Domestication and varietal diversification of Old World cultivated cottons (*Gossypium* sp.) in the Antiquity

Domestication et diversification variétale des cotons cultivés (Gossypium sp.) de l'Ancien Monde dans l'Antiquité

Christopher Viot

Electronic reference

Christopher Viot, « Domestication and varietal diversification of Old World cultivated cottons (*Gossypium* sp.) in the Antiquity », *Revue d'ethnoécologie*, 15 | 2019, Online since 30 June 2019. URL : http://journals.openedition.org/ethnoecologie/4404 ; DOI : 10.4000/ethnoecologie.4404

Author affiliation

Christopher Viot. CIRAD, UMR AGAP, F-34398 Montpellier, France. AGAP, Univ Montpellier, CIRAD, INRA, Montpellier SupAgro, France. christopher.viot@cirad.fr



Abstract

Textile use of the cotton fibres from the Old World species *Gossypium arboreum* (Tree Cotton) and *G. herbaceum* (Levant Cotton) started around eight thousand years ago and possibly earlier. During the third millennium before the present, cotton cultivation, textile use and trading developed strongly in many places of the Indian Subcontinent and the Near East, but the species involved are often undetermined. *Gossypium arboreum* and *G. herbaceum* are difficult to distinguish morphologically when dealing with archaeological remains. Many traditional varieties have been described for each of these two species; an evolution is apparent from perennial, rather primitive forms through to annual varieties and modern phenotypes, with adaptions to cultivation in diverse conditions, including cold climates and short summers. The present work examines which varieties could have been cultivated in the Antiquity and in which regions had cotton cultivation spread then. Some extant varieties are hypothesized as close to the cultivars of Antiquity. The geographic distribution of each species probably reflects the initial domestication region and the subsequent adaptations, particularly regarding climate. Recent progress on ancient DNA analysis should permit easier specific assignments of archaeological remains of cotton seeds, fibres, threads or fabrics.

Keywords: Gossypium arboreum, Gossypium herbaceum, cotton, domestication, varietal diversification, Old World, Antiquity

Résumé

L'usage textile des fibres de coton des espèces de l'Ancien Monde, Gossypium arboreum et G. herbaceum, a commencé il y a huit millénaires, et possiblement plus tôt. Au cours du troisième millénaire avant le présent, la culture, l'utilisation textile et le commerce du coton se sont fortement développés dans de nombreux sites du Subcontinent Indien et du Moyen-Orient, mais sans détermination, souvent, de l'espèce impliquée dans chaque cas. G. arboreum et G. herbaceum sont très difficilement distinguables morphologiquement dans les vestiges archéologiques. Plusieurs variétés traditionnelles ont été décrites pour chacune des deux espèces; une évolution apparaît depuis des formes pérennes relativement primitives jusqu'à des types annuels de morphologie moderne, avec l'adaptation à la culture en conditions très diverses incluant climats froids et étés courts. Le présent travail examine lesquelles de ces variétés ont pu être cultivées dans l'Antiquité et dans quelles régions de l'Ancien Monde avait alors diffusé la culture du coton. Certaines des variétés actuelles peuvent être supposées représentatives des cultivars de l'Antiquité. La répartition géographique des deux espèces résulte à la fois du lieu initial de domestication et de leurs adaptations agronomiques, particulièrement concernant le climat. Les progrès récents en analyse des ADN anciens devraient permettre de déterminer plus facilement l'espèce des graines, fibres et textiles archéologiques de coton.

Mots-clés : *Gossypium arboreum*, *Gossypium herbaceum*, coton, domestication, diversification variétale, Ancien Monde, Antiquité

Introduction

The Old World cultivated cotton species Gossypium arboreum and G. herbaceum give textile fibres whose use by humans stretches back thousands of years. Cotton fibres and a thread have been found in different burial contexts in Balochistan, Pakistan that could be dated to the 6th millennium BCE (Moulherat et al. 2002). Both cotton fibres and seeds appear as frequent in the second half of the 3rd millennium BCE in the Indian sub-continent which seems to have developed as a major centre for cotton production and exportation, in particular in the Mediterranean region. In Africa, the earliest cotton textiles are attested at a Nubian site in the 3rd millennium BCE (Chowdhury & Buth 1971), without certainty as to their date and origin. In the Classical Antiquity, cotton cultivation seems to have been widely established in the northeastern part of the African continent and the Arabian Peninsula (Bouchaud et al. 2018). In Central Asia, cotton textiles are known from the Later Han (25-220 AD) in Khotan and the Turfan basin, in N-W China. The cotton plant was present by 1st century AD in west Yunnan and Szechwan (Kuhn 1988) in Southern China and by Sung times (5th c. AD) had reached the Chiang-nan region, south of Shanghai (Chao 1977). Throughout the cotton growing regions of the Old World during the past two centuries, there has been a gradual partial or total replacement of both G. arboreum and G. herbaceum by the New World tetraploid species G. hirsutum and G. barbadense.

The textile fibres of cotton crop species are hairs that grow out of the seed coat. The harvested seed-cotton is de-seeded, that is, fibres are separated from seeds, and then the longest fibres or lint can be spun while the seeds are fed to animals or pressed for oil extraction. Thanks to its oil- and protein-rich seeds, cotton is also an oleoproteaginous crop, which potentially constitutes another trait that led to domestication and cultivation.

The diploid cottons *G. arboreum* and *G. herbaceum* are sister species with a most recent common ancestor approximately 1-2 Mya (Hinze *et al.* 2015; Renny-Byfield *et al.* 2016); they are rather similar with respect to plant morphology, lint characteristics and seed morphology. Archaeological remains of cotton textiles, fibres or seeds can only with much difficulty be assigned to one of these two cotton species. A precise specific assignment is crucial for the understanding of how cotton was domesticated, the agricultural, artisanal and commercial activities it was linked with and the cultural and trade networks it supported in Afro-Eurasia before the modern era.

This paper examines how *G. arboreum* and *G. herbaceum* are agronomically and morphologically differentiated, the timing of the geographic spread of these Old-World cottons in the Antiquity, which of the known cultivated types of each species were likely involved and which traits could be useful in distinguishing the two species in archaeological remains.

Taxonomy of Old World cultivated cottons

The scientific names *Gossypium herbaceum* L. and *G. arboreum* L. were given by Linnaeus (1753), who also named the two American cultivated cotton species, *Gossypium hirsutum* L. and *G. barbadense* L. Parlatore (1866) published the first relatively modern classification of *Gossypium*. Making the essential distinction between diploid and tetraploid species, Zaitzev (1928, cited in Hutchinson *et al.* 1947) contributed to our understanding of the cultivated cottons, partitioning them into the currently recognized four species. This insight has stood the test of time, as supported in all later work, including classical taxonomy (Hutchinson *et al.* 1947); Fryxell (1979)) and molecular phylogenetic and genomic studies (Wendel *et al.* 1989, 2010; Wendel & Albert 1992; Wendel & Grover 2015).

Hutchinson *et al.* (1947) defined the taxonomy for the two Old World cultivated cotton species as:

Gossypium arboreum L., Sp. Pl. 693. 1753.

- *G. arboreum* cv. *indicum*
- *G. arboreum* cv. *burmanicum*
- *G. arboreum* cv. *bengalense*
- *G. arboreum* cv. *cernuum*
- *G. arboreum* cv. *soudanense*
- *G. arboreum* cv. *sinense*

Gossypium herbaceum L., Sp. Pl. 693. 1753.

- o G. herbaceum L. var. africanum (Watt) Vollesen
- o *G. herbaceum* L. var. *acerifolium* (Guill. et Perr.) Chevalier (1939)

As for *G. arboreum*, these authors considered that no neatly differentiated subspecies could be defined. They also rejected a taxonomical distinction between perennial and annual types and they discussed the usefulness and practicability of the distinction of geographical races and listed the six geographical races or cultivars as above.

Fryxell (1979) cited one inferior subdivision, in only one of the two species, *G. herbaceum* var. *africanum*.

Most modern germplasm banks follow the taxonomic concepts of Hutchinson *et al.* (1951) and Fryxell (1979), notwithstanding the likely artificiality of the cultivar groups as meaningful biological or genetic entities (Wendel *et al.*, 1989). For example, the Germplasm Resources Information Network (GRIN, https://npgsweb.ars-grin.gov/, consulted 2018/11/18) uses the following classification system:

- *G. herbaceum* subsp. *africanum*, the wild, non-domesticated form in Southern Africa and possible ancestor of the domesticated variants;
- *G. herbaceum* subsp. *herbaceum*, the improved, hypothetically monophyletic group comprising the following four geographical cultivars, the same as identified by Kulkarni *et al.* (2009):
 - *G. herbaceum* subsp. *Herbaceum* cv. *acerifolium*
 - *G. herbaceum* subsp. *Herbaceum* cv. *persicum*
 - G. herbaceum subsp. Herbaceum cv. kuljianum
 - *G. herbaceum* subsp. *Herbaceum* cv. *wightianum*.

Vernacular names are numerous for these two species:

- *G. arboreum*: Tree cotton, Ceylon cotton, Indian Cotton Tree, Nankeen cotton, Oriental cotton, red-flowered cotton tree (English), cotonnier arborescent, cotonnier en arbre, cotonnier d'Inde, cotonnier rouge (French).
- *G. herbaceum*: Levant cotton, Syrian cotton, Arabian cotton, Maltese cotton, short-staple cotton (English); cotonnier herbacé, cotonnier africain (French).

