

# Multidisciplinary perspectives on banana (*Musa* spp.) domestication

Xavier Perrier<sup>a</sup>, Edmond De Langhe<sup>b</sup>, Mark Donohue<sup>c</sup>, Carol Lentfer<sup>d</sup>, Luc Vrydaghs<sup>e</sup>, Frédéric Bakry<sup>a</sup>, Françoise Carreel<sup>f</sup>, Isabelle Hippolyte<sup>a</sup>, Jean-Pierre Horry<sup>a</sup>, Christophe Jenny<sup>g</sup>, Vincent Lebot<sup>h</sup>, Ange-Marie Risterucci<sup>a</sup>, Kodjo Tomekpe<sup>a</sup>, Hugues Doutrelepont<sup>e</sup>, Terry Ball<sup>i</sup>, Jason Manwaring<sup>j</sup>, Pierre de Maret<sup>i</sup>, and Tim Denham<sup>k,1</sup>

<sup>a</sup>Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Unité Mixte de Recherche Amélioration Génétique et Adaptation des Plantes, F-34398 Montpellier, France; <sup>b</sup>Laboratory of Tropical Crop Improvement, Katholieke Universiteit, 3001 Leuven, Belgium; <sup>c</sup>Department of Linguistics, Australian National University, Canberra 0200, Australia; <sup>d</sup>School of Social Science, University of Queensland, St. Lucia 4072, Australia; <sup>e</sup>Research Team in Archaeo- and Palaeosciences, 1160 Brussels, Belgium; <sup>f</sup>Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Unité Mixte de Recherche Biologie et Génétique des Interactions Plantes-Parasites, F-34398 Montpellier, France; <sup>g</sup>Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Unité Mixte de Recherche Amélioration Génétique et Adaptation des Plantes, F-97130 Capesterre-Belle-Eau, Guadeloupe, France; <sup>h</sup>Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Unité Mixte de Recherche Amélioration Génétique et Adaptation des Plantes, Department of Agriculture, Centre Agronomique de Recherche et de Formation du Vanuatu, 946 Port Vila, Vanuatu; <sup>i</sup>Department of Ancient Scriptura, Brigham Young University, Provo, UT 84602; <sup>j</sup>Secrétariat, Centre d'Anthropologie Culturelle, Université Libre de Bruxelles, 1000 Brussels, Belgium; and <sup>k</sup>School of Geography and Environmental Science, Monash University, Victoria 3800, Australia

Edited by Dolores R. Piperno, Smithsonian National Museum of Natural History and Smithsonian Tropical Research Institute, Panama City, Panama, and approved May 25, 2011 (received for review March 5, 2011)

**Original multidisciplinary research hereby clarifies the complex geodomestication pathways that generated the vast range of banana cultivars (cvs). Genetic analyses identify the wild ancestors of modern-day cvs and elucidate several key stages of domestication for different cv groups. Archaeology and linguistics shed light on the historical roles of people in the movement and cultivation of bananas from New Guinea to West Africa during the Holocene. The historical reconstruction of domestication processes is essential for breeding programs seeking to diversify and improve banana cvs for the future.**

plant genetics | historical linguistics | archaeobotany | diploid banana cultivars | triploid banana cultivars

New multidisciplinary findings from archaeology, genetics, and linguistics clarify the complex geodomestication pathways—the geographical configurations of hybridization and dispersal—that generated the range of modern banana cultivars (cvs). Although recent molecular research, combined with the outcomes of previous genetic studies, elucidates major stages of banana domestication, such as the generation of edible diploids and triploids, it sheds only partial light on the historical and sociospatial contexts of domestication. The geographic distributions of genotypes involved in banana domestication require human translocations of plants, most likely under vegetative forms of cultivation, across vast regions. Linguistic analyses of (traditional) local terms for bananas reveal several striking regional-scale correspondences between genetic and linguistic patterns. These multidisciplinary findings enable the relative dating of the principal events in banana geodomestication and situate banana cultivation within broader sociospatial contexts. Archaeological findings provide a timeline to anchor and calibrate the relative chronology.

Banana ranks next to rice, wheat, and maize in terms of its importance as a food crop. In addition to being a major cash crop around the world, more than 85% of bananas are grown for local consumption in

tropical and subtropical regions (1). The hundreds of cultivated varieties are products of centuries—in some cases millennia—of clonal (vegetative) propagation. Banana fruits can be cooked, roasted, or even brewed (e.g., plantains; East African Highland cvs) or eaten raw (e.g., the yellow Cavendish banana sold in supermarkets globally).

## ***Musa* Genetics and Domestication Thresholds**

The monocotyledon Musaceae family includes the Asian and African genus *Ensete*, the genetically proximal Asian *Musella* genus, and the East Asian genus *Musa*, which is divided into sections with 22 (*Eumusa*, *Rhodochlamys*) and 20 chromosomes (*Australimusa*, *Callimusa*). Most edible bananas belong to the *Eumusa* section, and are diploid or triploid hybrids from *Musa acuminata* (A-genome) alone or from hybridization with *Musa balbisiana* (B-genome). A minor cv group, including Fe'i bananas, is confined to the Pacific region and is derived from *Australimusa* species.

The evolution from wild to edible bananas involved seed suppression and parthenocarpy development. Simmonds (2) proposed the first reconstruction of this complex process in the 1960s. Since then the spectrum of available genetic resources has been greatly expanded, notably with collections from New Guinea

(NG), the native area of the *M. acuminata* subsp. *banksii* (3), and through the application of molecular methods to banana diversity analysis (4–7).

In the present study, we present the conclusions of two recent analyses by using simple sequence repeats (8) and DarT (9) markers on the international collection in Guadeloupe (Centre de Coopération Internationale en Recherche Agronomique pour le Développement, France) which assembles more than 400 wild and cultivated accessions covering much of global *Musa* diversity. Analyses here include additional accessions from Cameroon and Nigeria to better represent African cv diversity (Table S1). The results of these analyses were integrated with previous results, including ploidy level and DNA content (10), anthocyanin patterns (11), isozymes (12), nuclear restriction (13) and amplified fragment length polymorphism (14), and chloroplastic and mitochondrial restriction fragment length

Author contributions: P.d.M. designed research; E.D.L., M.D., C.L., L.V., F.B., F.C., I.H., J.-P.H., C.J., L.V., A.-M.R., K.T., H.D., T.B., and J.M. performed research; X.P. analyzed data; and X.P., E.D.L., M.D., and T.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

<sup>1</sup>To whom correspondence should be addressed. E-mail: tim.denham@arts.monash.edu.au.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1102001108/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1102001108/-DCSupplemental).

polymorphism (15). The synthesis of these results sheds light upon the two founding events during banana domestication: the transition from wild to edible diploids and the formation of triploids from edible diploids (16) (Fig. S1).

