three burials. An adult burial lay within two parallel rows of large postholes that probably supported the roof. A hearth was situated against the southern wall and lidded vessels were placed in each corner. The town lane, unlike the rooms, contained occupation debris dominated by broken potsherds. Clay walls were large, with foundations 50 cm wide. The interior edge of the wall contains a line of postholes for timbers which would have supported the roof. A narrow channel running between the posts and the wall may well have been the foundation for the interior lining of the chamber.

A second excavation area provided a sequence of superimposed houses. It is possible to enter rooms and uncover cooking stoves, some still surrounded by chopped animal bones and rice grains (Fig. SI3). Fire, also a pAA etymon (*ʔu:ɛ ‘fire, firewood’), was a constant hazard, it seems, and this has led to the survival of collapsed wattle and daub walls in which parallel rows of the wooden framework together with blocks of burnt daub are visible (Fig. 5). Further excavation has revealed clay wall foundations for rectangular or square rooms each incorporating the post holes for wall



Fig. 4 The clay wall foundations, floor and massive post holes of this Iron Age house at Non Ban Jak are divided by a town lane from further houses beyond. Three graves were cut through the floor of this room



Fig. 5 The burnt framework of a wattle and daub wall at Non Ban Jak

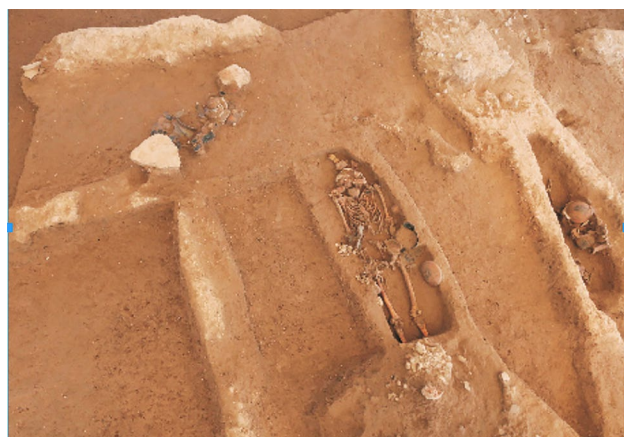


Fig. 6 At Non Ban Jak, graves cut through floors were aligned with the house walls

supports and evidence for the use of wattle and daub structures. The clay floors were severed with the interment of the dead. The regular placement of graves on the same orientation as the walls and often lying adjacent to them confirms that the dead were interred within residential houses (Fig. 6). This relationship is particularly clear for the third mortuary phase, where a house with interconnecting rooms has been identified (Fig. 7). The graves are closely aligned to the house walls.

4 WALLS and POSTS ('wall, fence' *Cə.'de:ŋ; 'walling/fencing materials' *Cə.'dan; 'post' *jə.'raŋ, *jə.'rəŋ; 'to fence, enclose' *Cə.'ga:r)

The AA etymon for wall *Cə.'de:ŋ 'walling/fencing materials' is attested in Katuic, Khmeric, and Monic, as well as the more peripheral Khasic and Nicobaric. The etymon

for post *jə.'raŋ, *jə.'rəŋ is widely attested, including the Bahnaric, Khmeric, Khmuic, Mangic, Palaungic, Pearic, and Khasic branches. In addition, a verb *Cə.'ga:r 'to fence, enclose' has been reconstructed, with lexical attestations in Bahnaric, Khmeric, Khmuic, and Khasic. While lacking fine detail in lexical terminology, these three lexical reconstructions constitute evidence of the basic elements of the early Austroasiatic house structural elements from the Neolithic.

Possibly the best example of a surviving Neolithic wall comes from Khok Phanom Di, where fragments of wattle and daub have been recovered, dating to at least 4000 BP (Fig. 8). One retained the impression of the wattle support, and the smoothed surface was painted red. Many posts were excavated at this site, one retaining the wooden post itself (Fig. 9). There were also straight rows of what would have been posts for a substantial structure, possibly a wall, and walls encircled sets of burials.