The taxonomic history of these two cotton species appears as complex (see Supplementary data S1) and confusing, species names *G. arboreum* and *G. herbaceum* having seemingly been used interchangeably by different authors. Moreover, it is apparent that interspecific gene flow has occurred (Wendel *et al.* 1989; Wendel, pers. comm.), either in farmers fields over the last couple of millennia, in germplasm banks during seed propagation, or intentionally in breeding programs. There is thus only a relative utility to the older and commonly cited cultivar groupings.

Phylogeny and the origins of the textile lint

In Figure 1 is presented a simplified general evolutionary tree of the *Gossypium* genus. *G. arboreum* and *G. herbaceum* are sister species, the only known A-genome species and the only diploids 2n=26) with spinnable fibres among the whole *Gossypium* genus.

The allotetraploid (2n=52) cottons combine complete A-genome and D-genome chromosome sets; they originated and were domesticated in the New World. Molecular genetics studies indicate that the At subgenome of the allotetraploid *Gossypium* species is nearly equidistant from *G. herbaceum* and *G. arboreum*. The D-genome diploids produce short, not-spinnable fibres and it can be guessed that it is their A-genome diploid ancestor that contributed the spinnable fibres of the cultivated tetraploid cottons.

In B-genome the group, G. anomalum Wawr. & Peyr. has been considered a good model for the African common ancestor of G. arboreum and G. herbaceum (Hutchinson et al. 1947); the seeds of G. anomalum bear fine rather long (a. 10 mm) hairs, which nevertheless are not spinnable as, on the contrary of lint hairs of the cultivated species, they do not flatten and develop the convolutions that permit that lint fibres cling to each other while twisted into a thread.

Even closer to the cultivated diploids than the B-genome is the F-genome (Wendel & Albert 1992; Wendel & Grover 2015), which includes only a single wild species, *G. longicalyx*.



Figure 1. Schematic evolutionary tree of the *Gossypium* genus. Adapted from: Senchina *et al.* (2003), Wendel & Grover (2015), Grover *et al.* (2018).

The realization that this species is the closest living relative of the A-genome diploids was an important discovery, as it sets the stage for understanding how long lint fibres evolved (Hovav *et al.* 2008).

Domestication and diversification of G. arboreum and G. herbaceum

Spread of cotton cultivation in the Antiquity

The map in Figure 2 indicates the earliest archaeological evidence of cotton cultivation in the Old World; the dates are given in years Before Present (BP). The map also features the perimeter of cottongrowing areas at around 2000 years BP, or around the beginning of the Current Era (CE), that can be hypothesized from the archaeological data of earliest cultivation (Sources: Table in Supplementary data S2) and the following considerations:

• The increased trade and cultural exchanges characterizing the Greco-Roman world introduced the use of cotton for clothing from India towards the Mediterranean and Europe in the second half of the first millennium BCE (Wild & Wild 2014a), but maybe cotton cultivation didn't disseminate



Figure 2. Main earliest archaeological evidence of cotton cultivation in the Old World. Dates in years BP. Blue color for *G. arboreum*, orange for *G. herbaceum*, when species identified. Estimated geographic distributions at two millennia BP.

rapidly, as there is no identified archaeological site with cotton cultivation in Ancient Greece or Turkey in the few centuries later.

- Many authors consider that the cotton first cultivated in Tylos (nowadays Bahrain) and in Nineveh or Babylon (Mesopotamia) around 2700-2400 BP must have been *G. arboreum* introduced from India (Muthukumaran 2016; see also Quillien, this volume).
- Cotton cultivation is not known before the first millennium CE in Ethiopia (Aksum), Southern Arabia (the ancient Yemenite civilizations) and the eastern coast of Africa (Jacke 2014, Crowther *et al.* 2016). Cotton cultivated in the first millennium CE along the coasts of Eastern Africa is most likely *G. arboreum*, "arrived from tropical Asia", through Austronesian (Southeast Asia origin) colonization which is estimated to have taken grounds around the seventh to eight century CE (Crowther *et al.* 2016).

Towards Central Asia north of Persia, towards South-eastern Asia and China, towards the Saharan oases and towards Sudan and Ethiopia, the available data permitted to delimit hypothetical geographical areas where *G. arboreum* and *G. herbaceum* were cultivated at the beginning of the Current Era, as shown in Figure 2.

Location and date of the original domestications

The earliest probable cotton cultivation is in Mehrgarh in Balochistan, nowadays Pakistan, in 7^{th} millennium before present (Costantini 1984 cited in Hildebrandt & Gillis 2017). The oldest archaeological cotton remains, nevertheless, are fibres and a thread from the 8^{th} millennium before the present (BP) in this same archaeological site (Moulherat *et al.* 2002). Cotton cultivation appears frequent at the end of the 5-4th millennia BP in archaeological sites of Balochistan, that is, more than two millennia later than the oldest archaeological seed remains in this region. It is assumed by all authors that this cotton is *G. arboreum*.

A growing body of evidence shows the presence of cotton in Central Sudan, Nubia and oases of Egypt and Libya around 2000 BP (Fuller 2014; Bouchaud *et al.* 2018). DNA studies (Palmer *et al.* 2012) of Nubian cotton seeds dated to 1600 years BP identified these seeds as *G. herbaceum*. It can thus be hypothesized that in the Antiquity the cotton cultivated in these regions was the African species *G. herbaceum*. Simultaneously with the local growing, spinning and weaving of an indigenous cotton, trading via the Red Sea or the Nile river could also introduce *G. arboreum* cotton from India (Kriger 2005; Wild & Wild 2014b). The development of cotton cultivation in Meroitic Nubia was presumably derived from an earlier cotton production in northern Sudan (Fuller 2014).

Species assignment of archaeological remains

As for which of the two Old World cotton species was involved in each archaeological site has seldom been determined, as distinguishing between the two Old World cultivated cottons on the basis of seed morphology or lint characteristics is still uncertain, this even more with archaeological remains. Textual accounts from the Antiquity do not distinguish the two species and the expression "cotton tree" frequent in ancient texts cannot be taken as a proof of *G. arboreum* cultivation. Probably the only reliable species assignment in Fig. 2 are the one based on seed DNA in Nubia by Palmer *et al.* (2012) and the one based on detailed fibre studies by Cao *et al.* (2009).

The domestication of *G. arboreum* begun around 8000 years BP, according to all authors nowadays, in the north-western part of the Indian Subcontinent, in Sindh in the Indus Valley or more probably in the Kacchi Plain in central Balochistan (Moulherat *et al.* 2002). The cotton remains until around 2000 years BP in this region most likely all are *G. arboreum*.

The domestication of *G. herbaceum* seems to have occurred much later than that of *G. arboreum*, between around 4500 years BP and the first half of the third millennium BP, if it took place in Nubia and nearby regions as is probable (Palmer et al. 2012). Other hypotheses for the region where *G. herbaceum* was domesticated have been the Sindh (Hutchinson *et al.* 1947) or Southern Arabia (Kulkarni *et al.* 2009).

Varietal diversification and geographical spread

Huge time spans separate the initial cotton remains from nearly 8000 years ago in Mehrgarh, those of the Mature Harappa and Mohenjo-Daro phase from around 5000 years ago and those corresponding to the spread of cotton cultivation in India, Mesopotamia, Persia, Arabia, Egypt, south India, central Asia and south China, beginning around 1000-1500 years later, that is, around 3000 to 2500 years ago. One hypothesis for the thousands of years before the spread of cotton cultivation is the need of genetic evolutions to adapt to cultivation outside of its natural range. Similarly, the spread of cotton cultivation in Central Asia seems to have occurred rather rapidly, which could mean some decisive adaptation had occurred.

The maps in Fig. 3 and 4 show very schematically the hypothetical sequences (Hutchinson & Ghose 1937; Silow 1944; Kulkarni *et al.* 2009) of the emergence of the different varieties of the two Old World cultivated cotton species, with possible timings (years BP) deduced from the considerations above and below. It is to be noted again that these "varieties" are not clear-cut biological entities, but are instead placeholders for geographical/morphological groupings at a certain point in time.

G. arboreum CV. indicum *G*. and herbaceum CV. acerifolium are considered the respective departure points for the presently observed geographic spreads and varietal diversifications of the Old World two cultivated cotton species (Hutchinson



Figure 3. Geographic spread of cultivation of *G. arboreum* and its diverse cultivated varieties. Hypothetical dates in years BP. Sources: see text.

& Ghose 1937; Silow 1944; Hutchinson *et al.* 1947; Kulkarni *et al.* 2009). These perennial varieties were abundantly cultivated until two centuries ago.

If the increase in cotton production in Baluchistan beginning somewhat before 4000 BP were related to the apparition of *G. arboreum* cv. *indicum*, then this permits to date approximatively

the emergence of this variety considered the most primitive inside *G. arboreum* (Kulkarni *et al.* 2009).