Wild *M. acuminata* species are diverse and have been differentiated into several subspecies (13, 16). This differentiation is clarified here with the identification of specific nuclear alleles, chromosomal rearrangements, and chloroplast and mitochondrial patterns. The exclusive natural ranges of the subspecies are also confirmed (Fig. 1).

The key result relevant to domestication history is that edible diploid cv subgroups (AACvs) derive from hybridizations between different *M. acuminata* subspecies in island Southeast Asia (SEA; ISEA) and western Melanesia (Fig. 2 and Fig. S2), which could only be brought into contact by human movement and propagation, most probably during the Holocene (i.e., the past 11,700 y). The

structural heterozygosity of these hybrid AACvs, caused by chromosomal rearrangements between parental subspecies of *M. acuminata* (17), contributed to gametic sterility (18). This sterility, in association with human selection for pulp enhancement, led to parthenocarpic fruits and edibility.

A further consequence of hybrid status was erratic meiosis in edible AACvs, thereby occasionally producing diploid gametes (2). The fusion of diploid gametes with haploid gametes generated sterile triploid genotypes. Spontaneous triploidizations involved almost all diploid cvs leading, under human selection via vegetative propagation, to the diversity of modern cultivated triploids, including pure *M. acuminata* varieties (AAA) and interspecific *M. acuminata* × *M. balbisiana* varieties (AAB, ABB) (19).

Molecular phylogenetic analyses have identified the diploid subgroups contributing to modern triploid cvs. Of significance, NG-derived *banksii* cvs played a

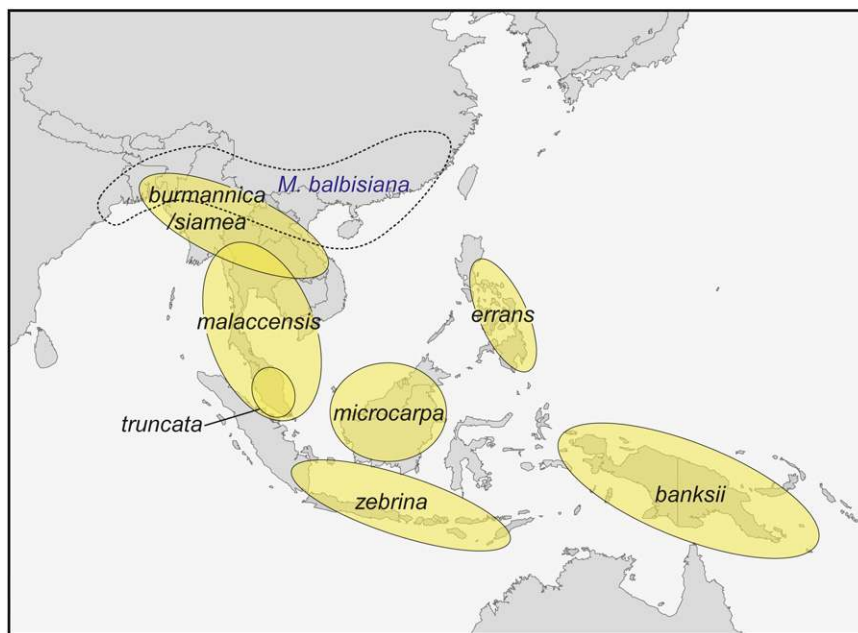
major role in the development of several important diploid and triploid cv groups that are now widely dispersed. For example, *M. acuminata* subsp. *zebrina*- and subsp. *banksii*-derived AACvs contributed to AAA Highland bananas of East Africa. Likewise, subsp. *banksii*-derived AACvs with the BB genome contributed to AAB plantains of West Africa and the Pacific. Moreover, in several triploid subgroups, specific accessions within AACv subgroups have been identified as potential ancestors, suggesting more recent formation (16, 20). For example, the 2N and N parents of the common dessert banana, the AAA “Cavendish,” are genetically close to “Akondro Mainty” of the AACv “Mlali” subgroup and “Pisang Pipit” of the AACv “Khai” subgroup, respectively (Fig. 2).

Strict vegetative propagation (i.e., cloning) over long periods of the most popular diploid and triploid varieties led to somaclonal variants, thus amplifying phenotypic diversity (14). The geographical ranges of diploid parents and triploid hybrids point to numerous long-distance movements of either one of the parents before hybridization, or of triploid cvs after hybridization. The geographical dimension of the banana domestication process requires a reconstruction of relevant human movements and interactions—the drivers of the geodomestication process.

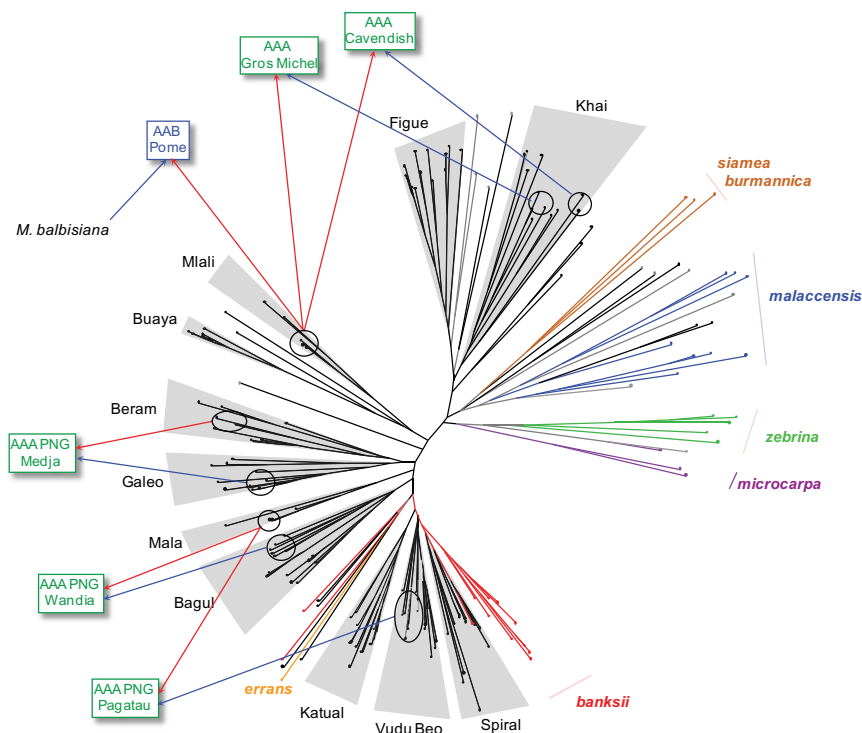
### Anchoring Banana Domestication in Time and Place

Microfossil (phytoliths, pollen) and macrofossil (seeds, pseudostem impressions) remains of bananas have been documented at numerous archaeological and paleoecological sites from Oceania to Africa (21), although their significance varies greatly for understanding the history of banana domestication and cultivation. Within the natural range of *Musa*, archaeobotanical and paleoecological finds require differentiation to a species’ or subspecies’ level to indicate potential human movement, and also require associated evidence of human exploitation or cultivation. Most identifications are to the genus level and not directly associated with archaeological remains; consequently, they are of limited value for understanding banana domestication because they may solely record wild bananas (21). Outside the natural range of *Musa*, bananas were exotic, dispersed under cultivation, predominantly sterile, and vegetatively propagated; consequently, any archaeobotanical find of *Musa* signifies the presence of agriculture.

