A global perspective on the origins of agriculture: the importance of unconscious selection

Thomas Kluyver

Department of Animal and Plant Sciences

A thesis submitted for the degree of Doctor of Philosophy

July 2013

Acknowledgements

Abstract

The protracted model of agricultural origins proposes that farming developed gradually in broad regions, rather than being invented and adopted rapidly in small 'core areas'. This view points to an important role for unconscious selection in the evolution of crop plants, wherein humans unintentionally modified the environment of plants in cultivation, setting up selection pressures different from those acting in the wild. This thesis examines the role of unconscious selection on seed mass and photosynthesis, especially in grass and legume crops. Domestication is known to have increased seed mass in many seed crops, while studies that have compared photosynthetic rate have mostly, but not universally, found no difference between wild and domestic forms. An important aspect of this work has been making comparisons among a range of crop species and geographic regions. This is not to presume that the same processes were acting in these different cases: it is important to study each crop and each region individually. However, it is also natural to look for larger patterns. Agriculture in widely separated parts of the world appears to have started more-or-less simultaneously, and in many cases to have used plants from the same families, especially the grasses and the legumes. I have therefore chosen to examine how far these similarities extend, and how great the differences between the regions and crop species are. Analysing seed mass data shows that crop progenitors already have large seeds in comparison to other wild species, but only the centre of agricultural origins in Western Asia has

an unusual abundance of large-seeded species. A variety of vegetable crops, including vegetatively propagated species, have larger seeds than their wild progenitors, providing evidence that unconscious selection has acted on seed mass; the difference, however, is smaller than that seen in seed crops. A comparative experiment tested the hypothesis that seed burial and the need to emerge from deeper in the soil drove the evolution of larger seeds, the results of which support this mechanism in some, but not all, of the grain legume species tested. Finally, photosynthetic rate has not changed in domestication in any of a range of grass and legume crops; possible explanations for this are discussed. In summary, I find broad taxonomic and geographic patterns in the seed mass of crop progenitors, increased seed mass in both seed and vegetable crops, and no change in photosynthesis during domestication, while the effect of seed mass on emergence depth differs among species, and the availability of wild large-seeded grass and legume species differs markedly among regions.

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Context

The transition from a hunter-gatherer lifestyle to farming was one of the most significant changes that human societies have undergone. It is widely thought that the food surpluses produced by agriculture were a key factor allowing urban civilisations to arise (Childe 1941; Flannery *et al.* 1967; Harlan 1992; Maisels 1993; Hanson 1999). On the other hand, agriculture was initially detrimental to health (Steckel *et al.* 2002; Larsen 2006; Starling & Stock 2007), and it may take more of a society's time than foraging (Lee 1968; Diamond 1987), although this point is disputed (Kaplan 2000). Because of this, and the increased inequality of complex societies, one author has dramatically, if controversially, called the switch to agriculture "a catastrophe from which we have never recovered" (Diamond 1987).

Although the first clear signs of cultivation are not until around 8,000 BC, a key precursor began some 40,000 years ago, in the late Palaeolithic. From around this time, archaeological evidence indicates that humans started to exploit a much wider range of food sources than they had previously—including smaller, faster animals which represented a worse trade-off of energy expended to nutritional value (Flannery 1969; Stiner 2001). Although Western Asia is the best studied region, a similar pattern is also evident elsewhere (Piperno & Dillehay 2008; Prendergast, Yuan & Bar-Yosef 2009). The "broad spectrum revolution" included plant species: at the Ohalo II site in

Israel (~21,000 BC), seeds of some 142 taxa were found, including wild cereals and small grained grasses (Weiss *et al.* 2004b). Both there and in Europe, there is evidence that humans were grinding seeds to produce flour, and perhaps cooking a primitive bread (Piperno *et al.* 2004; Aranguren *et al.* 2007).

Previously, the transition from a hunter-gatherer lifestyle to farming was thought to have been a relatively sudden affair, driven by specific human innovations and the evolution of domestication traits in crops. This view is perhaps most clearly stated in Childe's concept of the 'Neolithic Revolution' (1941), and more recently it has been promoted by Diamond (1997). The idea of rapid, localised transitions to agriculture is supported by genetic studies of crop plants which found that domestic forms had a single origin, e.g. Heun et al. (1997) for einkorn wheat, Matsuoka et al. (2002) for maize, and Guo et al. (2010) for soybean.

However, in recent years, a number of lines of evidence have challenged this view, and the dominant picture is now of a much more gradual shift from foraging to cultivation (Brown *et al.* 2009; Fuller, Asouti & Purugganan 2012a), although not everyone agrees with this view (Abbo, Lev-Yadun & Gopher 2010b; Peleg *et al.* 2011). Genetic studies do now indicate multiple independent domestication events for some crops: barley, for instance, has two different versions of the nonshattering allele, a key part of the domestication syndrome, and it is widely accepted that these relate to separate domestications (Takahashi 1955; Molina-Cano *et al.* 2005; Morrell & Clegg 2007;

Azhaguvel & Komatsuda 2007), although this conclusion is not universally held (Badr *et al.* 2000). Common beans and Lima beans were likely domesticated both in Central America and in the Andes (Gutiérrez Salgado, Gepts & Debouck 1995; Chacón S, Pickersgill & Debouck 2005), and the single origin of Asian rice is debated (Londo *et al.* 2006; Vaughan, Lu & Tomooka 2008; Fuller & Qin 2009). In addition, a computer simulation of evolution has shown that the apparent monophyly found in many studies is a likely result even for crops which were domesticated several times, due to the effects of genetic drift (Allaby, Fuller & Brown 2008). Alongside this, archaeological excavations have revealed that wild forms of many crops were farmed before domestication (Willcox 2005; Weiss, Kislev & Hartmann 2006; Willcox, Fornite & Herveux 2008). The feasibility of such *pre-domestication cultivation* has, however, been debated, especially for legume crops (Ladizinsky 1987, 1993; Zohary 1989; Weiss *et al.* 2006; Willcox *et al.* 2008; Abbo *et al.* 2011). Where wild and domesticated forms of a crop can be distinguished from archaeobotanic remains, the rise of domesticated strains may take several millennia (Tanno & Willcox 2006; Fuller *et al.* 2009), although this too is debated (Abbo *et al.* 2011).

1.2 Sources of Evidence

The primary evidence for work on the origins of agriculture comes from archaeobotanical remains. Gathered seeds and associated chaff can be preserved by charring in a fire, by desiccation, or occasionally by waterlogging (Kislev, Nadel & Carmi 1992). Although charring

distorts and shrinks seeds (Hubbard & al Azm 1990; Braadbaart 2008), they, and especially chaff, can often be identified to a genus or species level (e.g. Hillman *et al.* 1996; Jones, Valamoti & Charles 2000; Fuller & Harvey 2006). For many species, their size allows domesticated forms to be distinguished from wild forms. Mature grains may also be distinguished from immature grains, which may have been harvested from wild type plants before the seed dispersed (Hillman & Davies 1990a; Fuller *et al.* 2009). For cereals, the abscission scar where the ear separates distinguishes wild grains and spikelet bases (shattering, leaving a small, smooth abscission scar) from domesticated ones (non-shattering; threshing leaves a larger, jagged scar) (Tanno & Willcox 2006, 2012; Fuller *et al.* 2009). No seeds from the relevant period remain viable (the oldest recorded viable seed was some 2,000 years old (Sallon *et al.* 2008)), but it is possible to recover and study DNA (Schlumbaum, Neuhaus & Jacomet 1998; Freitas *et al.* 2003; Elbaum *et al.* 2006; Giles & Brown 2008). With sufficient quantities of seed preserved, statistical techniques can also be applied. For example, Colledge (2002) identified a suite of plant taxa which may represent an early arable weed assemblage.

Like other organic remains, preserved seeds and chaff are suitable for radiocarbon dating. By convention, all ages are now described in calibrated years unless otherwise noted, and this work follows this standard.

Other physical forms of evidence, which are not further discussed in this thesis, include:

- Microfossils such as starch grains and phytoliths (e.g. Piperno et al. 2009), and pollen (e.g. Meadows 2005).

- Biomarkers, stable chemicals indicative of certain species (Evershed 2008).

- Stable isotopes of carbon and nitrogen, which can be a proxy for environmental conditions and trophic networks (e.g. Araus *et al.* 1999; Barton *et al.* 2009).

Genetic evidence is another important strand, for example in illuminating the ancestry of a domesticated crop. This is necessary particularly for polyploid crops, which include wheat, peanuts, finger millet and potatoes, among many others (Singh & Smartt 1998; Bisht & Mukai 2001; Petersen *et al.* 2006; Ovchinnikova *et al.* 2011). Genetic studies have also been used to investigate whether a crop has been domesticated multiple times; in barley, the presence of two different mutations leading to non-shattering rachises, backed up by haplotype data, strongly suggests at least two separate domestications (Morrell & Clegg 2007).

Details of the genetic changes underlying components of the domestication syndrome (see below) can inform models of the domestication process. In lentils, for example, seed dormancy is controlled by one main locus. This has been used to suggest that a large enough wild population would include non-dormant mutants which could have been co-opted into agriculture (Ladizinsky 1993).

1.3 The domestication syndrome

Domestication refers to genetic changes in a species as it adapts to human cultivation rather than the wild environment. The details differ for each domesticated species, but a number of traits are commonly seen in domesticated plants. These constitute the 'domestication syndrome', which is described here (and in most sources) for seed crops, although the relevant parts also apply to other types of crops:

- Loss of natural dispersal (nonshattering rachis in cereals, and indehiscent pods in legumes). This would allow ripe seeds to be harvested by cutting plants with a sickle, although wild-type plants could have been harvested unripe, or by beating seeds into a basket. Loss of dispersal is the most important part of domestication, as it leaves the plant dependent on human activity to propagate it. For cereals in particular, natural dispersal can be distinguished from human separation of grains in archaeobotanical remains (see above).

- Increased seed size. This could have been consciously selected, as it has been more recently, or it could be an adaptation to deeper burial of seeds due to tillage, or to seedling competition (Harlan, de Wet & Price 1973). Selection could also have acted on traits correlated with seed size, such as resilience to damage (Cunniff 2009). See chapters 3 and 4 for further discussion.

- Loss of physical defences and dispersal aids such as hooks and awns

- Changes in composition in favour of compounds favoured by humans, including a loss of toxins

- Loss of seed dormancy

- Simultaneous production and ripening of seeds

- Reduced plant size, and increased proportion of biomass allocated to edible parts (harvest index)

- More determinate growth, including reduced branching

The classic description of these traits in cereals is that of Harlan et al. (1973), and it was Hammer (1984) who first used the term 'Domestikationssyndrom' in German.

1.4 Unconscious Selection

'Unconscious selection' refers to selective pressure effected by human activity on crops and weeds, but not by deliberately choosing which individuals to propagate, which is called 'conscious selection'.

The term was coined by Darwin, although his definition differs somewhat from the sense in which it is usually used today. According to Darwin, unconscious selection consists of deliberate efforts to preserve favourable characteristics in the next generation, in contrast to the long term modification of the breed intended in 'methodical selection':

> "*Unconscious selection* is that which follows from men naturally preserving the most valued and destroying the less valued individuals, without any thought of altering the breed; and

undoubtedly this process slowly works great changes. Unconscious selection graduates into methodical, and only extreme cases can be distinctly separated; for he who preserves a useful or perfect animal will generally breed from it with the hope of getting offspring of the same character; but as long as he has not a predetermined purpose to improve the breed, he may be said to be selecting unconsciously." (Darwin 1875 p. 140)

Darlington (1956) refers to Darwin, and goes on to define "a particular class of unconscious selection, which we may call operational selection". This includes selection pressures caused by sowing and harvesting conditions, acting without any deliberate human choice of which individuals reproduce. Similarly, Harlan et al. (1973) refer to 'automatic selection'. More recent authors, however, have mostly used the term 'unconscious selection' for this kind of selection (Heiser 1988; Zohary 2004), along with the contrasting phrase 'conscious selection' for what Harlan et al. called 'deliberate selection'.

The best example of unconscious selection is the loss of wild type dispersal mechanisms—the shattering ears of wild grasses and the dehiscent pods of wild legumes. Once cultivators began to collect seed in a particular way, e.g. by cutting the stalks, and replant that seed on newly cleared land, seeds which were retained on the plant until harvest would have comprised a greater proportion of the next generation, potentially setting up a strong selection pressure for seed retention (Harlan *et al.* 1973; Hillman & Davies 1990a; Zohary 2004). Another, less favoured hypothesis is that wild seeds were harvested

by beating seeds into a basket, leaving non-shattering mutants in the
field to produce the next generation (Blumler *et al.* 1991). The
archaeobotanical evidence of evolutionary rates indicates that the
actual selection pressure was considerably weaker than the potential
maximum; early cultivators may well have harvested before all the
seeds were mature, and mixed cultivated seed with seed gathered
from the wild, or resown on the same land where the seed of
shattering spikelets fell (Tanno & Willcox 2006; Fuller 2007).

1.5 Where agriculture started
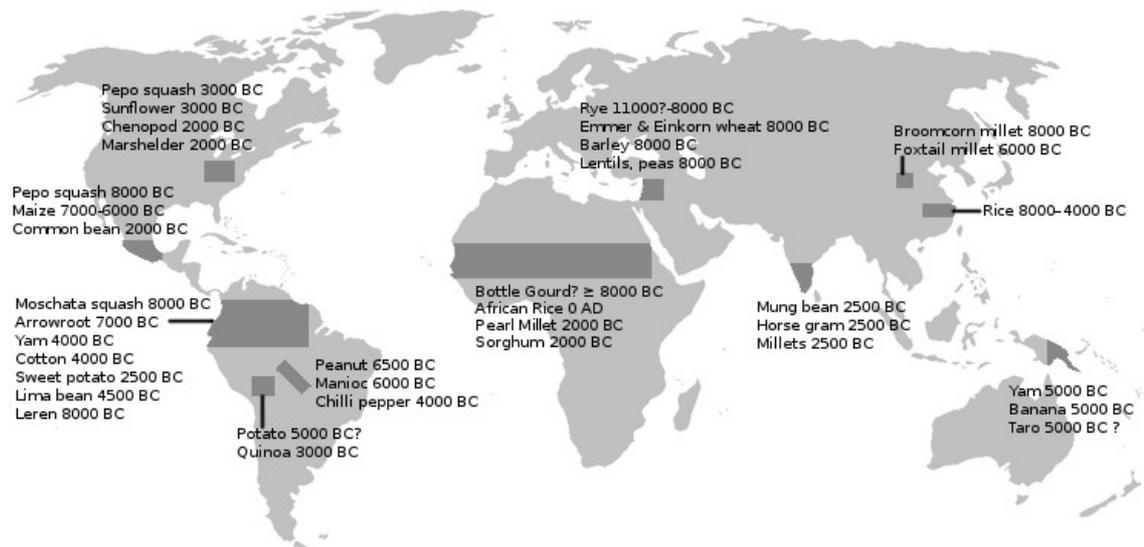
*Figure 1.1: One view of global centres of origin of agriculture, and
selected early crops. After Balter (2007), updated with dates from
Zhao (2010), Fuller et al. (2007) Lu et al. (2009) and Erickson et al.
(2005).*

De Candolle (1885 p. 448), while noting the unequal distribution of
the origins of cultivated species, described the "impossibility of
subdividing the continents and… islands in natural regions." However,
the Russian botanist Vavilov proposed seven basic centres of origin
(Vavilov 1931), in Southwestern Asia, India and Indo-china, China, the

Mediterranean, Ethiopia, Central America and part of highland South America. Most of these regions are still discussed today, although most researchers now think that agriculture diffused around the Mediterranean from Western Asia, and some also believe that there was no 'independent invention' of agriculture in Africa (Blumler 1992b). Eastern North America and New Guinea have also been proposed as independent centres (Denham *et al.* 2003; Smith 2006; Fuller 2006). Within and around these regions, there is disagreement over how many localities developed agriculture independently: Diamond refers to "at most nine areas of the world" (Diamond 2002), while other authors claim as many as 24 (Purugganan & Fuller 2009). Harlan (1971) proposed a system of three independent origins, each with a centre and a diffuse 'noncentre', but later (Harlan 1992 p. 53) decided to speak of broader regions rather than specific centres. It is also important to be clear what is meant by 'independence': the concepts and practices of agriculture may have spread to regions where crops were domesticated anew. Abbo et al. (2010a) argue that, even where crops have genetically independent origins, their cultivation may not be culturally independent, although Fuller et al. (2012b) hold that agriculture requires a set of knowledge and practices that are unlikely to diffuse rapidly via brief encounters for trade.

While it is generally held that plant cultivation began separately in the Old and New Worlds, even this may be challenged. Remains of an Old World domesticate (bottle gourd, *Lagenaria siceria*) have been

identified in the Americas from 7900 BC (Erickson *et al.* 2005), and the authors suggest that it was transported by humans along the coast as they migrated into the Americas, although they also consider the possibility of the tough, buoyant fruits floating across the ocean. If this species was widely cultivated at such an early date, many of the proposed centres of origin may not be 'pristine' (in the sense of Fuller (2011a)), in that cultivation of indigenous crops was initiated by people aware of cultivated plants introduced from elsewhere. The peculiar use of its fruits as containers may, however, have allowed bottle gourd to be domesticated without cultivation, if seeds were scraped out, and germinated on dumpheaps. Arguably, it is the decision to cultivate food plants, not the awareness of how to do so, which is critical in the origins of agriculture; many hunter gatherers may have been aware of how plants reproduced, but as discussed below, farming was not necessarily desirable.

That agriculture appears to have begun independently in a number of different regions is very valuable to modern researchers. Comparative work looking at multiple centres of origin can distinguish general patterns from specific historical circumstances. With a knowledge of general patterns, we can attempt to explain the causes and effects of the development of farming.

1.6 Causes of the Transition to Agriculture

It is easy now to see agriculture as a 'superior' option to foraging, but when it began, it probably was not beneficial to individuals. As mentioned above, it probably provided a less healthy diet, and may

have required more work. One member of an extant hunter-gatherer tribe in Africa is quoted as saying "Why should we [farm], when there are so many mongongo nuts in the world?" (Lee 1968). Harlan (1992 p. 41) cites records of Aboriginal Australians expressing similar sentiments.

One possibility is that pressure on resources, for example from an expanding population, required a shift to farming (Cohen 1977; Richerson, Boyd & Bettinger 2001). Even if it requires more work, cultivation allows more food to be produced from a given area of land, to support a greater population density. It has been suggested that the apparently different course of agriculture in Africa was driven by a requirement for stability in an unpredictable environment, rather than total food production (Marshall & Hildebrand 2002).

As described above, people in several regions independently began to farm around the same time. In particular, the first evidence of agriculture in Western Asia, in China, and in Central America, are all around 8,000 BC (Smith 1997; Brown *et al.* 2009; Piperno *et al.* 2009; Crawford 2009; Jones & Liu 2009), pointing to some common, global trigger (Blumler 1992b; Cohen 2009).

At this time, the cold Younger Dryas period had recently come to an end, marking the beginning of the Holocene, the current geological epoch (Gulliksen *et al.* 1998). The global extent of the Younger Dryas is debated (Rodbell 2000; Bertrand *et al.* 2008), but it seems to have affected at least a large part of the Northern Hemisphere (Andreev,

Klimanov & Sulerzhitsky 1997; Islebe & Hooghiemstra 2006). The warmer, wetter, more stable conditions of the Holocene may have made agriculture possible (Richerson *et al.* 2001; Gupta 2004; Willcox, Buxo & Herveux 2009). One author, however, has argued that the *start* of the Younger Dryas triggered the transition to agriculture, as the colder, dryer conditions made wild-growing food less reliable (Hillman *et al.* 2001).

Across the boundary between the Pleistocene and the Holocene, atmospheric carbon dioxide levels also rose over a few millennia, from around 190 parts per million (ppm) to about 265 ppm (current levels being some 380 ppm) (Monnin *et al.* 2001). The resulting increase in plant productivity may have been a prerequisite for successful farming (Sage 1995; Cunniff *et al.* 2008; Cunniff 2009).

Other writers have argued that more cultural pressures led people to intensify food production. The feasting model of domestication proposes that surpluses beyond what could be stored for later use were used in social events to form alliances (Hayden 2009). Bowles and Choi (2013) have recently claimed that the development of individual property was a necessary condition for agriculture, because the incentives to farm would have been insufficient if the products were shared out; interestingly, Bogaard et al. (2009) believe that plant foods at the Neolithic site of Çatalhöyük were stored in individual houses, but animal foods were shared. Such internal forces cannot, however, explain the apparently independent and synchronous origins of agriculture in widely separated regions.

Demographic pressure, climate change and social factors are the three most common kinds of explanation for the beginning of agriculture (Zeder 2006). Rindos (1984) proposed a more complex model in which humans and crop plants gradually co-evolved into a mutualistic relationship.

1.7 Groups of Cultivated Plants

A full account of every domesticated species is beyond the scope of this literature review. For more details, see Zohary & Hopf (2000) for Old World species, and Smartt & Simmonds (1995) for a worldwide list.

1.7.1 Grasses

Grasses are highly productive, were key early crops in several regions, and remain our most important crops today. They are commonly categorised into cereals (with large seeds), including emmer & einkorn wheats (*Triticum dicoccum* Schrank & *T. monococcum* L.), barley (*Hordeum vulgare* L.), oats (*Avena sativa* L.), rice (*Oryza sativa* L. & *O. glaberrima* Steud.) and maize (*Zea mays* L.), and small seeded grasses or millets, including broomcorn millet (*Panicum miliaceum* L.), foxtail millet (*Setaria italica* (L.) P. Beauv.), pearl millet (*Pennisetum glaucum* (L.) R. Br.), finger millet (*Eleusine coracana* Gaertn.), teff (*Eragrostis tef* (Zucc.) Trotter) and Sorghum (*Sorghum bicolor* (L.) Moench).

1.7.2 Legumes

After the grasses, the Fabaceae are the second most important family

of domesticates globally. In Western Asia, four pulses are found from around the time of the transition to agriculture: lentil (*Lens culinaris* Medik.), pea (*Pisum sativum* L.), chickpea (*Cicer arietinum* L.) and bitter vetch (*Vicia ervilia* (L.) Willd.). Other old-world domesticated pulses include other *Vicia* species (prominently *V. faba* L.), *Vigna* spp. and soybean (*Glycine max* (L.) Merr.).

In the new world, various species of *Phaseolus* have been domesticated, most notably *Phaseolus vulgaris* L., which supplies various types of bean now sold in western supermarkets, including the haricot beans used to produce 'baked beans'. Peanut (*Arachis hypogaea* L.) is also a legume, despite its obvious differences from peas and beans. Other species of legume are also cultivated for their fruit (e.g. Pacay, *Inga feuillei* DC.) or roots (e.g. Jícama, *Pachyrhizus erosus* (L.) Urb.).

The grasses and the legumes are particularly interesting for comparative work, because representatives of both families have been independently domesticated in multiple centres of origin of agriculture (Table 1.1). Much of this thesis is based on these two crop families.

	Grasses	**Legumes**
Western Asia	Einkorn wheat, *Triticum monococcum* L. Emmer wheat, *Triticum dicoccum* Schrank Barley, *Hordeum vulgare* L. Rye, *Secale cereale* L. Oat, *Avena sativa* L.	Pea, *Pisum sativum* L. Lentil, *Lens culinaris* Medik. Chickpea, *Cicer arietinum* L. Broad bean, *Vicia faba* L. Vetches, *Vicia* spp. Lupins, *Lupinus* spp.
Africa	Sorghum, *Sorghum bicolor* (L.) Moench Pearl millet, *Pennisetum glaucum* (L.) R.Br. Finger millet, *Eleusine coracana* Gaertn. African rice, *Oryza glaberrima* Steud. Teff, *Eragrostis tef* (Zucc.) Trotter	Cowpea, *Vigna unguiculata* (L.) Walp. Winged bean, *Psophocarpus tetragonolobus* (L.) DC. (origins uncertain, (Klu 1996 p. 10))
India	*Echinochloa frumentacea* Link	Mung bean, *Vigna radiata* (L.) R. Wilczek Urd bean, *Vigna mungo* (L.) Hepper Pigeon pea, *Cajanus cajan* (L.) Millsp.
China	Rice, *Oryza sativa* L. Broomcorn millet, *Panicum miliaceum* L. Foxtail millet, *Setaria italica* (L.) P. Beauv.	Soybean, *Glycine max* (L.) Merr.
Central/ South America	Maize, *Zea mays* L.	Common bean, *Phaseolus vulgaris* L. Lima bean, *Phaseolus lunatus* L. Jack bean, *Canavalia ensiformis* (L.) DC. Peanut, *Arachis hypogaea* L.

Table 1.1: Grass and legume species domesticated in different regions. This list is not exhaustive.

1.7.3 Pseudocereals

Certain dicotyledons have been grown as grain crops and prepared

and used in a similar manner to grasses. In China, buckwheat

(*Fagopyrum esculentum* Moench) is ground into flour, while in the New World, where only one grass, maize, was an important crop, a number of pseudocereals were cultivated, of which quinoa (*Chenopodium quinoa* Willd.) and amaranth (*Amaranthus* spp.) are now the most prominent.

1.7.4 'Root' crops

Many plants are cultivated for their edible roots, tubers or corms. Yams (*Dioscorea* spp.) have been domesticated in Africa, South-East Asia and South America. In New Guinea, taro (*Colocasia esculenta* (L.) Schott) was cultivated by around 4500 BC (Denham *et al.* 2003). In South-East Asia and the Pacific region, a number of plants provide food from underground organs, including *Amorphophallus konjac* K. Koch., *A. paeonifolius* (Dennst.) Nicolson and *Alocasia macrorrhizos* (L.) G. Don. The greatest profusion of root crops, however, is found in South America. Potato (*Solanum tuberosum* L.) and sweet potato (*Ipomoea batatas* (L.) Lam.) are now familiar to us, others, such as cassava (*Manihot esculenta* Crantz.), ulluco (*Ullucus tuberosus* Caldas), arrowroot (*Maranta arundinacea* L.) and lerén (*Calathea allouia* (Aubl.) Lindl.), less so (Martin & Cabanillas 1976; Piperno *et al.* 2000).

1.7.5 Other crops

Tidy categories can never account for the full range of cultivated species around the world. Other important early crops include squashes (*Cucurbita* spp.) and sunflower (*Helianthus annuus* L.) in the Americas (Smith 2006; Piperno *et al.* 2009), and banana (*Musa* spp.)

in New Guinea (Denham *et al.* 2003). Many other crops, including onion & garlic (*Allium* spp.) and vegetables in the genus *Brassica* are not recorded until later (Zohary & Hopf 2000). Herbs and spices, although used in smaller quantities, also have long histories; notably, chilli peppers (*Capsicum* spp.) have been used since at least 4,000 BC (Perry *et al.* 2007), and stores of wild mustard seeds (*Descurainia sophia* (L.) Webb ex Prantl) were found at a Neolithic site in Turkey (Twiss *et al.* 2008; Twiss 2011).

It is generally held that perennial crops, including fruit and nut trees as well as vines, were not cultivated until some millennia after farming began, as horticulturalists would need to be highly sedentary (Zohary & Hopf 2000). One find has been taken to imply that the fig (*Ficus carica* L.) was domesticated in Western Asia even before cereals (Kislev, Hartmann & Bar-Yosef 2006a), but this is, unsurprisingly, debated (Kislev, Hartmann & Bar-Yosef 2006b; Lev-Yadun *et al.* 2006). However, a different pattern may hold in tropical areas, where perennial root crops were important at an early time in South America and South-east Asia (Piperno *et al.* 2000; Denham *et al.* 2003; Fullagar *et al.* 2006). Sago, starch harvested from palms, may also have been an early food source, although it is not clear whether the palms were cultivated (Jones, Hunt & Reimer 2013; Yang *et al.* 2013).

Food, of course, is not the only use to which plant parts have been put. The bottle gourd (*Lagenaria siceraria* (Molina) Standl.) has been mentioned above for its apparent early spread to the new world. Flax

(*Linum usitatissimum* L.), an oil and fibre crop, is among the 'founder crops' of Western Asia, perhaps originally for its oil (Allaby *et al.* 2005). Hemp (*Cannabis sativa* L.) was domesticated in Asia, while cotton (*Gossypium* spp.) was domesticated independently from species in the old and new worlds (it is a new world species, *G. hirsutum* L., which is widely grown today). Dye plants, such as woad (*Isatis tinctoria* L.) and madder (*Rubia* spp.) have also been used for millennia.

1.8 Selection of Plants for Cultivation

Not all of the plants eaten by humans following the broad spectrum revolution (see above) were later cultivated. One prominent example is brome (*Bromus* sp.), which is represented by the greatest number of seeds from Ohalo II, a site dating from 21,000 BC (Weiss *et al.* 2004b). Brome is not thought to have ever been cultivated in that region.

Large-seeded food plants may be more likely to be selected for cultivation (Weiss *et al.* 2004b; Willcox *et al.* 2008). Many domesticated plants certainly evolved larger seeds, albeit at different rates and times (Purugganan & Fuller 2009), although this does not necessarily imply that the species with the largest seeds were initially chosen for cultivation. Growth experiments have found that three large-seeded crop progenitors (wild emmer & einkorn wheats, and wild barley) grow larger (although at the same relative rate) and tolerate defoliation better than several other, smaller-seeded, grass species found in Neolithic near-eastern seed assemblages (Cunniff

2009 chap. 6). Anecdotally, however, the cultivation of various small-seeded crops, notably teff (*Eragrostis tef* (Zuccagni) Trotter), suggests that the story is more complex. These ideas are explored in chapter 2.

In addition to studying plants grown in controlled environments, information has been gained from experimental harvests of stands of food plants using appropriate tools. Despite different possible techniques, harvesting wild cereals in the Fertile Crescent gives high yields of grain per hour spent (Harlan 1967, cited in 1992 p. 160; Ladizinsky 1975). In contrast, collection of wild pulses from the same region tends to give much lower yields (Abbo *et al.* 2008a; b). A simple 'return on effort' foraging model does not appear to account for the domestication of pulses, but their nutritional value, particularly their protein content, may have played a role (Murphy 2007; Abbo, Saranga & Peleg 2009). This, in turn, relates to their biology: pulses are members of the Fabaceae, which harbour mutualistic nitrogen-fixing bacteria in their roots. One study even suggested that a more specific nutrient, the amino acid tryptophan, was the pressure leading to chickpea domestication (Kerem *et al.* 2007).