Cultivated cottons in Mesopotamia and Persia seem to appear at around 2600 BP (Malatacca 2014; Muthukumaran 2016, Ouillien this volume). In southern Mesopotamia, in Babylon for instance, the climate permits the cultivation of perennial cotton plants. In northern Mesopotamia, in Nineveh for instance, and in Persia, winter temperatures are very cold and perennial cottons would often not survive through to a second cycle. As annual no G. arboreum existed then, the extension into Persia can be supposed to be based on the annual G.



Figure 4. Geographic spread of *G. herbaceum* and its diverse cultivated varieties. Sources: see text.

herbaceum cv. *persicum* which so should have emerged in the first half of the 3^{rd} millennium BP. *G. herbaceum* cv. *persicum* was supposedly derived from the primitive, mostly perennial *G. herbaceum* cv. *acerifolium* which thus should have appeared some millennia earlier.

Finally, in the Antiquity, the cultivated Old World cottons could have been, 1) *G. arboreum* cv. *indicum*, perennial types, in the Indian sub-continent and close surroundings including maybe some sites in the Persian Gulf and Assyria, *G. arboreum* cv. *burmanicum*, perennial types, in the north of the Indochinese Peninsula, east of the Indian subcontinent and south of China, and, 2) *G. herbaceum* var. *acerifolium* as perennial in Arabia and Egypt and *G. herbaceum* cv. *persicum* as annual cultivar in Mesopotamia and Persia. The emergence and dissemination of *G. arboreum* cv. *soudanense* is to be documented.

G. arboreum exists only as a cultigen, as there are nowadays no certain wild plants from this species and the geographic origin of domestication is unknown (Hutchinson *et al.* 1947). The archaeological findings in Baluchistan lead to hypothesize a wild *G. arboreum* population in this region or a nearby regions, created by a chance dispersal of a plant from the common African ancestor of *G. arboreum* and *G. herbaceum*, which disappeared during the domestication process. Diverse *Gossypium* species show evidence of transoceanic, long-

distance dispersals (Wendel & Grover 2015) and in the same time period, 1-2 Mya, when the split between the lineages of *G. arboreum* and *G. herbaceum* occurred, a plant which must have been close to their common ancestor migrated to South or Central America and gave the allotetraploid *Gossypium* species.

A wild population of G. herbaceum exists in southern Africa; this subspecies G. herbaceum L. ssp. africanum (Watt) is considered as "a reasonable model of the ancestor" (Hutchinson 1954) of the domesticated varieties of this species; it appears as having been neither domesticated nor used by local populations. Recent works agree on a domestication of G. herbaceum in northeastern Africa or in Arabia, which is greatly distant, around 5,000 km (Figure 5), of the southern-African wild population of G. herbaceum L. ssp. africanum. Could G. herbaceum be domesticated from a wild population in Northeastern Africa or Arabia? As shown in the map of Figure 5, the species G. anomalum shows a discontinuous geographic distribution area at symmetric latitudes both sides of the Equator. A reasonable hypothesis is that the natural wild population of G. herbaceum var. africanum could have had a similar disjunct geographic



Figure 5. Discontinuous geographic distribution of G. anomalum. Source: Saunders 1961.

distribution and that the population in Northeastern Africa was domesticated in Nubia and/or in a nearby region and later disappeared in the same way as the wild *G. arboreum* population disappeared after its domestication. The hypothesis of a wild ancestor in Northeastern Africa for the domesticated *G. herbaceum* has already been raised by authors such as Nicholson (1960), Moulherat *et al.* (2002), Bouchaud *et al.* (2018). Chowdhury & Buth (1970) concluded on the evidence of a wild ancestor in Nubia of Old World cultivated cottons. Earlier, Watt (1907, p. 157) cites Poiret, Rawlinson (1881) and Joret (1897) as expressing the opinion that *G. herbaceum* originated in Upper Egypt.

Loss of photoperiodism and adaptation to annual cultivation

Two essential evolutions occurred after the initial domestications of G. *arboreum* and G. *herbaceum*: 1) loss of photoperiodism and 2) transformation of shrubby perennial cotton plants to annual plants, and progressively through to compact plants adapted to row cultivation.

Photoperiodism in plants coordinates their reproductive cycle with the seasonal changes in day or night length. Most *Gossypium* cottons are short-day plants, that is, they flower when night becomes longer than some threshold varying according to species. Tropical plants which flower with decreasing day length would, out of the tropical regions, flower when the cold season begins and are thus unadapted to cultivation at high latitudes. The unconscious selection of "day-neutral" varieties permitted to cultivate *Gossypium* cottons north of the subtropical regions. The genetic control of the photoperiodic habit involves the circadian clock and several other integrated physiological pathways encoded by a large number of genes (Grover *et al.* 2015) but is not yet understood in cotton.

In plants of the genus *Gossypium*, the main branches (and the main stem) are monopodial, that is, their terminal meristem produces a continuous growth; they don't bear reproductive organs directly and these branches are called vegetative (Fig. 6). The fruiting branches, bearing the reproductive organs, are distal and grow on the monopodial branches; they have a discontinuous

growth, with nodes where reproductive organs develop; they are called sympodial branches. Cotton plants with developed monopodial branches are tall, have long annual cycles and produce few or no cotton on the first year of cultivation. The morphological evolution of the cotton plants for their cultivation has involved the dimorphic branching pattern characteristic of *Gossypium*. As shown in Figure 6, the annualized modern cottons develop mainly sympodial fruiting branches. The monopodial vegetative branches are iterations of the whole plant: they repeat its general pattern with basal vegetative branches and distal fruiting branches. Two genes have recently been identified (McGarry *et al.* 2016) as regulating the monopodial and sympodial branching architecture in cotton *Gossypium hirsutum*: 'Single Flower Truss' - regulator of floral initiation -, and 'Self-Pruning' – mediator of growth termination.



Figure 6. Schematic architectures comparing monopodial and sympodial branches of (left) a primitive perennial cotton plant, species *G. arboreum*, in India, about 2.5 m high in the original 1832 illustration, and (right) a modern cultivated annual cotton plant, about 1-1.5 m high.

These two evolutionary processes did not necessarily occur simultaneously. Nor did the two species necessarily evolve at the same pace. Hutchinson *et al* (1947) guessed that the spread of *G. herbaceum* to Persia and the harsh winters there resulted in annual habit achieved earlier in *G. herbaceum* than in *G. arboreum*.

Agricultural characteristics of G. arboreum and G. herbaceum

The overall agronomic adaptations of the two cultivated Old World cottons have been described as below:

- Generally, the diploid cultivated cottons show high levels of drought-resistance and are able to resist to the main sucking pests (Kulkarni *et al.* 2009)
- *Gossypium arboreum* varieties are often well adapted to monsoon climate (Valicek 1979)
- *G. arboreum* is generally grown in wetter, warmer environments while *G. herbaceum* var. *acerifolium* confined to more cooler and dryer cultivation (Hutchinson *et al.* 1947)
- *G. arboreum* has agronomic qualities that give it a notable rusticity and adaptability: drought and salinity resistance, suitability for low input conditions, tolerance and resistance to pests and diseases: bollworms, aphids, and leafhoppers, fungal and viral diseases (Kulkarni *et al.* 2009). The resistance to sucking pests originates in the presence of a lower palisade layer, that increases the distance between lower epidermis of midrib and phloem (Sehti *et al.* 2014).
- *G. herbaceum* varieties are often highly tolerant to high temperature and low relative air humidity, and some near-zero (°C) temperatures, while in humid environments it can be infested by bacteriosis and fungal diseases (Valicek 1979).
- Gossypium herbaceum in China is a short-season annual plant (Kuhn 1988).

G. arboreum and *G. herbaceum* have been thus generally described as being adapted to rather contrasted environments, humid and warm for the former, dryer conditions for the latter with tolerance to high as well as rather low temperatures. This seems in accordance with their hypothesized origins: tropical South Asia with monsoon for the first species, dry regions of Africa for the second species.

The varietal diversity in *G. arboreum* is much greater than the six geographic cultivars presented here, as numerous varieties were selected and are locally cultivated in India (Hutchinson *et al.* 1947, Kulkarni *et al.* 2009). Nevertheless, while this species shows a high morphological diversity, its area of dispersion appears mainly restricted to the monsoon regions in Southern Asia and similar tropical-humid regions of Africa, apart from the cv. *sinense*, considered the most divergent inside this species, selected around one thousand years ago in eastern China and cultivated as high as Korea.

In Table 1 are presented data that may characterize the older cultivars of the two species for agronomic adaptation and fibre quality; the indicated geographic distribution corresponds to nowadays area of dispersion.