Predominantly, the *Musa* genus has been identified in the archaeological record by using phytoliths (21). As a result of parthenocarpy in most cvs, seeds are unlikely to be found in prehistoric agricultural contexts, and usually leaf phytoliths



**Fig. 1.** Geographical distribution of *M. balbisiana* and subspecies of *M. acuminata*, the wild ancestors of cultivated bananas. Nuclear and cytoplasmic markers (8, 13, 15) differentiate *M. acuminata* wild diploids into clusters that fit with subspecies previously defined from morphological characters: *banksii* cluster from NG; *malaccensis* cluster from Malayan Peninsula; and *burmannica*, *burmannicoides*, and *siamea* cannot be discriminated by their nuclear genome and form a complex that covers South China, Thailand, Myanmar, Bangladesh, northeast India with sporadic populations southward to Sri Lanka. This complex is genetically closer to the *malaccensis* cluster than to other clusters, they are geographically overlapping, and several wild accessions from Thailand are identified as hybrids. *Zebrina* cluster from Java and its genome size is significantly (10%) higher (8). Its anthocyanin pathway is the most primeval (19), *banksii* is at an intermediate stage, and the other subspecies are at a more advanced step of evolution. *Microcarpa* was identified from morphological characters, isozymes, and chloroplastic genome; however, nuclear genome revealed similarities with other subspecies, particularly *zebrina*, instigating consideration of a *zebrina/microcarpa* complex. *Truncata* is endemic to the highlands of Peninsular Malaysia. *Errans* from the Philippines was given the status of a subspecies of *M. acuminata* (22), although only one accession has been studied. It is the only wild *M. acuminata* with the  $\alpha$ -mitochondrial type, an important type found in a lot of diploid and triploid cultivars. It shares similarities with *banksii* nuclear genome. *M. balbisiana* has a more northerly distribution and, although not domesticated, it has been widely translocated for its many uses, thereby founding small free-growing populations from NG in the east to Sri Lanka in the west.



**Fig. 2.** Phylogenetic relations between AA cvs and wild *acuminata* subspecies. NJtree on genetic dissimilarities from 22 simple sequence repeat markers, on 41 AAw and 131 AAcv (also refer to Fig. S2): *M. acuminata* subsp. are in color (in gray for unclassified AAw); clusters of AAcv, identified by the name of a representative accession, are in black. AAcv appeared as hybrids between *M. acuminata* subspecies as illustrated by the clusters from Spiral to Beram collected in PNG, the native area of subsp. *banksii*. If the contribution of the *banksii* genome was still found dominant for the first clusters, it decreased rapidly, balanced by an increasing contribution of *zebrina/microcarpa* genome, in parallel with an increasing heterozygosity. The frequency of *banksii* cytoplasmic type V $\phi$  (12) decreased to the benefit of hybrid forms V $\alpha$  (V $\phi$   $\times$  II $\alpha$ ) or specifically to the Mala cluster, II $\phi$  (II $\alpha$   $\times$  V $\phi$ ). Contributions of these AAcv to triploids, as 2N donor (red arrows) and N donor (blue arrows), are illustrated for some AAA and AAB.

have been used for identification to the genus level, although they are less readily differentiable to the species level than seed phytoliths (22–24). Here, discriminating morphological criteria of leaf and seed phytoliths are reexamined and augmented for reliable interspecific and intraspecific differentiation within *Musa* and between *Musa* and *Ensete* (Figs. S3 and S4 and Tables S2 and S3). These revised discriminatory criteria are applied to assemblages at two key archaeological sites that anchor thresholds of prehistoric *Musa* geodomination: cultivation of *M. acuminata* at 6,950 to 6,440 calibrated y (cal) BP at Kuk Swamp in the highlands of NG and the dispersal under cultivation of sterile AAB plantains by 2,750 to 2,300 cal BP to Nkang, Cameroon.

Abundant Musaceae phytoliths in feature fills associated with agricultural practices and dated to 6,950 to 6,440 cal BP at Kuk are indicative of banana cultivation (25, 26). Identifiable morphotypes with attributes matching the comparative reference samples of *M. acuminata* subsp. *banksii* from NG (Fig. 3 A–C), the only naturally occurring *M. acuminata* subspecies on the is-

land, can be differentiated from other Musaceae grown there (Fig. S3 and Table S2), although archaeobotanical occurrences are rare (Fig. 3D). So, the occurrence of *M. acuminata* morphotypes in an early agricultural context is highly significant and provides a minimum date for the cultivation of subsp. *banksii* derivatives in NG.

Volcaniform leaf phytoliths of *Musa* spp. were reported from two secure archaeological contexts at Nkang in Cameroon: from two horizons within pit 9 ( $n = 20$ , dated to 2,790–2,300 cal BP) and from a charred deposit adhered to a pottery sherd from pit 7 ( $n = 5$ , dated 2,750–2,100 cal BP; Fig. 3 E–H) (22, 23, 27). The original morphological criteria used to distinguish *Musa* and *Ensete* phytoliths at Nkang have been revised (Fig. S4 and Table S3) and confirm the phytoliths at Nkang to be *Musa*, as opposed to the native African genus *Ensete*. So, this reevaluation of the Nkang assemblages provides a secure archaeobotanical foundation and minimum date for the cultivation of exotic plantains in West Africa.

This verification of the archaeobotany of *Musa* at Kuk and Nkang is essential to the