Noen U-Loke is a moated site with a sequence spanning virtually the millennium of the Iron Age. The cultural remains included the remains of burnt clay floors, hearths and furnaces, and 114 fragments of wattle and daub structural material. These comprise the charred remains of clay, tempered with rice straw, over a framework of split bamboo laced together horizontally (Chetwin (2007). The exterior surfaces were nearly all flat and were smoothed by a wooden tool or by hand, since fingerprints have survived (Fig. 10). Chetwin (2007) has concluded that they were the burnt remains of domestic walls, associated with the burnt floors. The remains of a collapsed wall at Ban Non Wat, with the parallel rows of charred wattle posts, evidence the use of this building technique for the walls of domestic structures. An excavation of limited area at the particularly large moated site of Non Muang Kao uncovered successive floors so hard that excavation was challenging. However, their square form was very clear (O'Reilly 2007 Fig. SI4).

Some of the rows of posts at sites like Khok Phanom Di might have been used to enclose space. These look very similar to the slightly later range of posts seen at Rạch Núi in southern Việt Nam, interpreted by Oxenham et al. (2015) as a fence constructed possibly, to restrict the access of animals to domestic dwellings. Further archaeological evidence for an enclosure comes from Iron Age Ban Non Wat, where a palimpsest of water buffalo and bovid hoofprints may be interpreted as evidence for a corral or pound (Fig. SI5).

5 FLOORING ('floor joist, floor timber' *kə.'nəŋ; 'bamboo tube/joint' *(kə.')dɪŋ)

The etymon *kə.'nəŋ 'floor joist, floor timber' for this key component of a house is attested in Bahnaric, Katuic, and Monic in MSEA and Khasic outside MSEA. Despite the smaller number of lexical attestations, the geographic