	Subspecies	Cultivars	Main characteristics, origin and geographic distribution						
Gossypium	No	cv. indicum	"Most primitive race" of G. arboreum; very variable genetic						
arboreum	recognized		material; perennial or annual, monopodial or sympodial forms,						
L.	subspecies		broad-leaved; lint short in wild forms, medium long and						
			moderately fine in cultivated strains; the perennial, sympodial						
			forms might be the oldest cultivated types.						
			Numerous traditional local cultivars in the Indian subcontinent.						
			"Rozi" cotton of India probably the oldest form; Gaorani, Mathio						
			could be rather old forms.						
			Western India, Sri Lanka, coastal Tanzania, Madagascar.						
		cv. soudanense	Rather primitive; perennial, predominantly monopodial, broad-						
			leaved; lint short, usually white; low-ginning forms; nowhere						
			cultivated as a field crop, although grown for home use.						
			Egypt, Sudan, North, West, central and Southern Africa.						
		cv. burmanicum	Highly variable over its range; perennials and annuals; a very						
			heterogeneous germplasm with very variable lint characters, from						
			short and coarse to the longest and finest in this species.						
			North-eastern India, South-eastern Asia.						
Gossypium	ssp.	cv. acerifolium	Could correspond to the primitive cultivated forms of G .						
herbaceum	herbaceum		herbaceum; perennial; large shrubs with many vegetative						
L.			branches; stout stem; rather small leaves; some primitive						
			agronomic characters: indehiscent capsule and bud-shedding in						
			rainy season; small bolls and seeds; scanty coarse lint.						
			Distribution extremely wide in Asia, Arabia and Africa.						
		cv. persicum	Annual, small, early-maturing shrub; few or no vegetative						
			brancnes; stiff stem; large leaves almost fleshy; big, round bolls;						
			large seeds; copious fair quality or medium-quality lint.						
			Guessed to correspond to the type introduced to the						
			Nediterranean in the Alexander the Great era, c. 2350 BP.						
			Recently still cultivated in Iran in the form of the Boumi type.						
			Egypt, Mediterranean basin, southern Central Asia.						

Table 1. The cotton varieties possibly cultivated in the Old World in the Antiquity. Characteristics, origin and geographic distribution. Main sources: Silow (1944), Hutchinson *et al.* (1947), Raingeard (1966), Kulkarni *et al.* (2009), Menon & Uzramma (2017).

Concerning the general adaptation of *G. arboreum* and *G. herbaceum* to climate conditions and their ability to withstand cultivation outside of their primary centres of origin, it is necessary to turn to the climatic conditions prevailing in the latter (climate data from Wikipedia):

- Sindh is mostly arid with mild to warm temperatures in winter (falling to 2°C in plain, to 0°C on the coast line) and inundation of the Indus twice a year in the plain. Vegetation is scant except for the irrigated Indus valley. The coastline benefits in July–August from moderate monsoon rainfalls.
- The Kacchi plain of Baluchistan has cold to relatively mild winters, temperature never falling below freezing point, and its summers are hot and dry.
- Nubia in present Southern Egypt and Northern Sudan has nowadays a hot desert climate; it was somewhat less dry until 5000 years ago.
- In southern Africa, the wild subspecies *G. herbaceum* ssp. *africanum* is distributed over regions with climates from hot arid to hot semi-arid.

The two species thus appear to be naturally adapted to hot desert climates (BWh in Köppen classification) and maybe hot semi-arid climate (BSh), where summers are hot and dry or

moderately dry and temperatures in winter won't go below freezing point but nevertheless commonly fall below 5°C.

In the regions were cotton cultivation extended from India in the Antiquity, climates are:

- Bahrain: an arid climate with an extremely hot summer and a relatively mild winter, with very limited rainfall confined to winter months.
- Mesopotamia: the regions of Babylonia and Nineveh nowadays have a hot desert climate. Rainfall during the summer is extremely rare and irrigation of agriculture is essential. Winter temperatures are mild with night-time lows falling to 2 to 5 °C in Babylonia and much lower, frequently well below 0°C, in Nineveh.

Thus, climate conditions in Bahrain and southern Mesopotamia don't appear as markedly different from those in Sindh. *G. arboreum* could have been grown in southern Mesopotamia with irrigation water from the Euphrates or the Tigris in the same way as in Sindh irrigation water was taken from the Indus. In northern Mesopotamia, Nineveh for instance, winter temperatures fall much lower and frequently below 0°C, probably selecting for annualized cottons.

It was a hypothesis of Hutchinson *et al.* (1947, p. 88) that in the absence of irrigation facilities, Sindh and Rajputana (Northeast of Sindh) would be unhospitable to *G. arboreum* or *G. herbaceum* and constitute for the two species an impassable barrier between Persia and Peninsular India. Apart from being in contradiction with his hypothesis of a domestication of the Indian cotton in Sindh, it leads to question where precisely were growing the natural populations of wild *G. arboreum* plants in order to survive the summer droughts.

Morpho-agronomic and fibre quality differentiation between *G. arboreum* and *G. herbaceum*

Classical plant descriptors: leaf and bract morphology

The species identification in the field is classically based on plant morphological traits, including plant architecture, but the distinction between *G. arboreum* and *G. herbaceum* is considered difficult (Hutchinson *et al.* 1947). In Figure 7 are shown schematized drawings of leaves and bracts of the two species. Leaf length of *G. herbaceum* ssp. *africanum* is around 6 cm. Variability is higher in *G. arboreum* for leaf morphology as shown in the figure, but also for corolla colour which varies from yellow to red. Some traits characterize one or the other species - palmate leaves with fine lance-shaped lobes are characteristic of *G. arboreum*, finely dented bracts are characteristic of *G. herbaceum*, for instance - but generally diverse traits have to be considered jointly for a reliable specific assignment, according to Kulkarni *et al.* (2009). Even so, there are cases where it may be difficult to assign individuals to species on the basis of the sole plant morphology.



Figure 7. Schematized leaves and bracts of (left) *Gossypium arboreum* and (right) *G. herbaceum*. Drawings based on plants in greenhouse in CIRAD (Lavalette campus) and descriptions in Watt (1907), Hutchinson *et al.* (1947), Saunders (1954) and Fryxell (1979).

Quantitative plant and fibre descriptors of wild/feral gene bank accessions

Using the usual descriptors of the *Gossypium* cotton plants, *G. arboreum* and *G. herbaceum* were compared in order to search for traits efficient at differentiating the two species. The data correspond to plants from CIRAD's gene bank.

The map of Figure 8 shows the geographic origins of genotypes of wild and feral G. arboreum and G. herbaceum in our gene bank in CIRAD Montpellier. The wild plants found during the prospections can be in fact only feral plants, that is, descendants of "escapes", or cultivated genotypes returned to wild, and not truly wild plants. It was hypothesized that these feral plants could be somewhat representative of the traditional cultivated perennial varieties cultivated in the Antiquity, although this too is not clear. In the map, accessions were grouped by country and group size is indicated by a number when there was more than one plant. Globally, means could be calculated over 10 to 30



Figure 8. Geographic location of wild/feral *G. arboreum* and *G. herbaceum* accessions from CIRAD's gene bank. Abbreviations: wa= wild/feral *G. arboreum*, wh= wild/feral *G. herbaceum*; sample size indicated whenever more than one plant was sampled in the country.

genotypes for *G. arboreum* and 9 to 13 for *G. herbaceum*. The geographic areas of the two studied species were unequally covered; Central Asia in particular counts no accessions.

The dispersion areas figured for the two species are estimations of the present global areas (sources: Hutchinson *et al.* 1947, Valicek 1979, Kulkarni *et al.* 2009, geographic data from accessions of CIRAD's gene bank), mainly resulting from human dispersion.

These cotton plants have been studied when they were grown for seed renewal in Africa or South America, measurements and observations being standardized between years and sites with a reference cultivar. The graphs in Fig. 9 and 10 compare the means of *G. arboreum* and *G. herbaceum* genotypes over 28 agro-morphological traits and ten lint quantity and quality traits (see Supplementary data S3 for the complete data). The values are given as the percent of the overall mean over both species, in order that all these descriptors with very diverse measurement units and scale ranges be side by side in the same graphs. The charts permit comparisons between the two species for plant morphology and phenology in Fig. 9 and for fibre quality in Fig. 10. Explanations of the abbreviations for trait names are given in the legends.

Agronomic descriptors

In the graph for agronomic traits (Fig. 9), some notable differences appeared between *G. arboreum* and *G. herbaceum*: height of 1st sympodial branch (H1FruBr), Plant shape (PlaShap), Leaf shape (LfShap), numbers of bract dents and a few other traits. There is however an overlap of data distributions, except for height of 1st sympodial branch, a trait probably hardly satisfactory for species characterization.



Figure 9. Comparison of wild/feral *G. arboreum* and *G. herbaceum* over 28 agro-morphological traits. For each trait, the ranges compared are calculated as one standard deviation above and below mean and expressed in percent of the average mean over all plants. Traits: Da1Flo: time to 1st flower (days after emergence); Da1Bo: time to 1st boll; PlaShap: Plant height to width ratio; PlaHei: Plant height at harvest; H1FruBr: Height 1st sympodial branch; VegBrN: Vegetative branches number; StmHair: Main stem hairiness; LfShap: Leaf sinus depth; LfPlan: Leaf planarity; LfLobN: Leaf lobes number; LfHair: Leaf lower side hairiness; LfSiz: Leaf size; NecN: Leaf nectaries number; LfGla: Leaf gossypol glands density; BracSiz: Bracts size; BracFor: Bracts dents depth; BracDtN: Bracts dent number; Macula: Flower bottom macula size; BoSiz: Boll size; BoMuc: Boll tip size; BoLocN: Boll lobules number; BoGla: Boll gossypol glands density; BoAper: Boll aperture; BoStorm: Boll stormproof ; BoWei: Boll weight; SdWei: Seed weight (100 seed); SdFuz: Seed fuzz density; SdHair: Seed fuzz density.