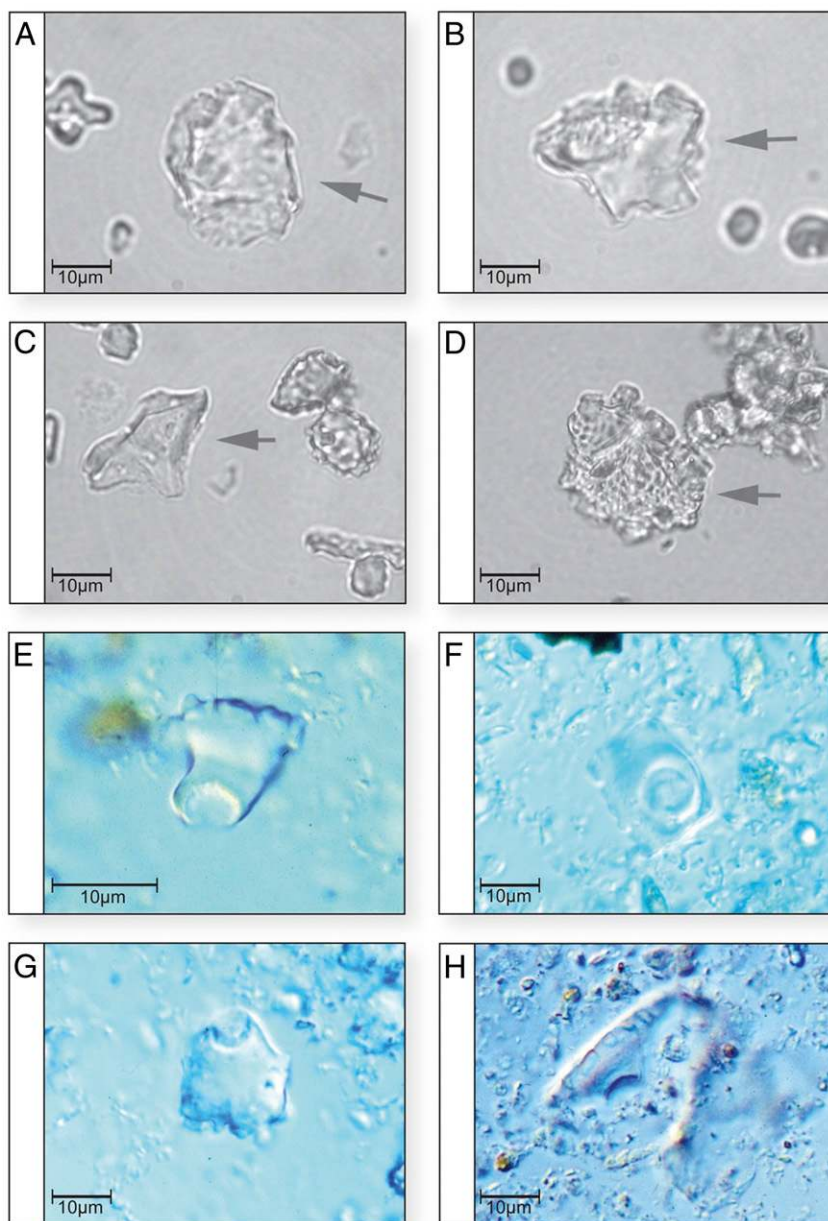
development of a chronological framework for the domestication phylogeny (from genetics) and dispersal of cvs (from linguistics). Other significant archaeobotanical finds of *Musa* are discussed with reference to other lines of evidence in the subsequent sections. However, differentiation below the genus level has not been possible at most archaeological or paleoecological sites; for example, *Musa* phytoliths date to 4,000 y ago at Kot Diji in Pakistan (28) and have also been associated with the post-3,500-y-old Lapita phenomenon in the western Pacific (21, 29). Exceptions include the identifications of *M. balbisiana* and cf. *M. acuminata* seeds from Terminal Pleistocene contexts at Beli-Lana on Sri Lanka (30) and the discrimination of *Ensete*, *M. acuminata* subsp. *banksii*, and Australimusa bananas in mid-Holocene contexts on central New Britain (31). The Sri Lanka evidence is consistent with modern, wild *M. acuminata* subsp. *burmannica* there, reflecting continuity with populations in East India and Myanmar. However, no significant cv (AA, AAA, or AAB) based on subsp. *burmannica* has been identified; consequently, this subspecies and this region are not significant for understanding banana domestication.

### Tracing Banana Dispersal Through Historical Linguistics

A cultivated plant often migrates with its name, and when this plant is culturally innovative, its name is often retained in the receiving language. Successive dispersals accumulate terminological changes from the original forms that enable the historical path of successive transformations to be reconstructed. Unlike names that frequently change when moving, the associated genotype, in case of banana, is relatively fixed by vegetative propagation. A geographical reading of the linguistic paths illustrates the movements of these terms and associated genotypes, providing rich evidence for the spread of *Musa* cvs.

More than 1,100 terms related to banana varieties were collected during a large survey focused on Melanesia and SEA (Table S4). The great variety of modern terms for bananas have been described in terms of a smaller number of cognate sets representing common origins with distinctive, although overlapping, distributions. Four of these cognate groups provide valuable cross-regional insights into ancient processes of banana domestication and dispersal (Fig. 4 and Fig. S5), those reflecting \*muku, \*punti, \*qaRutay, and \*baRat (32–34). Phylograms for these key terms have been reconstructed using the standard comparative method (35), implemented with bioinformatic software to determine likely developmental paths





**Fig. 3.** Ancient banana phytoliths recovered from archaeological excavations at Kuk Swamp, Papua New Guinea (SEM images, A–D), and at Nkang, Cameroon (optical images, E–H). For Kuk: (A and B) dorsal and lateral view of *Eumusa* seed phytolith recovered from sample 5, Kuk dated to 6,990 to 6,440 cal BP; (C) lateral view of another *Eumusa* seed phytolith from sample 5, Kuk; and (D) dorsal view of *Eumusa* seed phytolith recovered from sample 28, Kuk predating 3,000 cal BP. Morphotypes shown in A–C are specific to *M. acuminata* (Fig. S3 A, B and D). The morphotype with lobate margins shown in D occurs in *M. acuminata* (Fig. S3A) and a similar morphotype occurs in *Musa schizocarpa* (Fig. S3 E and G). For Nkang: (E–H) multidimensional diagnostics of *Musa*-type volcaniform phytolith from Pit F9, Horizon 7 at Nkang, dated to 2,750 to 2,100 cal BP; (E) small indentation on left side of crater rim; (F) rectangular base, psilate surface, eccentric cone, and continuous rim; (G) processes along the edge of the base; and (H) processes along the base.

(36). Each distribution is suggestive of different historic processes of dispersal; local diversity of other terminological sets, primarily in NG and near the Himalayas, is suggestive of long-term use of Musaceae.

The distribution of more than 40 surviving reflexes of the protoform \*muku is

concentrated in Wallacea and western NG (34) (Fig. 4A). This “cognate” is attested across a number of unrelated language families. It spreads west from NG across a range of modern Papuan languages in Wallacea, suggesting an early distribution before the purported Austronesian dispersal ca. 4,000 to 3,500 cal BP (34, 37). Reflexes

of \*muku are found in areas (38) where pre-Austronesian linguistic ecologies survive. In areas now dominated by Austronesian languages, \*muku represents terminological survival and strong resistance to change in the face of radical language shift.

The protoform \*qaRutay is evident in the Philippines with forms “\*agutay” and “kelutay” for wild bananas. Hence, it is likely that the initial reflexes of \*qaRutay were generated in the Philippines region (Fig. 4C). The linguistic data suggest an initial southward dispersal of \*qaRutay from the Philippines into Indonesia, with later “ripples” west into mainland Asia and east as far as NG. The initial southward dispersal to Indonesia saw the \*qaRutay term come into the attested range of \*muku. Early branches of the \*qaRutay dendrogram, representing reflexes of \*kaluay and \*kalu, are distributed east to Melanesia and west to the Asian mainland. An intermediate form is \*kelo, which is currently attested in multiple languages as far west as the Indian subcontinent and east on NG, but is now almost absent from ISEA. The last of these terminological transformations, \*loka (from \*kelo via metathesis), has only recently begun its spread and is restricted to locales across Sulawesi and eastern Indonesia.