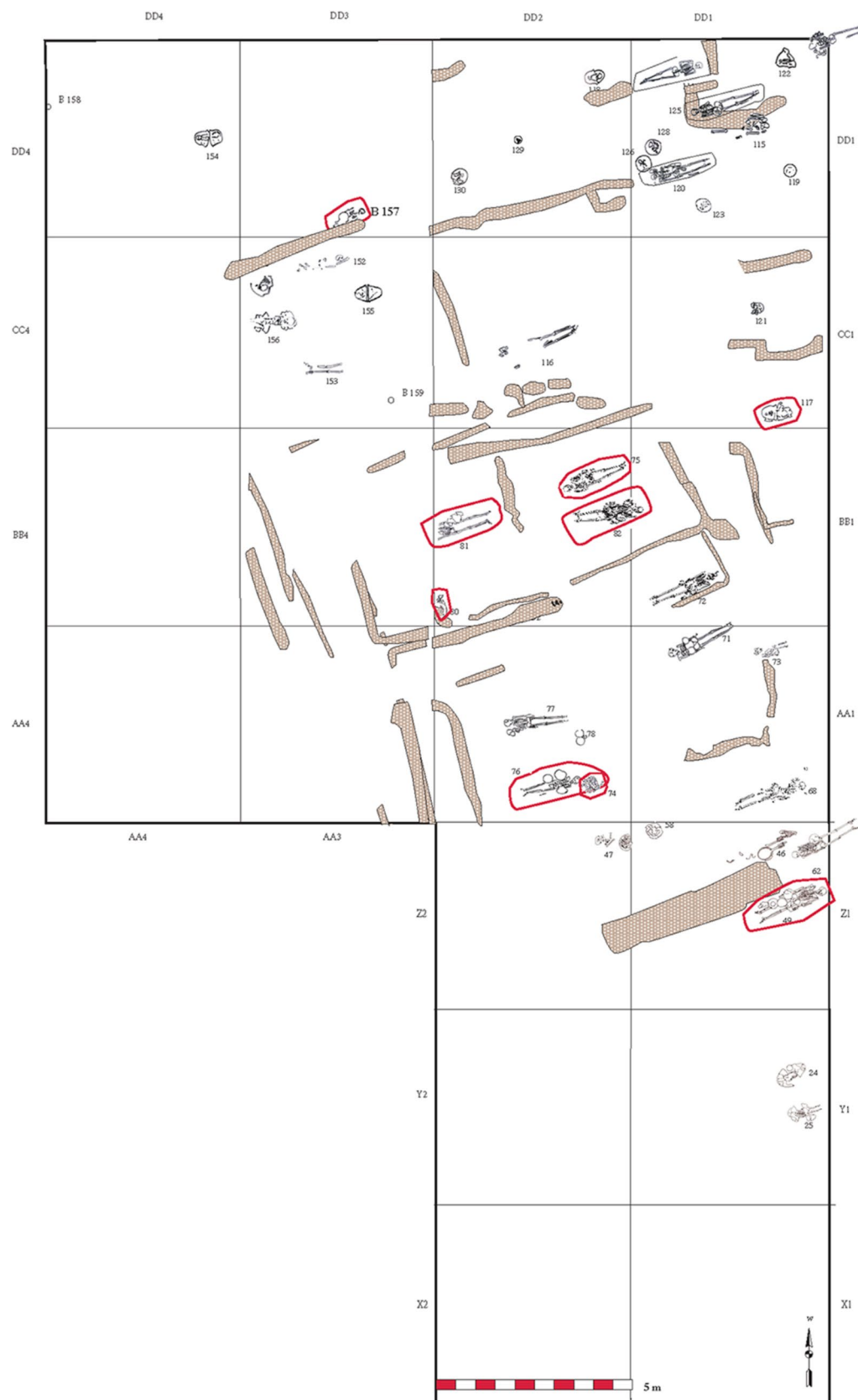


Fig. 7 By opening a large area at Iron Age Non Ban Jak, the plan houses with residential burials was obtained



Fig. 8 A burnt piece of wattle and daub wall from Khok Phanom Di, ca. 2000 BCE, shows two layers of daub, the outer surface of which was painted red



Fig. 9 The wood survived in the post hole on the left, and the row of posts on the right might have been for a fence or a house wall. Khok Phanom Di, ca. 4100 BP

distribution in Khasic makes this a valid pAA etymon, or at least one with deep history in pAA.

While floors themselves are well documented in prehistory, joists to support them are rare, the best instance of bamboo joists comes from Rạch Núi a Neolithic settlement in southern Việt Nam, dating ca. 3500–3300 BP (Fig. SI6). A configuration of charred joists running parallel with each other was identified there, associated with

lime-plastered floors (Oxenham et al. 2015). The remains of houses raised on posts are found at several sites in the low-lying terrain of the Dong Nai Delta in southern Việt Nam. At sites like Bưng Bắc, for example, wooden posts up to 2.2 m long, still bearing planks, indicate river-side villages comprising houses raised on stilts above the flood level (Bùi et al. 2012). At Iron Age Non Ban Jak, parallel impressions of probably bamboo spanning a room are interpreted as floor joists.

Another relevant etymon **(kə.)dɨŋ* ‘bamboo tube/joint’ is attested in the MSEA branches Bahnaric, Monic, and Palaungic and the MSEA-external branches of Aslian and Munda, establishing this as a solid pAA etymon.

The use of bamboo in prehistory is documented in the burnt wattle and daub fragments from several sites. The circular impressions in the daub make identification of the supporting structure as being made from bamboo. This is seen at Neolithic Khok Phanom Di and Rạch Núi, and the Iron Age sites of the upper Mun Valley region. Thus, while we cannot, of course, claim with certainty that this word is specifically associated with production of household structures, this AA etymon attests to bamboo in early AA culture and matches the archaeological evidence.

Vàm Cỏ Tây.

6 WEAVING (‘to weave, plait’ **ta:p*; ‘spread a mat’ **be:l*)

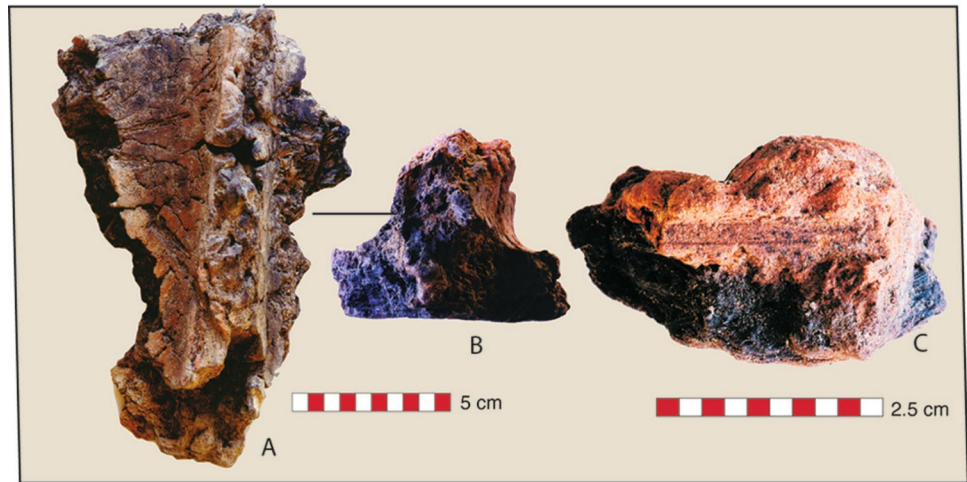
While proto-language words for ‘woven material’ or ‘mat’ are not available, two verbs imply the making and use of woven items. The etymon **ta:p* ‘to weave, plait’ is attested in all branches of pAA, dating this practice to well before the pAA dispersal and highlighting the significance of this word from that early period. The etymon **be:l* ‘to spread a mat’ is attested in Bahnaric, Katuic, Khmuic, Monic, and Palaungic in MSEA, and in the Aslian and Munda branches outside, the latter two demonstrating the pAA origins of this term.