This raises an obvious question. Neolithic man was not an expert on his own nutrition, and didn't have the facilities to analyse potential foodstuffs. How could he (or she) have selected crops for nutritional value? For tryptophan, it is possible that it has direct effects on the brain through conversion to serotonin (Kerem *et al.* 2007), but this would not necessarily have been sufficient for conscious selection.

Could a subconscious system lead humans (and other animals) to prefer the taste of foods containing nutrients they particularly need? This is discussed further in chapter 2.

The 'dumpheap model' of domestication involves seeds growing in disturbed, fertile ground near to human settlements. Unconscious selection would be expected to favour weedy, ruderal plants, but the Western Asian founder crops mostly do not fit the description, casting doubt on that idea (Abbo *et al.* 2005). Harlan (1992 p. 95) notes that it "fits some species, but not others." Secondary crops, however, may have originated as weeds in cultivated 'fields', a scenario thought to be the case for oats and rye (Sencer & Hawkes 1980; Zhou, Jellen & Murphy 1999). Zohary & Hopf (2000) mention that wild peas (peas are among the Western Asian founder crops) occasionally occur as weeds of cereal cultivation.

A substantial number of crop species are polyploid: wheat, finger millet and peanut are three of many examples. There are some theoretical reasons why polyploids might be pre-adapted for cultivation, but a statistical analysis showed no influence of ploidy on domestication (Hilu 1993).

What might explain the concentration of domesticates in certain families, such as grasses and legumes? Dempewolf et al. (2008) examined the Asteraceae, a large family containing only a handful of domesticates (including the sunflower and artichokes), and suggested that their chemical defences may have prevented humans from using

other species. However, the presence of chemical defences does not necessarily preclude domestication: many 'edible' legume species contain toxic substances which have to be removed by soaking or cooking (Gupta 1987)

1.9 Aims and objectives

The key idea driving this work is that, when we test big ideas about the early development of agriculture, we should look for similarities and differences between the various regions where agriculture began independently. A lot of work in this field focuses on Western Asia: conditions there have preserved many archaeobotanical remains, while the spread of agriculture from there to Western Europe meant that for many years, the people most likely to ask academic questions about human prehistory have been most familiar with the crops and practices developed from this centre. I deliberately include crop plants from China, Africa, India, and Central and Southern America, as well as Western Asia. New Guinea and Eastern North America have also been proposed as harbouring independent origins of agriculture, but grasses and legumes are not among their domesticates, so they do not feature much in this work.

This is not to deny the importance of research and hypotheses pertaining to a specific region. There is undoubtedly some truth in Harlan's 'no-model model' (1992 p. 46), in which he argues that we should not presume that agriculture began in the same way in different parts of the world. One of the key findings of chapter 2 is that the patterns of seed mass in wild grasses and legumes

distinguish Western Asia from all the other centres of domestication. As such, there are many good reasons to better understand a single system. But we all have an inexorable interest in patterns. Especially with the "remarkable synchronicity of agricultural origins around the world" (Blumler 1992b), and the repeated domestication of plants from certain families, it is fascinating to consider whether the same processes could have been at work in regions separated by thousands of miles.

Through the lens of global comparisons, this thesis examines unconscious selection in the domestication of crop plants. In particular, I test the idea that seed size has globally determined the geographic and taxonomic patterns of seed mass, and how the seed mass of crop plants increased under unconscious selection. Finally, I consider whether unconscious selection during domestication might have affected photosynthesis. In each case, I look at a broad range of crop species and geographic regions, to evaluate where general conclusions can be drawn, and where we must look at specific explanations.

1.9.1 Chapter 2: The taxonomic and geographic distributions of seed mass in wild legumes and grasses

This chapter explores the patterns of large-seededness in these two families, which both contain numerous important crops from separate regions. I test two hypotheses:

Firstly, that crop progenitors had larger seeds than other available wild species. It is known that smaller-grained grasses were collected

before agriculture started (Weiss *et al.* 2004a), and the difference suggests that large-seeded species were favoured for cultivation, but it is not clear whether this applies across different regions and different groups of crops.

Secondly, that the availability of large-seeded grasses and legumes was a key factor in determining the locations in which agriculture originated. This aims to robustly test Diamond's (2005) assertions, elaborating from Blumler's thesis (1992a) on seed weight, that the distribution of large-seeded grasses in particular determined the development of agriculture, and thus much of human history.

1.9.2 Chapter 3: Unconscious selection on seed size in vegetable crops

Increased seed size is a common trait of many seed crops, as described above, and authors have explained this in terms of unconscious selection, such as through seed burial or competition. But to my knowledge, there is no strong evidence that unconscious selection affected seed mass, or of how significant a role it played. I look for a change in seed size among crops where the seed is not the harvested part, including both species normally grown from seed, and species normally grown from a vegetative organ.

1.9.3 Chapter 4: Did greater burial depth increase the seed size of domesticated legumes?

It has been proposed that cultivation of fields led to seeds being buried deeper in the soil, and that this selected for larger seeds which, having more reserves, would be better able to emerge from

that depth. Using eight grain legume crops from several centres of origin of agriculture, I test three predictions of this hypothesis:

- That larger seeds of a given species are more likely to emerge from depth.

- That domestication, which is known to increase seed size, also increases emergence depth.

- That the difference in emergence depth between wild and domestic crops should be greater than expected from the change in seed mass alone.

1.9.4 Chapter 5: Comparison of photosynthesis among crop species and progenitors

Photosynthesis itself could not have been under conscious selection until quite recently, but unconscious selection could well have acted via connections between photosynthesis and more visible traits such as growth. I look for differences in photosynthetic rate between wild and landrace accessions of a variety of grass and legume crops.

2 The taxonomic and geographic distributions of seed mass in wild legumes and grasses

2.1 Introduction

As described in the general introduction, people were exploiting a broad range of wild plants since the 'broad spectrum revolution', millennia before agriculture began (Weiss *et al.* 2004b; Aranguren *et al.* 2007). Many wild plant species still play an important role in human diets today (Freedman 2012). What factors led our ancestors to cultivate and domesticate the species that they did? In some cases, it seems that species with larger seeds were preferentially domesticated (Weiss *et al.* 2004b; Willcox *et al.* 2008), although our familiarity with domesticated crops, which have undergone selection for still larger seeds, may make us quick to give credence to this hypothesis. The relevant comparison is that between the wild progenitors of crops and other species that were not domesticated.

The term that should be used for the process by which only certain species came to be cultivated and domesticated is not clear. 'Selection' would be an obvious candidate, but that term is familiar to biologists for another meaning. 'Choice' implies a conscious human act which need not have been involved. Below, I have tried to refer to this process as 'filtering', but this is not ideal either: whereas something passing through a filter does not affect that filter's selectivity, the use of one plant species may well affect the likelihood of using other species. For instance, only one or a few crops of a given type may be required, while the cultivation of a productive but

inconsistent species may lead to the use of a more robust species as a form of insurance.

In a controversial popular science book entitled *Guns, Germs and Steel*, Diamond (2005 p. 136) suggests that the distribution of large-seeded grass species was an important factor determining where agriculture began, building on results from Blumler's thesis on seed size (Blumler 1992a). Diamond's book has been criticised for its environmental determinism (e.g. Sluyter 2003; Judkins, Smith & Keys 2008), but could seed size have affected where agriculture began? Diamond's claims were based in part on the idea that the Fertile Crescent in Western Asia was the original cradle of agriculture, and farming did not start in other regions until much later. However, there is a good case that agriculture in China and in Central America is only a little younger than in Western Asia (Piperno *et al.* 2009; Crawford 2009), and the cultivation of root crops in New Guinea may also have begun at a similar time (Fullagar *et al.* 2006). The transition to agriculture has been described as 'synchronous' for some decades (e.g. Cohen 1977).

Two plant families, the grasses (Poaceae) and legumes (Fabaceae) have been the source of multiple domesticates from independent centres of agricultural origin. If seed size was a critical factor determining which wild species were cultivated and subsequently domesticated, then we might predict that:

- Crop progenitors will have larger seeds than other wild species,

when compared with any of a number of possible outgroups: all species in the same family, species congeneric with crop progenitors, or species occurring in the centres of agricultural origin.

- The centres of agricultural origin will have above average availability of large-seeded species which could be candidates for domestication.

I have therefore collected seed mass and distribution data for hundreds of species in each of these two families, to examine how seed mass in crop progenitors and the centres of agricultural origin compare to that in other species and regions.

2.2 Methods

2.2.1 Data Sources

Seed mass data was taken from the USDA GRIN/NPGS database (USDA Agricultural Research Service 2012), Kew's Seed Information Database (Royal Botanic Gardens, Kew 2008) and USDA PLANTS (USDA, NRCS 2013). This was supplemented by my own measurements of seed mass, and seed masses found from various literature sources (Janzen 1977; Davies 1977; Schnee & Waller 1986; Rees 1995; Abayomi *et al.* 2001; Williams *et al.* 2003; Parera & Ruiz 2003; Zhang, Du & Chen 2004; Menezes, Franke & Dall'agnol 2004; Bisht *et al.* 2005; Clements *et al.* 2005; Maass 2005; Mendoza & Dirzo 2007; Wu & Du 2007; Dirzo, Mendoza & Ortíz 2007; Bu *et al.* 2007; Sammour *et al.* 2007; Rowland 2008; Eastwood & Hughes 2008;

Vamosi, Mazer & Cornejo 2008; Borek *et al.* 2009; Wang *et al.* 2009, 2012; Teixeira-Sá *et al.* 2009; Yang *et al.* 2010; Eule-Nashoba 2010; Pulse Western Australia 2013; Schloen, Peters & Schultze-Kraft 2013; Hu *et al.* 2013). Seed mass is variously given for a single seed, a hundred or a thousand seeds. All values were converted to a per-seed mass in milligrams. Where a paper quoted a mass range, the midpoint of that range was used. Vamosi et al. (2008) listed seed volumes and a close correlation between mass and volume; the relationship in that study was used to transform their calculated volumes into masses.

Distribution information was sourced from the Kew grass synonymy database (Clayton *et al.* 2002) and the ILDIS legume database (International Legume Database & Information Service 2005). Life history data was taken from ILDIS for legumes, and from Kew GrassBase (Clayton *et al.* 2006) for grasses. Life history information was supplemented with data from the USDA PLANTS database (USDA, NRCS 2013).

2.2.2 Accessing DELTA Data

The data from Kew GrassBase (Clayton *et al.* 2006) was stored in the DELTA format (Dallwitz 1980). Searching the internet, I found that a software package had been written, but not released, which would allow me to convert this data into a format that could be combined with other data and statistically analysed. I contacted the author, and collaborated with him to improve this code and release it as an open source package under the name PyDelta (Cavalcanti & Kluyver 2010). Details of how to obtain and use this are available online at

http://freedelta.sourceforge.net/pydelta/

2.2.3 Combining Data

In order to combine data from the sources listed above, each dataset was mapped onto botanical names from the Kew grass synonymy database (Clayton *et al.* 2002) and the ILDIS legume database (International Legume Database & Information Service 2005). Automatically matching scientific names on this scale is not trivial: synonyms, homonyms, spelling variations and occasional spelling mistakes must all be handled. Small numbers of names can be matched up by hand, but this is frustrating and time-consuming, and the series of decisions involved cannot be easily reproduced. As such, an automated matching process is highly desirable, and essential for large datasets.

A software tool, Taxonome, has been developed to handle and match scientific names automatically, following standard taxonomic rules (Kluyver & Osborne 2013). It uses fuzzy matching to account for spelling variations or mistakes. While initial development focussed on plant nomenclature (McNeill *et al.* 2006; Miller *et al.* 2011), it is also flexible enough to deal with zoological names (International Commission on Zoological Nomenclature 1999), although the two systems use slightly different formats.

Taxonome treats a taxon as having one accepted name (as described by the chosen data source), and a number of synonyms. Each taxon may also have other associated information, such as its distribution

and data about biological traits. A group of taxa from one source are stored in a data structure (a TaxonSet) which indexes all the names, so that a taxon can be quickly found given a binomial name.

Where separate data sources have information on the same taxa, these are represented as two separate collections, and one may be matched against the other. Matching preserves the information attached to each taxon, but reassigns its name to the accepted name from the dataset against which it is matched. The matching process can also produce CSV files recording the matches made and the different steps taken. Several collections of taxa with matched names may then be combined into one set.

To match a name, a number of possibilities are tried, most of them user configurable:

- An exact match, including the authority, is always preferred.

- If a name matches but doesn't have a matching authority, this can be used unless the user has disabled such matches. However, if the authorities specifically indicate that the names refer to different taxa, the match is rejected (see below).

- Taxa below species level which do not have an exact match can be matched to the parent species. This can be done for all subspecies, only for nominal subspecies (e.g. *Zea mays* subsp. *mays*), or disabled.

- Where possible, fuzzy matching is used to account for spelling variations and errors in the data (see below).

31

In the case of homonyms, more than one match may be found. If one of the matches is an accepted name, Taxonome can accept it as the most likely option. This is done by default when the name being matched does not have author information. Otherwise, the matching process can be set to let the user decide in such cases. The user can pick from the available matches, enter a replacement name, or reject all the options.

Taxonome employs fuzzy string matching to account for differences in spelling. For binomial names, an approach based on Q-grams is used (Gravano *et al.* 2001). Each name is broken into overlapping chunks of three letters, including two padding characters at the beginning. The standard Q-gram algorithm also includes padding characters at the end, but Taxonome omits these to give less weight to the ending, where the spelling most often differs. The proportion of these chunks which another name has in common gives a similarity score. To speed up lookups, the first three characters of the name must match exactly. For example, if no exact match is found for *Mucuna holtoni*, it is broken down to '^^M', '^Mu', 'Muc', 'ucu', etc. The set of q-grams is then compared with those for each name beginning with 'Muc', finding a 93% overlap with the q-grams for *Mucuna holtonii* (with a double i). By contrast, *Mucuna restonii* only shares 60% with *Mucuna holtonii*, below the default acceptance threshold of 70%. This threshold can be altered by the user.

For author citations, which are typically very short strings, a more bespoke approach is used. Taxonome identifies components such as

initials, surnames and dates. This is particularly important when a name is qualified with a phrase like 'non Vahl', which means that it is not the name defined by Vahl. A simple string similarity test might erroneously match with 'Vahl', but Taxonome will recognise the word *non*, and exclude such matches.

Data can be read from CSV files, and the software is flexible enough to accept a range of possible structures. Output data are also written to CSV files. Data that are to be re-used within Taxonome can be saved in a simple format based on JSON (Crockford 2006), which can store structured data, such as nested lists, more conveniently than tabular CSV files.

Custom scripts were written to load taxonomic data from the Kew grass synonymy database (Clayton *et al.* 2002), and from the ILDIS legume database (International Legume Database & Information Service 2005). These scripts are available from Taxonome's website.

2.2.4 Data Preparation

Data was filtered to only annual species and, for legumes, a separate dataset of only herbaceous species. Although there is contemporary interest in perennial grains (Glover *et al.* 2010), most important grain crops are grown as annuals, suggesting that this is a prerequisite for domestication. Exceptions do exist, such as pigeonpea, *Cajanus cajan* (L.) Millsp., a perennial legume grown in the tropics, but even this is "normally grown as an annual" (van der Maesen 1995). Woody perennials in both families include much larger seeded species, such

as the bamboos and leguminous trees. Species were included if they were listed as having variable habits including annual (or herbaceous). Species for which life history information was not included in the database were excluded, as when details were found for a sample of these, many of them proved to be woody perennials.

For each species, the arithmetic mean of the seed mass measurements was calculated, so that the analysis was based on one mass value per species. For statistical tests, log-transformed values were used, as seed masses are log-normally distributed, both in this dataset and in the literature (Leishman, Westoby & Jurado 1995).

The International Taxonomic Database Working Group (TDWG) has defined a set of regions at four different scales, for which GIS data are available (Brummitt *et al.* 2001). These are pragmatically defined, following political boundaries, but biologically relevant: larger countries are separated into smaller regions. The distribution information for grass species was provided in the form of TDWG Level 3 regions. For legumes, distributions were given as lists of country and region names, and with a few minor adjustments, these could be successfully matched to TDWG regions, using functionality available in Taxonome. All geographical analysis was therefore based on TDWG Level 3 regions, which generally correspond to small countries (e.g. France), subdivisions of large countries (e.g. Brazil Southeast) or island groups (e.g. the Society Islands).

2.2.5 Domesticates and centres of origin

From my knowledge of the literature, I compiled lists of species known to be domesticated, along with their immediate wild progenitors, where known (Table 2.1 and Table 2.2). For some crops, such as *Panicum miliaceum*, a progenitor is not yet known. In others, such as *Vicia ervilia*, the wild forms are regarded as part of the same species, with no taxonomic classification separating wild and domestic. In these cases, the species was listed as domestic; since the primary aim of this chapter is to compare crop progenitors with other wild species, excluding these species from the set of crop progenitors is the conservative option. For *Vigna aconitifolia*, however, a datum was available from a wild accession, so this was included as a progenitor.

In the case of neopolyploid crops, the immediate progenitor of the same ploidy level was used, such as *Triticum dicoccoides* Koern. for emmer wheat, *Eleusine africana* Kenn.-O'Byrne for finger millet (Bisht & Mukai 2001) and *Arachis monticola* Krapov. & Rigoni for peanut (Seijo *et al.* 2007). Genome donors were not considered progenitors, as they do not represent the plants that early famers chose to cultivate. Therefore, no progenitor species is included for bread wheat, *Triticum aestivum* L., a hexaploid believed to have been arisen in cultivation.

Domesticates	Progenitors
Avena sativa L. *Avena byzantina* C. Koch. *Avena nuda* L. *Avena abyssinica* Hochst. *Avena strigosa* Schreb. *Avena brevis* Roth.	*Avena sterilis* L. *Avena fatua* L. *Avena barbata* Pott ex Link
Brachiaria ramosa (L.) Stapf *Brachiaria deflexa* (Schum.) C. E. Hubb. ex Robyns	
Coix lacryma-jobi L. *Coix lacryma-jobi* var. *ma-yuen* (Romanet) Stapf	*Coix lacryma-jobi* var. *stenocarpa* Oliver *Coix lacryma-jobi* var. *monilifer* Watt
Digitaria exilis (Kippist) Stapf *Digitaria iburua* Stapf	*Digitaria longiflora* (Retz.) Pers. *Digitaria barbinodis* Henrard
Echinochloa stagnina (Retz.) Beauv. *Echinochloa frumentacea* Link *Echinochloa esculenta* (A. Braun) H. Scholz	*Echinochloa colona* (L.) Link *Echinochloa crusgalli* (L.) Beauv.
Eleusine coracana (L.) Gaertn.	*Eleusine africana* Kennedy-O'Byrne
Eragrostis tef (Zucc.) Trotter	*Eragrostis pilosa* (L.) Beauv.
Hordeum vulgare L. *Hordeum distichon* L.	*Hordeum spontaneum* C. Koch
Oryza sativa L. *Oryza glaberrima* Steud.	*Oryza rufipogon* Griff. *Oryza barthii* A. Chevalier
Panicum miliaceum L. *Panicum sumatrense* Roth ex Roem. & Schult.	*Panicum psilopodium* Trin.
Paspalum scrobiculatum L.	
Pennisetum glaucum (L.) R. Br.	*Pennisetum violaceum* (Lam.) Rich.
Secale cereale L.	*Secale vavilovii* Grossheim
Setaria italica (L.) Beauv.	*Setaria viridis* (L.) Beauv.
Sorghum bicolor (L.) Moench	*Sorghum arundinaceum* (Desv.) Stapf
Triticum aestivum L. *Triticum macha* Dekaprel. & Menabde *Triticum compactum* Host *Triticum ×fungicidum* Zhuk.	*Triticum aegilopoides* Forssk. *Triticum boeoticum* Boiss. *Triticum dicoccoides* (Koern.) G. Schweinfurth *Triticum timopheevii* subsp.

Domesticates	Progenitors
Triticum ×timococcum Kostov *Triticum durum* Desf. *Triticum monococcum* L. *Triticum spelta* L. *Triticum turgidum* L. *Triticum turanicum* Jakubz. *Triticum polonicum* L. *Triticum carthlicum* Nevski *Triticum dicoccon* (Schrank) Schübl. *Triticum timopheevii* (Zhuk.) Zhuk.	*armeniacum* (Love) M.W. van Slageren
Zea mays L.	*Zea mexicana* (Schrad.) Kuntze *Zea mays* subsp. *parviglumis* H. H. Iltis & J. F. Doebley

Table 2.1: Grass species treated as domesticates and progenitors, with names following the Kew grass synonymy database (Clayton et al. 2002). Not all of these species have seed mass data.

Domesticates	Progenitors
Arachis hypogaea L. *Arachis villosulicarpa* Hoehne	*Arachis monticola* Krapov. & Rigoni *Arachis pietrarellii* Krapov. & W.C. Greg.
Canavalia ensiformis (L.) DC. *Canavalia gladiata* (Jacq.) DC.	
Cajanus cajan (L.) Millsp.	*Cajanus cajanifolius* (Haines) Maesen
Cicer arietinum L.	*Cicer reticulatum* Ladiz.
Glycine max (L.) Merr.	*Glycine max* subsp. *soja* (Siebold & Zucc.) H. Ohashi
Lablab purpureus (L.) Sweet	*Lablab purpureus* subsp. *uncinatus* Verdc.
Lathyrus sativus L. *Lathyrus cicera* L.	*Lathyrus clymenum* L.
Lens culinaris Medik. *Lens culinaris* subsp. *culinaris* Medik.	*Lens culinaris* subsp. *orientalis* (Boiss.) Ponert
Lupinus albus L. *Lupinus luteus* L. *Lupinus angustifolius* L. *Lupinus mutabilis* Sweet	*Lupinus albus* subsp. *graecus* (Boiss. & Spruner) Franco & P. Silva *Lupinus piurensis* C.P. Sm.
Macrotyloma uniflorum (Lam.) Verdc.	
Mucuna pruriens (L.) DC. *Mucuna pruriens* var. *utilis* (Wall.) L. H. Bailey	*Mucuna pruriens* var. *pruriens* (L.) DC.
Phaseolus acutifolius A.Gray *Phaseolus acutifolius* var. *acutifolius* A.Gray *Phaseolus coccineus* L. *Phaseolus lunatus* L. *Phaseolus vulgaris* L. *Phaseolus vulgaris* var. *vulgaris* L.	*Phaseolus acutifolius* var. *latifolius* F.L. Freeman *Phaseolus vulgaris* var. *aborigineus* (Burkart) Baudet
Pisum sativum L. *Pisum sativum* subsp. *sativum* L. *Pisum abyssinicum* A. Braun	*Pisum sativum* subsp. *elatius* (M. Bieb.) Asch. & Graebn.
Psophocarpus tetragonolobus (L.) DC.	
Vicia ervilia (L.) Willd. *Vicia faba* L. *Vicia faba* var. *minor* Beck	*Vicia sativa* subsp. *amphicarpa* (Dorthes) Asch. *Vicia sativa* subsp. *cordata*

Domesticates	Progenitors
Vicia sativa L. *Vicia sativa* subsp. *sativa* L. *Vicia sativa* subsp. *nigra* (L.) Ehrh.	(Hoppe) Asch. & Graebn.
Vigna aconitifolia (Jacq.) Marechal *Vigna angularis* (Willd.) Ohwi & H. Ohashi *Vigna mungo* (L.) Hepper *Vigna radiata* (L.) R. Wilczek *Vigna radiata* var. *radiata* (L.) R. Wilczek *Vigna subterranea* (L.) Verdc. *Vigna umbellata* (Thunb.) Ohwi & H. Ohashi *Vigna unguiculata* (L.) Walp. *Vigna unguiculata* subsp. *cylindrica* (L.) Verdc. *Vigna unguiculata* subsp. *pubescens* (R. Wilczek) Pasquet *Vigna unguiculata* subsp. *sesquipedalis* (L.) Verdc. *Vigna unguiculata* subsp. *stenophylla* (Harv.) Marechal & al. *Vigna unguiculata* subsp. *unguiculata* (L.) Walp.	*Vigna aconitifolia* (wild accessions) *Vigna angularis* var. *nipponensis* (Ohwi) Ohwi & H. Ohashi *Vigna mungo* var. *silvestris* Lukoki et al. *Vigna radiata* var. *sublobata* (Roxb.) Verdc. *Vigna umbellata* var. *gracilis* (Prain) Marechal et al. *Vigna unguiculata* subsp. *dekindtiana* (Harms) Verdc.

Table 2.2: Legume species treated as domesticates and progenitors, with names following the ILDIS database (International Legume Database & Information Service 2005). Not all of these species have seed mass data.

Similarly, I compiled a list of regions where agriculture began, including Western Asia, China, Cental America, South America and Africa, based on various literature sources (Kochert *et al.* 1996; Ba, Pasquet & Gepts 2004; Chacón S *et al.* 2005; Balter 2007; D'Andrea 2008; Brown *et al.* 2009; Piperno *et al.* 2009; Lu *et al.* 2009; Fuller 2011b; Fuller *et al.* 2012b). As it is often impossible to know exactly where a crop was domesticated, these were classified into two levels of confidence (Figure 2.1). The list does not include Eastern North

America or New Guinea: while they may have independently given rise to agriculture, their crops did not appear to include grasses or legumes (Smith 2006; Fullagar *et al.* 2006).
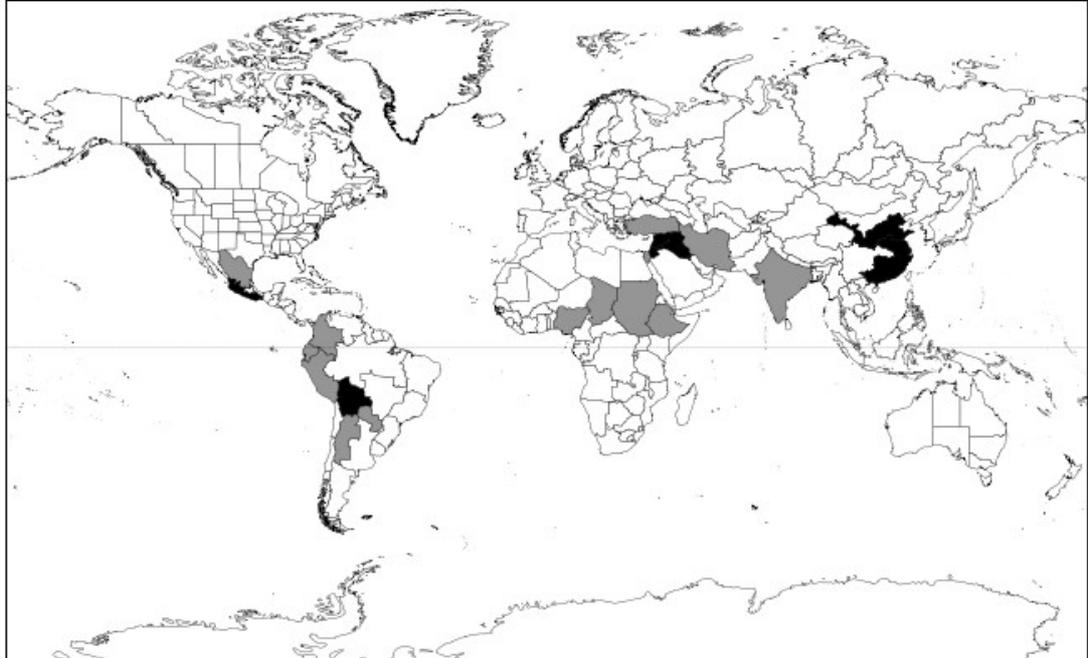
Figure 2.1: TDWG level 3 regions where agriculture is thought to have started. Black regions are assigned with higher confidence, and grey regions with lower confidence.

2.3 Results

2.3.1 Structure

In the grass family, classification to tribes explained 52% of the variance in species seed mass, and classification to genera explained 82%. In the legumes, classification to tribe explained 61% of seed mass variance, and to genus explained 82%.

2.3.2 Progenitors and congeners

In both annual grasses and herbaceous legumes, crop genera have large seeds relative to the remaining genera in the family (Figure 2.2; U=560, p=0.0065 for annual grasses; U=360, p=1.94×10^{-5} for

herbaceous legumes; Mann-Whitney U tests on genus geometric mean seed masses).