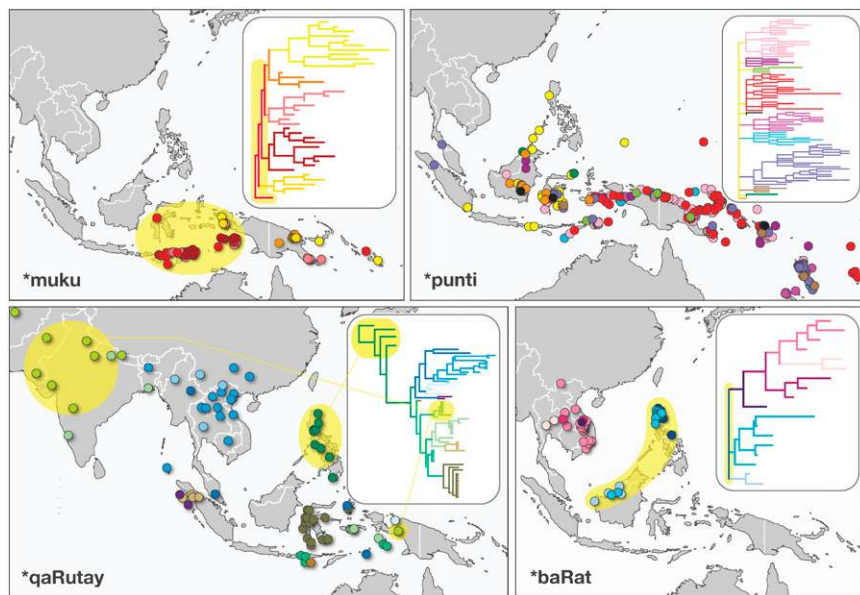
The analysis of \*baRat indicates a probable origin in the Philippines and later spread toward Borneo and SEA (Fig. 4D). Although not widespread, \*baRat dominates in those areas in which it is found.

In contrast to the phylogenetic depth of \*qaRutay, the lack of a hierarchy associated with \*punti reflects a relatively recent and rapid spread with Austronesian languages in ISEA (Fig. 4B) and attests multiple processes of local differentiation.

### Multidisciplinary Consilience on Banana Domestication

The integrated analysis of genetic, linguistic, and archaeological data enable a coherent reconstruction of the major events in banana domestication, including geographical occurrence, timeline, and cultural associations.

**Generation of Diploid cv Subgroups.** The first crucial step in banana domestication was the hybridization between geographically isolated subspecies of *M. acuminata*, which were brought into contact by people within ISEA and western Melanesia. Ethnobotany supports the human transport of these fertile genotypes because useful parts of banana plants were potentially exploited over millennia for food, fodder, medicine, fiber, domestic uses, or construction materials (39). The term “cultiwild” has been coined to denote these potentially pre-domesticated banana plants, “[a]ny



**Fig. 4.** Dendrograms and maps indicating the development and dispersal of the major banana cognate sets referred to here. Maps indicate different color-coded branches of each dendrogram; the identification of the four cognate sets was moderated by linguistically informed judgments on the likelihood of similar developments being independent. \*muku (A) is restricted to areas near NG, with the center of diversity along the west, whereas the \*punti terms (B) are widespread with little indication of a “homeland.” \*baRat (D) is ambiguous between a homeland in the Philippines, Borneo, or on mainland SEA. \*qaRutay (C) is the term with the most divergent dendrogram, with an unambiguous origin in the Philippines and many clades being attested to both the east and the west, including NG and South Asia. Likely homelands for the different terminologies are shown in yellow shading, and two locations for \*qaRutay are similarly shaded for the subclade containing \*kela, found in the extreme west and extreme east of the range. Phylogeographies of individual clades for the different cognate sets are presented in Fig. S5.

member of a wild population, or its direct fertile derivatives, whether cloned or not, growing outside of the natural habitat, or range, of the species [or subspecies] to which it belongs” (40). Diverse wild *M. acuminata* subspecies were probably subject to predomesticated exploitation in their native regions. In particular, cultiwilds derived from subsp. *banksii* in the NG region contributed genetically to numerous cv groups; their antiquity is attested by archaeobotanical evidence at Kuk by at least 6,950 to 6,440 cal BP.

Domesticated parthenocarpic diploids resulted from anthropic translocation of cultiwilds outside of their natural range and subsequent hybridization with local subspecies in at least three contact regions (Fig. 5A): between NG and Java (Southern), between NG and the Philippines (Eastern), and among the Philippines, Borneo, and mainland SEA (Northern). Each of these hybridization areas has a clear linguistic association (Fig. 5B).

In the Southern contact area, between NG and Java, initial hybridizations occurred between subspecies of *banksii* and the *zebrina/microcarpa* complex. This early dispersal can be traced linguistically by the distribution of \*muku reflexes in Wallacea and in western NG (34). This dispersal predates the purported arrival of Austro-nesian languages in the region at approx-

imately 3,500 cal BP. This antiquity would be consistent with the distribution of derived cvs AACvs or AAACvs that are not now found in the area of origin but are widely cultivated in Africa.

The Eastern contact area is genetically evidenced by the contribution of *M. acuminata* subsp. *errans* from Philippines to the cytoplasm of numerous AACvs. This Eastern contact area is linguistically attested by the Philippines’ derived Negrito term \*qaRutay, which has widely dispersed modern reflexes. This contact area accords with a prehistoric linguistic trail from the Philippines to northwestern NG (41). *M. balbisiana* probably followed this trail from South China before hybridizing with populations of *M. acuminata* subsp. *banksii*.

Within the Eastern contact area, the southward translocation of *M. balbisiana* is attested by small populations growing along the trail from South China to NG. Although not domesticated for edibility, *M. balbisiana* has been used and moved by people for multiple purposes (e.g., male bud, immature fruit, leaves). Moreover, several AAB and ABB triploid subgroups, with *banksii* genome as an A component, are absent from the northern end of this trail, thereby indicating a southern origin for the contact area and availability of the B genome in this region. The reverse hy-

pothesis, a northward translocation of *M. acuminata* subsp. *banksii* and subsequent hybridization with *M. balbisiana* populations, is discounted because no genetic trace of the subspecies is found in the Philippines and further north.