Impressions of matting on clay floors have been identified at the site of Lò Gạch, on the Vàm Cỏ Tây.

River in southern Việt Nam (Cameron 2017). These have been dated to ca. 2750 BP. Mat impressions also survive as impressions on clay at Noen U-Loke and as pseudomorphs on iron artefacts at Ban Non Wat (Fig. SI7). Evidence for woven fabric comes from pseudomorphs in the rust on iron mortuary offerings (Fig. SI7). The manufacture of the threads is universally present from the Neolithic occupation, seen in the cord marks impressed on the surface of ceramic vessels. These are produced by employing a spindle and weighted whorl, the latter artefact and universally from the Bronze Age, often among the grave goods of a weaver.

Lò Gạch.

Fig. 10 These fragments of wattle and daub walling from Noen U-Loke, ca. 100 CE, reveal the impressions of bamboo and a cross piece of wood



7 WOOD ('tree, wood' *jə.ɛiːʔ, *jə.ɛeːʔ)

The pAA etymon *jə.ɛiːʔ, *jə.ɛeːʔ 'tree, wood' is widely attested, including Bahnaric, Katuic, Khmeric, Khmuic, Mangic, Monic, Palaungic and Vietic inside MSEA, and the Aslian and Nicobaric branches outside. While 'tree' and 'wood' are generic concepts, it is still necessary to present available relevant data. The employment of wood in the construction of houses is evidenced in the many post holes that are encountered from the early Neolithic onwards (Fig. 9). Wood was also used with clay daub for house walls, and to enclose ceramic kilns. Pseudomorphs of wooden handles have been found on Iron Age sickles, and large charred wooden wall foundations, along with slots in wall foundations to take wooden beams, have been uncovered at Iron Age Non Ban Jak (Fig. SI8).

8 LIME (MINERAL) ('lime (mineral)' *k(m)pur)

The etymon *k(m)pur 'lime (mineral)' is attested in Bahnaric, Katuic, Khmeric, Khmuic, Monic, Palaungic and Vietic branches. It is not present beyond MSEA. In the approach of Alves (2023), this rules out this word as a pAA etymon sensu lato, even if it still may have a deep history in AA. A more significant complication is the widespread distribution of a possible shared wordform, namely, proto-Austronesian *qapuR 'lime (mineral)' (a reconstruction of Blust and Trussel 2010), which is also associated with the practice of betel chewing. While it is not possible yet to determine the time or even direction of the borrowing, it is attested in Old Khmer inscriptions in the 900 s CE (Jenner 2009), giving it some historical depth. In addition, the use of lime for floor construction dates back millennia, so even if it

is a loanword into Austroasiatic, the practice of using lime, whether for floor treatment, betel chewing, or both, appears to be quite deep in AA history.

It might be that the use of lime mortar in house construction was an innovation post-dating the migrations into the Nicobar Islands and South Asia. However, that it was employed in the early Neolithic of Southern Việt Nam is definite. Elle Grono et al. (2022a, b) have examined a stratified sequence of floors at the southern Việt Name Neolithic site of Lộc Giang, that reflect consistent rebuilding of a structure over a period of between 50–360 years. These are a close match with those identified at about the same period at nearby sites of An Sơn and Rạch Núi. Dating between ca. 3980 and 3270 cal BP, the Lộc Giang floors are early Neolithic examples of domestic structures paralleled at the same period at Khok Phanom Di. The application of microstratigraphic techniques at Lộc Giang identified 41 hard surfaces in three phases. The inhabitants manufactured a lime mortar from shells for floors, varying between 10–50 mm thick, that were coated with a thin layer of lime plaster. At Rạch Núi 13 floors were superimposed over a period of about two centuries, suggesting the consistent replacement of houses due to the rapid degeneration of organic building materials, such as timber wall supports (Oxenham et al. 2015).