However, there is no consistent difference in seed mass between crop progenitors and congeneric wild species. In herbaceous legumes, the graph hints at a pattern (Figure 2.3), but the repeated domestications in *Vicia* and *Vigna* counteract this, and the difference is not significant overall (p=0.150, combining results for individual genera using the weighted-Z method). This remains true if *Mucuna*, a genus of tropical vines containing one relatively minor domesticate, *Mucuna pruriens* var. *utilis* (Wall. ex Wight) Baker ex Burck, is excluded from the comparison (p=0.147). In annual grasses, no clear pattern is evident (Figure 2.4; p=0.188). The difference between barley's progenitor and other species in the genus *Hordeum* is striking, but in this species in particular, there are concerns that wild forms may actually be weedy escapees from cultivation (Harlan 1992 p. 93), or at least have been subject to considerable introgression with domestic forms (Fuller 2007), although some genetic studies do find that *spontaneum* forms a separate clade from the cultivated *H. vulgare* (Kochieva, Goryunova & Pomortsev 2001).
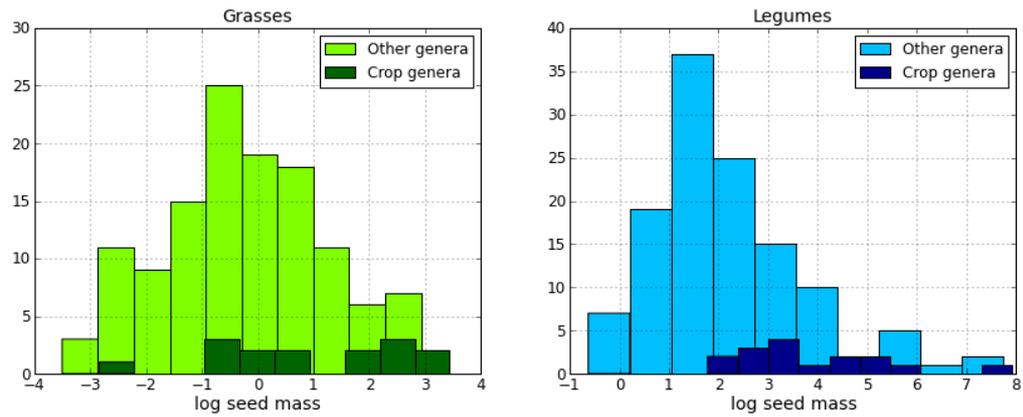
Figure 2.2: Distribution of genus average seed masses in annual grasses and herbaceous legumes, highlighting genera containing crops, but excluding the crop species themselves.
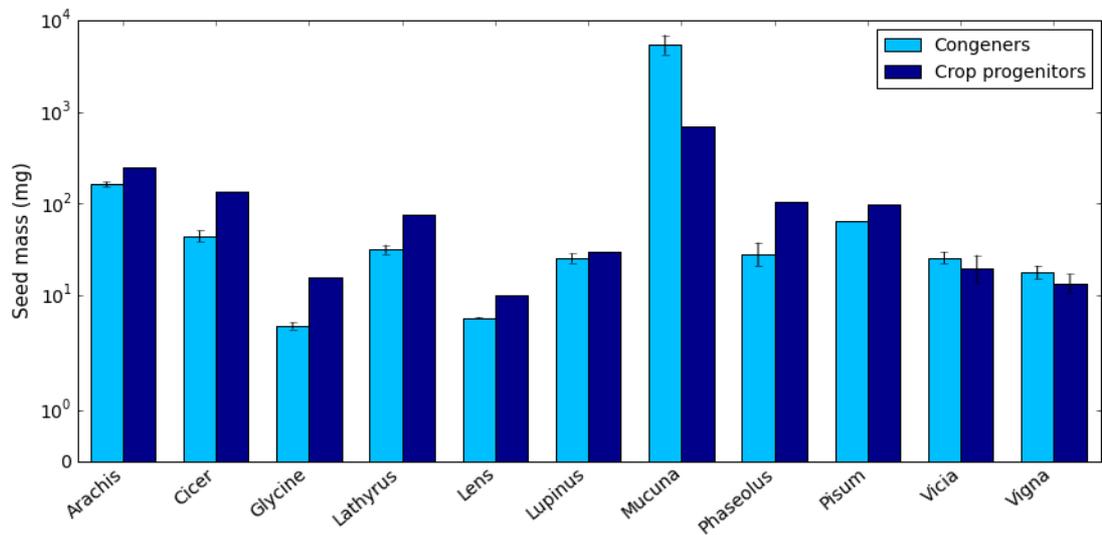
Figure 2.3: Seed masses of herbaceous legume crop progenitors in comparison with congeneric wild species.
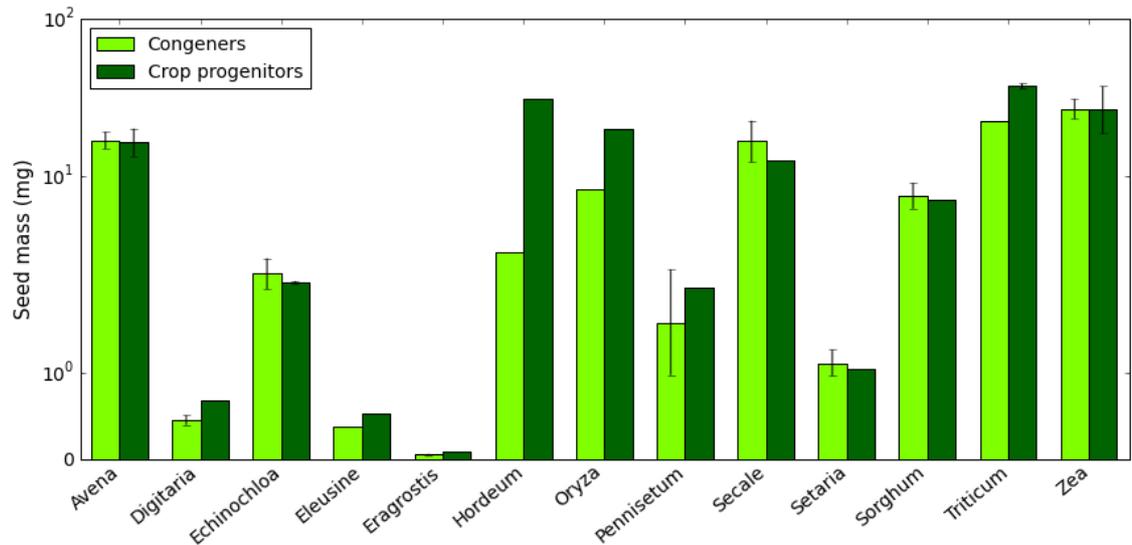
Figure 2.4: Seed masses in progenitors of annual grass crops in comparison with wild congeneric species.

2.3.3 Species from the centres of origin of agriculture

In the Western Asian centre of origin, crop progenitors in both annual grasses and annual/herbaceous legumes have larger seeds than other species occurring in the region (t=4.33, p=1.29×10^{-5}<0.01 for grasses; t=3.49, p=2.74×10^{-4}<0.0083 for herbaceous legumes; t=3.31, p=5.44×10^{-4}<0.01 for annual legumes; t-tests with the alpha value of 0.05 adjusted by the Holm-Bonferroni correction for the different regions tested with each species group).

In Central America, the progenitors of maize also have large seeds compared to other annual grasses in the region (t=3.87, p=1.24×10^{-4}<0.0125), and the various legume crop progenitors are large-seeded relative to other annual (t=2.89, p=0.00342<0.0125) but not herbaceous species (t=1.49, p=0.0697>0.0125). Legume crop progenitors in Africa have large seeds relative to both herbaceous (t=3.37, p=4.46×10^{-4}<0.01) and annual (t=4.97,

p=1.15×10^{-6}<0.0083) species in the region.

In other regions (grasses and legumes in South Asia, China and South America, and grasses in Africa) , crop progenitors do not have significantly larger seeds than other available species. However, this may be partly a consequence of the small number of progenitors: in all cases, the average of the progenitors is larger than the average of the other wild species (Figure 2.5, Figure 2.6), and combining the regions using the weighted-Z method (Whitlock 2005), the differences are highly significant (p=2.28×10^{-7} for annual grasses, 2.76×10^{-7} for herbaceous legumes, 2.18×10^{-9} for annual legumes).

Figure 2.7 and Figure 2.8 put the magnitude of these differences in context, showing the seed masses of the pooled crop species and pooled progenitor species in comparison with wild species.
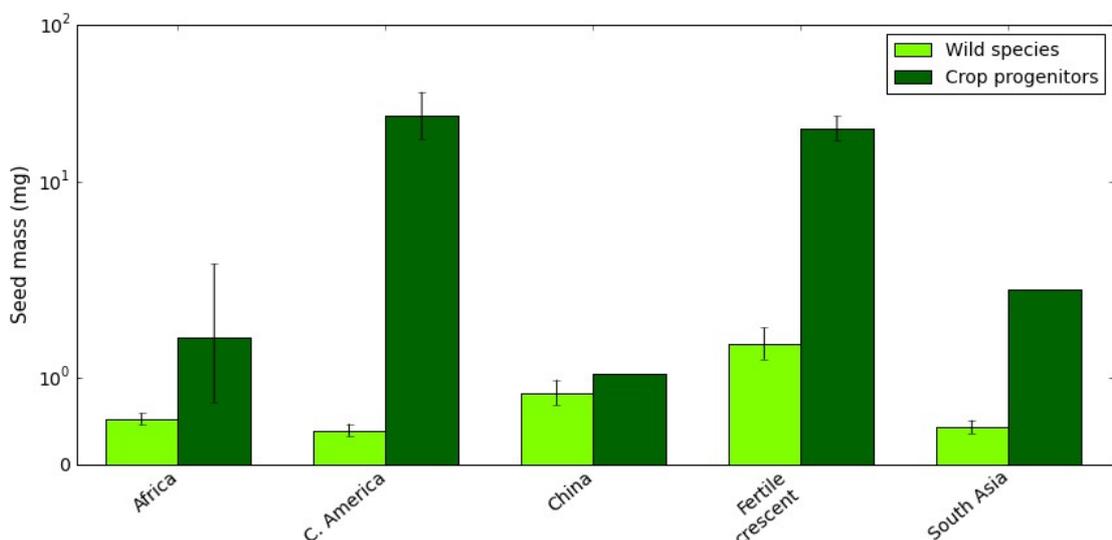
Figure 2.5: Comparison of seed masses between annual grass crop progenitors and other wild species in centres of origin of agriculture. Geometric means and standard errors of the geometric mean are shown.
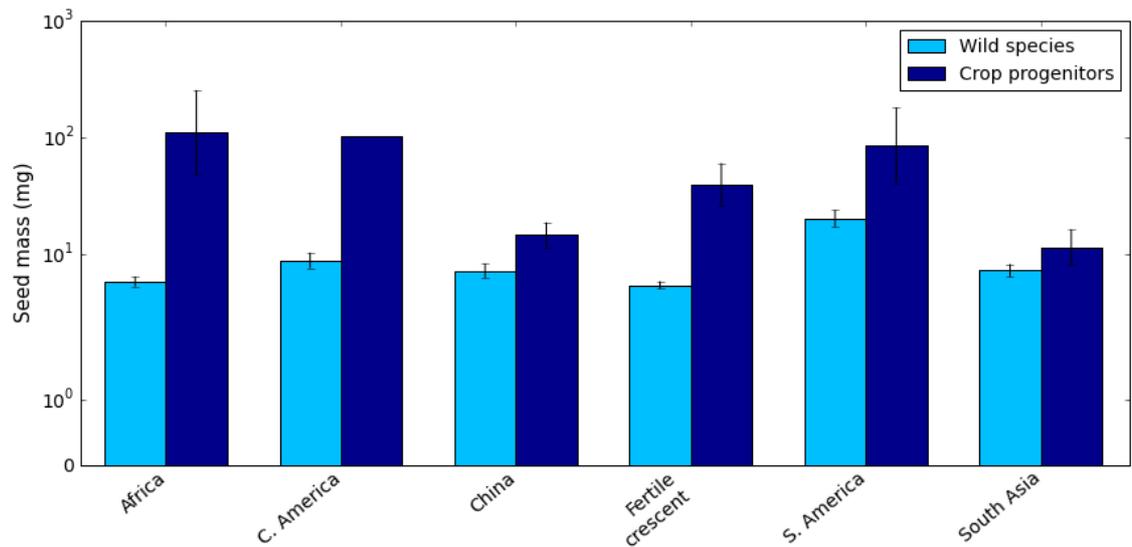
Figure 2.6: Comparison of seed masses between herbaceous legume crop progenitors and other wild species in centres of origin of agriculture. Geometric means and standard errors of the geometric mean are shown.
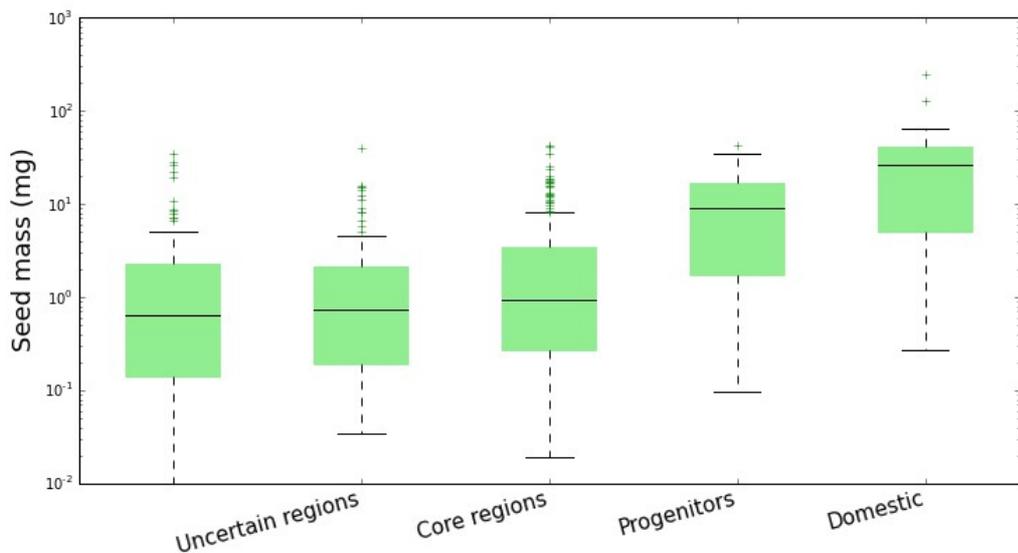
Figure 2.7: Seed masses in annual grasses related to their occurrence in centres of origin of agriculture. The column on the left contains all the species not in another group, and the leftmost three columns all contain wild, non-progenitor species.
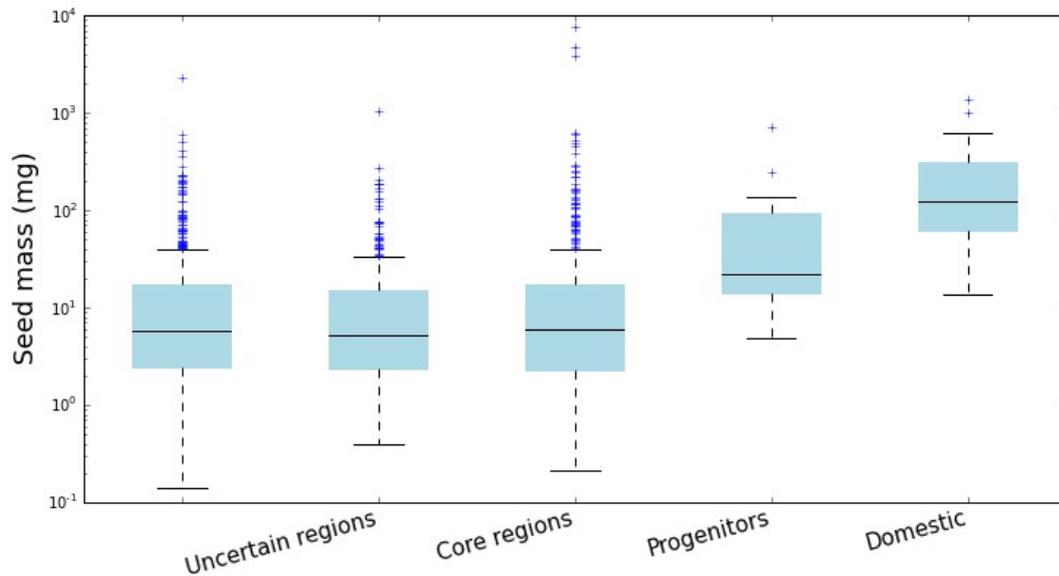
Figure 2.8: Seed masses in herbaceous legumes related to their occurrence in centres of origin of agriculture. The column on the left contains all the species not in another group, and the leftmost three columns all contain wild, non-progenitor species.

2.3.4 Geography and seed size distribution

In wild annual grasses, species occurring in the regions of Western

Asia where agriculture is thought to have started (see Figure 2.1)

have larger seeds than species not present in any of the centres of

origin (Figure 2.9; t=3.82, p=8.53×10^{-4}, planned comparisons using

Dunnett's test). However, this was not true of any of the other centres

of origin (p=0.66–1.00). The same pattern holds using only the

regions with higher confidence (black regions in Figure 2.1): the

Western Asian region differs from the remainder of species with

t=4.04, p=2.46×10^{-4}.

This analysis is potentially confounded if seed mass is correlated with

the extent of a species' distribution, as species with a wider

distribution are more likely to be included in the defined set of

regions. However, there is no correlation between the number of

regions in which a wild annual grass species occurred and its seed mass (F=0.0495, p=0.82, both variables log-transformed).

Among herbaceous legumes, seed mass *is* correlated with distribution, with more widely distributed species having slightly smaller seeds (F=13.1, p=3.02×10^{-4}; Figure 2.10). Using the residuals from this relationship, only seed masses in the South American centre are significantly larger than expected (Figure 2.11; t=7.65, p<10^{-4}, Dunnett's test); this is partially due to a number of species in the genus *Arachis*, relatives of peanuts. The differences for Central America and South Asia are also approaching significance (p=0.050, 0.060 respectively). Again, the same pattern appears when using only the regions with higher confidence (t=6.58, p<10^{-5} for South American centre). Considering annual legume species rather than herbaceous species gives similar results, with the differences in South Asia and China also marginally significant (p=0.035, 0.046 respectively).
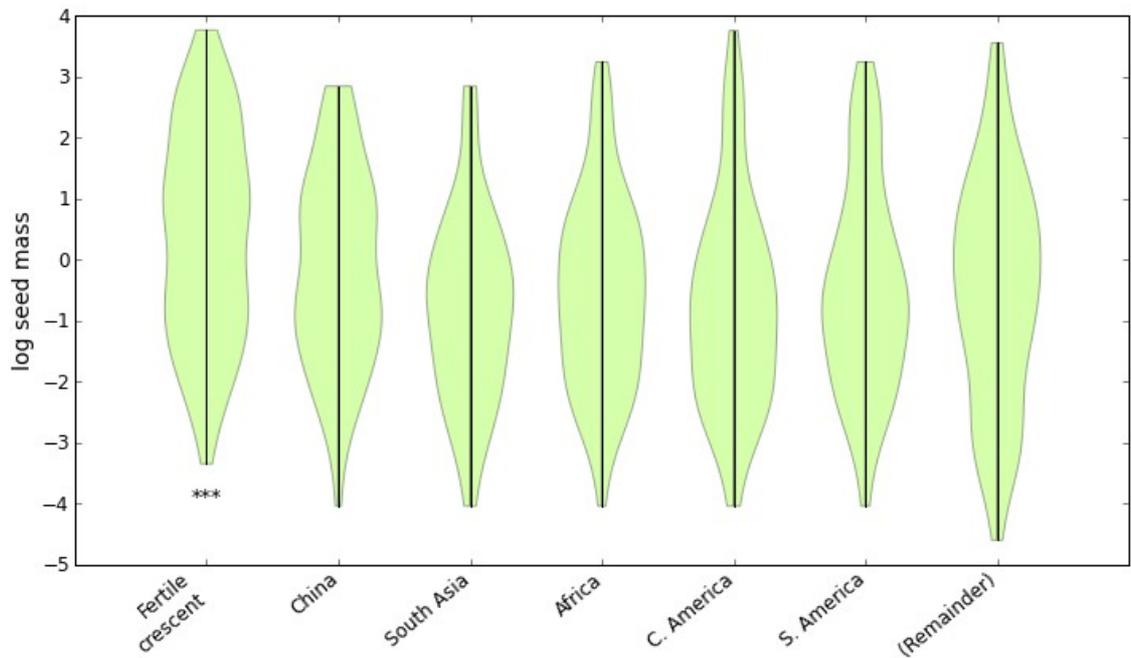
Figure 2.9: Relative frequency distributions of seed masses of wild annual grass species in regions where agriculture is proposed to have originated, compared with the remainder of species. Asterisks represent standard significance levels in comparison to the remainder of species, using Dunnett's test.
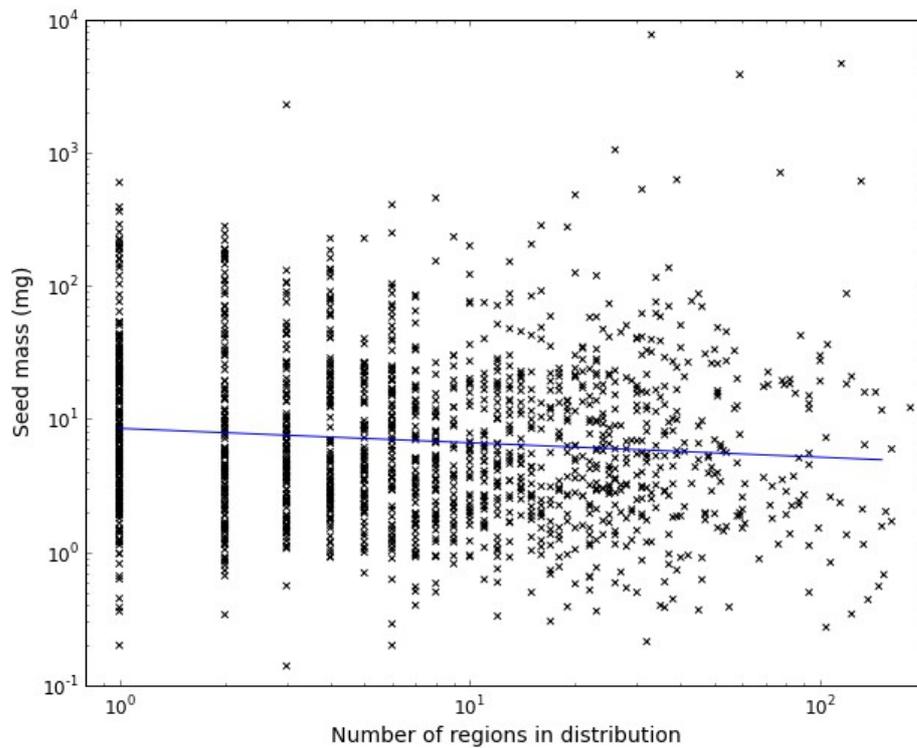
Figure 2.10: The relationship between seed mass and recorded distribution among herbaceous legume species.
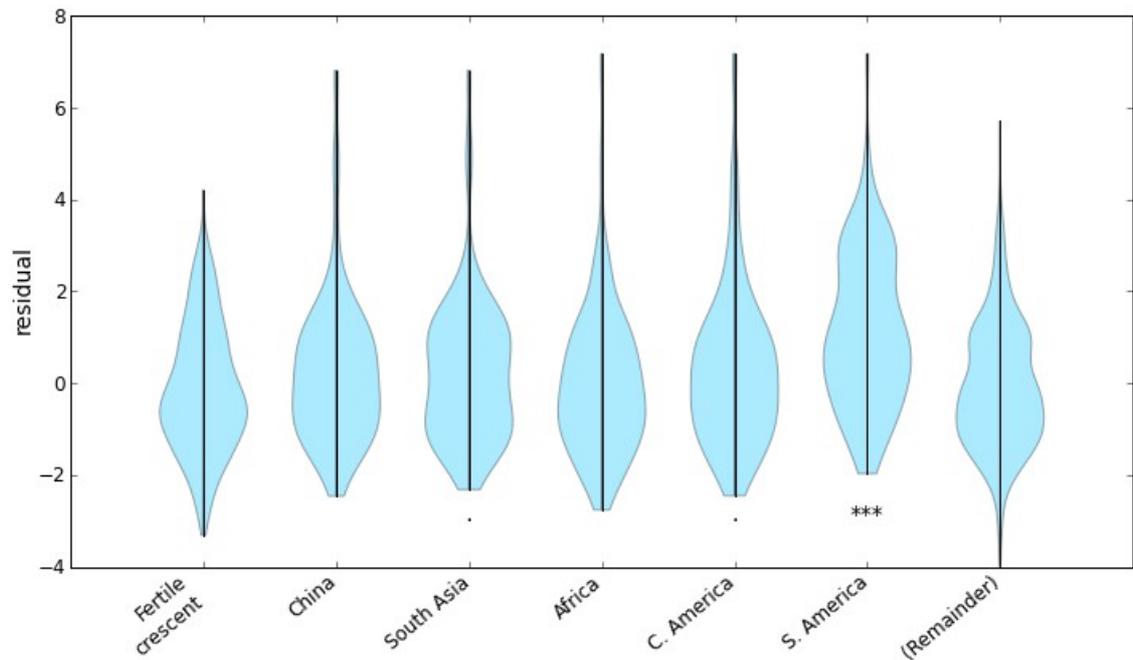
Figure 2.11: Distributions of the residuals from Figure 2.10 in different centres of origin of agriculture and the remainder of species. Asterisks and dots represent standard significance levels for the difference with the remainder of species, using Dunnett's test.

2.3.5 Geography and large-seeded species

The abundance of especially large-seeded species, which are proposed as the key factor for the development of agriculture (Blumler 1992a), does not necessarily correspond to the general distribution of seed masses. There may be more large-seeded species in a region simply because there are more species present there in total, or because of a small group of atypically large seeded species.
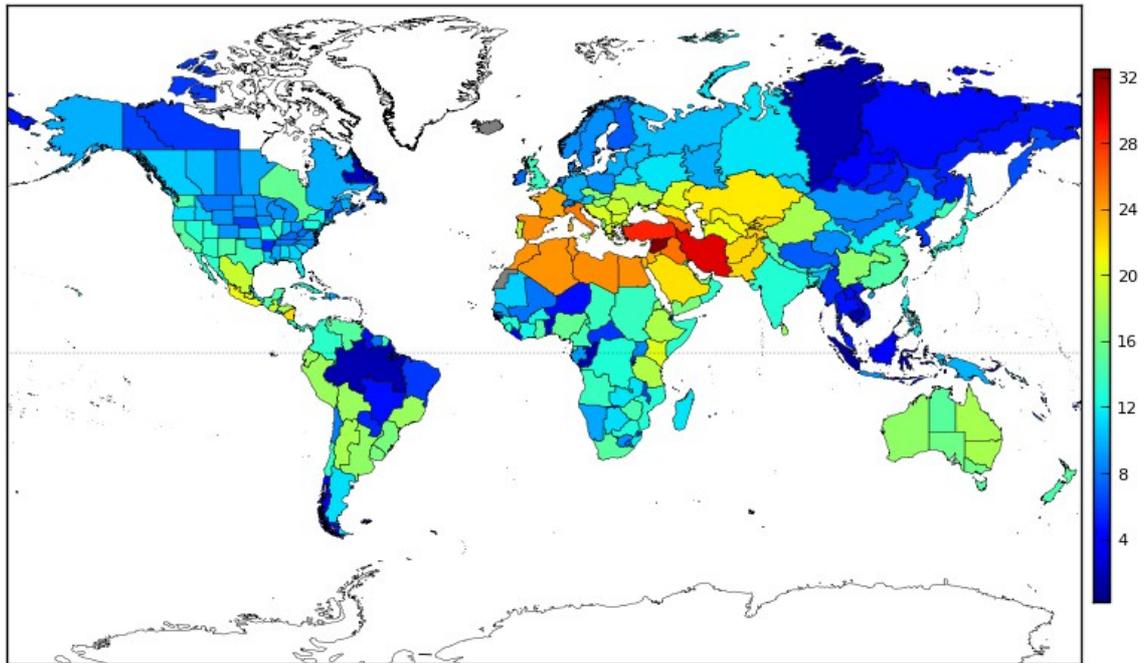
Figure 2.12: Mean of the five largest seed masses of annual grass species in each region in milligrams.

Visually, mapping the mean of the five largest seed masses for each region shows a clear hotspot in Western Asia for annual grasses (Figure 2.12). In legumes, the pattern is much less clear. Large-seeded herbaceous legumes (Figure 2.13) appear especially in Africa, Central America and Northern South America, while large-seeded annual species (Figure 2.14) are most prominent in Bolivia and East Africa. In both groups, a lesser region of large-seededness extends from the Indian subcontinent through Southeast Asia.

We might expect that extreme values would be captured more often in better sampled regions. To control for this potential bias, these values were compared with the proportion of species in an area for which mass data are available, obtaining weak correlations (F=3.75, p=0.0537 for annual grasses; F=8.92, p=0.0031 for annual legumes, F=4.22, p=0.0408 for herbaceous legumes). However, the patterns

with the residuals from these models are similar to those with the original data, so the patterns seen are unlikely to be the results of sampling bias.