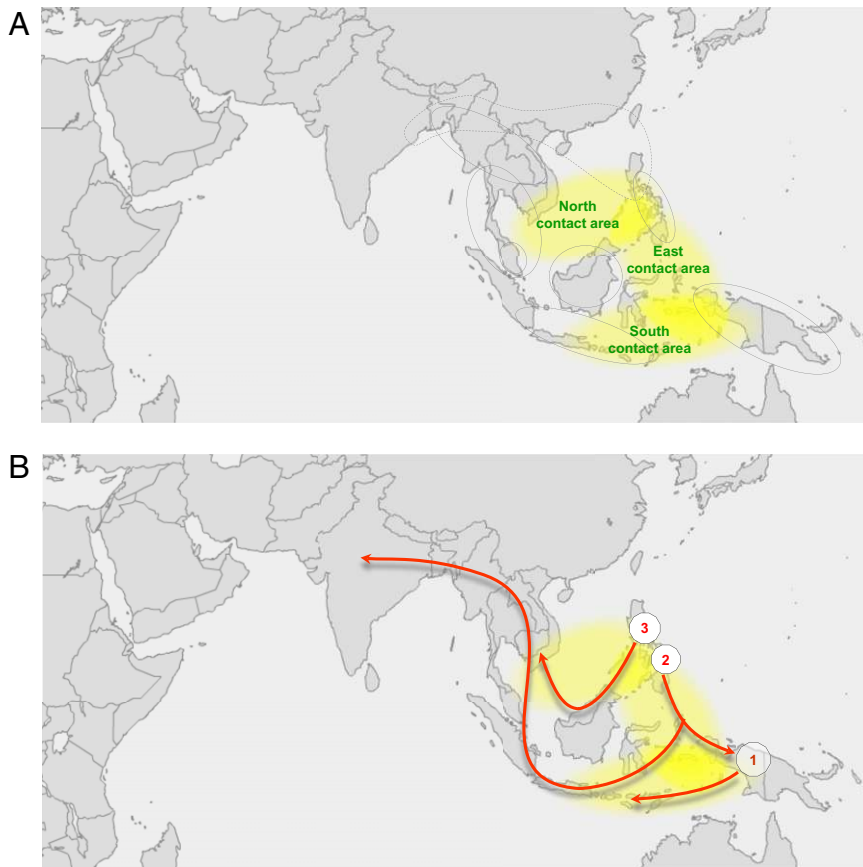
The Northern contact area is hypothesized for the formation of several AACvs, which contributed to important commercial AAA crops. Their genomes represent contributions of *M. acuminata* subsp. *malaccensis* or *microcarpa*, derived from mainland SEA and Borneo, and of subspecies *errans* from the Philippines. The dispersal of the reflex \*baRak in North Borneo from the prototerm \*baRat in the Philippines clearly corresponds to the genetic hypothesis.

These hybrid AACvs were widely adopted within the primary area of *M. acuminata* subspecies diversity, where they are still commonly cultivated. Very few dispersed outside this primary area. A striking exception is the AA Mlali subgroup, which provided the parents of several popular triploid subgroups. Allelic frequencies point to this subgroup being *M. acuminata* subsp. *banksii* × *zebrina* hybrids, with the Southern contact area as probable origin. Interestingly, this AACv subgroup is now only grown along the east coast of Africa and proximal islands (i.e., Madagascar, Zanzibar, Comoros). The isolated occurrence of these cultivated AA diploids far away from the region of presumed origin, in which they no longer occur, suggests an ancient transfer across the Indian Ocean.

**Emergence of Triploid cv Subgroups.** A direct consequence of the perturbed gamete formation in hybrid AACvs was the emergence of triploid genotypes AAA and AAB or ABB by interspecific hybridizations. Triploidization occurred independently in various contact areas between diploids and from different parental combinations (Fig. 6). Some of the triploids selected for cultivation underwent somaclonal variation under vegetative propagation to produce the modern phenotypically characterized triploid subgroups. Triploidization probably started early after the AACv emergence and is certainly still ongoing, as attested by the recent origin inferred for the genome of several AAA in NG.

Of the numerous triploid subgroups, three are remarkable because they are largely cultivated far from their region of generation: the African AAA “Mutika Lujugira,” AAB “African Plantains,” and AAB “Pacific Plantains.” These triploids are not cultivated in Asia, with the exception of a few African Plantain cvs that were probably introduced to South India with the East African slave trade. The antiquity of each subgroup is attested by the extraordinarily large number of cultivated

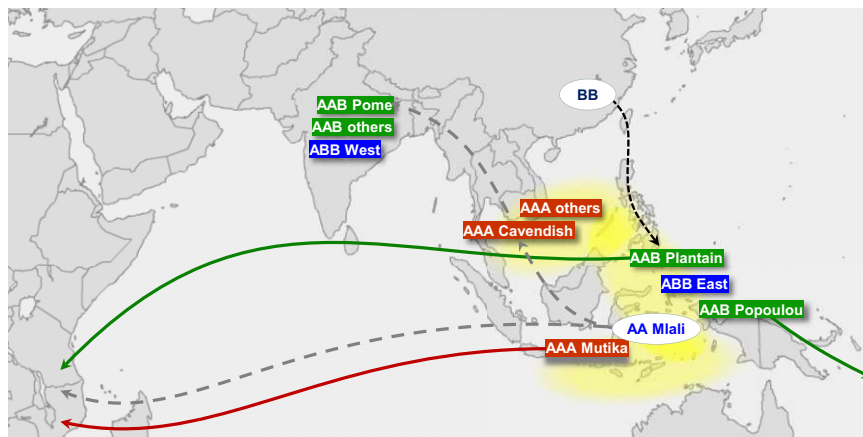




**Fig. 5.** Genetically derived contact areas between *M. acuminata* subsp. at the origin of cultivated diploids and corresponding linguistic evidence. (A) The three main contact areas: north among *malaccensis*, *microcarpa*, and *errans*; east between *errans* and *banksii*; and south among *banksii*, *zebrina*, and *microcarpa*. (B) Linguistic extension for terms \*muku ("1"), \*qaRutay ("2"), \*baRat ("3"), and their derivatives (Fig. 4).

morphotypes (42, 43) from only very few imported introductions, which indicates extremely long periods of somaclonal mutation within their regions of cultivation.

The AAA Mutika Lujugira subgroup is widely cultivated in the East African highlands for cooking or brewing beer. The constitutive *banksii* and *zebrina*



**Fig. 6.** Origins and migrations of the main triploid subgroups. Plain arrows indicate long-term prehistoric migrations of triploid cvs to Africa and Pacific islands. Gray dotted arrows indicate (i) the migrations of Mlali AAcv subgroup, which is not found in ISEA today, to mainland southeast Asia, where it contributed to AAA Cavendish, then to India, where it met *M. balbisiana* to give AAB Pome; and (ii) migrations of the Mlali subgroup to the East African coast. Black dotted arrows indicate the route of *M. balbisiana* from south China to NG over Taiwan and the Philippines, if Austronesian speakers were instrumental in the dispersal of this species.

genomes indicate the eastern side of the Southern contact area as origin, where genetically close AAAs are still cultivated.