9 THATCHING ('thatching grass' *pə.'lan; 'thatching grass' *sə.'puːʔ, *sə.'pɔːʔ)

The two etyma *pə.'lan and *sə.'puːʔ/*sə.'pɔːʔ meaning 'thatching grass' are attested in several branches inside MSEA and two branches each outside MSEA, making both valid pAA terms. Whether these words were completely synonymous or represented different categories of thatching grass of differing functions cannot be determined. Nevertheless, that two terms have been retained highlights their

importance in early pAA practices. While archaeological evidence for some of the etyma for houses is either absent or fugitive, in the case of thatch, it is noted that rice straw was readily available and was used liberally as an additive to the daub used in wall construction. At Rạch Núi, impressions of palm leaves on floors might indicate its use in thatching. Regardless, the lexical data, two widely attested etyma in this case, are valuable evidence in reconstructing this element of the early AA house.

10 Conclusions

Combined archaeogenetic, linguistic and archaeological research has identified the expansion of rice and millet farmers into Southeast Asia from the north. Incoming farmers encountered permanently occupied hunter-gatherer settlements that led to their integration. This seminal change took place during the 5th millennium BP.

Apart from Vietnamese and Khmer, languages of the Austroasiatic family (AA) are distributed today in enclaves from Central India to Việt Nam (Fig. 1), and some display divergence measured in millennia. It has long been recognised that these languages include cognate words for rice, millet and aspects of their cultivation, providing *prima facie* evidence that the pAA language was spoken by the incoming farmers. A detailed examination for pAA etyma for the house, aspects of its structure and activities of a domestic nature has identified 17 that are distributed across mainland Southeast Asian AA languages, some of which extend into the AA languages in India, peninsular Malaysia and the Nicobar Islands.

This paper has identified archaeological correlates for these words in sites dating to the initial ingress of farmers, and beyond into the Bronze and Iron Ages, a span of about three millennia. This is justified by the growing evidence for demographic continuity that characterised this sequence until the expansion of Han and South Asian populations into Southeast Asia, events that led to the introduction of new building traditions such as the use of bricks and tiles.²

This assessment of the archaeological evidence weighed against the proposed etyma supports their validity, and when the evidence from both is combined, it provides an invaluable key to understanding the physical characteristics of prehistoric settlements. This is important when it is appreciated that sites are often deeply stratified, and as yet none has been excavated over a sufficient area to trace the plan of a village or community, as is seen at Çatalhöyük and Maidanetske

(Hodder 2016, Hofmann et al. 2019). At Khok Phanom Di and contemporary early Neolithic sites in southern Việt Nam, multiple house floors associated with wall foundations come from houses built at ground level. This is confirmed in later sites, dating to the Iron Age in Northeast Thailand, where conflagrations in prehistory have preserved almost complete house plans with internal rooms, clay floors and wall foundations, collapsed wattle and daub walls and the holes within which the structural posts were inserted. At Non Ban Jak, a conflagration has preserved a kitchen with a pottery vessel still on the hearth and rice scattered over the floor (Fig. 3).

The pAA word for lime is one of the etyma identified. Manufactured from shells, it was used both in the matrix of the floor and as a thin veneer on the floor surface, making them rock hard. The word for ‘bamboo tube’ is another etymon in pAA, and it was widely employed in constructing clay daub walls. This was a favoured building method for centuries, seen for example in Iron Age town sites located in Northeast Thailand. Some posthole configurations are also likely to have been for fences to control or restrict access as, for example, in a water buffalo pound from Iron Age Ban Non Wat. We do not have any evidence for the use of thatch roofing, but this seems likely given the abundance of rice straw that tempered daub walls, and now, two strong pAA reconstructions for thatching grass provide further evidence that it was part of early pAA culture.