As already noted by Hutchinson *et al.* (1947) and Kulkarni *et al.* (2009), when considering that for many traits the ranges of variation overlap between the two species, then no single morphological trait is sufficient to distinguish very firmly the species *G. arboreum* and *G. herbaceum*. Nevertheless, the same authors add, reliable specific determinations are possible when many morphological traits are considered simultaneously.

Lint descriptors

Concerning fibre quality comparisons, the data indicated an inferiority of G. herbaceum for lint strength (LinStr), maturity (LinMat) and fineness (LinFin), while lint length (LinLen), length uniformity (LinLenU), elasticity (LintElo) and colorimetry (reflectance, LinRef and yellowness, LinYel) looked identical in the two species. Ginning outturn, which measures the percent of lint fibres obtained from the harvested seed-cotton, appeared rather identical in the two species.

As shown by the present data, coarse fibres above of 300 mtex in ancient textiles are indicative of a G. herbaceum origin, while fine fibres below 250 mtex correspond more to G. arboreum. The fibres with highest fineness measures were for both species those of the African wild/feral genotypes. Much overlapping of fibre fineness distributions is observed between the two species, as shown by the

histogram of Fig. 11, so that fibre fineness could most often only contribute to a species assignation, in a similar manner as above when dealing with morphological traits. The fibre fineness used here (called Hs, standard fineness) is a measure of the mass per unit length (in mg/m), which can be rather unpractical for archaeological textiles for diverse reasons; Hs is well correlated with the fibre diameter (in μ m) which is currently used for archaeological cottons.



Figure 10. Graphics comparing wild/feral *G. arboreum* and *G. herbaceum* over ten lint quantity and quality traits. For each trait, the ranges compared are calculated as one standard deviation above and below mean and expressed in percent of the average mean over all plants. Traits: RGOT: Roller-gin ginning outturn, i.e. % fibre on roller gin; LinLen: Lint length: HVI UHML; LinLenU: Lint length uniformity; LinStr: Lint strength HVI; LinElo: Lint elongation; LinMat: Lint maturity; LinFin: Lint fineness; LinRef: Lint reflectance; LinYel: Lint yellowness.



Figure 11. Fineness (Hs in mtex) distribution in wild/feral *G. arboreum* and *G. herbaceum* genotypes from CIRAD's gene bank.

The similarity of the two species for lint length and colorimetry probably reflects simultaneously the importance of these traits for the textile craftsmanship, the relative ease with which they can be assessed visually and selected for in traditional breeding to adjust to the standards of craftsmanship and commerce, and the physiologic limits of these diploid genomes.

The numbers of gene bank accessions used for this comparison of the two species are rather modest and only a part of their geographic areas was covered. Nevertheless, the data in these graphs are coherent with previous knowledge about the morphological differences between the two species, in particular from the preceding study by Stanton *et al.* (1994). These data indicate that finer fibres are expected more from *G. arboreum* than from *G. herbaceum*.

New methodologies for cotton lint characterization

Cao *et al.* (2009) experimented new methodologies for the characterization of archaeological lint fibres between diverse *Gossypium* species. They realized an in-depth study of cotton fibres dated to A.D. 1161–1255, unearthed from Yingpan in Xinjiang, western China, in comparison with modern samples of *G. arboreum*, *G. herbaceum* and *G. hirsutum* fibres. In addition to the common morphological parameters fibre length, fineness, strength and elongation, they assessed with a video microscope the width of the fibre ribbon and its convolutions (number of 180° reverses per centimeter in length) and with scanning electron microscope images the shapes of the fibres; they also measured the degree of polymerization of the cellulose and the fibre density according to ASTM International standard test methods, the IR spectra with a spectrometer and the wide-angle X-ray diffraction with a diffractometer for cellulose crystallinity parameters.

Cao *et al.* (2009) concluded that the lint fibres of the two cultivated diploids appeared very similar for all the parameters, slight differences being seen only for the crystallinity orientation angle and crystallite %, with the archaeological fibres from Yingpan appearing similar to the modern *G. herbaceum* fibres for the two latter parameters.

There were clear differences in Cao *et al.* (2009) between the lint fibres of the Old-World cotton species and those of the American species *G. hirsutum* for the fibre ribbon width, convolutions numbers and crystallinity orientation parameters, apart from the classical differences for lint length, fineness and strength; the *G. hirsutum* fibres appeared longer, finer and more convoluted. The authors consider that the spiral angle φ obtained by X-ray diffraction is a particularly effective way for distinguishing between lint fibres of the diploid *Gossypium* species from those of the tetraploids.

The archaeological fibres studied by Cao *et al.* (2009) showed no difference for crystallinity parameters relative to modern samples, except for a larger crystallite size along the longitudinal axis, this to be confirmed nevertheless. The density remained similar to modern samples. The fibre strength and elongation were nevertheless neatly inferior, as well as the degree of polymerization of the cellulose, all of which being interpreted as a natural deterioration of the supermolecular structure over time and in relation to soil contact.

Thus, the in-depth study of lint fibres by Cao *et al.* (2009) showed moderate differences for the crystallinity orientation angle and crystallite % between the two cultivated diploid *Gossypium* species.

Cytogenetic and genomic differences

Although natural hybridizations between *Gossypium arboreum* and *G. herbaceum* do exist and the resulting interspecific hybrids show fertility, germination deficiencies appear in the progeny and a rapid return to one or the other parental type occurs in the following generations (Silow 1944). The two species could thus coexist and even be cultivated side by side without losing their respective characteristics, even if introgressions occurred. This phenomenon was explained when a reciprocal translocation differentiating their genomes was discovered by Gerstel (1953). This translocation occurred in *G. arboreum*, the ancestral chromosomes being those of *G. herbaceum* (Desai *et al.* 2006). Notwithstanding this apparent barrier to interspecific gene flow, genetic surveys show unequivocal evidence that it occurs (Wendel *et al.* 1989).

The last shared common ancestor of *G. arboreum* and *G. herbaceum* was recently estimated at 0.4–2.5 million years ago through rates of synonymous substitution on shared genes (Renny-Byfield *et al.* 2016). This speciation event is sufficiently old to exclude that the species *G. arboreum* diverged from *G. herbaceum* a few thousand years ago, as had been hypothesized by some authors (see for example Hutchinson *et al.* 1947).

Simple molecular genetics methodologies allow species assignment among the *Gossypium* genus (Wendel *et al.* 1989; Hinze *et al.* 2015; Renny-Byfield *et al.* 2016). The extent of genetic differentiation between *G. arboreum* and *G. herbaceum* permits unambiguous species assignations by extraction of DNA from archaeological remains such as well-preserved seeds (Palmer *et al.* 2012). Each cotton fibre is a unicellular hair that elongates from the seed epidermis and it contains in its central lumen the nuclear and cytoplasmic DNAs, which can be extracted (Ibrahim *et al.* 2019), although the quantities are small. DNA extracts from raw fibres, yarns or fabrics could permit species assignations, if no excessive degradation by mineralization, chemical treatments or humidity occurred.

Acknowledgments

The author expresses his deepest gratitude to Charlene Bouchaud for inviting him to contribute to this special edition on Old World cottons and to Jonathan F. Wendel, Charlène Bouchaud and Elsa Ybanez for useful advice and suggestions.

References

Bouchaud C., Tengberg M. & Dal Prà P. 2011. Cotton Cultivation and Textile Production in the Arabian Peninsula During Antiquity; the Evidence from Madâ'in Sâlih (Saudi Arabia) and Qal'at Al-Bahrain (Bahrain). *Vegetation History and Archaeobotany* 20.5 (2011): 405–17.

Bouchaud C., Clapham A., Newton C., Thanheiser U. & Tallet G. 2018. Cottoning on to cotton (Gossypium spp.) in Arabia and Africa during Antiquity. *In*: Mercuri AM, D'Andrea AC, Fornaciari R, Höhn A (eds). *Plants and People in the African Past: Progress in African Archaeobotany*. Springer International Publishing, Cham, pp 380–426.

Bowen G. 2010. Texts and Textiles: A Study of the Textile Industry at Ancient Kellis. http://arts.monash.edu.au/archaeology/excavations/dakhleh/ismant-el-kharab/ismant-texts-textiles.php). In Brite & Marston 2013.

Brite E.B. 2011. The Archaeology of the Aral Sea Crisis: Environmental Change and Human Adaptation in the Khorezm Region of Uzbekistan ca. AD 300–800. Unpublished Ph.D. Dissertation, Department of Anthropology, University of California, Los Angeles.

Brite E.B. & Marston J.M., 2013. Environmental Change, Agricultural Innovation, and the Spread of Cotton Agriculture in the Old World. *Journal of Anthropological Archaeology* 32.1 (2013): 39–53.

Cao Q., Zhu S., Pan N., Zhu Y. & Tu H. 2009. Characterization of Archaeological Cotton (*G. herbaceum*) Fibres from Yingpan. *Technical Briefs in historical archaeology*, 2009, 4: 18–28.

Chao K. 1977. The Development of Cotton Textile Production in China (Cambridge, MA: Harvard University Press). Cited in Zurndorfer Harriet T., 2011. Cotton Textile Manufacture and Marketing in Late Imperial China and the 'Great Divergence'. *Journal of the Economic and Social History of the Orient*, Volume 54: Issue 5. DOI: https://doi.org/10.1163/156852011X614028.