Likewise, larger regions would on average be expected to contain more species, and thus more large-seeded species, without all those species necessarily being available to early farmers in a given part of such a region. In annual grasses and legumes, the values are correlated with the regions' areas (F=26.2, p=5.30×10^{-7} for grasses, F=5.91, p=0.0157 for legumes; using the square root of area), but again, mapping the residuals produces the same pattern, indicating that the patterns are not driven by differences in area.
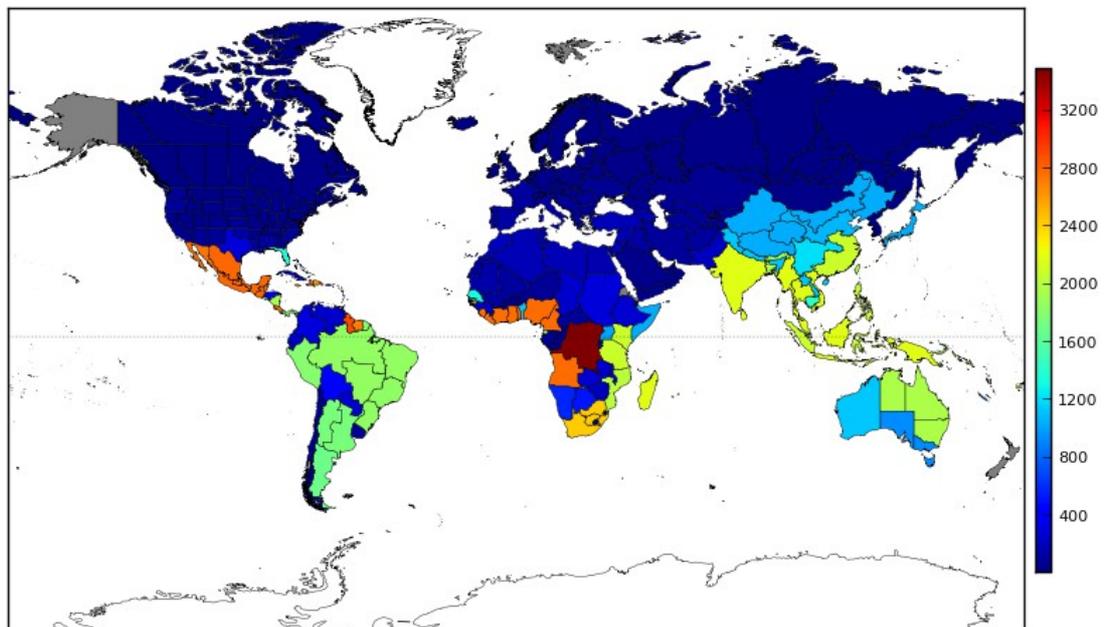
Figure 2.13: Mean of the five largest seed masses of herbaceous legume species in each region in milligrams.
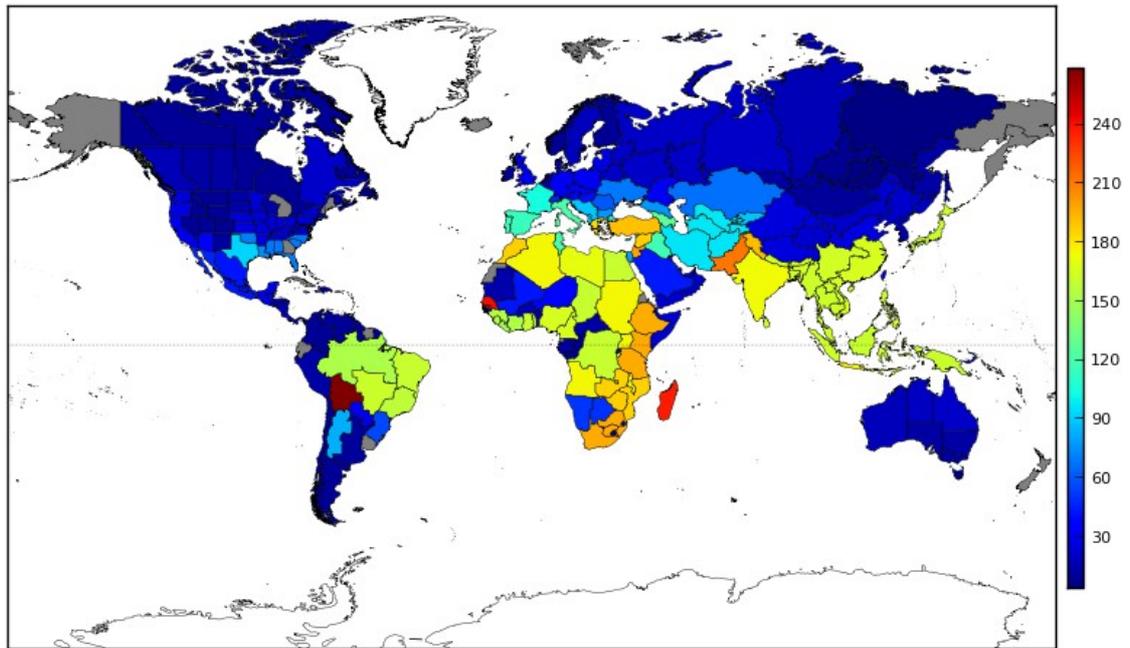
Figure 2.14: Mean of the five largest seed masses in annual legume species in each region in milligrams.

I have used the same data to calculate the number of annual legume species in each region with a seed mass above 10 mg—an arbitrary threshold borrowed from Blumler (1992a), who used it for grasses. 10 mg is quite small relative to legume crops: lentils weigh some 30–50 mg. This view gives quite a different result: annual legumes in particular show an abundance of such large-seeded forms in Western Asia and around the Mediterranean (Figure 2.15).
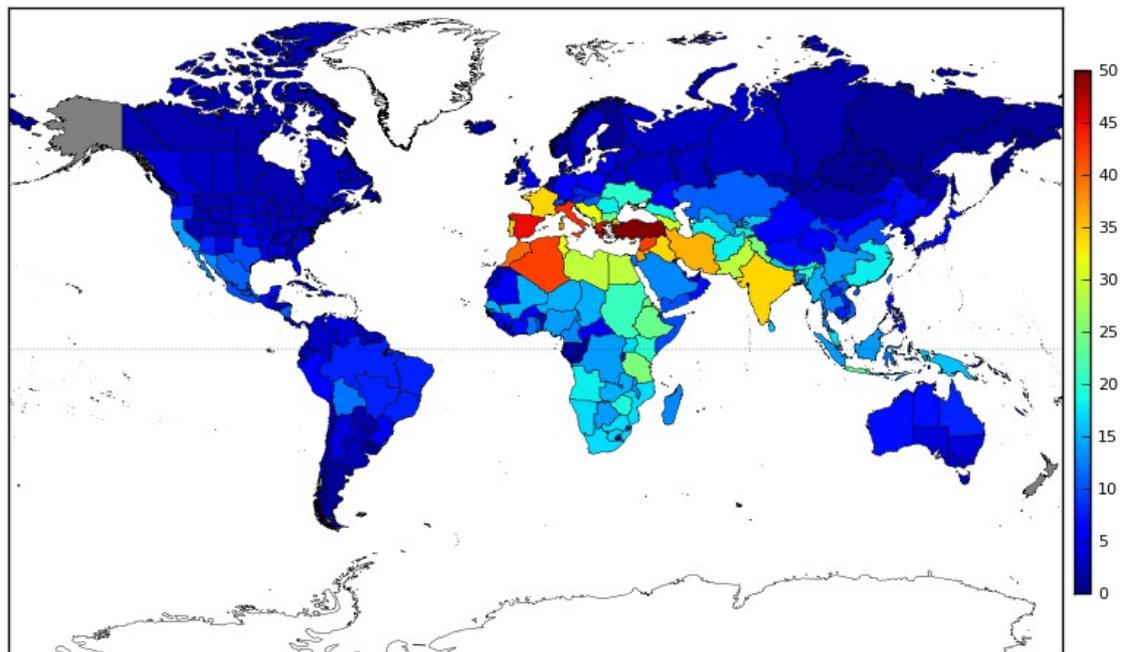
Figure 2.15: The number of wild annual legume species in each region with a seed mass above 10 mg.

The mean of the five largest seed masses in each region was strongly correlated with the total number of species in the region—including species without seed mass data—in annual grasses (F=263, p=2.76×10^{-44}; Figure 2.16) and in annual legumes (F=374, p=9.50×10^{-53}, calculated on log-log axes; Figure 2.17). The correlation was weaker in herbaceous legumes (F=13.0, p=3.62×10^{-4}), where the distribution was sharply bimodal (Figure 2.18).

This relationship with species number explains part, but not all, of the pattern: mapping the residuals from this relationship for annual grasses (Figure 2.19) still shows a concentration of large-seeded species in Western Asia.

The five largest seed masses are also correlated with the median seed mass in each region, in annual grasses (F=73.0, p=6.25×10^{-16}; medians log-transformed), annual legumes (F=50.8, p=1.82×10^{-11};

both axes log-transformed) and herbaceous legumes (F=123, p=3.48×10^{-24}; both axes log-transformed).
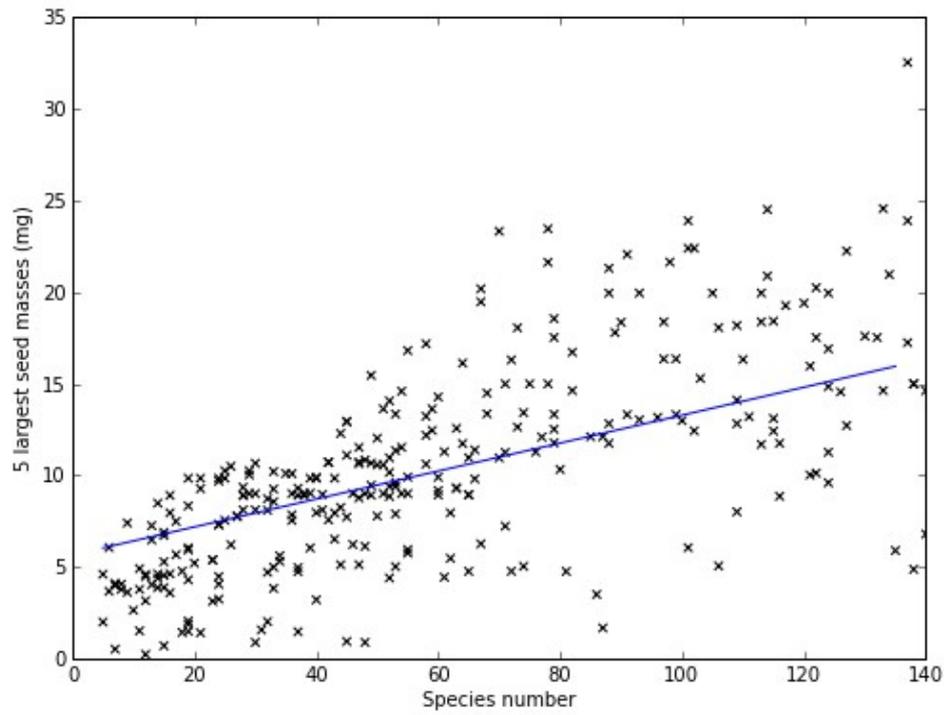
Figure 2.16: Correlation between the mass of the five largest wild annual grass species in a region and the number of wild annual grass species in that region.
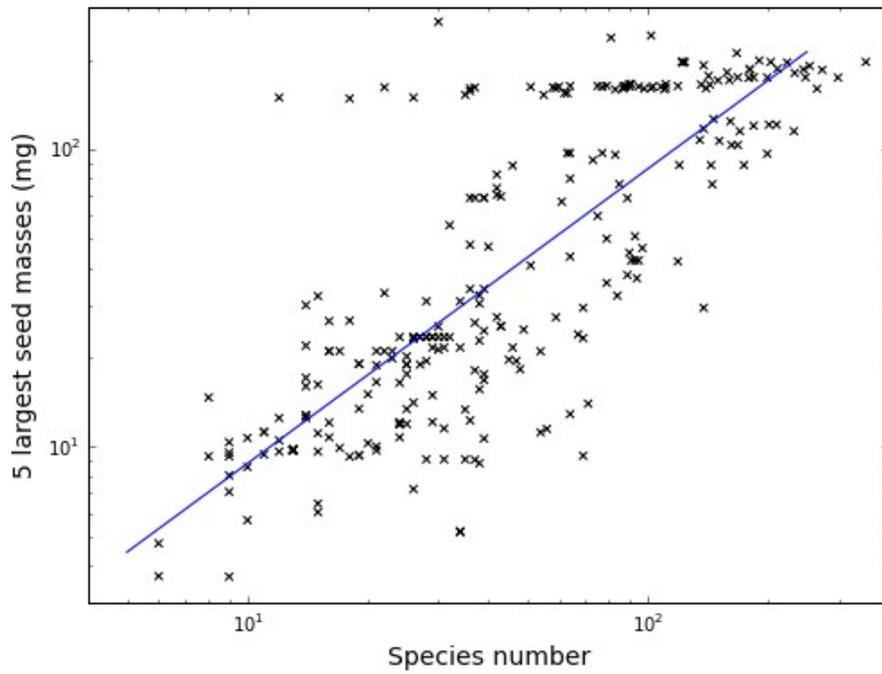
Figure 2.17: Correlation between the mass of the five largest wild annual legume species in a region and the number of wild annual legumes species in that region, on log-log axes.
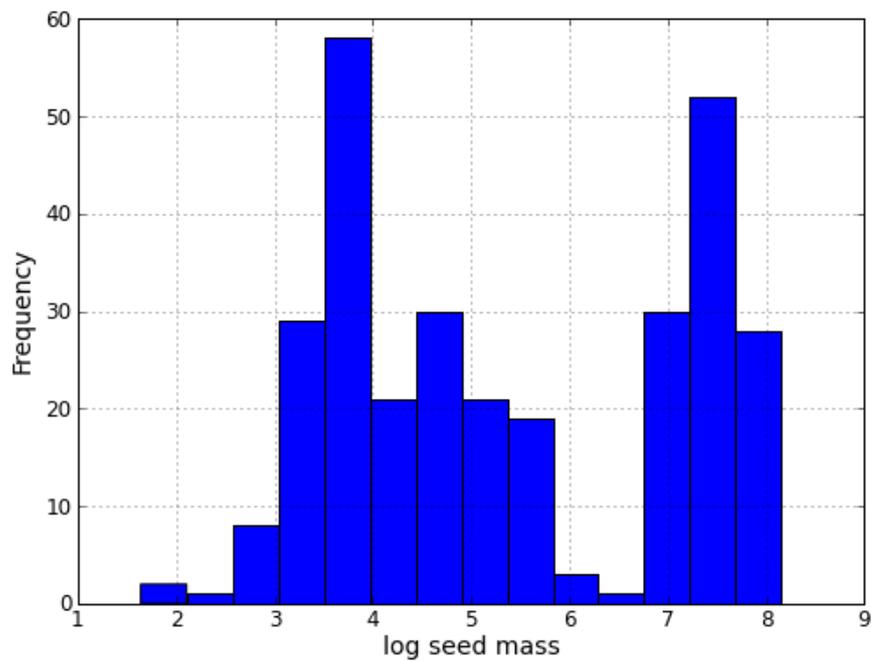
Figure 2.18: Distribution of the mean of the five largest seed masses of herbaceous legumes in each region.
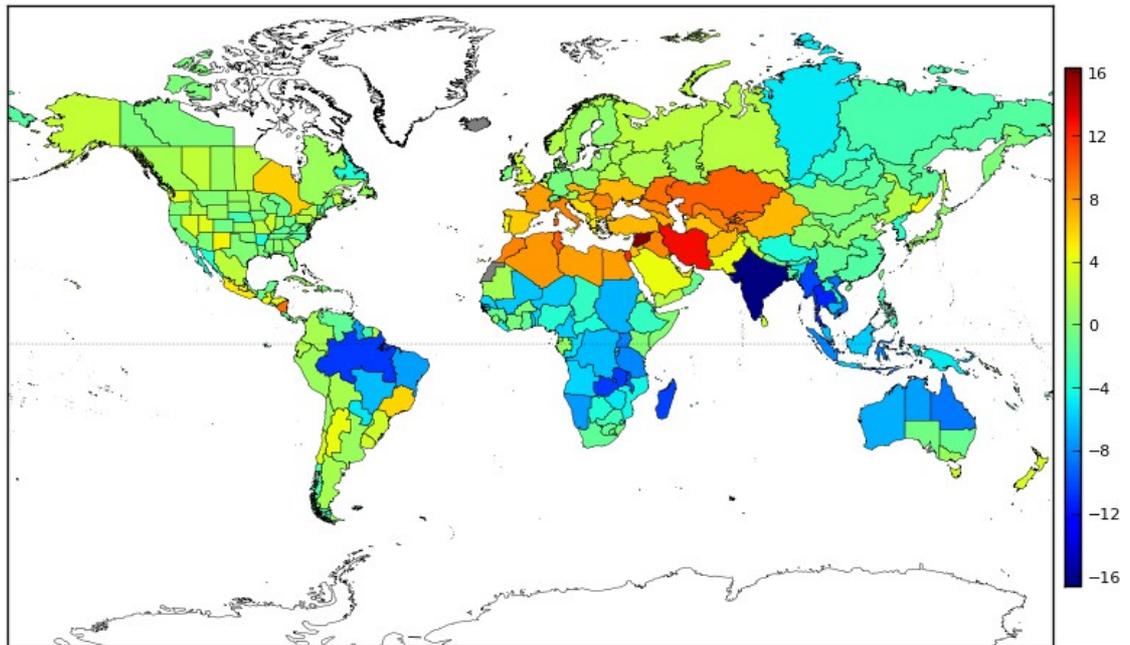
Figure 2.19: The residuals from the relationship between the seed mass of the five largest seeded species and the total number of species in a region, in annual grasses.

2.4 Discussion

Geographically, the concentration of large-seeded species in Southwestern Asia which Blumler (1992a) found is confirmed with a broader dataset of annual grass species, which was not based on selecting only species believed to be large-seeded for measurement. A similar concentration is also found in legume species, although this is less clear cut than for grasses.

However, the other centres of origin of major grass crops (China and Central America) do not have especially large-seeded species. There are a number of possible explanations for this. Data in those other regions is less complete (Figure 2.20), but we might expect large-seeded species to be better represented: for instance, the dataset includes five wild annual taxa from the small genus *Zea* L.,

including two wild subspecies of *Zea mays* L. Sampling bias is not an

adequate explanation for the patterns seen, as described in section

2.3.5.
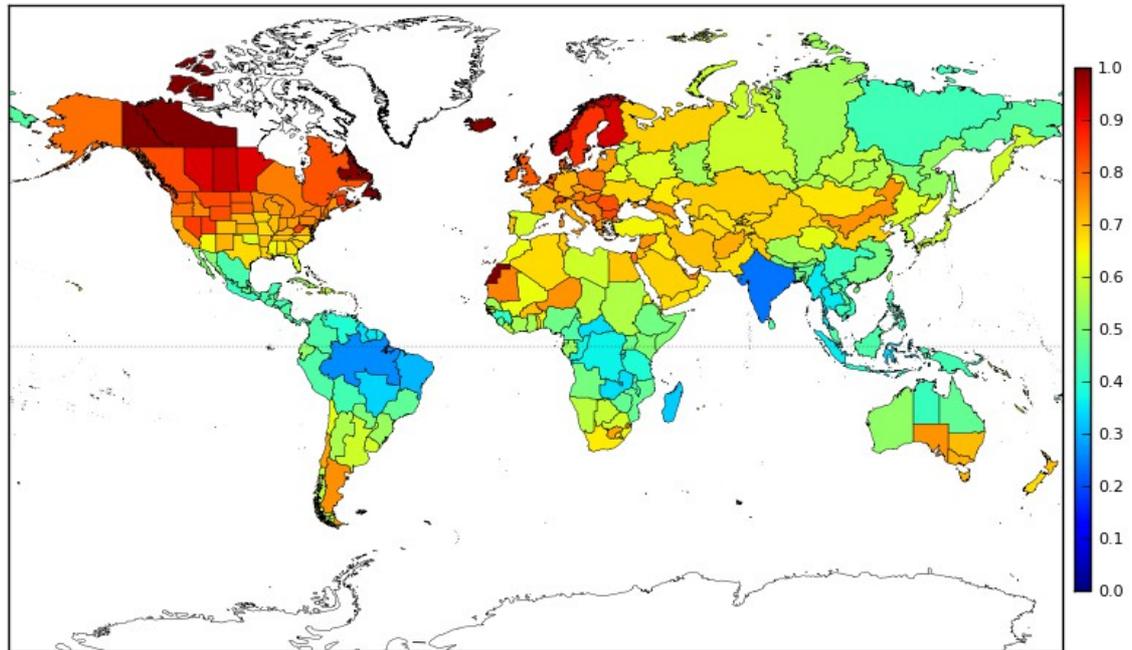
Figure 2.20: The proportion of wild annual grass species recorded in each region for which seed mass data is available.

It may be that large-seeded species are present, but there are only

one or two in each region, so they don't show up in these analyses.

Blumler (1992a) takes this position by focussing on individual

large-seeded species which were domesticated in each region, but the

clear patterns in Figure 2.12 and Figure 2.15 argue that an

explanation constructed from patterns observed in the Fertile

Crescent may not explain the location of other centres of origin of

agriculture, even though some of the species chosen for cultivation in

those centres had similar properties. There are also large seeded and

edible wild species in other regions, such as wild rice (*Zizania* spp.) in

North America (Eule-Nashoba 2010), and some extremely

large-seeded legumes (discussed below), so the availability of such species alone does not explain why agriculture began where it did.

Finally, the focus on the beginnings of agriculture in Western Asia may have given us a distorted view of the importance of large-seeded grass species. Small seeded grasses, often lumped together under the name 'millet', have been important in diverse regions (Weber & Fuller 2008). In Northern China, broomcorn millet, *Panicum miliaceum* L. was probably the first species cultivated (Lu *et al.* 2009). Small seeded grasses were also domesticated in India (Kimata, Ashok & Seetharam 2000; Fuller 2006) and Africa (Portères 1976; Kimber 2000; Blench 2012). In the New World, *Bromus mango* É. Desv., *Setaria parviflora* (Poir.) Kerguélen and *Panicum hirticaule* J. Presl. may have been locally important domesticates (Nabhan & de Wet 1984; Hammer 2004 sec. 5.5.1.2; Austin 2006), and small seeds from various wild grasses were certainly collected for human consumption (Doebley 1984; Austin 2006).

Among the legumes, the distribution of large seeded species is more complex. Species in the genus *Arachis* L., the wild relatives of peanut (*Arachis hypogaea* L.) make a hotspot of large-seeded legumes in South America. *Arachis* comprises some 70 species, including both annuals and perennials. Besides the globally important *A. hypogaea*, the indigenous inhabitants of Matto Grosso, Brazil, also cultivate the perennial species *A. villosulicarpa* Hoehne (Galgaro, Valls & Lopes 1997).

A second group of large-seeded legumes is defined functionally, not taxonomically, being represented in all three subfamilies. Drift seeds or sea beans are adapted for long-distance waterborne dispersal, and consequently have extensive distributions, generally in the tropics. These are responsible for the peak of regions with very large-seeded legumes in Figure 2.18, although that arguably overstates their global significance, as many of those regions are small tropical islands. While none of these species are familiar crops, it is not the case that they are universally inedible or impossible to cultivate. Sea beans are used medicinally, such as the grey nickernut, *Caesalpinia bonduc* (L.) Roxb. (Sharma, Dwivedi & Swarup 1997), collected for food, such as Gila bean, *Entada phaseoloides* (L.) Merr. (Siddhuraju, Becker & Makkar 2002), and at least one, *Mucuna pruriens* (L.) DC., is a minor domesticate (Siddhuraju & Becker 2005). The phenomenon is not limited to the Fabaceae, and intriguingly, some close relatives of sweet potato (*Ipomoea batatas* (L.) Lam., Convulvulaceae) produce drift seeds (Hemsley 1892).

Other large-seeded legume species mapped include Marama bean, *Tylosema esculentum* (Burch.) A.Schreib., an herbaceous perennial which is collected as food in southern Africa. Its self incompatibility (Hartley, Tshamekeng & Thomas 2002) may be one factor that has kept it from being domesticated. The inclusion of the Madagascan species *Neoapaloxylon tuberosum* (R. Vig.) Rauschert may be erroneous—sources list it as annual or perennial, citing Du Puy et al. (2002), but its two congeners are both listed as perennial, and little

information is available.

The results reveal a pattern which appears to be similar in grass and legume crops. In both families, crop progenitors have unusually large seeds, relative to their whole taxonomic group, and relative to the subsets of species from the same regions. This is driven by the domestication of species from especially large-seeded genera, while among the close relatives within those genera, it is not consistently large-seeded species which are crop progenitors. In some senses, this is surprising: given the markedly larger seeds of legumes in general (Figure 2.21), we might expect seed mass to play a somewhat different role in the two groups. For instance, if there were a minimum seed size for the 'filtering' (see the note on terminology in section 2.1), perhaps related to the dexterity of human hands, more legume species would exceed that threshold. The fact that the pattern is similar suggests that the preference for larger seeds continues across some range of sizes.
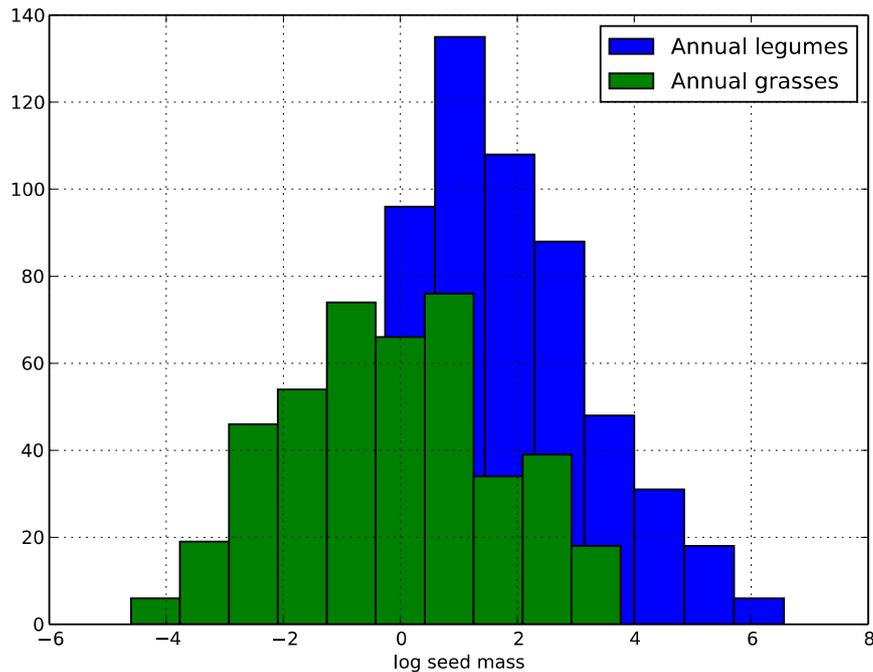
Figure 2.21: Seed mass distributions for annual legumes and grasses, both excluding domesticated species.

Why might crop progenitors not have larger seeds than other species in their genera? It may simply be that the largest seeded species were not always available in the regions where agriculture began. If this were the case, it would be further evidence against the theory that the regions where agriculture began were determined largely by the distribution of large-seeded species.

Another possibility is that nutritional factors were more important in the 'filtering' of cultivated species. This could be especially important for legume crops: cereal crops, which are primary caloric staples, mostly have low levels of the essential amino acid lysine, but this can be complemented with legumes, which have a different amino acid profile, typically low in the sulphur-containing amino acids cysteine and methionine (Iqbal *et al.* 2006). The filtering may have favoured

61

species with an optimal nutritional content over close relatives with larger seeds. In a similar vein, Kerem et al. (2007) postulated that chickpea may have undergone selection for increased levels of tryptophan, another essential amino acid. This relies on there being a mechanism by which an individual deficient in a particular amino acid has a heightened preference for food containing that amino acid, to counter the deficiency. There is evidence for such a mechanism in rats (Markison *et al.* 2000) and piglets (Ettle & Roth 2004). There is also archaeobotanical evidence, in the form of large quantities of wild mustard seeds found stored in a house at Çatalhöyük, that some people in the Neolithic could afford to collect food for flavour, not just for survival value (Twiss 2011).

However, legumes also contain an array of toxic substances and factors which inhibit digestion (Gupta 1987). Pichare and Kachole (1996) showed that there is considerable variation in protease inhibitors among wild members of the subtribe Cajaninae, close relatives of the pigeonpea, *Cajanus cajan* (L.) Millsp. It is plausible, therefore, that individual species differ markedly in seed palatability or processing requirements, which may have been more important filtering criteria than seed size.

In both annual grasses and annual legumes, the availability of especially large-seeded species is related to both the median seed mass and the number of species in a region, but the stronger correlation is with the number of species.

A promising avenue for further work would be to explore these patterns in a phylogenetic framework. Crop species appear to be phylogenetically clustered; for instance, the tribe Triticeae includes the wheats, barley and rye, while in the legumes, cowpea, mung bean, urd bean, moth bean and rice bean are all domesticated members of the genus *Vigna* Savi. By mapping the species here onto a phylogeny, such as the grass phylogeny constructed by Edwards and Smith (2010), the seed mass data could be analysed using phylogenetically independent contrasts.

3 Unconscious selection on seed size in vegetable crops

3.1 Introduction

It is widely observed that domestic seed crops—particularly cereals and grain legumes—have larger seeds than their wild progenitors (Harlan *et al.* 1973; Smith 2006; Purugganan & Fuller 2009; Lee *et al.* 2011). There is interest in unconscious selection mechanisms for this increase in seed size. A current hypothesis is that deeper burial in cultivated fields favoured larger seeds better able to emerge from depth (Fuller 2007); this is evaluated in more detail in Chapter 4.

However, conscious selection could easily have played a role as well: early cultivators may have preferred larger seeds which were easier to handle or gave a greater return on harvesting effort. In some crop species, larger seeds also increase total yield (Shomura *et al.* 2008), although in other cases the relationship is reversed (White & González 1990). Many people with whom I have discussed the subject express surprise at our interest in unconscious selection, and voice an assumption that conscious selection is a sufficient explanation for increased seed size. I therefore looked for a way to distinguish the effects of unconscious selection from those of conscious selection.

Many crops are grown for vegetative rather than reproductive parts of the plant. Underground storage organs are most often the part of interest: this category includes tubers as in potato (*Solanum tuberosum* L.) and sweet potato (*Ipomoea batatas* (L.) Lam.), corms as in taro (*Colocasia esculenta* (L.) Schott) and several other

cultivated aroids, and enlarged roots as in carrot (*Daucus carota* L.) and beet (*Beta vulgaris* L.). Root crops were particularly important in the development of agriculture in South America: besides potato and sweet potato, this was the origin of cassava (*Manihot esculenta* Crantz) and an array of Andean crops such as ullucu (*Ullucus tuberosus* Caldas) and oca (*Oxalis tuberosa* Molina). Other harvested vegetative plant parts around the world include:

- Leaves, e.g. lettuce (*Lactuca sativa* L.), cabbage (*Brassica oleracea* var. *capitata* L.)

- Petioles, e.g. rhubarb (*Rheum rhabarbarum* L.) and celery (*Apium graveolens* L.)

- Shoots, e.g. bamboo shoots (various species)

- Stems, e.g. Manchurian wild rice (*Zizania latifolia* (Griseb.) Turcz. ex Stapf) and sugar cane (*Saccharum* spp.)

I hypothesise that, if seed size is subject to unconscious selection under cultivation, we will see larger seeds in some of these crops where neither the seed nor the fruit is of agronomic interest. We can further divide these crops into two categories: those which are grown from seed, and those which are vegetatively propagated, including many tuber crops. In the former category, selection would have regularly acted on seeds at sowing time. In the latter category, selection likely had less opportunity to act on seeds, although sexual reproduction and growth from seed must have occurred at intervals to generate genetic diversity.

65

To establish a baseline for the magnitude and generality of increases in seed mass, I also examined seed masses from wild and landrace accessions of important seed crops in the grass and legume families.

3.2 Materials and Methods

I looked for data for many species of crops thought to have been domesticated in antiquity (Ugent, Pozorski & Pozorski 1982; Zohary & Hopf 2000; Piperno *et al.* 2000, 2009; Lebot 2009 p. 91). For seed crops, I used a range of annual grasses and legumes domesticated in different parts of the world. For vegetable crops, I looked for both species which are typically grown from seed, and species which are vegetatively propagated. Fruit crops were not included in this, although we may expect that selection for larger fruits would also have increased seed size.