This concordance between linguistic and archaeological evidence illuminates the nature of early Neolithic domestic architecture and its continuity until the introduction of Han Chinese and South Asian building techniques in metropolitan regions. It emphasises the value of incorporating inputs from palaeogenetics, proto-languages and archaeology in reconstructing vital chapters in the human past. For Southeast Asia, this concordance provides the strongest possible support for the expansion of Neolithic rice farmers from the Yangtze River region during the 5th millennium BP. This involved interaction with the indigenous and long-established affluent hunter gatherers, the establishment of permanent farming communities and in due course the foundation of the first indigenous state societies.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s41826-025-00107-0>.

Acknowledgements We wish to acknowledge the permissions given by the Fine Arts Department of Thailand and the National Research Council of Thailand to undertake archaeological fieldwork. We are grateful to three anonymous referees for their constructive comments.

Author contributions CFWH wrote the archaeological sections of this paper and MJA wrote the linguistics sections. Both wrote the conclusions.

² Alves (2023: 4) notes the Indic origin of ‘brick’ in Khmer and the Old Chinese origin of ‘tile’ in Vietnamese, both with corresponding ancient archaeological evidence.

Funding Open access funding provided by The University of Otago. Funding was received from the Marsden Fund of New Zealand under Grant UOO1113.

Data availability Not Applicable.

Code availability Not Applicable.

Declarations

Ethics approval Not Applicable.

Compliance with ethical standards Not applicable.

Conflict of interest Not applicable.

Consent to participate Not applicable.