The AAB African Plantains are a staple crop in the rainforest zone of West and Central Africa. The constitutive A genomes were identified as *banksii*. An AAB similar to African Plantains has been discovered under traditional cultivation by Negrito tribes in the Philippines (44). Another AAB subgroup, Laknau, with analogous morphology and genetic similarity has been documented in the Philippines' vicinity. These findings suggest a probable origin for the African Plantains in the Eastern contact area between the Philippines and NG, possibly around the Celebes Sea (45).

Two independent introduction events are proposed for these triploid subgroups to Africa: AAA Mutika Lujugira and associated AACv to East Africa from the South contact area between Java and NG and AAB African Plantains from the East contact area between the Philippines and NG. The arrival of bananas to East Africa is not documented archaeologically; the reported finds at Munsu, Uganda, are problematic, primarily in terms of chronology (46, 47). The generic names for African Plantain are mostly reflexes of the proto-Bantu term \*kondo (48), indicating that they arrived in Africa before the expansion of Bantu-associated cultures from ca. 3,000 cal BP (49). The archaeological finds at Nkang (Cameroon) 2,750 to 2,100 BP conform to this timeframe and predate the spread of Austronesian languages from Borneo to Madagascar. Consequently, the mode of dispersal from Asia to Africa is uncertain; terrestrial (42) and maritime routes (48) have been proposed. It is currently unclear how, or if, the banana phytoliths dating to ca. 4,000 y ago at Kot Diji in Pakistan (28) are relevant to understanding the dispersal of bananas to Africa.

AAB Pacific Plantains are spread over the whole Pacific and include the closely related Maia Maoli, Popoulou, and Iholena. Like African Plantains, they are comprised of *M. balbisiana* and *M. acuminata* subsp. *banksii* genomes. The absence of genetic relationships with Philippine cvs and the persistence in NG of genetically similar AAB forms indicate an origin in the Eastern contact area, but east of and distinct from that of AAB African Plantains. AAB Pacific Plantains potentially originate along the north coast of NG or in the Bismarck Archipelago (29) and were dispersed across the Pacific by Austronesian-speaking colonists (30).

The continual relocation of AACvs by people contributed to the development of numerous other triploid subgroups

through hybridization with local diploids. For example, the globally distributed commercial AAAs, “Gros Michel” and “Cavendish,” derive genetically from a 2N gamete of the AACv Mlali subgroup and an N gamete of the AACv Khai. The northern origin of these AAA—around the Gulf of Thailand or the South China Sea—implies an earlier northwestward dispersal of Mlali from its original Southern contact zone (Fig. 6). Further westward, the Mlali genome reached India and is found in the Indian AAB cv subgroups “Pome” and the genetically proximal “Nendra Padathi” and “Nadan” (Fig. 6). The long dispersal route of these AA cvs from Wallacea to India is supported linguistically by the continuous series of \*qaRutay reflexes from Indonesia over mainland SEA to India (Fig. 5B).

## Conclusions

The integration of archaeology, genetics, and linguistics provides robust insights to the history of banana domestication. Foremost, it allows us to clearly identify several major stages in the domestication history of *Musa* bananas. Early contacts between different *Musa* genotypes were created by human interactions, whether

by migration or exchange. The resultant interspecific and intersubspecific hybridizations generated both parthenocarpic diploids and triploids. Some of these hybrid cvs, whether by preference or by chance, were widely adopted and dispersed. A small number of triploids, perhaps because of the environmental adaptability conferred by triploidy, have been dispersed by clonal propagation across vast areas.

Research on banana cultivation and domestication is a window on often poorly known human/environment interactions within tropical and subtropical rainforests in the past. Additionally, the dispersals of bananas through wet tropical and subtropical regions, including around or across the Indian Ocean, are indicators of interlinked, yet predominantly local, social networks extending from NG to West Africa; these networks are at least 2,500 y old. Research on banana domestication demonstrates how multidisciplinary teams, as well as the application of new methods within each discipline, are essential for unraveling complex social processes in the past.

Bananas have contributed over several millennia to the staple diets of numerous

peoples in the tropics and subtropics. Current global production of more than 100 million tons is based on large-scale vegetative propagation of a small number of genotypes, which derive from only a few ancient sexual recombination events. These genetically restricted and inflexible clones are particularly susceptible to diseases, pests, and current ecological changes. The challenge for banana improvement is to produce resistant and sterile polyploid hybrids through genetic recombinations of fertile diploids that meet consumer expectations for each cv type. The required breeding strategy will need to reproduce the sequence of crossings and selections that occurred minimally during the past 6,500 y, while substituting punctually some genitors from closely related genomes selected for their level of resistance to biotic and abiotic stresses. Hence, a prerequisite for banana improvement is to reconstruct as precisely as possible the domestication pathways of the major cv groups.

**ACKNOWLEDGMENTS.** The authors thank Toby Wood (Monash University) and Kay Dancy (Australian National University) for their assistance with figure production.