The range of crops I could compare was limited by the species for which seed mass data is available, especially in vegetatively propagated crops where seeds are used more rarely. All three vegetatively propagated species tested originate from South America: my efforts to obtain true seed of taro, *Colocasia esculenta* (L.) Schott, which was domesticated in the Asia-Pacific region, were unfortunately unsuccessful.

For each species of interest, I initially used a custom script to download data from the USDA GRIN germplasm database (for which I obtained permission). Where there were multiple weight measurements for one accession, I found the arithmetic mean, so that

each datum represents a single accession. Where there were insufficient data to allow a comparison, I ordered seed accessions and weighed them. Seed crops were ordered from GRIN, except for mung bean (*Vigna radiata* (L.) R. Wilczek), which was ordered from the Australian AusPGRIS database. Data and seeds of root crops were sourced from:

- The USDA GRIN/NPGS database (http://www.ars-grin.gov/npgs/)

- IPK Gatersleben (http://gbis.ipk-gatersleben.de)

- The International Potato Centre (CIP) in Peru (http://cipotato.org/)

- Cassava seed masses collected by Pujol et al. (2005b)

- Garden seed was ordered from various UK suppliers: Nicky's Nursery, D.T. Brown, Thompson & Morgan, Unwins, Marshalls, vegetableseeds.net, Dobies of Devon

The sources used for each species are detailed in Table 3.1.

I was able to source sufficient true seed or true seed weight data for seven root crop species to compare wild and domestic forms (table 1). Of these, four are crops typically grown from seed, and three are vegetatively propagated tuber crops.

The domestic forms I used in the comparison were landrace accessions, in order to exclude any effect of modern commercial breeding. For the crops grown from seed, I also compared these landrace accessions with cultivars from commercial breeding (seed

sold for gardeners, and accessions listed as 'cultivars'), to see

whether any pattern was continued in modern selection.

Crop	Improvement	Data sources
Beet	Wild	GRIN (d)
	Landrace	GRIN (d), IPK (s)
	Modern	Garden seed (s)
Lettuce	Wild	GRIN (d)
	Landrace	GRIN (d)
	Modern	Garden seed (s)
Carrot	Wild	GRIN (d)
	Landrace	GRIN (d)
	Modern	GRIN (d), Garden seed (s)
Parsnip	Wild	GRIN (d)
	Landrace	GRIN (d), IPK (s)
	Modern	GRIN (d), Garden seed (s)
Potato	Wild	IPK (s)
	Landrace	IPK (s)
Sweet potato	Wild	GRIN (s), CIP (s)
	Landrace	GRIN (s), CIP (s)
Cassava	Wild	Pujol (d)
	Landrace	Pujol (d)

Table 3.1: Seed and data sources for each crop group. Sources marked d supplied seed mass data, those marked s supplied seed samples which were weighed in Sheffield. Full details of suppliers above.

For all crops, only wild and landrace seed that were collected from the

broad region in which the crop originated were included, to limit the

inclusion of feral accessions of varieties developed by modern

breeding.

Both the seeds and the seed mass data available for beet are actually

seed capsules, each containing one or two seeds in a tough, woody

structure. While the capsule masses were analysed, we also dissected a subset of the capsules we had in Sheffield, by soaking the capsules in water for half an hour to soften them, and weighed 5–10 true seeds per accession together. These measurements of true seeds were analysed separately from the masses of seed capsules.

Wild and cultivated potatoes (*Solanum* sect. *Petota*) include an array of introgressing species with a range of ploidy levels, and treatments of their taxonomy differ substantially (Ovchinnikova *et al.* 2011). The landrace accessions I considered included representatives of the subspecies cultivated worldwide, *Solanum tuberosum* subsp. *tuberosum* L., as well as Andigena potatoes, *S. tuberosum* subsp. *andigena* (Juz. & Bukasov) Hawkes, and two diploid cultivated species, *S. stenotomum* and *S. phureja*. The wild accessions included six species from the *Solanum brevicaule* complex, from which cultivated potatoes are thought to have originated (Spooner *et al.* 2005): *S. brevicaule* Bitter, *S. bukasovii* Juz. ex Rybin, *S. canasense* Hawkes, *S. candolleanum* P. Berthault, *S. gourlayi* Hawkes and *S. spegazzinii* Bitter. We also included four accessions of the wild species *S. acaule* Bitter, although it is probably more distantly related to cultivated potatoes, as a control for an effect of polyploidy on seed size: like the predominant cultivated potatoes and Andigena potatoes, *S. acaule* is a tetraploid (Iwanaga, Freyre & Watanabe 1991).

Seed masses typically follow a log-normal distribution (Leishman *et al.* 1995; Khazaei, Jafari & Noorolah 2008). In the species that fit this pattern, log-transformed data for each wild/domestic species pair

were compared using a t-test. Where data did not appear to fit any particular distribution, the non-parametric Mann-Whitney U test was used instead.

3.3 Results

3.3.1 Seed Crops

Seeds in landrace forms of these crops are between 14% heavier and 15 times heavier than seeds from their respective wild progenitors (Figure 3.1 and Table 3.2).

Species	df	t	p	Increase factor
Barley	96	-9.75	5.11e-16	1.66
Einkorn wheat	46	-5.68	8.78e-07	2.27
Emmer wheat	55	-2.2	0.0322	1.14
Foxtail millet	98	-6.63	1.82e-09	1.58
Maize	85	-38	4.11e-55	15.2
Oats	178	-12.3	1.12e-25	2.26
Pearl millet	11746	-14.7	1.74e-48	4.73
Rice	32	-8.94	3.24e-10	1.91
Rye	146	-10.5	1.16e-19	2.37
Sorghum	21034	-13.8	3.42e-43	2.61
Chickpea	32	-5.74	2.28e-06	1.91
Common bean	498	-24	3.91e-85	3.61
Cowpea	88	-26.7	5.69e-44	6.82
Lentil	130	-23.2	4.27e-48	4.06
Lima bean	302	-8.88	6.11e-17	3.23
Mung bean	17	-3.54	0.00252	2.23
Pea	52	-2.08	0.0427	1.51
Peanut	140	-8.83	3.79e-15	2.36
Soybean	44	-21.2	1.1e-24	7.57

Table 3.2: Differences in seed mass between wild and landrace seed crops. Log-transformed data were compared using t-tests.
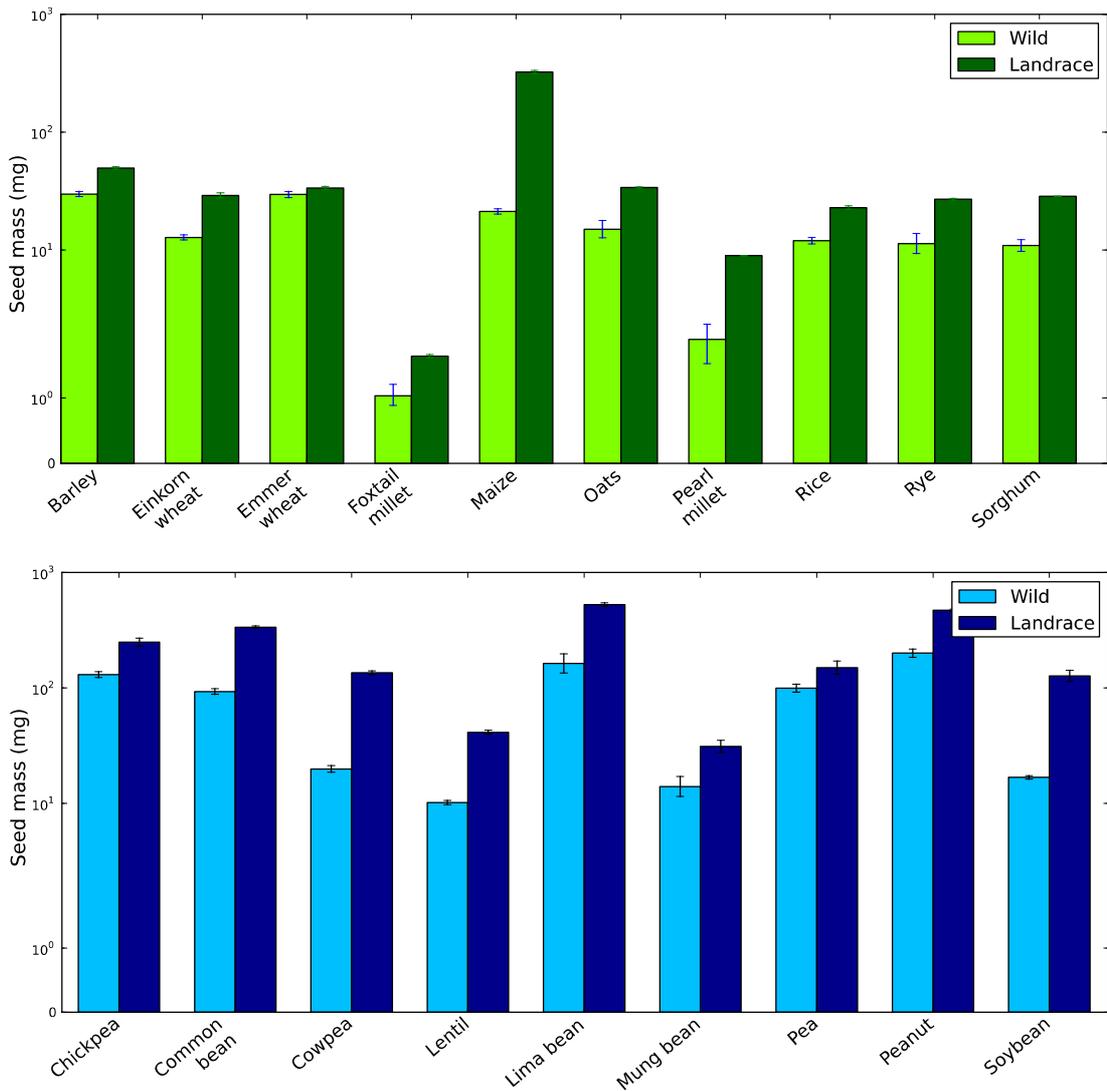
Figure 3.1: Seed masses in wild and landrace grasses (top) and legumes (bottom), plotted on a log scale. Each bar is the geometric mean of at least five accessions, and the error bars indicate the standard error of the mean.

3.3.2 Vegetable crops

In all seven of the species tested, landrace seeds were significantly

larger than their wild counterparts (Table 3.3 and Figure 3.2).

Crop	Test used	n	Test statistic	p value	Increase factor
Beet (capsules)	t-test	195, 109	-3.46	6.21×10^{-4}	1.19
Beet (seeds)	M-W U	10, 14	33	0.0162	1.33
Lettuce	t-test	175, 124	-14.8	1.97×10^{-37}	1.75
Carrot	t-test	121, 191	-17.1	7.04×10^{-47}	1.60
Parsnip	M-W U	16, 12	19	1.90×10^{-4}	1.38
Potato	M-W U	17, 13	45	0.0033	1.21
Sweet potato	M-W U	13, 4	2	0.0039	2.17
Cassava	M-W U	79, 231	7426	0.0067	1.08

Table 3.3: Comparisons of wild and landrace seed accessions. The test statistic is the value of the t or U statistic. The increase factor is the ratio of geometric means (where the t-test was used) or medians (where the Mann-Whitney U test was used) of the landrace and wild accessions.
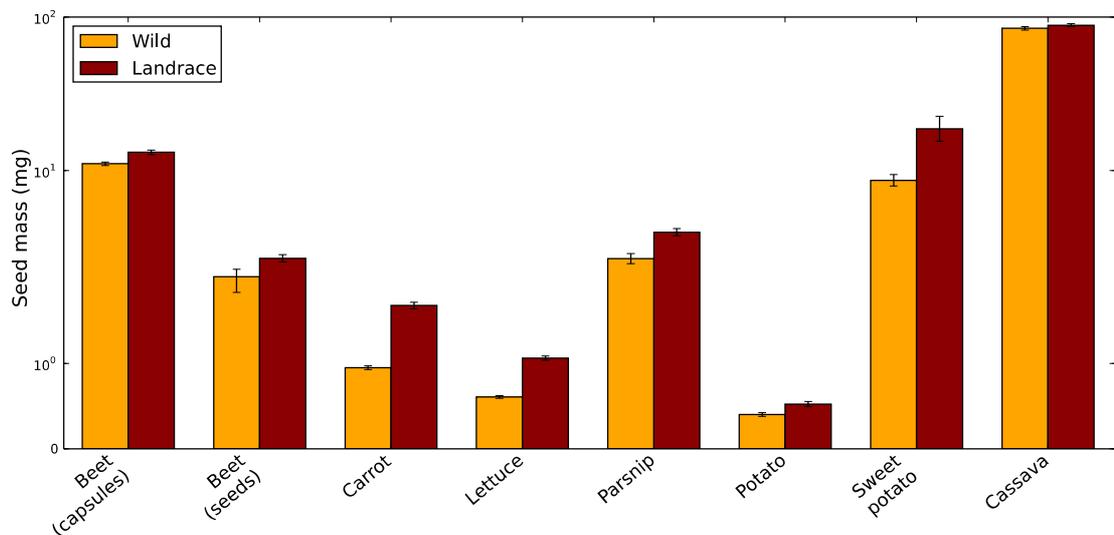
Figure 3.2: Seed masses in wild and landrace vegetable crops, on a logarithmic scale. Bars show geometric means, and error bars show the standard error of the mean.

The values given in table 2 for the difference in potato seed mass are calculated from all the values available, including both the *Solanum brevicaule* complex and *S. acaule*. The difference remains significant if we exclude *S. acaule*, thought to be more distantly related

(p=0.0038, Mann-Whitney U test), and if we compare only the likely tetraploids, *S. acaule* and *S. tuberosum* (p=0.036, Mann-Whitney U test).

I also attempted to ascertain the provenance of the seed, to check whether the observed differences could be merely due to an environmental effect. Environmental effects on seed size are well known for many species (Fenner 1992), including effects on true potato seed (Pallais *et al.* 1987; Roy, Nishizawa & Ali 2007). For instance, plants grown in more fertile soil might produce larger seeds, and if the landrace seeds measured here were collected from plants grown in more fertile soils, the observed difference in phenotype may not reflect any underlying genotypic effect.

- All seed from IPK Gatersleben was regenerated, either in common field conditions at Gatersleben, or for potatoes in greenhouse conditions in Groß Lüsewitz.

- The sweet potato accessions ordered from GRIN were all seed lots which had been regenerated in greenhouse conditions in Georgia, USA.

- The carrot and parsnip seed data obtained from GRIN were a mixture of 'original lots' directly obtained by collection and 'increase lots' from regeneration in Iowa, USA. I received additional data which distinguished seed lots by origin, and showed that seed from increase lots is on average slightly larger (p=4.4×10^{-41}, 0.0058 for carrot, parsnip; paired t-test)

and has smaller variance ($p=1.74\times10^{-5}$, 0.0017; Bartlett's test on log-transformed data). However, the difference between wild and landrace seeds was still found using only the data from increase lots, both for carrot ($p=7.92\times10^{-29}$, t-test on log-transformed data) and parsnip ($p=0.023$, Mann-Whitney U test).

- The beet and lettuce seed data obtained from GRIN were likewise a mixture of directly collected and increased lots. I received additional data distinguishing these. The beet seed capsules from increase lots were on average smaller ($t=11.9$, $p=3.20\times10^{-29}$; paired t-test) and less variable in size ($T=64.3$, $p=1.09\times10^{-15}$; Bartlett's test). In lettuce, seeds did not significantly change in size ($t=1.06$, $p=0.28$; paired t-test), but were considerably less variable, due to the presence of a few implausible outliers in the seed masses recorded on collection ($T=911$, $p=4.19\times10^{-200}$; Bartlett's test). The data was insufficient to confirm a difference between wild and landrace accessions from the masses of increased seed.

- The wild and landrace cassava seed were collected in separate sites: wild seed was from a site in Rondonia, Brazil, while landrace seed was from a separate site in Rondonia, and two further sites in French Guiana (Pujol *et al.* 2005b). Some additional seed mass data was obtained from EMBRAPA, Brazil, in which sample the wild and domestic accessions did not differ

(U=90.5, p=0.440), but the statistical power was limited, as this included only five domestic accessions. The wild accessions from EMBRAPA had significantly larger seeds than both the wild (U=138, p=1.07×10^{-15}) and landrace (U=837.5, p=6.73×10^{-16}) accessions collected by Pujol et al., so data from these sources were not pooled.

The comparisons of modern cultivars with landrace accessions were less consistent. In parsnip, the seeds of modern varieties were larger than those of landrace varieties at a level approaching significance (U=200, p=0.053, Mann-Whitney U test). In carrot, the modern varieties had significantly smaller seeds than the landraces (t=8.64, p=1.7×10^{-6}, t-test on log-transformed data). In beet and lettuce, there was no significant difference between landrace and modern accessions (t=0.85,1.26; p=0.42,0.20; t-test on log-transformed data).

3.4 Discussion

In the vegetable crops I have examined, landrace accessions consistently have larger seeds than their wild relatives, providing strong support for the hypothesis that unconscious selection acted on seed size during their domestication. Before we discuss these results in more detail, three minor caveats should be mentioned.

Firstly, some plants are grown for both vegetative and reproductive parts. Flax, *Linum usitatissimum* L., a plant often grown for its fibre, is also grown for its seed, linseed. The bulbs, leaves and seeds of fennel

(*Foeniculum vulgare* Mill.) are all used. It has even been suggested that maize (*Zea mays* L.) was initially domesticated not as a grain but for sugar extracted from its stem (Smalley & Blake 2003). We should not overlook the possibility that some species now regarded as vegetable crops may in earlier times have been seed crops, and subject to different selection pressures. Of the crops considered here, archaeological evidence suggests that the carrot in Europe and lettuce in Egypt may have been used millennia ago as seed crops, although it is not clear whether either use was important enough to much affect the crops' evolution (Andrews 1949; de Vries 1997; Iorizzo *et al.* 2013). The literature does not indicate that any of the other species in this study were ever grown for seed (Ugent *et al.* 1982; Smartt & Simmonds 1995; Zohary & Hopf 2000; Lebot 2009).

Secondly, the classification of accessions as wild or landrace may not be entirely reliable. In lettuce, for example, some accessions of the domestic taxon, *Lactuca sativa* L., were listed as wild material. These were excluded from the main analysis, but their seed masses were more similar to landrace accessions (t-test on log-transformed data: t=1.14, p=0.25) than to accessions of the wild taxon (*Lactuca serriola* L.; t=4.25, p=3.28×10^{-5}), suggesting that they represent feral plants from cultivated lineages. It is likely that some of the accessions included in the analysis as wild are also feral, or may have interbred with cultivated varieties. However, any such confusion would reduce the differences between wild and landrace accessions, and since I found a significant difference in all cases, it does not affect my

conclusions.

Thirdly, the accessions within each species are not completely independent: they will have somewhat similar seed sizes because accessions have not evolved much since they shared common ancestors. However, even if we disregard the statistical significance within each species, the different wild/domestic pairs are independent, as each pair has diverged more recently than it has separated from any of the other species. The probability of the seven pairs all differing in the same direction if there were no underlying difference is 2×0.5^7, or 0.0156.

The vegetatively propagated tuber crops are of particular interest, since selection could not act annually on seed collected and replanted by cultivators. There are two possible mechanisms for the evolutionary change in seed size. First, volunteers may frequently grow from seed and be incorporated in the crop gene pool, allowing selection to act directly on seed traits affecting natural dispersal, germination, seedling growth and survival in cultivated environments. This is on the border between what we call unconscious selection and natural selection. Ethnographic evidence for several vegetatively propagated crops supports this hypothesis. Cassava crops grown under swidden cultivation often include volunteer seedlings from dormant seeds in the soil, and traditional farmers include these volunteers in later vegetative propagation (Salick, Cellinese & Knapp 1997; Elias, Rival & McKey 2000; Pujol *et al.* 2002), influencing the

crop's evolution (Pujol, David & McKey 2005a; Pujol *et al.* 2005b). Some Andean farmers deliberately save and plant potato true seed, in part to eliminate viral diseases affecting the tubers (Malagamba & Monares 1988; Quiros *et al.* 1992), while traditional farming practices also make recruitment of volunteer seedlings likely, as fields typically contain a mixture of varieties rather than a clonal monoculture (Jackson, Hawkes & Rowe 1980; Johns & Keen 1986). There are a few reports of the preservation of volunteer seedlings of sweet potato in New Guinea, a secondary centre of diversity (Yen 1960; Bulmer 1965), of taro (*Colocasia esculenta* (L.) Schott; a vegetatively propagated crop domesticated in the Asia-Pacific region) in Vanuatu (Caillon & Lanouguère-Bruneau 2005; Caillon *et al.* 2006), and of ensete (*Ensete ventricosum* (Welw.) Cheesman; a multi-purpose African crop) in Ethiopia (Shigeta 1990). In contrast, there is no evidence that traditional cultivation of three minor Andean tuber crops (Oca, *Oxalis tuberosa* Molina; Ulluco, *Ullucus tuberosus* Caldas; Mashua, *Tropaeolum tuberosum* Ruiz & Pav.) recruits volunteer seedlings (Lempiäinen 1989; Malice 2009 p. 4), although the use of plants originating from seedlings in the past has been posited as an explanation for the observed diversity in all three species (Rousi *et al.* 1989; Malice 2009 p. 166; Malice *et al.* 2009).

If selection is able to act on volunteer seedlings, why might it favour larger seeds? Larger true seeds of potato germinate faster and more reliably than small seeds (Bhatt *et al.* 1989). Larger seeds of sweet potato are also more likely to germinate (Martin & Cabanillas 1966).

Likewise, Strauss et al. (1979) found that smaller seeds of taro were less likely to germinate at all, although the results of Tyagi et al. (2004) did not corroborate this. The broader ecological literature also indicates that plants from larger seeds have a greater growth rate at a given size (Turnbull *et al.* 2008). Fast germination and rapid early growth may be especially advantageous when in competition with a crop growing from tubers, which can contain many times more resources than do seeds.

The second possible mechanism is selection acting on other traits which are connected with seed size, either by common genetic mechanisms (pleiotropy) or allometric constraints. For instance, true seed weight in potatoes is genotypically correlated with tuber yield and harvest index, among other characteristics (Dayal, Upadhya & Chaturvedi 1984; Upadhya & Cabello 2000). Dignat et al (2013) found that leaf growth and ear growth in maize share part of their genetic control. Across species, a correlation between seed size and mature plant size is well known, but most explanations focus on the effects of seed size on survival and growth rate (Rees & Venable 2007; Falster, Moles & Westoby 2008), which would only apply when plants are regenerated from seed.

An allometric link between plant size and seed size has been suggested by some sources. Maximum seed size may be constrained by the need to support the weight of the seed on terminal branches (Aarssen 2005; Grubb, Coomes & Metcalfe 2005), but this is "likely to apply only when there is one seed per fruit… and when there is little

flexibility in number of fruits per inflorescence" (Grubb *et al.* 2005). Similarly, developmental constraints may prevent seed number from increasing in proportion to available resources (Vega *et al.* 2001), potentially pushing extra resources into larger seeds. These mechanisms may act in cassava, where seed capsules have a fixed three seeds per capsule (FAO), and sweet potato, where capsules are limited to at most four seeds, and normally hold one or two (Martin & Cabanillas 1966). In contrast, they are unlikely to be relevant in potato, where there are often over 100 seeds per berry, and the number varies within and between cultivars (Almekinders, Neuteboom & Struik 1995), or taro, with 15-2300 seeds per inflorescence from one cultivar (Tyagi *et al.* 2004).

The increase in seed size in these vegetable crops ranges from 1.15 times larger in parsnip to 2.17 times larger in sweet potato (Table 3.3). The grass and legume seed crops we examined tend to show a larger increase (Mann-Whitney U test on increase ratios: U=16, p=0.0019; Table 3.2). This suggests that the total selection pressure on seed size is greater in seed crops—either because of conscious selection for larger seeds, or additional forms of unconscious selection related to harvesting seed. Unconscious selection pressures on seedlings would also have more opportunity to act on crops grown from seed annually than on vegetatively propagated crops where plants grown from seed are only occasionally incorporated into the gene pool.

On the other hand, the increase in seed size stands in stark contrast

to the seedless varieties of vegetatively propagated fruit crops such as the banana. This, however, is a failure to form seeds, not necessarily a reduction in seed size (Simmonds 1962 p. 81). Conscious selection for seedless fruit seems the most likely explanation, as "wild banana fruits are full of hard, black seeds and are quite inedible." (Simmonds 1962 p. 76). Cultivators may have switched to vegetative propagation in such cases once seed production was too low to allow regular seed propagation.

3.5 Conclusion

I have provided evidence that seed size has increased during domestication in a number of vegetable crops where seed is not normally harvested, including some which are mainly vegetatively propagated. This may be due to unconscious or natural selection for larger seeds on the occasions when plants grow from seed and are integrated into the crop gene pool, or due to selection on traits linked to seed size genetically or allometrically, such as whole plant size. However, the change in seed size is less marked than in seed crops, indicating that the selection pressure on seed size is stronger in the latter group.

4 Did greater burial depth increase the seed size of domesticated legumes?

4.1 Introduction

Various possible explanations, which are not mutually exclusive, have been advanced to explain the increase in seed size observed during domestication. As described above (section 3.4), selection for plants which grow faster or reach a greater size at maturity could increase seed size, as could intense seedling competition. Conscious selection by early farmers for larger seeds may also have played a role.

One current hypothesis for the increase in seed size focuses on the processes of germination and seedling emergence. It holds that seeds were generally buried deeper by deliberate human planting, than by dispersal in a natural environment. The need to emerge from a greater depth in the soil would have selected for seeds with larger reserves (Harlan *et al.* 1973; Zohary 2004; Purugganan & Fuller 2009). In particular, the development of simple animal-pulled ploughs some time after domestication (Lal, Reicosky & Hanson 2007) has been proposed as an explanation for the late increase in seed size observed in grain legumes in the archaeological record (Fuller 2007). Archaeological evidence also suggests that early cultivation may have been in small-scale, intensively managed 'gardens', where seeds could have been sown by dibbling, dropping seeds into individually made holes, rather than broadcasting, scattering seed over a tilled plot (Bogaard 2005; Jones 2005). There is some archaeobotanical evidence supporting the burial hypothesis, primarily from Indian

Vigna species (mung and urd bean), where the seed size starts to increase approximately contemporaneously with the development of the ard plough (Fuller & Harvey 2006). Ecological experiments have demonstrated repeatedly that larger seeded species are able to emerge from greater depths than species with smaller seeds (Bond, Honig & Maze 1999; Benvenuti, MacChia & Miele 2001; Pearson *et al.* 2002). However, experiments that have compared seeds within species have produced more mixed results (Townsend 1992; Qiu & Mosjidis 1993; Chen & Maun 1999; Gan, Miller & McDonald 2003; Li, Zhao & Fang 2006). My literature search found a single study comparing wild and domestic forms of a crop species, namely cowpea (*Vigna unguiculata*) (Lush & Wien 1980). In line with the burial hypothesis, this found that the larger seeds of the domesticated subspecies were more likely to emerge from 12 cm burial depth, although it tested just one wild and two domesticated accessions.

I used a comparative experimental approach to test three predictions of the burial hypothesis in eight legume crop species, domesticated in six regions on different continents. Current thinking is that agriculture could have begun independently in all six of these regions (Diamond 2002; Purugganan & Fuller 2009), but even the most conservative estimates accept three separate origins, all of which are represented here (Harlan 1971). The first prediction is that, within crop species, emergence depth is positively correlated with seed size. Secondly, since seed size increases with domestication, domestic accessions should be able to survive deeper burial than wild accessions.

Finally, I predict that the effect of domestication on emergence depth exceeds that expected based on seed mass alone. If a selective pressure was favouring seedlings able to emerge from greater depths, various traits besides seed size could respond to that, using the available resources more efficiently to grow upwards to the surface. For instance, seedlings could invest a greater fraction of their reserves in shoot growth, rather than root growth, or produce a proportionately thinner hypocotyl or epicotyl. The effect of such changes would be that crop seedlings are better able to emerge from depth than wild seedlings, even if they had seeds of the same size. However, caution is required in interpreting this, as selection for other factors, such as growth rate, may also have affected these traits.

4.2 Materials and methods

4.2.1 Plant material

Eight legume crop species were chosen, representing several geographical centres of agricultural origins (Table 4.1). The sampling strategy was not designed to be exhaustive; for example, I sampled only two of the grain legumes domesticated in the Fertile Crescent. Instead, within logistic constraints of the experimental set-up, I aimed to cover a broad range of geographical regions, and different sized grains spanning lentil to Lima bean.

To confirm the expectation that domestic forms have larger seeds, seed mass data for each species were collected from the US GRIN/NPGS germplasm database (http://www.ars-grin.gov/npgs/), and supplemented with my own weight measurements. These data were

filtered to include only the accessions collected in the region where the crop was domesticated and, where possible, wild and domestic accessions were filtered by species or subspecies as well as the recorded improvement status, to exclude feral (weedy) accessions of domesticated crops.

Origin	Crop	Domestic	Progenitor
Western Asia	Lentil	*Lens culinaris*	*L. culinaris subsp. orientalis*
	Pea	*Pisum sativum**	
Africa	Cowpea	*Vigna unguiculata*	*Vigna unguiculata* subsp. *dekindtiana*
South Asia	Mung bean	*Vigna radiata*	*V. radiata* var. *sublobata*
China	Soybean	*Glycine max*	*G. soja*
Central/South America	Common bean	*Phaseolus vulgaris*	*P. vulgaris* var. *aborigineus*
	Lima bean	*Phaseolus lunatus*	
South America	Peanut	*Arachis hypogaea*	*A. monticola*

Table 4.1: Legume crop species used. Names follow GRIN taxonomy (USDA, ARS, National Genetic Resources Program). *Wild peas includes accessions of Pisum sativum, P. sativum subsp. elatius, P. sativum var. arvense and P. sativum var. pumilio, but in each case had improvement status recorded as 'wild material'.