Consent for publication Not Applicable.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alves, Mark. 2023. Proto-Austroasiatic etymologies of words related to household structures. In *Papers from the Ninth and Tenth International Conferences on Austroasiatic Linguistics*, ed. Paul Sidwell and Mark Alves, 1–18. Journal of the Southeast Asian Linguistic Society Special Publication No. 12.
- Blust, Robert and Stephen Trussel. 2010 (revision 2016). Austronesian Comparative Dictionary (web edition). <http://www.trussel2.com/ACD>. Accessed 15 Oct 2024.
- Bùi Chí Hoàng, Phạm Chí Thân, and Nguyễn Khánh Trung Kiên. 2012. *Khảo cổ học Bà Rịa–Vũng Tàu từ tiền sử đến sơ sử* [Archaeology of Bà Rịa–Vũng Tàu Province from Prehistory to Protohistory]. Hà Nội: Nhà xuất bản Khoa học Xã hội.
- Cameron, Judith. 2017. Matting impressions from Lo Gach: Materiality at floor level. In *New Perspectives in Southeast Asian and Pacific Prehistory, Terra Australis 45*, ed. Philip Piper, Hiro Matsumura, and David Bulbeck, 357–367. Canberra: ANU Press.
- Carter, Alison K. 2023. Building from the Ground Up: The Archaeology of Residential Spaces and Communities in Southeast Asia. *Journal of Archaeological Research* 31: 1–54. <https://doi.org/10.1007/s10814-021-09170-4>.
- Chaubey, Gyaneshwer, Mait Metspalu, Ying Choi, et al. 2011. Population genetic structure in Indian Austroasiatic speakers: The role of landscape barriers and sex-specific admixture. *Molecular Biology and Evolution* 28: 1013–1024. <https://doi.org/10.1093/molbev/msq288>.
- Chetwin, James N. 2007. The remains of wattle and daub structures. In *The Origins of Civilization of Angkor Volume II. The Excavation of Noen U-Loke and Non Muang Kao*, ed. Charles F.W. Higham, Amphan Kijngam and Sarah Talbot, 447–463. Bangkok: The Fine Arts Department.
- Diamond, Jared, and Peter Bellwood. 2003. Farmers and their languages: The first expansions. *Science* 300 (5619): 597–603. <https://doi.org/10.1126/science.1078208>.
- Eberhard, David M., Gary F. Simons, and Charles D. Fennig (eds.). 2024. *Ethnologue: Languages of the World*. Twenty-seventh edition. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com.montgomerycollege.idm.oclc.org>. Accessed 15 Oct 2024.
- Grono, Elle, Philip Piper, Dang Ngoc Kinh, et al. 2022a. Early settlement construction in Southeast Asia: Lime mortar floor sequences at Lộc Giang, southern Việt Nam. *Antiquity* 96: 1538–1554. <https://doi.org/10.15184/aqy.2022.139>.
- Grono, Elle, Philip Piper, Kien Nguyễn, et al. 2022b. The identification of dwellings and site formation processes at archaeological settlements in the tropics: A micro-geoarchaeological case study from Neolithic Lộc Giang, southern Việt Nam. *Quaternary Science Reviews* 291: 10764. <https://doi.org/10.1016/j.quascirev.2022.107654>.
- Hán, V.K. 2009. *Xóm Rền, An Especially Important Key Site of the Bronze Age in Việt Nam*. National University Publishing House, Hanoi (in Việt Name).
- Heggarty, Paul, Cormac Anderson, Matthew Scarborough, et al. 2023. Language trees with sampled ancestors support a hybrid model for the origin of Indo-European languages. *Science* 381: 6656. <https://doi.org/10.1126/science.abg0818>.
- Higham, Charles. F. W. 1996. *The Bronze Age of Southeast Asia*. Cambridge: Cambridge University Press.
- Higham, Charles. F.W.. 2002. Languages and farming dispersals: Austroasiatic languages and rice cultivation. In *Examining the Farming/Language Dispersal Hypothesis*, ed. Peter Bellwood and Colin Renfrew, 223–232. Cambridge: The McDonald Institute.
- Higham, Charles F.W.. 2017. The prehistoric house: A missing factor in Southeast Asia. In *New Perspectives in Southeast Asian and Pacific Prehistory, Terra Australis 45*, ed. P.J. Piper, H. Matsumura, and D. Bulbeck, 369–384. Canberra: ANU Press.
- Higham, Charles F.W. and Rachanie Bannanurag. 1990. *The excavation of Khok Phanom Di: Volume I, The excavation, chronology and human burials*. London: The Society of Antiquaries.
- Higham, Charles .F.W. and Amphan Kijngam (eds.). 2020. *The origins of the civilization of Angkor volume VII. The excavation of Non Ban Jak*. Bangkok: Fine Arts Department.
- Higham, Charles F.W. 2024. Mortuary rituals and social change from the Neolithic to the Iron Age in Thailand. *Asian Archaeology*; <https://doi.org/10.1007/s41826-024-00094-8>
- Hodder, Ian. 2016. More on history houses at Çatalhöyük: A response to Carleton et al. *Journal of Archaeological Science* 67: 1–6. <https://doi.org/10.1016/j.jas.2015.10.010>.
- Hofmann, Robert, Johannes Müller, Liudmyla Shatilo, et al. 2019. Governing Tripolye: Integrative architecture in Tripolye settlements. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0222243>.
- Jenner, Philip N. 2009. *A Dictionary of Pre-Angkorian Khmer*. Edited by Doug Cooper. The Australian National University. Canberra: Pacific Linguistics. <http://hdl.handle.net/1885/146750>.
- Kanthilatha, Nelum, William Boyd, and Nigel Chang. 2014. Multi-element characterization of archaeological floors at the prehistoric archaeological sites at Ban Non Wat and Nong Hua Raet in north-east Thailand. *Quaternary International* 432: 66–78.
- Kanthilatha, Nelum, William Boyd, Parr N. Jeffery, et al. 2017. Implications of phytolith and diatom assemblages in the cultural layers of prehistoric archaeological sites of Ban Non Wat and Nong