- Frison E, Sharrock S (1998) *Musa* production around the world - trends, varieties and regional importance. *Focus Paper 2, INIBAP Annual Report* (INIBAP, Montpellier, France), pp 42–47.
- Simmonds NW (1962) *The Evolution of the Bananas* (Longmans, London).
- Sharrock S (1990) Collecting *Musa* in Papua New Guinea. *Identification of Genetic Diversity in the Genus Musa*, ed Jarret RL (International Network for the Improvement of Banana and Plantain, Montpellier, France), pp 140–157.
- Jarret RL, Gawel N, Whitemore A, Sharrock S (1992) RFLP based phylogeny of *Musa* species in Papua New Guinea. *Theor Appl Genet* 84:579–584.
- Ude G, Pillay M, Nwakanma D, Tenkouano A (2002) Analysis of genetic diversity and sectional relationships in *Musa* using AFLP markers. *Theor Appl Genet* 104: 1239–1245.
- Creste S, Tulmann NA, De Oliveira SS, Figueira A (2003) Genetic characterization of banana cultivars (*Musa* spp.) from Brazil using microsatellite markers. *Euphytica* 132:259–268.
- Heslop-Harrison JS, Schwarzacher T (2007) Domestication, genomics and the future for banana. *Ann Bot (Lond)* 100: 1073–1084.
- Hippolyte I, et al. (2010) A saturated SSR/DAT linkage map of *Musa acuminata* addressing genome rearrangements among banana. *BMC Plant Biol* 10:65–83.
- Risterucci A-M, et al. (2009) Development and assessment of Diversity Arrays Technology for high-throughput DNA analyses in *Musa*. *Theor Appl Genet* 119:1093–1103.
- Jenny C, et al. (2003) Tracking the banana: Its significance in Early Agriculture. *Banana in Genetic Diversity of Cultivated Tropical Plants* (Enfield Science, Montpellier, France), pp 99–124.
- Horry JP, Jay M (1988) Distribution of anthocyanins in wild and cultivated banana varieties. *Phytochemistry* 27:2667–2672.
- Lebot V, Aradhya KM, Manshardt R, Meilleur B (1993) Genetic relationships among cultivated bananas and plantains from Asia and the Pacific. *Euphytica* 67:163–175.
- Carreel F, et al. (1994) Evaluation de la diversité génétique chez les bananiers diploïdes (*Musa* sp.). *Genet Sel Evol* 26:125–136.
- Noyer JL, Causse S, Tomekpe K, Bouet A, Baurens FC (2005) A new image of plantain diversity assessed by SSR, AFLP and MSAP markers. *Genetica* 124:61–69.
- Carreel F, et al. (2002) Ascertaining maternal and paternal lineage within *Musa* by chloroplast and mitochondrial DNA RFLP analyses. *Genome* 45: 679–692.
- Perrier X, et al. (2009) Combining biological approaches to shed light on the evolution of edible bananas. *Ethnobot Res Appl* 7:199–216.
- Shepherd K (1999) *Cytogenetics of the genus Musa*. 160 (International Network for the Improvement of Banana and Plantain, Montpellier, France).
- Dessauw D (1988) Etude des facteurs de la stérilité du bananier (*Musa* spp.) et des relations cytotaxinomiques entre *M. acuminata* Colla et *M. balbisiana* Colla. *Fruits* 27:539–558.
- D’Hont A, Paget-Goy A, Escoute J, Carreel F (2000) The interspecific genome structure of cultivated banana, *Musa* spp. revealed by genomic DNA in situ hybridization. *Theor Appl Genet* 100:177–183.
- Raboin LM, et al. (2005) Diploid ancestors of triploid export banana cultivars: Molecular identification of 2n restitution gamete donors and n gamete donors. *Mol Breed* 16:333–341.
- Horrocks M, Bedford S, Spriggs M (2009) A short note on banana phytoliths in Lapita, immediately post-Lapita and modern period archaeological deposits from Vanuatu. *J Archaeol Sci* 36:2048–2054.
- Ball T, Vrydaghs L, van den Houwe I, Manwaring J, De Langhe E (2006) Differentiating banana phytoliths, wild and edible: *Musa acuminata* and *Musa balbisiana*. *J Archaeol Sci* 33:1228–1236.
- Vrydaghs L, et al. (2009) Differentiating the volcaniform phytoliths of bananas: *Musa acuminata*. *Ethnobot Res Appl* 7:239–246.
- Lentfer C (2009) Tracing domestication and cultivation of bananas from phytoliths: An update from Papua New Guinea. *Ethnobot Res Appl* 7:247–270.
- Denham TP, et al. (2003) Origins of agriculture at Kuk Swamp in the highlands of New Guinea. *Science* 301: 189–193.
- Denham TP, Haberle S, Lentfer C (2004) New evidence and interpretations for early agriculture in Highland New Guinea. *Antiquity* 78:839–857.
- Mbida CM, et al. (2001) First archaeological evidence of banana cultivation in central Africa during the third millennium before present. *Veget Hist Archaeobot* 10: 1–6.
- Fuller D, Madella M (2009) Banana cultivation in South Asia and East Asia. *Ethnobot. Res. Appl.* 7: 333–351.
- Lentfer C, Green R (2004) Phytoliths and the evidence for banana cultivation at the Lapita Reber-Rakival Site on Watom Island, Papua New Guinea. *A Pacific Odyssey: Archaeology and Anthropology in the Western Pacific: Records of the Australian Museum* 29, eds Attenbrow V, Fullagar R (Australian Museum Sydney), pp 75–88.
- Kajale MD (1989) Mesolithic exploitation of wild plants in Sri Lanka: Archaeobotanical study at the cave site of Beli-Lena. *Foraging and Farming: The Evolution of Plant Exploitation*, eds Harris D, Hillman G (Unwin Hyman, London), pp 269–281.
- Lentfer C, Pavlides C, Specht J (2010) Natural and human impacts in a 35 000-year vegetation history in central New Britain, Papua New Guinea. *Quat Sci Rev* 29:3750–3767.
- Donohue M, Denham TP (2009) Banana (*Musa* spp.) domestication in the Asia-Pacific region: Linguistic and archaeological perspectives. *Ethnobot Res Appl* 7:293–332.
- Zorc RD (1995) A glossary of Austronesian reconstructions. *Comparative Austronesian Dictionary*, ed Tryon DT (Mouton de Gruyter, Berlin), Part I, pp 1105–1197.
- Denham TP, Donohue M (2009) Pre-Austronesian dispersal of banana cultivars west from New Guinea. *Archaeol Ocean* 44:18–28.
- Joseph BD, Janda RD, eds. (2003) *Handbook of Historical Linguistics* (Blackwell, London).
- Saitou N, Nei M (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425.
- Blust R (2009) *The Austronesian Languages* (Pacific Linguistics, Canberra, Australia).
- Nichols J (1992) *Linguistic Diversity in Space and Time* (Univ Chicago, Chicago).

39. Kennedy J (2009) Bananas and people in the homeland of genus *Musa*: Not just pretty fruit. *Ethnobot Res Appl* 7:179–197.
40. De Langhe E, Vrydaghs L, de Maret P, Perrier X, Denham TP (2009) Why bananas matter. *Ethnobot Res Appl* 7:165–177.
41. Mahdi W (1994) Some Austronesian maverick protoforms with culture-historical implications: I. *Ocean Linguist* 33:167–229.
42. De Langhe E (2007) The establishment of traditional plantain cultivation in the African rain forest: A working hypothesis. *Rethinking Agriculture*, eds Denham TP, Irarte J, Vrydaghs L (Left Coast Press, Walnut Creek, CA), pp 361–370.
43. Ploetz RC, Kepler AK, Daniells J, Nelson SC (2007) Banana and plantain – an overview with an emphasis on Pacific island cultivars [Musaceae (banana family)]. *Species Profiles for Pacific Island Agroforestry* (Permanent Agriculture Resources, Ho-lualoa, Hawaii).
44. De Langhe E, Valmayor RV (1979) French plantains in Southeast Asia. *IBPGR Newsletter for SE Asia* 3–4.
45. De Langhe E (2009) Relevance of banana seeds in archaeology. *Ethnobot. Res. Appl.* 7:271–281.
46. Lejju BJ, Robertshaw P, Taylor D (2006) Africa's earliest bananas? *J Archaeol Sci* 33:102–113.
47. Neumann K, Hildebrand E (2009) Early bananas in Africa: The state of the art. *Ethnobot. Res. Appl.* 7: 353–362.
48. Blench R (2009) Bananas and plantains in Africa: Re-interpreting the linguistic evidence. *Ethnobot. Res. Appl.* 7:363–380.
49. Vansina J (1990) *Paths in the Rainforests* (Univ Wisconsin, Madison, WI).