For seed burial experiments, I obtained accessions of each species from GRIN/NPGS, except for mung bean, which came from the Australian AusPGRIS collection (http://www2.dpi.qld.gov.au/extra/asp/auspgris/). As with the seed mass data described above, all accessions were originally collected from the region where the crop originated, avoiding feral accessions where possible. If seed mass data were provided by the germplasm

database, accessions were chosen to represent as wide a range of seed sizes as possible; otherwise they were chosen at random. Seed listed as landrace accessions was used for the domesticated samples, to minimise the effects of modern commercial crop breeding. For those crop progenitors which do not have a distinct taxonomic name, accessions listed as wild material were used. One Lima bean accession was redesignated from landrace to wild on the basis of evidence of dormancy that I uncovered during my experiment after the experiment, so was counted as wild in the analysis.

4.2.2 Emergence depth trial

Accessions were randomly allocated to four blocks, established sequentially, and each containing one wild and one domesticated accession of each crop. Pea (*Pisum sativum*) was excluded from the last two blocks, as data showed that it could consistently emerge from the greatest depths used in the experiment.

Polythene 'layflat' tubing (postpack.co.uk) wrapped with aluminium foil was used to make containers: using 5 cm width tubing (approximate diameter 32 mm) for lentil, pea, cowpea, mung bean and soybean; 7.5 cm width tubing (approximate diameter 48mm) for common bean, Lima bean and peanut, since pilot trials showed that the larger seedlings of these species were constrained by the narrower tubes. While using two different diameters of tubing restricted the possibility of direct comparison between species, the principal aim of the experiment was to compare emergence within species. Tubes were 40 cm long, and were loosely fixed at the bottom

to allow drainage.

Up to 20 seeds of each accession were weighed individually, with the exception of wild peanut (*Arachis monticola*), for which only 12 seeds per accession were available. Seeds other than peanuts were scarified with medium grit sandpaper to expose part of the cotyledons. Tubes were packed to a constant density with a soil mix comprising 2:1:1 (by volume) M3 compost (East Riding Horticulture, Yorkshire, UK): Chelford 52 silica sand (Sibelco, Cheshire, UK): perlite (East Riding Horticulture), intended to provide a well draining medium suitable for a lab screen of seedling traits, and to be easy to pour into narrow tubes. A pilot experiment was done with seeds planted at between 2 and 18 cm below the soil surface, to determine the approximate emergence depths of the eight species. The results were used to choose five evenly spaced depths for each species, ranging between 2 and 28 cm, such that the deepest-planted seedlings would be unlikely to emerge, although pea proved consistently able to emerge from all depths used. Within each accession, seeds were assigned randomly among these depths.

Tubes were watered thoroughly, then placed in a growth room (MTPS 120, Conviron, Winnipeg, Canada), with a 12 hour day, 22/20°C day/night temperature regime, and constant 50% relative humidity. They were subsequently watered at 1-week intervals to maintain a moist but not waterlogged soil medium, and emergence recorded daily for 5 weeks. After this period, tubes where a seedling had not emerged were emptied to check for the presence of a seedling; where

a seedling was not found, the seed was taken to have not germinated, and the sample was excluded from subsequent analysis.

4.2.3 Statistical analysis

The seed masses of wild and landrace accessions were statistically compared using t-tests of log-transformed data, as seed mass data typically follows a log-normal distribution (Leishman *et al.* 1995; Khazaei *et al.* 2008).

To analyse the emergence data, generalised linear mixed effects models were built using the R package 'lme4', treating each species separately rather than including species as a factor in a combined model. Seedling emergence from the soil surface was the binomial dependent variable, and the independent variables were seed mass and domestication, modelled first separately, then together in an additive model. In each case, accession was included as a non-interacting random effect (allowing accession to interact with depth did not significantly improve model fit). The statistical power of this model was evaluated with simulated data, wherein emergence depth was proportional to the cube root of seed mass, plus a constant factor for domestication, and a normally distributed error term, with a standard deviation of 4cm. With data equivalent to a single species (four accessions of each of wild and domestic, five depths, four replicates at each depth), an effect size of 3cm from domestication was detected as significant in 69% of simulated samples, and an effect size of 4cm in 91%. With a coefficient of 2 in the seed mass term (selected to bring the simulated emergence depths roughly in

line with the results), log seed mass was detected as a significant factor in over 95% of samples in both cases. The experimental design therefore gave sufficient statistical power to detect the hypothesised effect of seed size on seedling emergence.

Survival analysis of the time from sowing to emergence was also performed, using mixed-effects Cox proportional hazard models. Again, accession was included as a random effect, while a separate model was fitted for each crop species. This uses more of the available information than condensing the data to binomial emergence, but it is a more complicated technique, and is normally applied to events which must occur eventually, unlike seedling emergence. We therefore saw this as complementary to the binomial analysis.

4.3 Results

4.3.1 Seed mass

In each species, seed mass data collected from germplasm databases and observations confirmed the expected increase in seed mass with domestication (see Figure 3.1). On average, the landrace accessions of a species had seeds that were 3.9 times heavier than the wild accessions; the ratios for individual species are shown in Table 4.2, ranging from 1.5 for pea to 7.8 for soybean.

Species	Seed mass multiple	Significance of		Predicted emergence depth change (cm)
		Seed mass	Domestication	
Lentil	4.1	*** z=4.41, p=10^{-5}	** z=2.65 p=0.0080	10.1
Mung bean	2.2	** z=3.00, p=0.0027	** z=2.61 p=0.0088	2.2
Lima bean	3.2	*** z=5.00, p=5.87×10^{-7}	N.S.	3.6
Cowpea	6.8	** z=3.16, p=0.0016	N.S.	5.4
Pea	1.5	* z=2.14 p=0.032	N.S.	5.4
Soybean	7.8	N.S.	N.S.	-
Common bean	3.6	N.S.	N.S.	-
Peanut	2.4	N.S.	N.S.	-

Table 4.2: Significance levels (asterisks indicate standard p-value thresholds) for factors predicting emergence in each species. The seed mass multiples compare landrace accessions to wild, based on data from germplasm databases as well as my own measurements. Changes in emergence depth were predicted from these using the fitted models of emergence depth on seed size.

4.3.2 Emergence depth

Of 1159 seeds planted, 952 germinated (82%), of which 593 (62%) emerged within five weeks (Figure 4.1). Seeds which did not germinate were ignored in further analysis. The probability of seeds germinating was affected by depth in only two species (mung bean, p=3.34x10-4; cowpea, p=0.0223; logistic regression).

In all species besides pea, depth had a highly significant effect on emergence (p < 10^{-3}; figure 2). Pea seedlings consistently emerged

from even the greatest depth used in the experiment (28 cm), so the models only detected a very weak depth effect. However, emergence was less than 50% at the greatest depths tested in all of the other species (Figure 4.1). For every accession of species except pea, I therefore used the generalised linear model fits of logistic curves to predict the depth at which 50% of individuals failed to emerge, and plotted these against seed mass (Figure 4.2).
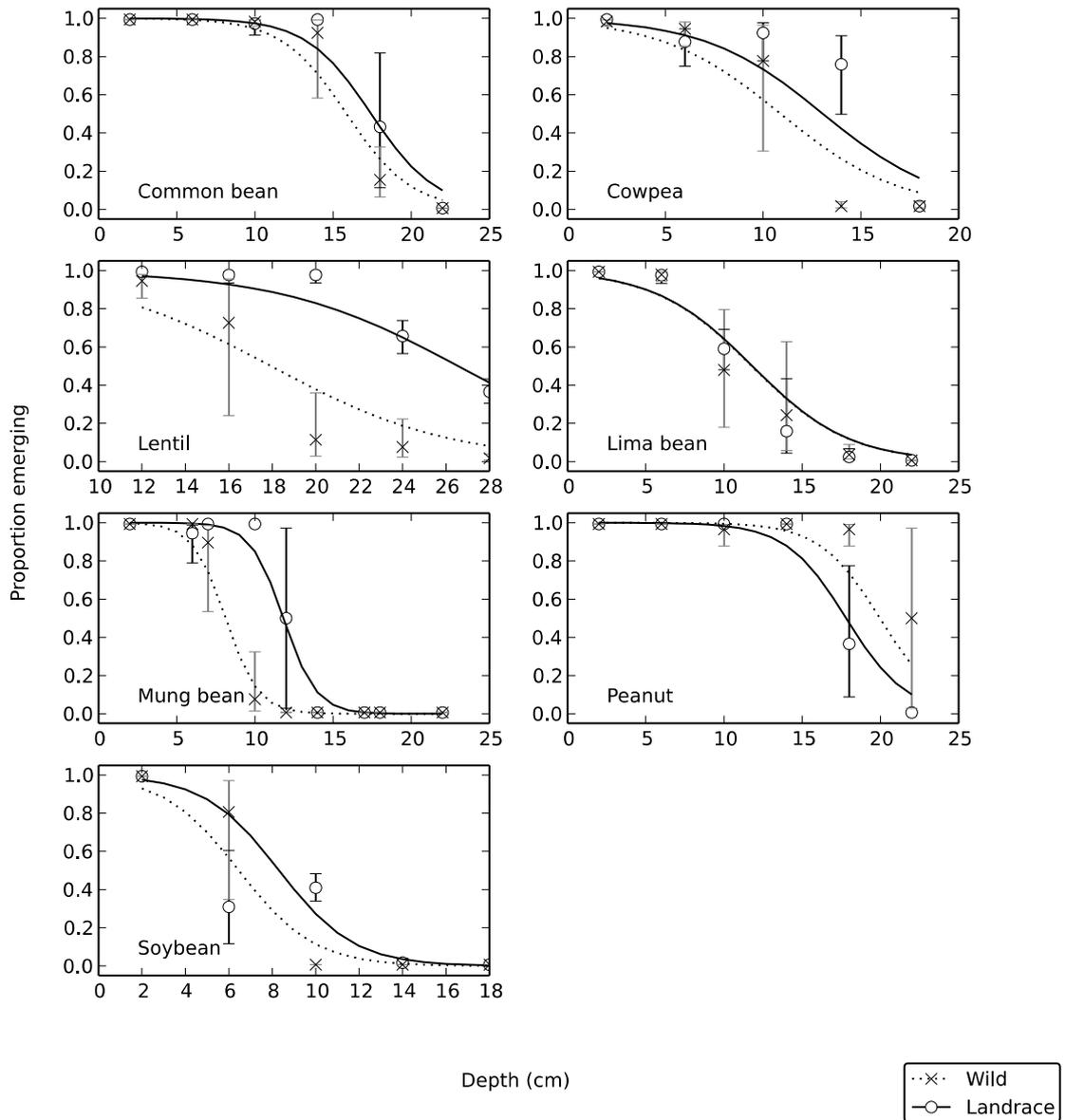
Figure 4.1: GLM predictions of emergence probability against depth, according to domestication. Models fitted in R. Error bars represent the standard error of the mean as calculated on a logistic scale.
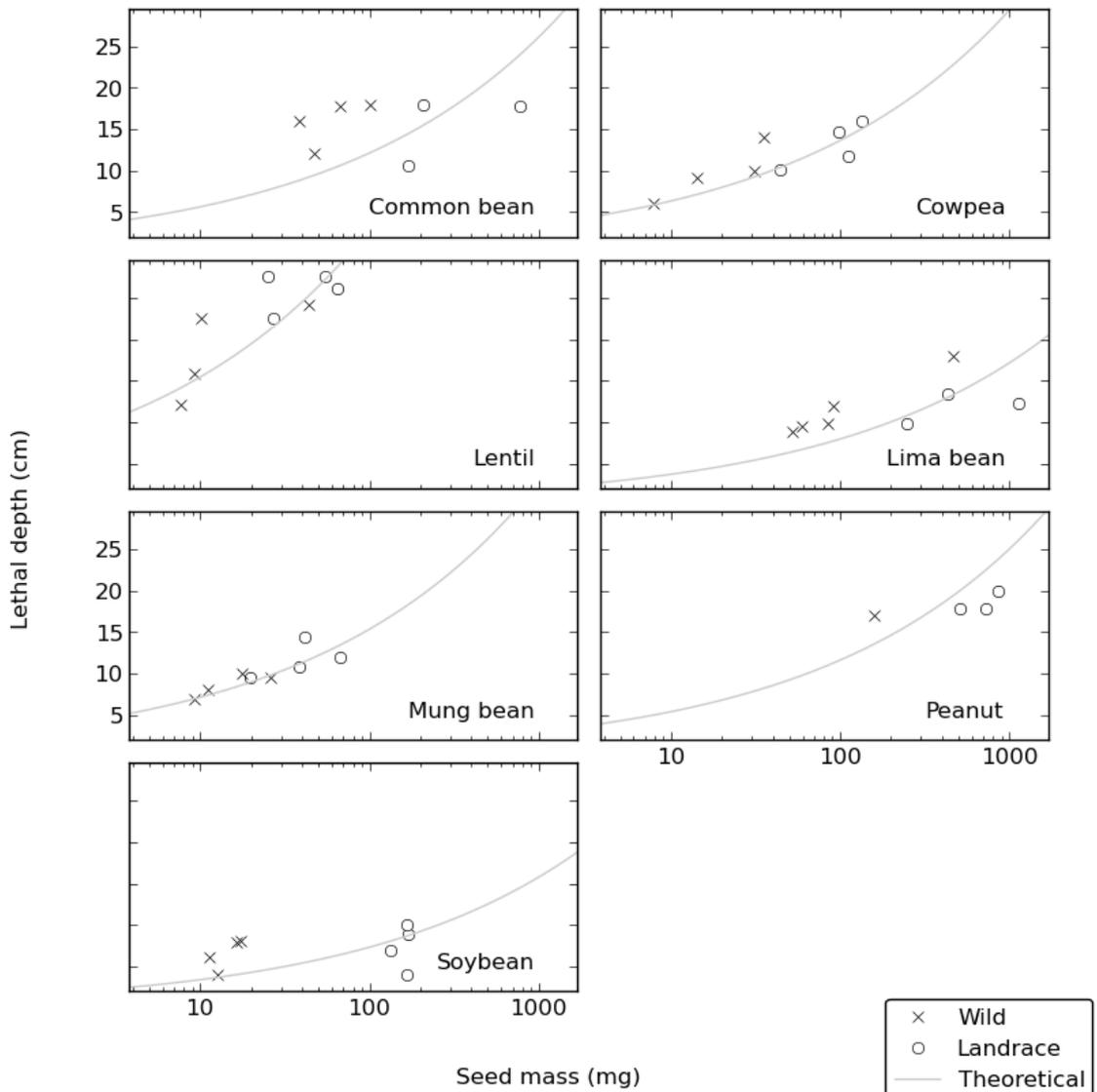
Figure 4.2: The depth at which 50% of seeds for each accession are expected to emerge (from fitted GLMs), against the average mass of each accession, on logarithmic axes. Data are excluded where it was not possible to fit a realistic lethal depth for an accession. The grey lines indicate the shape of a relationship of the form $depth \propto \sqrt[3]{mass}$ (predicted by theory), drawn through the centre of the points on each plot. In lentil, cowpea & mung bean, the 95% confidence interval for the gradient on log-log axes includes 1/3 (corresponding to the cube-root relationship) and excludes 0. In common bean, peanut & soybean, it includes 0 and excludes 1/3. In lima bean, the upper and lower bounds of the confidence interval are respectively just below 0 and 1/3.

In five of the species (lentil, Lima bean, mung bean, cowpea and pea; approaching significance in soybean), log seed mass was a significant predictor of emergence (Figure 4.2; Table 4.2). However, log seed

93

mass did not predict emergence in soybean, common bean or peanut (Figure 4.2; Table 4.2). Domestication was a significant predictor of emergence in only two species (lentil & mung bean; Figure 4.1; Table 4.2). However, with an additive model including seed mass and domestication, domestication did not significantly increase the likelihood of emergence in any species; in two species (cowpea & soybean), domestication significantly decreased emergence probability (z=2.20, 1.98; p=0.028, 0.048 respectively).

To estimate the effect of domestication on emergence depth via changes in seed size, I combined all significant within-species relationships between seed mass and emergence depth (Figure 4.2) with the effects of domestication on seed mass observed across a large number of accessions (Figure 3.1). The increase in emergence depth predicted from increased seed size during domestication varied markedly across species, from 2.2 cm in mung bean to 10.1 cm in lentil (Table 4.2).

4.3.3 Survival

The survival analysis used time-to-emergence to investigate the effects of burial depth, seed mass and domestication. It produced similar results to those of the simple binary (emerged/not emerged) analysis. Proportional hazard models assume a baseline "hazard function"—in this case, the probability of a seedling emerging on any given day—which is multiplicatively affected by "risk factors", such as seed size. Larger seeds had a higher likelihood of emergence in three species (lentil, cowpea and mung bean; z=2.83–4.89,

p=0.0046–1.1×10^{-5}; Lima bean was approaching significance). Domesticated seeds had a higher likelihood of emergence in just one species (lentil; z=2.74, p=0.0062).

4.4 Discussion

This work has provided the first general experimental test of the hypothesis that seed burial during early cultivation exerted unconscious selection for larger seeds. Seed mass data for grain legumes spanning a global sample of independent centres of crop domestication conformed to the widely cited observation that larger seeds are one of the hallmarks of domestication (Harlan *et al.* 1973; Smith 2006; Purugganan & Fuller 2009; Lee *et al.* 2011) (figure 1). However, my experimental results only offer limited general support for the burial depth hypothesis, finding a relationship between seed mass and emergence depth in some but not all of the species tested. The species in which neither seed size nor domestication affected emergence depth (common bean, peanut and soybean) indicate that selection on emergence depth cannot have been a general phenomenon in cultivated grain legumes.

The archaeobotanical evidence for a number of species—including mung and urd beans, lentil, pea, soybean and adzuki bean—suggests that a delay of some millennia between the earliest evidence of cultivation and an increase in seed size is a common pattern in legumes (Fuller 2007). My data suggest that this pattern cannot be explained by a common mechanism. Mung bean is one of two Indian *Vigna* species that have been studied to provide archaeobotanical

95

evidence for the hypothesis (Fuller & Harvey 2006), and in this case I found that both seed mass and domestication are predictors of emergence depth. The same is true of lentil, another species where archaeobotanical evidence has been interpreted in favour of the burial hypothesis (Fuller 2007). The experimental data thus support the archaeobotanical evidence in these cases: any change in cultivation practices that led to the deeper burial of seeds, such as the introduction of animal-drawn ploughs, would have been able to drive some degree of selection on seed mass in mung bean and lentil, and perhaps also in Lima bean, cowpea and pea. However, soybean, common bean and peanut showed no size-dependent response to depth, indicating that the increases in seed size associated with domestication in these species has another cause.

If selection had acted via burial depth, we might expect mechanisms other than seed size to also respond, increasing emergence from depth beyond that expected from increased seed size alone. For example, increased allocation of resources to seedling shoot (versus root) growth can allow emergence from deeper burial (Seiwa *et al.* 2002). There is evidence from cassava (*Manihot esculenta* Crantz) that seeds can change between epigeal germination, where the cotyledons are lifted out of the soil, and hypogeal germination, where the cotyledons remain in the soil where the seed was planted, within the timescale of domestication (Pujol *et al.* 2005b); the significance of this for emergence depth is discussed below. However, this prediction was not borne out in any of the species tested: additive models

including seed mass and domestication did not show an increase in emergence associated with domestication. In fact, domestication had the opposite effect for two species, with landrace seedlings less likely to emerge from a given depth than predicted on the basis of seed size.

Among the species tested, there was a marked difference between those with hypogeal germination, where the cotyledons remain in the soil as storage organs, and those germinating epigeally, raising the cotyledons to the soil surface where they have a photosynthetic role. Hypogeal species (lentils and peas) were best able to emerge from depth ($p=4\times10^{-3}$, adding germination type term to a mixed effects generalised linear model), even though lentils were among the smaller seeds used. A likely explanation is that, not needing to pull their cotyledons through the soil, they could produce a thinner shoot, requiring a smaller investment of resources per unit depth.

Theoretically, maximum emergence depth is the length the shoot can grow from seed reserves, which is expected to be proportional to the cube root of seed mass (Bond *et al.* 1999). While some species, such as cowpea and lentil, appear to fit this pattern, others show a smaller than expected change in emergence depth, such as common bean and soybean (Figure 4.2). Most of the species tested are epigeal, and the cotyledons have a role in photosynthesis as well as storage. In those species with relatively modest increases in emergence depth, selection may be producing a greater allocation of resources to the photosynthetic role (i.e. cotyledon area), driving faster initial growth.

Seeds had a surprising ability to emerge from depth under the experimental conditions, with some hypogeal seedlings growing through 28cm of soil (the greatest depth tested) to reach the surface. Human disturbance of the soil is unlikely to bury seeds to such depths. However, the conditions in this experiment (high moisture, homogeneous stone-free soil, small variance in temperature, and no competition) are expected to permit emergence from a greater depth than in the field. Few field experiments in the literature have tested such depths, although some tests on legume crops have shown no significant disadvantages to sowing depths down to 10cm (Siddique *et al.* 1997; Siddique & Loss 1999). A study of weedy grasses found that the median lethal depth in the field was about 30% shallower than in the greenhouse (Dawson & Bruns 1962). It is reasonable to assume that the differences in emergence depth which were the focus of this study would be similar, albeit of smaller magnitude, in the environment where selection could have acted on them.

4.5 Conclusion

Emergence depth increased with seed size in some crop species but not others, indicating that selection did not act generally on emergence depth during the domestication of grain legumes. In lentil and mung bean especially, the results offer some support for the hypothesis arising from archaeobotanical data that deeper burial in cultivated fields was a selective pressure on seed size. In other species, particularly common bean, soybean and peanut, the hypothesis is not supported. I therefore conclude that other selection

pressures were involved in the evolution of larger seeds during the domestication of grain legumes. Either another shared selection pressure drove the increase in seed size or, more plausibly, different crops may have been subject to different selection pressures, and even multiple selection pressures acting in concert. In this respect, the results suggest that the mechanism for increasing seed mass may be something which is not general to all crops and agricultural origins.

5 Comparison of photosynthesis among crop species and their progenitors

5.1 Introduction

In addition to seed size, total crop yield has increased through domestication (Evans 1980; Buckler, Thornsberry & Kresovich 2001; Peng *et al.* 2003), and we might expect selection to have favoured increased photosynthetic rates, as a factor contributing to yield. People were unable to directly ascertain the rate of photosynthesis until Ingenhousz's experiments in the 18th Century (Ingenhousz 1779), so any change in photosynthesis during domestication would constitute a form of unconscious selection, driven by some kind of selection on related traits such as yield. There is evidence for considerable variation in photosynthetic rates within domesticated crop species (Dwelle, Hurley & Pavek 1983; Evans & Seemann 1984; Peng, Krieg & Girma 1991), and experimental evolution in maize has demonstrated that it is possible to select for faster photosynthesis (Crosbie, Pearce & Mock 1981). Today, there is interest in increasing crop yields by improving their photosynthesis (e.g. Horton 2000; Long *et al.* 2006).

However, comparisons of crops with their wild progenitors often find little difference in the rate of photosynthesis (references in Gifford & Evans 1981, mostly for grass crops), and some studies of wheat and soybean have even found a lower rate in the domestic form (Evans & Dunstone 1970; Johnson *et al.* 1987; Kokubun & Wardlaw 1988). In many cases, crop plants have substantially larger leaves, giving them

a greater total carbon assimilation rate (Evans & Dunstone 1970; Sobrado & Turner 1986). Khan and Tsunoda (1970) suggested that the change in leaf morphology might be linked to irrigation, while Chapin et al. (1989) implicated seed size, although the relationship between seed size and leaf size does not appear to hold in rice (Cook & Evans 1983a).

There are exceptions to this pattern. Domestic cassava has higher photosynthetic rates than its progenitor, both per unit area and per unit leaf mass (Pujol *et al.* 2008). Cook and Evans (1983a) noted that in rice, "there appears to have been a tendency for photosynthetic rate to increase through domestication from the annual progenitor, associated with increasing [specific leaf weight] and N content." Also in rice, Kuroda and Kumura (1990) found a higher rate of photosynthesis during ripening among varieties released after 1950 than in earlier varieties.

While in most cases studied to date the maximum rate of photosynthesis has not increased in domestication, there is evidence that the leaves of domestic crops including wheat, rice and soybean retain a greater photosynthetic ability for a longer time (Evans & Dunstone 1970; Evans 1976; Lush & Rawson 1979; Cook & Evans 1983a). This change would increase the total photosynthetic production over the life of the plant. However, one evolutionary change does not preclude another change effected by the same selection pressure.

I therefore aimed to test for differences in potential light-saturated photosynthetic rate among a range of grass and legume crops, to see whether any exhibited a difference between wild and domestic forms, which would be an exception to the trend previously observed. As in other parts of this thesis, I used a comparative approach to test whether domestication changed photosynthetic rates in different human and biological contexts. Specifically, I chose a grass crop and a legume crop with origins in each of four regions where agriculture is thought to have begun independently (Balter 2007; Piperno *et al.* 2009; Crawford 2009; Fuller *et al.* 2012b), although other authors argue that agriculture diffused into Africa from Western Asia (Blumler 1992b). These are listed in Table 5.1. The study included some crop species which have been tested before, such as wheat (Evans & Dunstone 1970) and cowpea (Lush & Rawson 1979), as well as others which do not appear to have been tested in this way, such as foxtail millet and peas.

Region	Grass	Legume
Western Asia	Einkorn wheat, *Triticum monococcum* L.	Pea, *Pisum sativum* L.
China	Rice, *Oryza sativa* L. & *O. nivara* S. D. Sharma & Shastry; Foxtail millet, *Setaria italica* (L.) P. Beauv. and *S. viridis* (L.) P. Beauv.[a]	Soybean, *Glycine max* (L.) Merr. & *G. soja* Siebold & Zucc.
Central America	Maize, *Zea mays* L.	Common bean, *Phaseolus vulgaris* L.[b]
Africa	Sorghum, *Sorghum bicolor* (L.) Moench	Cowpea, *Vigna unguiculata* (L.) Walp.

Table 5.1: Species used in the experiment. Names follow GRIN taxonomy, and wild taxa are only listed where they are considered separate species. [a] Foxtail millet was added to the experiment after most rice plants failed to thrive. [b] The strong climbing habit of common bean made measurement difficult, and it was dropped from the experiment.

5.2 Methods

5.2.1 Plant Material

Three accessions of each species and domestication status were planted, except for landraces of foxtail millet, where only two accessions were available at the time. Of each accession, up to four replicates were grown for the experiment (additional seeds were germinated and grown as seedlings to allow for failures). The majority of the accessions were from the US GRIN/NPGS database, along with:

- Wild rice accessions from the Australian Tropical Crops & Forages collection.

- Maize from CYMMIT, Mexico

- Foxtail millet from IPK Gatersleben, Germany

- Wild foxtail millet from Herbiseed, Twyford, UK

5.2.2 Growth conditions

Seeds that required scarification to germinate were scarified using sandpaper. Seeds were then germinated on filter paper. Seedlings were initially potted in seedling trays in a 2:3 mixture of M3 compost (East Riding Horticulture, Yorkshire, UK) and vermiculite (East Riding Horticulture), and allowed to grow in a shaded environment, before being transferred to 13 cm round pots containing M3 compost and 1 g/l of a slow-release fertiliser (Osmocote Exact Standard 5-6M, Everris International B.V.), and exposed to full light.

Plants were grown in a controlled environment, with a 14 hour day, stepping up to maximum light intensity over 4 hours, and down again over 2 hours. Temperature was 24°C during the day, and 20°C during the night. Relative humidity was maintained at 60%. Maximum light intensity at canopy height was around 1200 $\mu mol/m^2/s$.

5.2.3 Measurements

All initial measurements were made on young, fully expanded leaves which had developed since the plants were in full light. Measurements were made using an LI-6400XT infra-red gas analyser (LI-COR, Lincoln, Nebraska, USA), with a 6400-02B LED head. The plants were allowed to equilibrate at a photosynthetic photon flux density (PPFD) of 2000 $\mu mol/m^2/s$ (a high level likely to be saturating or nearly saturating), reference CO_2 of 400 ppm, and block temperature of 26°C. Then the instrument was programmed to produce a light curve, decreasing

PPFD in steps from 2000 to 0 µmol/m²/s, while logging data at five-second intervals. At each step, the leaf was allowed to stabilise for at least two minutes, and until the sample CO_2 and air flow rate were changing at not more than 1 ppm/minute and 1 µmol/s/minute respectively.

From each dataset, the last four data at each light intensity were averaged to find the stable assimilation rate at that light intensity. A model was fitted to these based on the theoretical light curve equation from Long and Hällgren (1993):

$$A = \frac{A_{sat} + Q\Phi - \sqrt{-4 A_{sat} Q\Phi\theta + (A_{sat} + Q\Phi)^2}}{2\theta} - R_d$$

Equation 5.1: Carbon assimilation (A) rate relative to light intensity (Q). A_{sat} is the assimilation at saturating light, Φ (phi) is the quantum yield, R_d is dark respiration, and θ (theta) is a convexity parameter.

5.2.4 Statistical tests

The modelled asymptotic assimilation rates (A_{sat}) were analysed in an ANOVA, with domestication status nested inside species.

Repeated measurements on individual leaves of both wild and landrace plants were made on wheat, maize, sorghum, pea and cowpea. The decline in photosynthesis with leaf age was assessed as a gradient, percentage of the initial photosynthetic rate lost per day. A linear model of the decline in photosynthesis was also used by Kitajima et al. (2002) in tropical trees. These gradients were also evaluated by an ANOVA, with domestication nested inside species.