- Hua Raet in northeast Thailand. *Environmental Archaeology* 22: 15–27. <https://doi.org/10.1179/1749631415Y.00000000026>.
- King, Charlotte, R. Alexander Bentley, Nancy Tayles, et al. 2013. Moving peoples, changing diets: Isotopic differences highlight migration and subsistence changes in the Upper Mun River Valley, Thailand. *Journal of Archaeological Science* 40: 1681–1688. <https://doi.org/10.1016/j.jas.2012.11.013>.
- King, Charlotte, Nancy G. Tayles, Charles F.W. Higham, et al. 2015. Using isotopic evidence to assess the impact of migration and the two-layer hypothesis in prehistoric Northeast Thailand. *American Journal of Physical Anthropology* 158: 141–150. <https://doi.org/10.1002/ajpa.22772>.
- Li, Anjun. 2009. *Tianluoshan site. A new window of Hemudu culture*. Nanjing: Liu Xing.
- Lipson, Mark, Olivia Cheronet, Swapan Mallick, et al. 2018. Ancient genomes document multiple waves of migration in Southeast Asian pre-history. *Science* 361: 92–95. <https://doi.org/10.1126/science.aat3188>.
- Mason, Francis. 1854. The Talaeng language. *Journal of the American Oriental Society* 4: 277–288.
- McColl, Hugh, Fernando Racimo, Lasse Vinner, et al. 2018. The prehistoric peopling of Southeast Asia. *Science* 361: 88–92. <https://doi.org/10.1126/science.aat3628>.
- O'Reilly, Dougald J.W. 2007. The excavation of Non Muang Kao. In *The origins of the civilization of Angkor volume II. The excavation of Noen U-Loke and Non Muang Kao*, ed. Charles F.W. Higham, Amphan Kijngam and Sarah Talbot, 545–559. Bangkok: Fine Arts Department.
- Oxenham, Marc F., Philip J. Piper, Peter Bellwood, et al. 2015. Emergence and diversification of the Neolithic in southern Việt Nam: Insights from coastal Rạch Núi. *The Journal of Island and Coastal Archaeology* 10: 309–338. <https://doi.org/10.1080/15564894.2014.980473>.
- Rau, Felix and Paul Sidwell. 2019. The Munda maritime hypothesis. *Journal of the Southeast Asian Linguistics Society* 12.2: 35–57. <http://hdl.handle.net/10524/52454>.
- Ryan, William B.F., Suzanne Carbotte, Justin Coplan, et al. 2009. Global multiresolution topography (GMRT) synthesis data set. *Geochemistry Geophysics Geosystems* 10 (3): Q03014. <https://doi.org/10.1029/2008GC002332>.
- Sidwell, Paul and Mark Alves. 2023. Re-evaluating Shorto's MKCD reconstructions. In *Papers from the ninth and tenth international conferences on Austroasiatic linguistics*, ed. by Paul Sidwell and Mark Alves, 98–126. Journal of the Southeast Asian Linguistic Society Special Publication No. 12. Honolulu: University of Hawai'i Press. <https://hdl.handle.net/10524/73781>.
- Sidwell, Paul. 2021. History of MSEA Austroasiatic studies. In *The languages and linguistics of mainland Southeast Asia. The world of linguistics*, Volume 8, ed. Paul Sidwell and Mathias Jenny, 61–92. Berlin: de Gruyter Mouton.
- Sidwell, Paul. 2024. 500 Proto Austroasiatic Etyma: Version 1.0. *Journal of the Southeast Asian Linguistics Society* 17.1: i–xxiii. <https://hdl.handle.net/10524/52519>.
- Sun, Guoping. 2013. Recent research on the Hemudu culture and the Tianluoshan site. In *A Companion to Chinese Archaeology*, ed. Anne Underhill, 555–573. Hoboken: Wiley-Blackwell.
- Tätte, Kai, Luca Pagani, Ajai Pathak, et al. 2019. The genetic legacy of continental scale admixture in Indian Austroasiatic speakers. *Science Reports* 9: 3818. <https://doi.org/10.1038/s41598-019-40399-8>.
- Tayles, Nancy G. 1999. *The excavation of Khok Phanom Di. A prehistoric site in central Thailand. Vol. V: The people*. Research Reports of the Society of Antiquaries of London No. LXI. London: Society of Antiquaries.
- Nguyễn Thị Thúy, Lâm Thị Mỹ Dung, Nguyễn Thị Mai Hương, et al. 2024. Understanding settlement construction and chronology at Đền Đồi, Nghệ An province, Central Việt Nam. *Archaeological Research in Asia* 40: 100568. <https://doi.org/10.1016/j.ara.2024.100568>.
- ZPICA Zhejiang Provincial Institute of Cultural Relics and Archaeology et al. 2008. The Tianluoshan Neolithic site at Yuyao, Zhejiang. *Chinese Archaeology* 8: 44–54. <https://doi.org/10.1515/CHAR.2008.8.1.44>.
- ZPICA Zhejiang Provincial Institute of Cultural Relics and Archaeology et al. 2022. The Excavation of a Neolithic site at Jingtoushan in Yuyao, Zhejiang. *Chinese Archaeology* 22: 1–16. <https://doi.org/10.1515/char-2022-0001>.