Chowdhury K.A. & Buth G.M. 1970. 4500 year old seeds suggest that true cotton is indigenous to Nubia. *Nature* 227, July 4, 1970: 85-6.

Chowdhury K.A. & Buth G.M. 1971. Cotton seeds from the Neolithic in Egyptian Nubia and the origin of Old World cotton. *Biological Journal of the Linnean Society*, Volume 3, Issue 4, 1 December 1971, Pages 303–312, <u>https://doi.org/10.1111/j.1095-8312.1971.tb00524.x</u>

Chowdhury K.A. & Buth G.M. 2005. Plant Remains from Excavation of Terraces of the Nile at Afyeh, Nubia and Egypt. *Purattatva* 35, 154-9.

K.A. Chowdhury and G.M. Buth

Costantini L. 1984. The beginning of agriculture in the Kachi Plain: the evidence of Mehrgarh. In Bridget Allchin (Ed), *South Asian archaeology 1981: proceedings of the Sixth International Conference of the Association of South Asian Archaeologists in Western Europe*. Cambridge University, 5-10 July 1981.

Crowther A., Lucas L., Helm R., Horton M., Shipton C., Wright H.T., Walshaw S., Pawlowicz M., Radimilahy C., Douka K., Picornell-Gelabert L., Fuller D.Q. & Boivin N.L. 2016. Ancient crops provide first archaeological signature of the westward Austronesian expansion. Proceedings of the National Academy of Sciences, 113 (24) 6635-6640. DOI: 10.1073/pnas.1522714113.

Desai A., Chee P.W., Rong J., May O.L. & Paterson A.H. 2006. Chromosome structural changes in diploid and tetraploid A genomes of Gossypium. *Genome* 49: 336–345 (2006).

Fryxell P.A. 1976. A nomenclator of Gossypium: the botanical names of cotton. Techn. Bull. N° 1491, United States Agricultural Research Service, Texas A & M University, Texas. https://books.google.fr/

Fryxell P.A. 1979. The natural history of the cotton tribe. College Station : Texas A&m Univ. Press. 245 p.

Fuller D.Q. 2008. The spread of textile production and textile crops in India beyond the Harappan zone: an aspect of the emergence of craft specialization and systematic trade. In: Osada T, Uesugi A (Eds), 2008. *Linguistics and archaeology and the human past, occasional paper 3*:1–26. Research Institute for Humanity and Nature, Kyoto, Japan.

Fuller D.Q. 2014. Agriculture innovation and state collapse in Meroitic Nubia. In: Stevens C.J., Nixon S., Murray M.A., Fuller D.Q. (Eds). *Archaeology of African Plant Use*. Left Coast Press, Walnut Creek, pp 165–177.

Gerstel D.U. 1953. Chromosomal translocations in interspecific hybrids of the genus Gossypium. *Evolution* 7: 234-244.

Grover C.E., Arick II M.A, Thrash A., Conover J.L., Sanders W.S., Peterson D.G., Frelichowski J.E., Scheffler J.A., Scheffler B.E. & Wendel J.F. 2018. Insights into the Evolution of the New World Diploid Cottons (Gossypium, Subgenus Houzingenia) Based on Genome Sequencing. Genome Biology and Evolution, 11 (1): 53–71. DOI: 10.1093/gbe/evy256.

Grover C.E., Gallagher J.P. & Wendel J.F. 2015. Candidate gene identification of flowering time genes in cotton. *The Plant Genome* 8.2.

Gulati N. & Turner J. 1929. A note on the early history of cotton. Journal of the Textile Institute Transactions, 20:1, T1-T9, DOI: 10.1080/19447022908661470.

Herrmann G., Masson V.M. & Kurbasakhatov K. 1993. The International Merv Project, preliminary report on the first season (1992). *Iran* 31, 39–62.

Hildebrandt B. & Gillis C., 2017. *Silk: Trade & Exchange along the Silk Roads between Rome and China in Antiquity*. Oxbow Books, Oxford and Philadelphia.

Hinze L.L., Fang D.D., Gore M.A., Scheffler B.E., Yu J.Z., Frelichowski J. & Percy R.G. 2015. Molecular characterization of the Gossypium Diversity Reference Set of the US National Cotton Germplasm Collection. *Theor Appl Genet*. 128:313–327.

Hovav R., Udall J.A., Chaudhary B., Hovav E., Flagel L., Hu G. & Wendel J.F. 2008. The Evolution of Spinnable Cotton Fibre Entailed Prolonged Development and a Novel Metabolism. *PLoS Genet* 4(2): e25. https://doi.org/10.1371/journal.pgen.0040025

Hutchinson J. B. 1954. New evidence on the origin of the Old World cottons. *Heredity* 8:225–241.

Hutchinson J.B. & Ghose R. 1937. The classification of the cottons of Asia and Africa. *Indian J Agric Sci.* 7: 233–257.

Hutchinson J.B., Silow R.A. & Stephens S.G. 1947. *The Evolution of* Gossypium *and the Differentiation of The Cultivated Cottons*. Oxford University Press, London, 160p. (1947).

Ibrahim S.D., Alsamman A.M. & Khalifa K. 2019. A rapid method for DNA extraction of cotton mature fibre suitable for PCR fingerprinting. *bioRxiv preprint*, Jan. 23, 2019; doi: <u>http://dx.doi.org/10.1101/529305</u>.

Jacke P. 2014. The Foreign Contacts of Ancient Aksum: New finds and some random thoughts pp. 253-268. In Lohwasser Angelika, Wolf Pawel, 2014. *Der antike Sudan. Mitteilungen der Sudanarchäologischen Gesellschaft zu Berlin e.V. Sonderheft In Ein Forscherleben zwischen den Welten Zum 80. Geburtstag von Steffen Wenig.* Sudanarchäologischen Gesellschaft zu Berlin e.V. Sonderheft 2014. https://www.sag-online.de/wp-

 $content/uploads/2018/06/Phillips 2014_ForeignContactsOfAncientAksum_FSWenig.pdf$

Kriger C.E. 2005. Mapping the History of Cotton Textile Production in Precolonial West Africa. *African Economic History* 33: 87-116.

Kuhn D. 1988. Part 9, Textile Technology: Spinning and Reeling. In Joseph Needham, *Science and Civilisation in China: Volume 5, Chemistry and Chemical Technology*. Cambridge University Press; First Edition edition (July 29, 1988).

Kulkarni V.N., Khadi B.M., Maralappanavar M.S., Deshapande L.A. & Narayanan S.S. 2009. *The worlwide gene pools of Gossypium arboreum and G herbaceum L., and their improvement. In* Paterson A. (Ed.). *Genetics and genomics of cotton.* Pp . 69-97. Springer (2009)

Linnaeus C. 1753. Species Plantarum 2:693.

McGarry R.C., Prewitt S.F, Culpepper S., Eshed Y., Lifschitz E. & Ayre B.G. 2016. Monopodial and sympodial branching architecture in cotton is differentially regulated by the Gossypium hirsutum SINGLE FLOWER TRUSS and SELF-PRUNING orthologs. *The New phytologist* 212 (1): 244-58.

Malatacca L. 2016. Movements of Fibres, Dyes and Textiles in First Millennium BC Babylonia (Neo- and Late-Babylonian Periods). In Foietta, E. et al (eds.), 2016. Cultural and Material Contacts in the Ancient Near East. Proceedings of the International Workshop, 1-2 December 2014, Torino, 2016: 91-97.

Menon M. & Uzramma. 2017. *A Frayed History: The Journey of Cotton In India, a monograph*. vol. 1. New Delhi: Oxford University Press.

Moulherat C., Tengberg M., Haquet J. F. & Mille B. 2002, First Evidence of Cotton at Neolithic Mehrgarh, Pakistan: Analysis of Mineralized Fibres from a Copper Bead. December 2002, *Journal of Archaeological Science* 29(12):1393-1401. DOI: 10.1006/jasc.2001.0779

Muthukumaran S. 2016. Tree Cotton (*G. arboreum*) in Babylonia. *In* Foietta, E. et al (eds.), 2016. *Cultural and Material Contacts in the Ancient Near East. Proceedings of the International Workshop*. 1-2 December 2014, Torino, Italy: 98-105.

Nicholson, G.E. 1960. The production, history, uses and relationships of cotton (*Gossypium* spp.) in Ethiopia. Econ Bot (1960) 14: 3. https://doi.org/10.1007/BF02859364.

Palmer S.A., Clapham A.J., Rose P., Freitas F.O., Owen B.D., Beresford-Jones D., Moore J.D., Kitchen J.L. & Allaby R.G. 2012. Archaeogenomic evidence of punctuated genome evolution in Gossypium. *Mol. Biol. Evol.* 29(8):2031-8. doi: 10.1093/molbev/mss070.

Parlatore F. 1866. Le specie dei cotoni. Stamperia Reale, Firenze

Pelling R.E. 2007. Agriculture and trade amongst the Garamantes and the Fezzanese: 3000 years of archaeobotanical data from the Sahara and its margins. Doctoral thesis, UCL (University College London).

Raingeard J. 1966. Le coton en Iran. Cot. Fib. Trop. 15: 367-375.