5.3 Results

Light curves were measured for 129 leaves. Data fitted closely to the

theoretical expectation, with R^2 values greater than 0.99 in 91% of

cases.
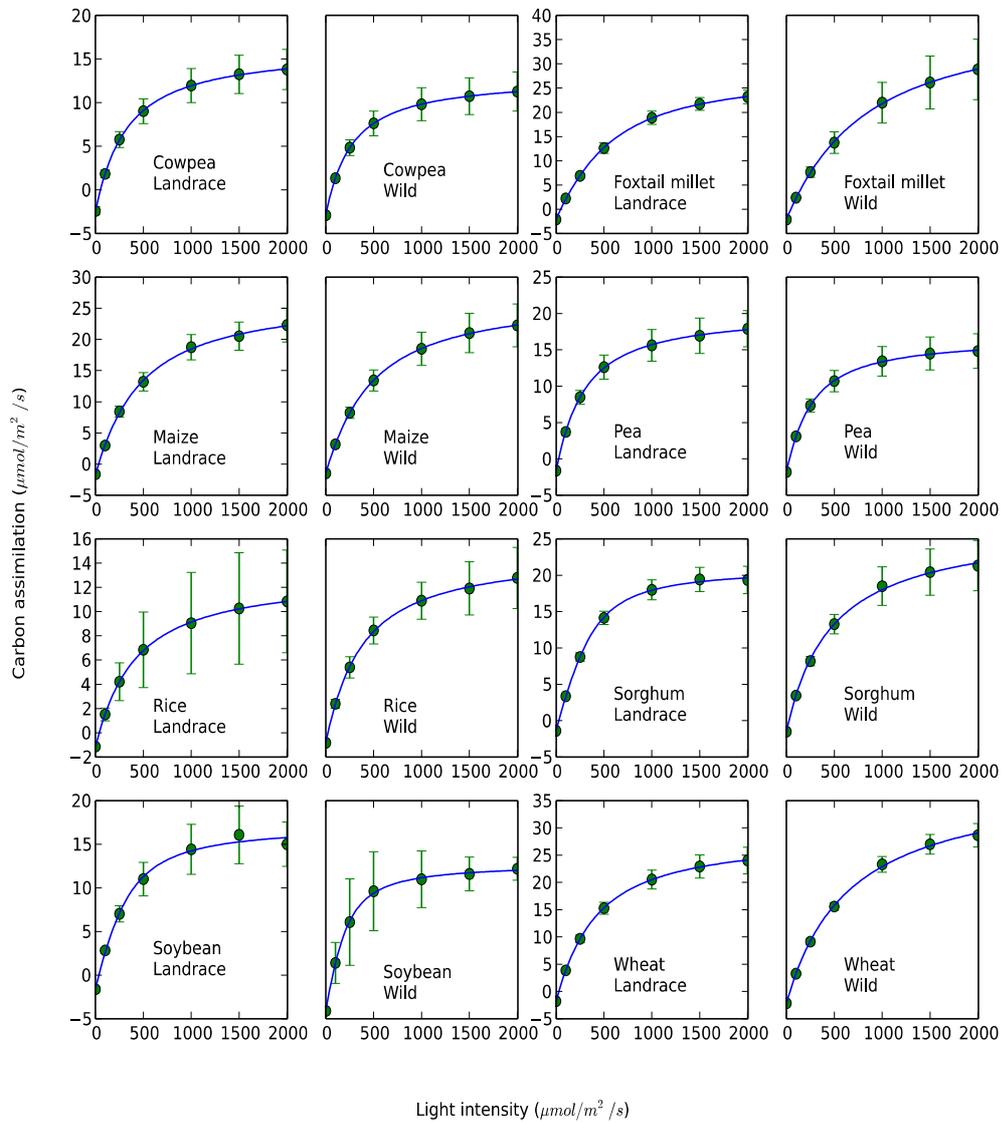
Figure 5.1: Light curves from average assimilation readings of leaves measured for the first time. Error bars show standard error of the mean. One low outlier for wild soybean was excluded. Curves are fitted using Equation 5.1.

There were significant differences in light-saturated assimilation rate

between the species tested, but no difference between wild and

106

landrace accessions (Figure 5.1, Figure 5.2, Table 5.2). Only in einkorn wheat did the difference in saturated assimilation rate even approach significance (t=1.67, p=0.097, with a higher rate in the wild form).
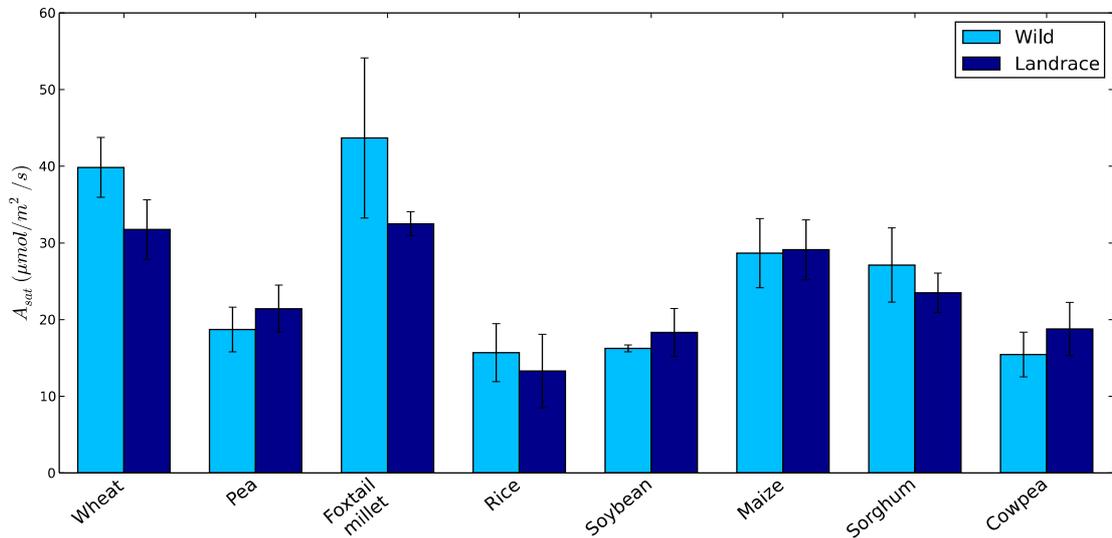
Figure 5.2: Modelled asymptotic photosynthesis at saturating light. Error bars show standard errors of the mean.

	df	sum_sq	mean_sq	F	PR(>F)
Species	7	6820.16	974.31	6.677	0.000001
Species: Domestication	8	883.21	110.40	0.757	0.641494
Residual	111	16196.06	145.91		

Table 5.2: ANOVA table of a model for assimilation at saturating light, by domestication nested within species, for measurements on new leaves.

The decline in photosynthetic rate varied considerably within species (Figure 5.3). The rate of the decline also differed among species, but not according to domestication (Table 5.3). The difference was not significant in any of the individual species.
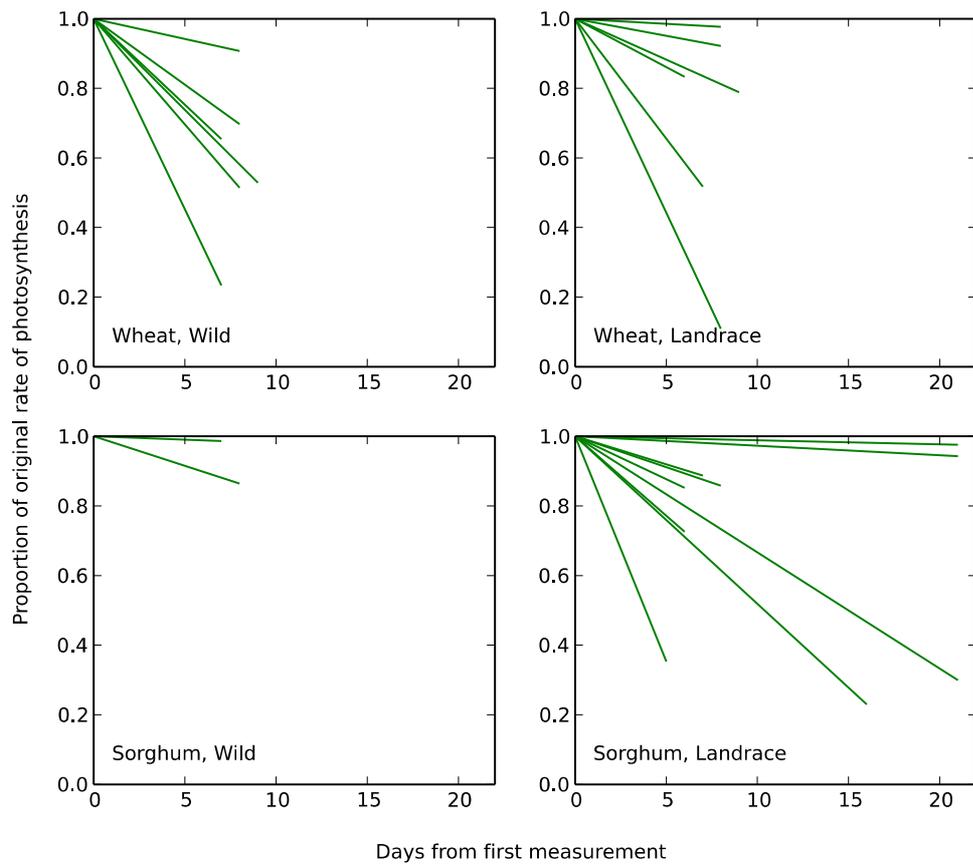
Figure 5.3: Declines in photosynthetic rate at saturating light in wheat and sorghum.

	df	sum_sq	mean_sq	F	PR(>F)
Species	4	0.014005	0.003501	2.884739	0.046285
Species: Domestication	5	0.004477	0.000895	0.737794	0.603190
Residual	22	0.026701	0.001214		

Table 5.3: ANOVA table of a model for the rate of decline of light-saturated photosynthesis in individual leaves, by domestication nested within species.

5.4 Discussion

The results demonstrate the generality of the pattern noted by Gifford

& Evans (1981): domestication has not selected for a greater rate of

photosynthesis. Even in recent times when we have better understood photosynthesis, breeding has not increased its rate (Richards 2000). Intuitively, this is somewhat surprising: if selection has increased crop yields, we might expect the carbon supply from photosynthesis to be boosted.

Why might photosynthetic rates have remained unchanged? Natural selection may have already increased photosynthesis to a physiological maximum in the wild, leaving no room for selection pressures in the cultivated environment to increase it further. On the other hand, if photosynthesis is limited by a trade-off with other traits such as defence (Massad, Dyer & Vega C. 2012; Nabity, Zavala & DeLucia 2013) or stress tolerance (Chapin, Autumn & Pugnaire 1993; Fernández & Reynolds 2000), we might expect relaxation of these constraints in cultivation to allow photosynthesis to increase. Faster growth in domesticated species has been linked to a reduction in both plant defence (Rosenthal & Dirzo 1997; Massei & Hartley 2000) and drought tolerance (Koziol *et al.* 2012).

There is an important general point here: the traits in the domestication syndrome cannot be explained solely in terms of selection pressures acting on crop plants. We must consider the *differences* in the selection regime acting on crop plants and on wild plants. For instance, crop plants may benefit from larger leaf area, but other factors being equal, larger leaves able to absorb more energy would also be advantageous in the wild. However, the development of irrigation relaxed selection for drought tolerance in cultivated plants,

releasing crops to evolve traits which are involved in trade-offs with drought tolerance, such as increased specific leaf area (Nautiyal, Rachaputi & Joshi 2002; Liu & Stützel 2004).

That increase in leaf area during domestication may have offset the lack of a change in photosynthetic rate per unit leaf area (Evans & Dunstone 1970; Khan & Tsunoda 1970; Sobrado & Turner 1986; Chapin *et al.* 1989). Moreover, the broader, thinner leaves seen in domesticates could have *constrained* the rate per unit area, because less photosynthetic machinery is present within a given area. Pujol et al. (2008) argued that we should measure photosynthesis per unit leaf mass, rather than leaf area, to account for this, although their results for cassava show a difference between wild and domestic on both the area and mass bases. Theoretically, neither basis completely represents the internal photosynthetic capacity: with no other difference, a thinner leaf will have a lower photosynthetic rate per unit area, because it captures less of the incident light, but a higher rate per unit mass, because its chloroplasts are on average less shaded. Some recent studies have found that expressing fundamentally area-proportional measurements, such as photosynthetic rate, on a mass basis can generate apparent correlations even from random data (Lloyd *et al.* 2013; Osnas *et al.* 2013). However, they can answer different questions: the mass basis better represents the return on resource investment from photosynthesis.

In rice, studies have found that the rate of photosynthesis depends on

leaf nitrogen content (Takano & Tsunoda 1971; Cook & Evans 1983a; b; Kuroda & Kumura 1990). If one particular resource limits photosynthesis, it may be informative to measure the assimilation rate relative to that resource. However, nitrogen allocation is itself an interesting question in crop domestication. In grain crops such as the species in this study, a high seed protein content is often desirable, at least in modern crops (Ries & Everson 1973; Mosse 1990; Diers *et al.* 1992), although the effect of domestication is not clear: domestic durum wheat carries a mutation increasing protein content compared to its wild relative (Uauy *et al.* 2006), but in finger millet and tepary beans, the wild progenitors had higher protein contents than domestic cultivars (Waines 1978; Barbeau & Hilu 1993). If selection did favour maintaining or increasing protein content while seed size increased (see chapter 3), the developing seed would require more nitrogen, which may limit its availability for photosynthesis.

Some studies have also found that yield in crop species such as wheat, maize, common bean and soybean is limited primarily by the strength of the carbon sink, not by the carbon source (photosynthesis) (Nakano *et al.* 2000; Borrás, Slafer & Otegui 2004; Miralles & Slafer 2007). There is evidence that older cultivars are more sink-limited than modern ones (Álvaro *et al.* 2008; Acreche & Slafer 2009), so progenitors would likely also have been sink-limited. Interestingly, rice may be an exception: one study which removed developing panicles found that it did not cause a decline in photosynthesis, indicating that the plants were not sink-limited

(Nakano, Makino & Mae 1995). This is a possible explanation for the observation that domestication *does* appear to have increased the photosynthetic rate in rice (Cook & Evans 1983a). However, other studies have shown that non-structural carbohydrates remain in rice stems after harvest, indicating that grain filling is sink-limited (Slewinski 2012). Paul and Foyer (2001) note that potato and citrus crops can easily increase sink strength, so are less sink-limited; as a root crop, cassava may have similar flexibility, which could explain the higher photosynthetic rate seen in domestic cassava compared to its wild progenitor (Pujol *et al.* 2008). In general, however, it has been argued that sources and sinks are co-limiting over the lifetime of a plant (Slewinski 2012), in which case we might expect domestication to increase both the source, photosynthesis, and the sinks, including developing seeds.

This study did not find the slower loss of photosynthetic capacity in leaves of domestic plants which some authors have observed (Evans & Dunstone 1970; Evans 1976; Lush & Rawson 1979; Cook & Evans 1983a). However, relatively few repeated measurements were made on individual leaves. It is possible that the difference only occurs in particular conditions or at particular times, such as during flower or seed development. If leaf lifespan is greater in domestic forms, this would fit with findings that yield is closely related to total plant photosynthesis, but not to instantaneously measured photosynthetic rates per unit area (Zelitch 1982; Richards 2000).

5.5 Conclusions

In none of the species has the light-saturated rate of photosynthesis increased in domestication; the broad global sample of grass and legume crop species investigated here conform to the patterns seen in earlier studies on individual species. This may be explained by trade-offs allowing a greater increase in net photosynthesis through increased total leaf area with less resources invested per unit area, and a greater duration of leaf photosynthetic capacity. I found no evidence that selection for increased yield during or after crop domestication effected unconscious selection on photosynthesis.

6 General Discussion

6.1 Synthesis of results

6.1.1 The value of large seed size in crops

The larger seeds of many grain crops compared to their progenitors constitute one of the major traits of the domestication syndrome (Harlan *et al.* 1973; Hammer 1984; Brown *et al.* 2009). Chapters 2 and 3 extend our knowledge of this difference: crop progenitors themselves have relatively large seeds compared with other wild species in the same families, while the increase in seed size associated with domestication extends to crops in which a vegetative organ is harvested, and even to crops which are normally vegetatively propagated, so the seed is neither harvested nor planted.

Why seed size should be so important for agriculture is not entirely clear. Naively, it may seem that larger seeds offer more energy per plant harvested. However, ecologists are familiar with a trade-off between seed size and number (Smith & Fretwell 1974; Venable 1992; Turnbull, Rees & Crawley 1999; Jakobsson & Eriksson 2000), meaning that much of the variation in seed size does not represent differences in the total production of seed biomass, but merely the difference between many smaller seeds and fewer larger ones. Within this trade-off, Evans (1996 p. 5) has argued that for a long time after cultivation started, farmers would have been most interested in the ratio of seed planted to seed harvested, and that this would be highest in plants producing numerous small grains.

In the ecological literature, large seeds have been associated with drought tolerance (Baker 1972; Schimpf 1977; Leishman & Westoby 1994; Manga & Yadav 1995) (although Blumler (1992a p. 455) notes that drought tolerant crop plants tend to have relatively small seeds), shade tolerance (Westoby *et al.* 1996; Walters & Reich 2000; Khurana, Sagar & Singh 2006) and competitive ability (Black 1958; Howe & Richter 1982; Geritz, van der Meijden & Metz 1999). These correspond to the 'competition' and 'stress-tolerance' poles of C-S-R theory (Grime 1974; Hodgson *et al.* 1999), while small seeds are a characteristic of ruderal species (Westoby 1998).

Arguably, the conditions in cultivation have some similarities with the ruderal pole of C-S-R theory: seeds are sown in fertile ground following disturbance (by tillage), weeding limits interspecific competition, at least relative to the wild state, and abiotic stress may be reduced by artificial irrigation. Grime (1977) placed annual plants in general in the ruderal category, and many crop plants in the secondary 'competitive-ruderal' category (Grime 2001 p. 119), although others have suggested that annuals in mediterranean grasslands are more akin to stress tolerators (Madon & Médail 1997). Intraspecific competition is a significant factor in arable fields (Fuller & Allaby 2009 p. 262), but if there is intense competition at the seedling stage, it suggests that the farmer could save seed by sowing at a lower density, and may even get a better stand yield by doing so (Lonsdale 1990). It has been suggested that tillage imposed a greater abiotic stress in the form of seed burial (Harlan *et al.* 1973; Zohary

2004; Fuller 2007; Purugganan & Fuller 2009), but the experimental results described in chapter 4 of this thesis do not support this as a general mechanism, at least in grain legumes. Burial depth and competition could also interact to favour seedlings which emerged sooner and were able to overtop their neighbours (Fuller & Allaby 2009).

Other selection pressures act on seed size in the opposite direction. Smaller seeds may be both better able to disperse and more likely to escape seed predation (Reader 1993; Crawley 2000 p. 170; Gómez 2004). Deliberate sowing of seeds removes the dispersal requirement, as evidenced by the loss of structures aiding dispersal, such as hooks and awns (discussed in detail by Fuller & Allaby 2009), while the storage of harvested seeds may have reduced the impact of seed predators (Hillman & Davies 1990b)—other than humans themselves, who presumably did not select the smallest seeds for replanting. So part of the effect of domestication may have been to reduce the importance of evolutionary forces which favour smaller seeds in the wild.

Not knowing what early farmers were thinking, we can never rule out that they consciously chose species to cultivate or individuals to propagate on the basis of seed size. Farmers may have preferred larger seeds as easier to handle, especially if they were sowing seeds individually by dibbling rather than broadcasting (Jones 2005). We might even speculate about a spiritual aspect, with the best grains from the harvest being returned to the soil as a kind of offering.

However, the archaeobotanical evidence indicates that the increase in seed size can be quite gradual, and occurs in different crops before or after the evolution of other domestication traits, patterns which fit more parsimoniously with unconscious selection (Fuller & Allaby 2009). The results in chapter 3 of this thesis reinforce this: seed size has increased in crops grown for vegetative parts, including crops which are typically vegetatively propagated, in which it is not plausible that seed size was under conscious selection; hence, unconscious selection must have acted on seed size.

One possible mode of unconscious selection on seed size might be selection for traits which are correlated with seed size. To achieve similar results in such diverse crops, the connection would likely have to be pleiotropy—multiple phenotypic effects of the same genes— rather than genetic linkage. Traditionally, relative growth rate has been considered to be negatively correlated with seed size, but there is ongoing debate over how to calculate representative figures for plant growth rates, and the relationship between seed size and growth rate now appears to be more variable than was previously realised (Turnbull *et al.* 2012). A recent analysis suggests that short-lived species—such as, potentially, annual crops—are more likely to have a positive relationship between seed size and relative growth rate at a given plant size (Turnbull *et al.* 2012). There is also a correlation between seed size and final plant size (Venable & Rees 2009), which may mean that despite the trade-off between seed size and number discussed above, a plant grown from a large seed does

produce a greater mass of seeds, because the whole plant is larger.

A final twist is the domestication of numerous small-seeded species. In the grasses, these are often collectively termed 'millet', but the term hides considerable diversity. Blumler (1992a p. 460) contends that most of these are probably secondary domesticates, although he acknowledges that the domestication of broomcorn millet and foxtail millet in Northern China is likely an exception. More recent finds support the theory that broomcorn millet, *Panicum miliaceum* L., was a primary domesticate in China (Lu *et al.* 2009). Other authors are also more enthusiastic about the independence of agricultural origins in India and Africa (Ehret 1979; Fuller 2006, 2007), where several small-grained grasses were domesticated, and Weber & Fuller (2008) point out that archaeobotanical remains of millets have often been overlooked, leading to underestimates of their significance. Additionally, in the legumes, the small-seeded lentil is held to be one of the Western Asian 'founder crops', alongside the considerably larger pea and chickpea, and the slightly larger bitter vetch. Thus, while most major crops originated from large-seeded progenitors, large seeds are not a hard requirement for domestication, nor is the presence of large-seeded species necessary for the transition to agriculture.

As discussed in chapter 2, a focus on agricultural origins in the Fertile Crescent, where several large-seeded grass species were domesticated, may exaggerate the importance of seed size globally. On the other hand, the results suggest that within each region, the

species which entered cultivation had large seeds in comparison to the pool of available species, suggesting that seed size was a 'filter' (see note on terminology, section 2.1), even though not all of the progenitors have large seeds on a global scale. However, certain crops, such as teff, would still appear to need a separate explanation.

6.1.2 Similarities between grasses and legumes

The results in chapter 2 and chapter 5 show similar patterns for grass and legume crops. In both cases, crop progenitors have larger seeds than other wild species in the region, and crop genera have larger seeds than other wild genera. Western Asia has an unmatched abundance of large-seeded grasses and legumes. Finally, domestication has not increased photosynthetic rate in either group of crops.

In some ways, this is surprising: grasses and legumes have different growth habits and different nutritional contents, and legumes typically have larger seeds than grasses in general. We might predict that their domestication would show different patterns. For instance, Fuller (2007), found that seed size did not increase in various pulse species until some time after domestication, while in at least some grasses, the increase in seed size precedes the evolution of the non-shattering phenotype that is usually used to define domestication.

On the other hand, these are part of a larger pattern of similarities in grass and legume domestication. Early farmers in multiple independent regions domesticated both grass and legume crops

(Table 1.1), so they may have applied similar harvesting and sowing techniques, despite the physical differences between grass and legume plants. Many traits of the domestication syndrome are also similar between the two groups, even where the mechanism is different: both have lost their natural means of seed dispersal, although that means is infructescence shattering in grasses and pod dehiscence in legumes.

This has implications for the interpretation of the results in chapter 4. That experiment was done entirely on legume crop species. If evolution in cultivation shows similar patterns, we might expect that the selection pressure effected by burial depth was similar in grasses and legumes, in which case burial would not consistently drive the evolution of larger grass crops. A similar experiment in grasses found that all but the smallest seeded species reliably survived burial to at least 8cm in M3 compost (unpublished results from Sarah Wilkinson, in Sheffield), which offers some support for this.

6.1.3 Comparison of centres of origin of agriculture

The maps in chapter 2 show clear clusters of large seeded grass and legume species around the Mediterranean and the Fertile Crescent in Western Asia. This supports Blumler's findings that 32 out of 48 wild, non-aquatic, non-bamboo grass species with seeds above 10 mg in mass occur in the Mediterranean/Near East (Blumler 1992a p. 451). He considers large-seeded annual species to be a product of mediterranean climates, which feature characteristic summer drought and winter rainfall. These occur in several regions around the world,

but Blumler notes that the Mediterranean/Near East region is much larger, better connected to adjacent biomes, and has a drier summer than any of the other comparable regions (Blumler 1992a pp. 16–21). However, without any possible replication, it is impossible to determine what combination of geological and biological factors are responsible for this region's distinct abundance of large-seeded species.

The seedling emergence results in chapter 4 also hint at a difference between the domestication centre in Western Asia and some other centres. Of the species tested, lentil, a primary domesticate from Western Asia, was one of two species where emergence depth clearly increased with domestication, the other being mung bean, from India, where the independence of agricultural origins is not certain (Fuller 2006). The other Western Asian domesticate tested, pea, did not produce satisfactory results, as it consistently emerged from all the depths tested. It is possible that burial depth was a more significant selection pressure on seed size in Western Asia than in other centres of origin such as China and Central/South America.

This leads to an interesting question. In chapter 4, I found that the hypogeally germinating species, those whose cotyledons remain in the soil, are better able to emerge from depth than epigeal species, which raise their cotyledons to the soil surface. This makes intuitive sense: epigeal species, having to pull their cotyledons up through the soil, presumably use more reserves to emerge from a given depth. If burial depth was a more significant factor in the Fertile Crescent, were

more of the crops domesticated there hypogeal?

Table 6.1 shows that, although there are hypogeal and epigeal legume crop species which originated in each region, in Western Asia there are most hypogeal species, including all four of the legumes (lentil, pea, chickpea, bitter vetch) in the set of eight 'founder crops' (see e.g. Weiss & Zohary 2011). In the other regions, most crops and the most significant crops are epigeal. However, this pattern is conflated with phylogeny: three of those species—pea, lentil and bitter vetch—along with broad bean, common vetch and grass pea, fall into the Vicieae, where hypogeal germination is a synapomorphy; chickpea is also closely related to this clade (Steele & Wojciechowski 2003).

Region of origin	Hypogeal	Epigeal
Western Asia	Pea Lentil Chickpea Broad bean Common vetch Bitter vetch Grass pea	Blue lupine Yellow lupine White lupine
East/South-East Asia	Adzuki bean Rice bean	Soybean
India	Pigeon pea	Mung bean Urd bean Horse gram Moth bean
Africa	Velvet bean	Cowpea Hyacinth bean
South/Central America	Runner bean	Common bean Lima bean Tepary bean Jack bean Tarwi
	Peanut (intermediate)	

Table 6.1: Epigeal and hypogeal germination in legumes

domesticated in different regions (Baudet 1974; Robinson 1975; Putnam et al. 1991; Steele & Wojciechowski 2003).

In itself, the fact that all of the major near-eastern legume crops come from a small clade of just a few genera is fascinating. This is perhaps the clearest example of phylogenetic clustering of domesticated crops, but there are a number of other instances:

- The three cereal founder crops from Western Asia—einkorn wheat, emmer wheat and barley—are all from the tribe Triticeae, as is rye. The Triticeae, however, is not as limited as the Vicieae, including a number of genera which were not domesticated.

- Five Asian pulses—mung, urd, adzuki, rice and moth bean—are all members of *Vigna* subgenus *Ceratotropis.* Three other species, *Vigna trilobata*, *V. trinervia* and *V. reflexopilosa* var. *glabra* are also cultivated as forage crops or as minor human food sources (Tomooka 2002). The African species cowpea and Bambara groundnut are also in *Vigna* sect. C*atiang*.

- In Central and South America, a number of species of *Phaseolus* have been domesticated, including common bean, Lima bean, runner bean and tepary bean. Evidence also suggests that there were two independent domestications within each of the common bean and Lima bean species (Gutiérrez Salgado *et al.* 1995; Chacón S *et al.* 2005).

- The genus *Canavalia* includes the sword bean, domesticated in

123

the old world, and the jack bean, from the new world (Smartt 1985). There may have been two independent domestications in the Americas, although this is uncertain (Pickersgill 2007).

- Also in South America, besides the common, tetraploid peanut, *Arachis hypogaea*, a related diploid species, *A. villosulicarpa* is a local crop (Galgaro *et al.* 1997).

- Cotton (*Gossypium* spp.) was domesticated for its fibre multiple times in both the old world and the new world.

- Similarly, there are local cultivars of different ploidy levels from the diverse species complex of *Solanum* sect. *Petota*, which includes the tetraploid potato grown internationally (Huamán & Spooner 2002). However, some or all of these may be the products of hybridisation between tetraploid cultivars and wild species (Ovchinnikova *et al.* 2011).

- Pickersgill (2007) lists a number of other American genera containing more than one domesticate: *Amaranthus*, *Chenopodium*, *Cucurbita*, *Pachyrhizus*, *Physalis*, *Gossypium* (fibre crop), *Capsicum* (spice) and *Nicotiana* (tobacco).

- Two species of *Echinochloa*, known as barnyard millets, have been domesticated in Asia: *E. frumentacea* from *E. colona*, and *E. esculenta* (syn. *E. utilis*) from *E. crus-galli (Hilu 1994)*.

- Horse gram and Kersting's groundnut, minor legume crops of India and Africa respectively, both belong to the genus

Macrotyloma, which contains some 25 species (International Legume Database & Information Service 2005).

- Broomcorn millet, *Panicum miliaceum,* was an early domesticate in China (Lu *et al.* 2009), and little millet, *P. sumatrense*, is cultivated in India (Kimata *et al.* 2000). Nabhan and de Wet (1984) found some evidence that a variety of *P. hirticaule,* called *P. sonorum* by some authors, was domesticated in prehistoric North America.

- Austin (2006) lists various species of small grained grasses in the genus *Setaria* which are used for food, and claims that three were domesticated, one in the New World. Of these, foxtail millet, *S. italica*, is the only one of any significance today.

- *Brachiaria deflexa* is grown in Guinea (Harlan 1976 p. 10), and *B. ramosa* is cultivated in Southern India (Kimata *et al.* 2000).

- *Digitaria exilis* and *D. iburua*, known as fonio, are grown in West Africa (Portères 1976), while *D. cruciata*, known as raishan, is a minor domesticate in Northeast India (Singh & Arora 1972).

In most of these cases, it is not clear whether multiple closely related species were independently domesticated, suggesting some form of pre-adaptation favouring their cultivation, or whether early farmers took into cultivation species similar to those with which they were already familiar.

6.2 Caveats and potential future work

This thesis has deliberately taken a global perspective, comparing

crop plants with origins in very different parts of the world. This has both advantages and disadvantages relative to focussing on a single domestication centre. It is essential to study individual regions in depth, to properly understand the unique human and biological processes at work in each. But there are fascinating similarities: people in several parts of the world independently began farming within a relatively short space of time, and certain plant taxa, such as grasses and legumes, were common sources of crops. There are also differences, such as the importance of root crops rather than grain crops in South America. If we are to explain these patterns, we need rigorous global comparative studies.