Renny-Byfield S., Page J T., Udall J.A., Sanders W.S., Peterson D.G., Arick M.A., Grover C.E. & Wendel J.F. 2016. Independent domestication of two Old World cotton species. *Genome Biology and Evolution* 8, 1940-47 (2016).

Saunders J.H. 1961. *The wild species of Gossypium and their evolutionary history*. Oxford University Press, New York.

Sethi K., Priyanka Siwach P. & Kumar Verma S. 2014. Genetic improvement of *Gossypium arboreum* L. using molecular markers: Status and development needs. *African journal of agricultural research* 9(29):2238-2249. DOI: 10.5897/AJAR2014.8682

Senchina D.S., Alvarez I., Cronn R.C., Liu B., Rong J., Noyes R.D., Paterson A.H., Wing R.A., Wilkins T.A. & Wendel J.F. 2003. Rate variation among nuclear genes and the age of polyploidy in *Gossypium*. *Mol Biol Evol.* 20 : 633-643.

Silow R.A. 1944. The genetics of species development in the Old World cottons. J. Genet. 46, 62.

Stanton M.A., Stewart J. McD., Percival A.E. & Wendel J.F. 1994. Morphological Diversity and Relationships in the A-Genome Cottons, *Gossypium arboreum* and *G. herbaceum*. *Botany Publication and Papers*. Paper 7.

Valicek P. 1979. Wild and cultivated cottons. Cot. Fib. Trop. 34, 2:239-264 (1979).

Watt G. 1907. The wild and cultivated cotton plants of the world. Longmanns, London.

Wendel J.F. & Albert V.A. 1992. Phylogenetics of the Cotton Genus (Gossypium): Character-State Weighted Parsimony Analysis of Chloroplast-DNA Restriction Site Data and Its Systematic and Biogeographic Implications. *Systematic Botany* 17 (1): 115-143. https://www.jstor.org/stable/2419069

Wendel J.F., Brubaker C.L. & Seelanan T. 2010. *The origin and evolution of Gossypium*. Chap. 1, pp 1-18. In Stewart J.McD., Heitholt J.J., Mauney J.R., Oosterhuis D. (Eds.), *Physiology of Cotton*. Springer Science+Business Media B.V.

Wendel J.F. & Grover C.E. 2015. *Taxonomy and Evolution of the Cotton Genus, Gossypium*. In: *Cotton*, 2nd ed., Agron. Monogr. 57. ASA, CSSA, and SSSA, Madison, WI. p. 25-44.

Wendel J.F., Olson P.D. & Stewart McD.J. 1989. Genetic Diversity, Introgression, and Independent Domestication of Old World Cultivated Cottons. *American Journal of Botany*, Vol. 76 (12): 1795-1806.

Wikipedia, https://en.wikipedia.org: articles "Sindh", "Hyderabad", "Karachi", "Bahrain", "Iraq", "Baghdad", "Nubia", consulted on June 15, 2019.

Wild J.P. & Wild F.C. 2014a. Berenike and textile trade on the Indian ocean. In: Droß-Krüpe K. (Ed.) *Textile trade and distribution in Antiquity*. Harrassowitz, Wiesbaden, pp 91–110.

Wild J.P. & Wild F.C. 2014b. Qasr Ibrim: New perspectives on the changing textile cultures of Lower Nubia. In O'Connell E.R. 2014. Egypt in the first millennium AD, Perspectives from new fieldwork. British Museum publications on egypt and Sudan 2. Peeters, Leuven – Paris – Walpole, MA: 71-80.

Zaitzev G.S. 1928. A contribution to the classification of the genus Gossypium L. *Bulletin Appl. Bot. Genet. Plant Breeding* 18: 39-65.

Zurndorfer H.T. 2011. Cotton Textile Manufacture and Marketing in Late Imperial China and the 'Great Divergence'. *Journal of the Economic and Social History of the Orient* 54 (2011) 701-738 brill.nl/jesh.

Christopher Viot, Domestication and varietal diversification of Old World cultivated cottons (*Gossypium* sp.) in the Antiquity *»*, *Revue d'ethnoécologie*, 15 | 2019, Online 30 June 2019.

Supplementary data

Supplementary data 1

Major synonyms of Gossypium herbaceum L. and Gossypium arboreum L.

1 - Source: Fryxell (1979), Fryxell (1976). Synonyms and subspecies for *G. arboreum* and *G. herbaceum* ordered by publication date.

Gossypium herbaceum L., Sp. Pl. 693. 1753.

Species synonyms (ordered by publication date): Gossypium frutescens Lasteyrie, Du Cottonier 435, t. 1. 1808. Gossypium album Buchanan-Hamilton, Trans. Linn. Soc. 13:494. 1822. Gossypium africanum (Watt) Watt, Kew Bull. 5:205. 1926. Gossypium transvaalense Watt, op. cit., 207. Gossypium abyssinicum. Watt, op. cit., 208. Gossypium zaitzevii Prokhanov, Bot. Zhurn. 32:70. 1947. "Described from Chinese Turkestan (Sinkiang province)."

Gossypium arboreum L., Sp. Pl. 693. 1753.

Species synonyms (ordered by publication date): Gossypium. rubrum Forskål, Fl. Aegypt.-Arab. 125. 1775. Gossypium indicum Medikus, Bot. Beobacht. Jahr. 1783 197. 1784. Gossypium rufum Scopoli, Delioc. Insub. 3:70. 1788. Gossypium nigrum Buchanan-Hamilton, Trans. Linn. Soc. 13:494. 1822. Gossypium croceum Buchanan-Hamilton, loc. cit. Gossypium obtusifolium. Roxburgh ex G. Don, Gen. Hist. 1:487. 1831. Gossypium nanking Meyen, Verh. Ver. Beförd. Gartenb. Königl. Preuss. Staat. 11:258, t. 3.1835. Gossypium speciosum Rafinesque, Sylva Tellur. 18. 1838. Gossypium puniceum Fenzl ex Jacquin, Ecolog. Pl. Rar. 2:7, t. 134. 1844. Gossypium sanguineum Hasskarl, Cat. Hort. Bogor. 200. 1844. Gossypium albiflorum Todaro, Giorn. R. Ist. Incoragg. Agric. Arti Manifatture Sicil., ser. 3, 1(2, 3):42. 1863. Gossypium cernuum Todaro, op. cit., 47. Gossypium neglectum Todaro, op. cit., 51. Gossypium royleanum Todaro, op. cit., 57. Gossypium intermedium Todaro, op cit., 58. Gossypium roxburghii Todaro, op. cit., 61. Gossypium wightianum Todaro, op. cit., 63.

Domestication and varietal diversification of Old World cultivated cottons (Gossypium sp.) in the Antiquity. *Revue d'ethnoécologie*, 15 | 2019.

Hibiscus cernuus (Todaro) Kuntze, Rev. Gen. Pl. 1:68. 1891.
Hibiscus nanking (Meyen) Kuntze, loc. cit.
<u>Gossypium soudanense</u> (Watt) Watt, Kew Bull. 5:201. 1926.
Gossypium bani (Watt) Prokhanov, Bot. Zhurn. 32:69. 1947.
<u>Gossypium perrieri</u> (Hochreutiner) Prokhanov, op. cit., 71.
Gossypium wattianum Hu, Fl. China, fam. 153 (Malv.), 65. 1955.

2 - Source: Wikispecies, U.S. National Plant Germplasm System (<u>https://npgsweb.ars-grin.gov/</u>)

Gossypium herbaceum L. (1753) Species synonyms (alphabetical order): Gossypium abyssinicum Watt. (1926) Gossypium africanum (Watt) Watt (1926) Gossypium album Buchana-Hamilton (1822) Gossypium frutescens Lasteyrie (1808) Gossypium herbaceum var. africanum Gossypium herbaceum var. herbaceum Gossypium prostratum A. Schumach. & Thonn. Gossypium punctatum A. Rich., Guill. & H. Perrier Gossypium transvaalense Watt. (1926) Gossypium zaitsevii Prokhanov (1947)

Gossypium arboreum L. (1753) Species synonyms (alphabetical order): Gossypium anomalum Watt Gossypium arboreum f. indicum J. B. Hutch. & Ghose, nom. inval. Gossypium arboreum f. soudanense (G. Watt) J. B. Hutch. & Ghose Gossypium arboreum subsp. cernuum (Tod.) Roberty Gossypium arboreum var. cernuum (Tod.) J. B. Hutch. & Ghose Gossypium arboreum var. obtusifolium (Roxb. ex G. Don) Roberty Gossypium cernuum Tod. Gossypium cernuum Tod. var. sylhetensis Gammie Gossypium herbaceum L. var. perrieri Hochreutiner Gossypium indicum Medik Gossypium intermedium Tod. Gossypium nanking Meyen Gossypium nanking var. soudanensis G. Watt Gossypium neglectum Tod. Gossypium obtusifolium Roxb. ex G. Don Gossypium sanguineum Hassk. Gossypium soudanense (G. Watt) G. Watt

Supplementary data 2

Table S2. Data sources for Fig. 2: Main earliest archaeological evidence of cotton cultivation in the Old World in the Antiquity. Dates in years before the present (BP).

Location	Period (BP)	LongLat.	Species	Precisions	Source
Mergarh, Kacchi Plain, Baluchistan	6400	29N-68E		Seeds	Costantini (1983), Fuller (2008)
Afyeh, Lower Nubia	4500	23N-32E		Seeds	Chowdhury & Buth (1970, 1971)
Mohenjo-Daro, Sindh	4500–3700	27N-68E	<i>G. arboreum</i> (hairs) *	Seeds	Gulati & Turner (1929) in Chowdhury and Buth (1971)
Kanmer, Kacchh	4000–3700	23N-71E		Seeds	Fuller (2008)
Hallur, Karnataka	2950–2900	14N-76E		Seeds	Fuller (2008)
Nineveh, Mesopotamia	2700	33N-44E		Text	Malatacca (2014), Muthukumaran (2016)
Sippar, Babylonia	2500	33N-44E		Text	Muthukumaran (2016)
Qal'at al-Bahrain, Arabia	2500	26N-50E		Seeds	Bouchaud et al. (2011)
Egypt	2500			Text	Herodotus in Malatacca (2016)
Arabia	2250			Text	Theophrastus (2300 BP) in Bouchaud et al. (2011)
Mada'in Salih, Saudia Arabia	1900	27N-38E		Seeds	Bouchaud et al. (2011)
Yunnan, China	2150	23N-102E	G. arboreum **	Text	Chao (1977) in Zurndorfer (2011)
Sichuan, China	2050	28N-102E	G. arboreum **	Text	Chao (1977) in Zurndorfer (2011)
Yunnan, China	1825	23N-102E	G. arboreum **	Text	Kuhn (1988)
Old Jarma (Germa), Fazzan, Libyan Sahara	1800	27N-13E	G. herbaceum	Seeds	Pelling (2007)
Kellis, Upper Egypt	1750–1550	25N-29E		Seeds, Text	Bowen (2010) in Brite & Marston (2013)
Kara-tepe, Khorezm, Uzbekistan	1660–1580	42N-59E		Seeds	Brite (2011) in Brite & Marston (2013)
Upper Egypt	1600	32E-22N	G. herbaceum (DNA)	Seeds	Palmer <i>et al.</i> (2012)
Turfan basin & Khotan basin, Xinjiang	1500 (Liang dynasty)	43N-89E	G. herbaceum **	Text	Kuhn (1988), Zurndorfer (2011)
Merv, Turkmenistan	1400–1500	N38-62E		Seeds	Herrmann et al. (1993) in Brite & Marston (2013)
Yingpan, Yuli County, Xiniiang	800	41N-87E	<i>G. herbaceum</i> (fibres)	Fibre	Cao (2009)

*: dubious identification as lint is indistinguishable morphologically with traditional instruments

** : no real species assignment ; it is admitted – and reasonable – that the cotton cultivated in the past in western China (XinJiang) was *G. herbaceum* and that the cotton in eastern China was *G. arboreum*.

Supplementary data 3

The table S3 below gives distribution statistics of the studied traits of wild/feral *Gossypium arboreum* and *G. herbaceum* accessions from CIRAD's gene bank; traits are 28 agronomic, phenologic and morphometric traits and 10 fibre quantity and quality traits.

Table S3. Comparison of means of phenotypic traits between wild and improved	Gossypium
herbaceum L. and Gossypium arboreum L.	

	Wild	G.	arboreum			Wild	G.	herbaceum				
	Nb	Mean	Max	Min	S-Dev	Nb	Mean	Max	Min	S-Dev		
Da1Flo	23	112	131	93	8.3	9	86	123	64	21.9	days*	Fleur1
Da1Bo	8	153	176	135	12.0	9	134	158	112	18.4	days*	Caps1
PlaShap	23	6.6	7.0	1.0	1.5	9	3.2	7.0	1.0	1.9	scale	Port
PlaHei	25	183	282	113	40.0	12	177	284	113	49.6	cm	Hauteur
H1FruBr	22	33.2	41.0	11.7	7.3	9	11.0	29.0	5.0	7.5	cm	H1bf
VegBrN	25	18.5	39.0	5.0	10.8	12	12.6	36.2	1.5	12.9	count	Bv
StmHair	29	2.9	6.0	1.0	1.5	12	5.3	8.0	1.0	2.7	scale	Piltige
LfShap	29	6.0	7.0	3.0	1.3	12	3.8	5.0	3.0	1.0	sinus	FormFeuil
LfPlan	29	4.7	7.0	2.0	1.2	12	5.3	7.0	3.0	1.6	scale	SurFeuil
LfLobN	29	5.9	7.0	4.0	1.0	12	5.2	7.0	5.0	0.6	count	Lobes
LfHair	29	2.8	7.0	1.0	2.5	12	4.9	7.0	1.0	2.5	scale	PilFeuil
LfSiz	29	4.4	6.0	2.0	1.5	12	3.8	5.0	3.0	0.6	scale	TailleFeuil
NecN	29	1.8	3.0	1.0	0.8	12	1.5	3.0	1.0	0.9	count	Nectaires
LfGla	29	4.7	7.0	2.0	1.5	12	5.6	7.0	3.0	1.1	scale	GosFeuil
BracSiz	29	3.2	7.0	2.0	0.9	12	3.3	5.0	2.0	0.9	scale	TailBract
BracFor	28	5.9	7.0	1.0	2.0	12	3.5	5.0	2.0	1.2	scale	FormBract
BracDtN	29	5.0	11.0	0.0	2.0	12	8.8	12.4	5.0	2.4	count	NbDentBrac
Macula	28	5.0	9.0	1.0	2.3	12	5.8	8.0	1.0	2.4	scale	Macule
BoSiz	29	3.0	4.0	2.0	0.3	12	3.2	4.0	2.0	0.7	scale	TailCaps
BoMuc	29	3.4	6.0	2.0	1.2	12	3.8	7.0	2.0	1.6	scale	MucronCaps
BoLocN	29	3.0	4.0	3.0	0.2	12	3.3	4.0	3.0	0.5	scale	Loges
BoGla	29	5.3	8.0	3.0	1.2	12	5.0	8.0	2.0	1.9	scale	Goscaps
BoAper	29	3.3	7.0	2.0	1.4	12	5.0	8.0	3.0	1.4	scale	Dehiscence
BoStorm	29	5.8	8.0	3.0	1.6	12	3.8	7.0	2.0	1.6	scale	Adherence
BoWei	24	1.4	2.5	1.1	0.4	9	1.8	2.6	1.1	0.6	g	PMC
SdWei	29	6.1	8.5	4.5	0.8	11	7.3	10.6	5.3	1.6	g	Si
SdFuz	25	8.1	15.7	1.7	3.9	11	7.9	16.4	2.7	4.8	scale	Linter
SdHair	21	5.0	8.0	3.0	2.3	10	6.5	8.0	3.0	1.9	scale	Pilgr
GOT/R	29	24.6	39.8	21.3	3.5	13	26.9	35.9	22.5	4.3	%	RDTegrer
LinLen	12	23.5	26.8	19.3	2.3	9	23.0	25.4	20.5	1.6	mm	Uhml
LinLenU	12	81	85	77	2.3	9	81	85	79	2.2	%	UI
LinStr	12	26.4	30.7	20.9	2.8	9	22.1	28.7	16.2	4.4	g/tex	T1Hvi
LinElo	12	5.9	8.3	4.4	1.3	9	5.6	6.6	5.0	0.4	%	E1Hvi
LinMic	28	6.5	7.5	4.9	0.6	9	5.2	6.3	3.9	0.8	scale	IM
LinMat	28	98.3	107.5	81.0	7.1	9	78.8	98.6	55.1	13.1	%	PM
LinFin	28	210	326	172	34	9	265	360	197	59	mtex	HS
LinRef	28	69	79	62	5	9	78	82	68	4	scale	RD
LinYel	28	10.4	11.7	8.2	0.9	9	9.3	13.8	8.0	1.8	scale	В

* days: days after emergence

Abbreviations for trait names:

Agro-morphological traits: Da1Flo: time to 1st flower (days after emergence); Da1Bo: time to 1st boll; PlaShap: Plant height to width ratio; PlaHei: Plant height at harvest; H1FruBr: Height 1st sympodial branch; VegBrN: Vegetative branches number; StmHair: Main stem hairiness; LfShap: Leaf sinus depth; LfPlan: Leaf planarity; LfLobN: Leaf lobes number; LfHair: Leaf lower side hairiness; LfSiz: Leaf size; NecN: Leaf nectaries number; LfGla: Leaf gossypol glands density; BracSiz: Bracts size; BracFor: Bracts dents depth; BracDtN: Bracts dent number; Macula: Flower bottom macula size; BoSiz: Boll size; BoMuc: Boll tip size; BoLocN: Boll lobules number; BoGla: Boll gossypol glands density; BoAper: Boll aperture; BoStorm: Boll stormproof; BoWei: Boll weight; SdWei: Seed weight (100 seed); SdFuz: Seed fuzz density; SdHair: Seed fuzz density.

Lint traits: RGOT: Roller-gin ginning outturn, i.e. % fibre on roller gin; LinLen: Lint length: HVI UHML; LinLenU: Lint length uniformity; LinStr: Lint strength HVI; LinElo: Lint elongation; LinMic: Lint micronaire; LinMat: Lint maturity; LinFin: Lint fineness; LinRef: Lint reflectance; LinYel: Lint yellowness.