Chapters 3-5 hinge on the distinction between wild and landrace forms of the crop species tested. It is not always easy to distinguish truly wild forms from weedy races that have escaped from cultivation, and the potential for interbreeding between wild and domestic plants further complicates the picture (Harlan 1965; Linder *et al.* 1998; Bartsch *et al.* 1999; Ellstrand, Prentice & Hancock 1999). The comparisons are still valid: the traits of the domesticated accessions have been influenced to a greater extent by cultivation, but any interbreeding may have reduced the magnitude of the difference. We should also remember that wild plants have continued to evolve over the last ten thousand years, so they are not exactly the same as the plants growing when agriculture began. However, this time span is quite short in evolutionary terms, so we can probably take their characteristics to represent those of their ancestors when some

species were taken into cultivation.

In a similar vein, the comparisons between crop progenitors and other wild species in chapter 2 rely on having an accurate picture of which species are crop progenitors. This is confounded by two issues. Firstly, there are crops for which an immediate progenitor is not known, such as broomcorn millet, *Panicum miliaceum* L. (Hunt *et al.* 2011). There may be uncertainty over whether a wild form is a feral escape from cultivation (also discussed below), the progenitor may be unclear within a pool of wild species, or the crop may be derived from a complex of interbreeding species, as in potato. Finally, some progenitors may be extinct.

Secondly, there may be progenitors of species which have been cultivated, but which I did not include as crops. My literature search turned up numerous minor millets and pulses, many of them only cultivated in a specific region, such as *Arachis villosulicarpa* Hoehne, a peanut grown by the Nambiquara in Matto Grosso, Brazil (Galgaro *et al.* 1997), and *Triticum timopheevi* (Zhuk.) Zhuk, a tetraploid wheat grown in parts of Georgia (Zohary & Hopf 2000; Jones *et al.* 2000). Some are known to have been grown in the past, but are no longer cultivated, such as *Bromus mango* E. Desv., a Chilean cultivated grass believed to have gone extinct in the 19th or 20th Century (Hammer 2004 sec. 5.5.1.2). In general, there is no sharp distinction between crops and wild food plants: Freedman's (2012) list of famine foods includes various close relatives of cultivated species, such as *Oryza punctata* Kotschy ex Steud. and *Lathyrus* spp. Given these

constraints, the list of crops and progenitors is inevitably incomplete. If those minor crops remained minor because they were in some sense less 'domesticable', then their omission merely demarks a boundary on a continuum of domesticability, and the comparison of more domesticable progenitors with other wild species remains valid. If they remained minor due to historical factors unrelated to their biology, then the progenitors listed are an approximately random sample of crop progenitors, which can still support the conclusions. However, if those species were equally suitable for domestication, but less suited to subsequent changes in agriculture, then the comparison conflates the factors predisposing species to enter cultivation with those which allowed species to survive and spread in later cultivation.

It would be particularly interesting to combine the seed mass database assembled for chapter 2 with phylogenies of grasses and legumes, such as the grass phylogeny prepared by Edwards and Smith (2010), along with data on other traits such as growth rate, if available. As described above, the phylogenetic distribution of crop plants appears to be strongly clustered. A trait such as seed size could distinguish the clades containing crop species from sister clades without domesticated species. Liu et al. (2012) found a phylogenetic signal in seed mass down to, but not within, subfamilies of the grasses.

Another exciting avenue would be to follow up the investigation of vegetable crops, especially tuber crops, in chapter 3. Besides reinforcing the conclusions with more data and more species,

researchers could explore the correlations between seed mass and other traits. The work of Dayal et al. (1984) in potatoes provides a starting point for this, but it may be especially interesting to look for differences evident when the plants are propagated vegetatively, which are linked with seed size pleiotropically rather than physiologically. Additionally, while I listed studies showing the incorporation of volunteer true seedlings into a number of vegetatively propagated crops, genetic and ethnographic work could constrain estimates of how common this has been for different species.

There is also the potential to combine measurements of photosynthetic rate, as in chapter 5, with measurements of specific leaf area, to compare photosynthesis per unit mass in crops and progenitors, as suggested by Pujol et al (2008). Alternatively, researchers could measure leaf nitrogen content, and compare photosynthesis per unit nitrogen, following studies which show that photosynthetic rate depends on nitrogen, both within and between species (Takano & Tsunoda 1971; Cook & Evans 1983a; b; Field & Mooney 1986; Kuroda & Kumura 1990). It might also be informative to study the structure of canopies in crops and progenitors, to consider the selective environment for photosynthesis. For instance, if denser crop canopies result in more shading of leaves, selection might favour more efficient photosynthesis at lower light intensities, rather than acting on the light-saturated photosynthetic rate.

6.3 Conclusions

This thesis has examined unconscious selection on seed size and photosynthesis during crop plant domestication by comparing crop species from around the world. While I feel that it would not be honest to make some sweeping statement about the disparate results, a number of important conclusions are warranted. Firstly, that the notable prevalence of large-seeded annual grasses and legumes in Western Asia is not clearly replicated in other regions where agriculture began, indicating that we should be careful when generalising conclusions about how agriculture began. Secondly, the patterns of seed size in grass and legume crop progenitors are broadly similar, both with respect to their close relatives and to species from the regions where agriculture began.

Looking at the selection pressures on crop species during domestication, unconscious selection does not appear to have affected photosynthetic rate, although alternative ways to consider this may yield different results. Burial depth does not appear to have been a mechanism generally selecting for larger seeds, at least among legumes. However, the results for vegetable crops in chapter 3 provide the first direct evidence that unconscious selection has played a role in increasing seed mass during domestication.

References

Aarssen, L.W. (2005) Why don't bigger plants have proportionately bigger seeds? *Oikos*, **111**, 199–207.

Abayomi, Y.A., Fadayomi, O., Babatola, J.O. & Tian, G. (2001) Evaluation of selected legume cover crops for biomass production, dry season survival and soil fertility improvement in a moist savanna location in Nigeria. *African Crop Science Journal*, **9**, 615–627.

Abbo, S., Gopher, A., Rubin, B. & Lev-Yadun, S. (2005) On the origin of Near Eastern founder crops and the "dump-heap hypothesis." *Genetic Resources and Crop Evolution*, **52**, 491–495.

Abbo, S., Rachamim, E., Zehavi, Y., Zezak, I., Lev-Yadun, S. & Gopher, A. (2011) Experimental growing of wild pea in Israel and its bearing on Near Eastern plant domestication. *Annals of Botany*, **107**, 1399 –1404.

Abbo, S., Saranga, Y. & Peleg, Z. (2009) Reconsidering domestication of legumes versus cereals in the ancient Near East. *The Quarterly Review of Biology*, **84**, 29–50.

Abbo, S., Lev-Yadun, S. & Gopher, A. (2010a) Agricultural origins: centers and noncenters; a Near Eastern reappraisal. *Critical Reviews in Plant Sciences*, **29**, 317.

Abbo, S., Lev-Yadun, S. & Gopher, A. (2010b) Origin of Near Eastern plant domestication: homage to Claude Levi-Strauss and "La Pensée Sauvage." *Genetic Resources and Crop Evolution*, **58**, 175–179.

Abbo, S., Zezak, I., Schwartz, E., Lev-Yadun, S. & Gopher, A. (2008a) Experimental harvesting of wild peas in Israel: implications for the origins of Near East farming. *Journal of Archaeological Science*, **35**, 922–929.

Abbo, S., Zezak, I., Schwartz, E., Lev-Yadun, S., Kerem, Z. & Gopher, A. (2008b) Wild lentil and chickpea harvest in Israel: bearing on the origins of Near Eastern farming. *Journal of Archaeological Science*, **35**, 3172–3177.

Acreche, M.M. & Slafer, G.A. (2009) Grain weight, radiation interception and use efficiency as affected by sink-strength in Mediterranean wheats released from 1940 to 2005. *Field Crops Research*, **110**, 98–105.

Allaby, R.G., Fuller, D.Q. & Brown, T.A. (2008) The genetic expectations of a protracted model for the origins of

domesticated crops. *Proceedings of the National Academy of Sciences*, **105**, 13982–13986.

Allaby, R., Peterson, G., Merriwether, D. & Fu, Y.-B. (2005) Evidence of the domestication history of flax (Linum usitatissimum L.) from genetic diversity of the sad2 locus. *TAG Theoretical and Applied Genetics*, **112**, 58–65.

Almekinders, C.J.M., Neuteboom, J.H. & Struik, P.C. (1995) Relation between berry weight, number of seeds per berry and 100-seed weight in potato inflorescences. *Scientia Horticulturae*, **61**, 177–184.

Álvaro, F., Royo, C., García del Moral, L.F. & Villegas, D. (2008) Grain filling and dry matter translocation responses to source–sink modifications in a historical series of durum wheat. *Crop Science*, **48**, 1523.

Andreev, A.A., Klimanov, V.A. & Sulerzhitsky, L.D. (1997) Younger Dryas pollen records from central and southern Yakutia. *Quaternary International*, **41-42**, 111–117.

Andrews, A.C. (1949) The carrot as a food in the classical era. *Classical Philology*, **44**, 182–196.

Aranguren, B., Becattini, R., Mariotti Lippi, M. & Revedin, A. (2007) Grinding flour in Upper Palaeolithic Europe (25000 years bp). *Antiquity*, **81**, 845–855.

Araus, J.L., Febrero, A., Catala, M., Molist, M., Voltas, J. & Romagosa, I. (1999) Crop water availability in early agriculture: evidence from carbon isotope discrimination of seeds from a tenth millennium BP site on the Euphrates. *Global Change Biology*, **5**, 201–212.

Austin, D. (2006) Fox-tail millets (Setaria: Poaceae)—Abandoned food in two hemispheres. *Economic Botany*, **60**, 143–158.

Azhaguvel, P. & Komatsuda, T. (2007) A phylogenetic analysis based on nucleotide sequence of a marker linked to the brittle rachis locus indicates a diphyletic origin of barley. *Annals of Botany*, **100**, 1009–1015.

Badr, A., M, K., Sch, R., Rabey, H.E., Effgen, S., Ibrahim, H.H., Pozzi, C., Rohde, W. & Salamini, F. (2000) On the origin and domestication history of barley (Hordeum vulgare). *Molecular Biology and Evolution*, **17**, 499–510.

Baker, H.G. (1972) Seed weight in relation to environmental conditions in California. *Ecology*, **53**, 997–1010.

Balter, M. (2007) Seeking agriculture's ancient roots. *Science*, **316**, 1830–1835.

Ba, F.S., Pasquet, R.S. & Gepts, P. (2004) Genetic diversity in cowpea [Vigna unguiculata (L.) Walp.] as revealed by RAPD markers. *Genetic Resources and Crop Evolution*, **51**, 539–550.

Barbeau, W.E. & Hilu, K.W. (1993) Protein, calcium, iron, and amino acid content of selected wild and domesticated cultivars of finger millet. *Plant Foods for Human Nutrition*, **43**, 97–104.

Barton, L., Newsome, S.D., Chen, F.-H., Wang, H., Guilderson, T.P. & Bettinger, R.L. (2009) Agricultural origins and the isotopic identity of domestication in northern China. *Proceedings of the National Academy of Sciences*, **106**, 5523–5528.

Bartsch, D., Lehnen, M., Clegg, J., Pohl-Orf, M., Schuphan, I. & Ellstrand, N.C. (1999) Impact of gene flow from cultivated beet on genetic diversity of wild sea beet populations. *Molecular Ecology*, **8**, 1733–1741.

Baudet, J.C. (1974) Signification taxonomique des caractères blastogéniques dans la tribu des Papilionaceae-Phaseoleae. *Bulletin du Jardin botanique national de Belgique / Bulletin van de National Plantentuin van België*, **44**, 259–293.

Benvenuti, S., MacChia, M. & Miele, S. (2001) Quantitative analysis of emergence of seedlings from buried weed seeds with increasing soil depth. *Weed Science*, **49**, 528–535.

Bertrand, S., Charlet, F., Charlier, B., Renson, V. & Fagel, N. (2008) Climate variability of southern Chile since the Last Glacial Maximum: a continuous sedimentological record from Lago Puyehue (40°S). *Journal of Paleolimnology*, **39**, 179–195.

Bhatt, A.K., Bhalla, T.C., Agrawal, H.O. & Upadhya, M.D. (1989) Effect of seed size on protein and lipid contents, germination and imbibition in true potato seeds. *Potato Research*, **32**, 477–481.

Bisht, I.S., Bhat, K.V., Lakhanpaul, S., Latha, M., Jayan, P.K., Biswas, B.K. & Singh, A.K. (2005) Diversity and genetic resources of wild Vigna species in India. *Genetic Resources and Crop Evolution*, **52**, 53–68.

Bisht, M.S. & Mukai, Y. (2001) Genomic in situ hybridization identifies genome donor of finger millet (Eleusine coracana). *TAG Theoretical and Applied Genetics*, **102**, 825–832.

Black, J. (1958) Competition between plants of different initial seed sizes in swards of subterranean clover (Trifolium subterraneum L.) with particular reference to leaf area and the light

microclimate. *Australian Journal of Agricultural Research*, **9**, 299–318.

Blench, R.M. (2012) Vernacular names for African millets and other minor cereals and their significance for agricultural history. *Archaeological and Anthropological Sciences*.

Blumler, M.A. (1992a) *Seed Weight and Environment in Mediterranean-type Grasslands in California and Israel*. Ph.D. Thesis, University of California, Berkeley, CA.

Blumler, M.A. (1992b) Independent inventionism and recent genetic evidence on plant domestication. *Economic Botany*, **46**, 98–111.

Blumler, M.A., Byrne, R., Belfer-Cohen, A., Bird, R.M., Bohrer, V.L., Byrd, B.F., Dunnell, R.C., Gordon Hillman, Moore, A.M.T., Olszewski, D.I., Redding, R.W. & Riley, T.J. (1991) The ecological genetics of domestication and the origins of agriculture [and comments and reply]. *Current Anthropology*, **32**, 23–54.

Bogaard, A. (2005) "Garden agriculture" and the nature of early farming in Europe and the Near East. *World Archaeology*, **37**, 177–196.

Bogaard, A., Charles, M., Twiss, K.C., Fairbairn, A., Yalman, N., Filipović, D., Demirergi, G.A., Ertuğ, F., Russell, N. & Henecke, J. (2009) Private pantries and celebrated surplus: storing and sharing food at Neolithic Çatalhöyük, Central Anatolia. *Antiquity*, **83**, 649–668.

Bond, W.J., Honig, M. & Maze, K.E. (1999) Seed size and seedling emergence: an allometric relationship and some ecological implications. *Oecologia*, **120**, 132–136.

Borek, S., Pukacka, S., Michalski, K. & Ratajczak, L. (2009) Lipid and protein accumulation in developing seeds of three lupine species: Lupinus luteus L., Lupinus albus L., and Lupinus mutabilis Sweet. *Journal of Experimental Botany*, **60**, 3453–3466.

Borrás, L., Slafer, G.A. & Otegui, M.E. (2004) Seed dry weight response to source–sink manipulations in wheat, maize and soybean: a quantitative reappraisal. *Field Crops Research*, **86**, 131–146.

Bowles, S. & Choi, J.-K. (2013) Coevolution of farming and private property during the early Holocene. *Proceedings of the National Academy of Sciences*.

Braadbaart, F. (2008) Carbonisation and morphological changes in modern dehusked and husked Triticum dicoccum and Triticum

aestivum grains. *Vegetation History and Archaeobotany*, **17**, 155–166.

Brown, T.A., Jones, M.K., Powell, W. & Allaby, R.G. (2009) The complex origins of domesticated crops in the Fertile Crescent. *Trends in Ecology & Evolution*, **24**, 103–109.

Brummitt, R.K., Pando, F., Hollis, S. & Brummitt, N.A. (2001) *World Geographical Scheme for Recording Plant Distributions*, 2nd ed.

Bu, H., Chen, X., Xu, X., Liu, K., Jia, P. & Du, G. (2007) Seed mass and germination in an alpine meadow on the eastern Tsinghai–Tibet plateau. *Plant Ecology*, **191**, 127–149.

Buckler, E.S., Thornsberry, J.M. & Kresovich, S. (2001) Molecular diversity, structure and domestication of grasses. *Genetics Research*, **77**, 213–218.

Bulmer, R. (1965) Beliefs concerning the propagation of new varieties of sweet potato in two New Guinea highlands societies. *The Journal of the Polynesian Society*, **74**, 237–239.

Caillon, S. & Lanouguère-Bruneau, V. (2005) Gestion de l'agrobiodiversité dans un village de Vanua Lava (Vanuatu) : stratégies de sélection et enjeux sociaux. *Le Journal de la Société des Océanistes*, 129–148.

Caillon, S., Quero-Garcia, J., Lescure, J.-P. & Lebot, V. (2006) Nature of taro (Colocasia esculenta (L.) Schott) genetic diversity prevalent in a Pacific Ocean island, Vanua Lava, Vanuatu. *Genetic Resources and Crop Evolution*, **53**, 1273–1289.

De Candolle, A. (1885) *Origin of Cultivated Plants*. D. Appleton and Company, New York.

Cavalcanti, M.J. & Kluyver, T. (2010) *PyDelta*.

Chacón S, M.I., Pickersgill, B. & Debouck, D.G. (2005) Domestication patterns in common bean (Phaseolus vulgaris L.) and the origin of the Mesoamerican and Andean cultivated races. *TAG Theoretical and Applied Genetics*, **110**, 432–444.

Chapin, F.S., Autumn, K. & Pugnaire, F. (1993) Evolution of suites of traits in response to environmental stress. *The American Naturalist*, **142**, S78–S92.

Chapin, F.S., Groves, R.H. & Evans, L.T. (1989) Physiological determinants of growth rate in response to phosphorus supply in wild and cultivated Hordeum species. *Oecologia*, **79**, 96–105.

Chen, H. & Maun, M.A. (1999) Effects of sand burial depth on seed

germination and seedling emergence of Cirsium pitcheri. *Plant Ecology*, **140**, 53–60.

Childe, V.G. (1941) *Man Makes Himself*, 2nd ed. Watts & co., London.

Clayton, W.D., Vorontsova, M.S., Harman, K.T. & Williamson, H. (2002) World Grass Species: Synonymy. URL http://www.kew.org/data/grasses-syn/ [accessed 6 July 2011]

Clayton, W.D., Vorontsova, M.S., Harman, K.T. & Williamson, H. (2006) GrassBase - The online world grass flora. URL http://www.kew.org/data/grasses-db/cite.htm [accessed 17 May 2013]

Clements, J.C., Buirchel, B.J., Yang, H., Smith, P.M.C., Sweetingham, M.W. & Smith, C.G. (2005) Lupin. *Genetic Resources, Chromosome Engineering, and Crop Improvement: Grain Legumes* (eds R.J. Singh & P.P. Jauhar), CRC Press, Boca Raton, Florida.

Cohen, M.N. (1977) *The Food Crisis in Prehistory: Overpopulation and the Origins of Agriculture*. Yale University Press, New Haven and London.

Cohen, M.N. (2009) Introduction: Rethinking the origins of agriculture. *Current Anthropology*, **50**, 591–595.

Colledge, S. (2002) Identifying pre-domestication cultivation in the archaeobotanical record using multivariate analysis: presenting the case for quantification. *The Dawn of Farming in the Near East* Studies in Early Near Eastern Production, Subsistence, and Environment. (eds R.T.J. Cappers & S. Bottema), pp. 141–151. ex oriente, Berlin, Germany.

Cook, M.G. & Evans, L.T. (1983a) Some physiological aspects of the domestication and improvement of rice (Oryza spp.). *Field Crops Research*, **6**, 219–238.

Cook, M.G. & Evans, L.T. (1983b) Nutrient responses of seedlings of wild and cultivated Oryza species. *Field Crops Research*, **6**, 205–218.

Crawford, G.W. (2009) Agricultural origins in North China pushed back to the Pleistocene–Holocene boundary. *Proceedings of the National Academy of Sciences*, **106**, 7271–7272.

Crawley, M.J. (2000) Seed predators and plant population dynamics. *Seeds: The Ecology of Regeneration in Plant Communities* (ed M. Fenner), CABI Publishing, Wallingford, UK.

Crockford, D. (2006) The application/json media type for JavaScript

Object Notation (JSON). URL http://tools.ietf.org/html/rfc4627 [accessed 6 February 2012]

Crosbie, T.M., Pearce, R.B. & Mock, J.J. (1981) Recurrent phenotypic selection for high and low photosynthesis in two maize populations. *Crop Science*, **21**, 736–740.

Cunniff, J. (2009) *The Roles of Atmospheric Carbon Dioxide and Plant Ecological Traits in the Origins of Agriculture.* University of Sheffield.

Cunniff, J., Osborne, C.P., Ripley, B.S., Charles, M. & Jones, G. (2008) Response of wild C4 crop progenitors to subambient CO2 highlights a possible role in the origin of agriculture. *Global Change Biology*, **14**, 576–587.

Dallwitz, M.J. (1980) A general system for coding taxonomic descriptions. *Taxon*, **29**, 41–46.

D'Andrea, A.C. (2008) T'ef (Eragrostis tef) in ancient agricultural systems of highland Ethiopia. *Economic Botany*, **62**, 547–566.

Darlington, C.D. (1956) *Chromosome Botany and the Origins of Cultivated Plants*. George Allen & Unwin Ltd., London.

Darwin, C. (1875) *Variation of Plants and Animals Under Domestication*, 2nd ed. John Murray, London.

Davies, D.R. (1977) DNA contents and cell number in relation to seed size in the genus Vicia. *Heredity*, **39**, 153–163.

Dawson, J.H. & Bruns, V.F. (1962) Emergence of barnyardgrass, green foxtail, and yellow foxtail seedlings from various soil depths. *Weeds*, **10**, 136–139.

Dayal, T.R., Upadhya, M.D. & Chaturvedi, S.N. (1984) Correlation studies on 1000-true-seed weight, tuber yield and other morphological traits in potato (Solanum tuberosum L.). *Potato Research*, **27**, 185–188.

Dempewolf, H., Rieseberg, L. & Cronk, Q. (2008) Crop domestication in the Compositae: a family-wide trait assessment. *Genetic Resources and Crop Evolution*, **55**, 1141–1157.

Denham, T.P., Haberle, S.G., Lentfer, C., Fullagar, R., Field, J., Therin, M., Porch, N. & Winsborough, B. (2003) Origins of agriculture at Kuk Swamp in the highlands of New Guinea. *Science*, **301**, 189–193.

Diamond, J.M. (1987) The worst mistake in the history of the human race. *Discover Magazine*, 64–66.

Diamond, J. (1997) Location, location, location: the first farmers. *Science*, **278**, 1243–1244.

Diamond, J.M. (2002) Evolution, consequences and future of plant and animal domestication. *Nature*, **418**, 700–707.

Diamond, J. (2005) *Guns, Germs and Steel: A Short History of Everybody for the Last 13,000 Years*, New edition. Vintage.

Diers, B.W., Keim, P., Fehr, W.R. & Shoemaker, R.C. (1992) RFLP analysis of soybean seed protein and oil content. *Theoretical and Applied Genetics*, **83**, 608–612.

Dignat, G., Welcker, C., Sawkins, M., Ribaut, J.M. & Tardieu, F. (2013) The growths of leaves, shoots, roots and reproductive organs partly share their genetic control in maize plants. *Plant, Cell & Environment*, **36**, 1105–1119.

Dirzo, R., Mendoza, E. & Ortíz, P. (2007) Size-related differential seed predation in a heavily defaunated neotropical rain forest. *Biotropica*, **39**, 355–362.

Doebley, J.F. (1984) "Seeds" of wild grasses: a major food of southwestern Indians. *Economic Botany*, **38**, 52–64.

Dwelle, R.B., Hurley, P.J. & Pavek, J.J. (1983) Photosynthesis and stomatal conductance of potato clones (Solanum tuberosum L.). *Plant Physiology*, **72**, 172–176.

Eastwood, R.J. & Hughes, C.E. (2008) Origins of domestication of lupinus mutabilis in the Andes. *Lupins for health and wealth. Proceedings of the 12th international lupin conference* pp. 14–18.

Edwards, E.J. & Smith, S.A. (2010) Phylogenetic analyses reveal the shady history of C4 grasses. *Proceedings of the National Academy of Sciences*.

Ehret, C. (1979) On the antiquity of agriculture in Ethiopia. *The Journal of African History*, **20**, 161–177.

Elbaum, R., Melamed-Bessudo, C., Boaretto, E., Galili, E., Lev-Yadun, S., Levy, A.A. & Weiner, S. (2006) Ancient olive DNA in pits: preservation, amplification and sequence analysis. *Journal of Archaeological Science*, **33**, 77–88.

Elias, M., Rival, L. & McKey, D. (2000) Perception and management of cassava (Manihot esculenta Crantz) diversity among Makushi Amerindians of Guyana (South America). *Journal of Ethnobiology*, **20**, 239–265.

Ellstrand, N.C., Prentice, H.C. & Hancock, J.F. (1999) Gene flow and introgression from domesticated plants into their wild relatives. *Annual Review of Ecology and Systematics*, **30**, 539–563.

Erickson, D.L., Smith, B.D., Clarke, A.C., Sandweiss, D.H. & Tuross, N. (2005) An Asian origin for a 10,000-year-old domesticated plant in the Americas. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 18315–18320.

Ettle, T. & Roth, F.X. (2004) Specific dietary selection for tryptophan by the piglet. *Journal of Animal Science*, **82**, 1115–1121.

Eule-Nashoba, A.R. (2010) *Seed Size in Lacustrine and Riverine Populations of Wild Rice (Zizania Palustris)*. M.S. Thesis, University of Minnesota, Minnesota.

Evans, L.T. (1976) Physiological adaptation to performace as crop plants. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, **275**, 71–83.

Evans, L.T. (1980) The natural history of crop yield. *American Scientist*, **68**, 388–397.

Evans, L.T. (1996) *Crop Evolution, Adaptation and Yield*. Cambridge University Press.

Evans, L. & Dunstone, R. (1970) Some physiological aspects of evolution in wheat. *Aust. Jnl. Of Bio. Sci.*, **23**, 725–742.

Evans, J.R. & Seemann, J.R. (1984) Differences between wheat genotypes in specific activity of ribulose-1,5-bisphosphate carboxylase and the relationship to photosynthesis. *Plant Physiology*, **74**, 759–765.

Evershed, R.P. (2008) Organic residue analysis in archaeology: the archaeological biomarker revolution. *Archaeometry*, **50**, 895–924.

Falster, D.S., Moles, A.T. & Westoby, M. (2008) A general model for the scaling of offspring size and adult size. *The American Naturalist*, **172**, 299–317.

FAO. Cassava. URL http://ecocrop.fao.org/ecocrop/srv/en/cropView?id=1420 [accessed 23 April 2013]

Fenner, M. (1992) Environmental influences on seed size and composition. *Horticultural Reviews Volume 13* Horticultural Reviews. (eds R.L. Fery, J.A. Flore & E.G. Williams), John Wiley & Sons.

Fernández, R.J. & Reynolds, J.F. (2000) Potential growth and drought

tolerance of eight desert grasses: lack of a trade-off? *Oecologia*, **123**, 90–98.

Field, C. & Mooney, H.A. (1986) The photosynthesis-nitrogen relationship in wild plants. *On the Economy of Plant Form and Function: Proceedings of the Sixth Maria Moors Cabot Symposium, Evolutionary Constraints on Primary Productivity, Adaptative Patterns of Energy Capture in Plants, Harvard Forest, August 1983* (ed T.J. Givnish), pp. 25–56. Cambridge University Press, Cambridge.

Flannery, K.V. (1969) Origins and ecological effects of early domestication in Iran and the Near East. *The Domestication and Exploitation of Plants and Animals* (eds P. Ucko & G. Dimbleby), pp. 73–100. Aldine Transaction, Chicago, IL.

Flannery, K.V., Kirkby, A.V.T., Kirkby, M.J. & Williams, A.W. (1967) Farming Systems and Political Growth in Ancient Oaxaca. *Science*, **158**, 445–454.

Freedman, R.L. (2012) Famine Foods. URL http://www.hort.purdue.edu/newcrop/FamineFoods/ff_home.html [accessed 20 May 2013]

Freitas, F.O., Bendel, G., Allaby, R.G. & Brown, T.A. (2003) DNA from primitive maize landraces and archaeological remains: implications for the domestication of maize and its expansion into South America. *Journal of Archaeological Science*, **30**, 901–908.

Fullagar, R., Field, J., Denham, T. & Lentfer, C. (2006) Early and mid Holocene tool-use and processing of taro (Colocasia esculenta), yam (Dioscorea sp.) and other plants at Kuk Swamp in the highlands of Papua New Guinea. *Journal of Archaeological Science*, **33**, 595–614.

Fuller, D.Q. (2006) Agricultural origins and frontiers in South Asia: a working synthesis. *Journal of World Prehistory*, **20**, 1–86.

Fuller, D.Q. (2007) Contrasting patterns in crop domestication and domestication rates: recent archaeobotanical insights from the old world. *Annals of Botany*, **100**, 903–924.

Fuller, D.Q. (2011a) Finding Plant Domestication in the Indian Subcontinent. *Current Anthropology*, **52**, S347–S362.

Fuller, D. (2011b) Pathways to Asian civilizations: tracing the origins and spread of rice and rice cultures. *Rice*, **4**, 78–92.

Fuller, D.Q. & Allaby, R.G. (2009) Seed dispersal and crop domestication: shattering, germination and seasonality in

evolution under cultivation. *Annual Plant Reviews Volume 38: Fruit Development and Seed Dispersal* (ed L. Østergaard), pp. 238–295. Wiley-Blackwell, Oxford.

Fuller, D.Q., Asouti, E. & Purugganan, M.D. (2012a) Cultivation as slow evolutionary entanglement: comparative data on rate and sequence of domestication. *Vegetation History and Archaeobotany*, **21**, 131–145.

Fuller, D.Q. & Harvey, E.L. (2006) The archaeobotany of Indian pulses: identification, processing and evidence for cultivation. *Environmental Archaeology*, **11**, 219–246.

Fuller, D.Q., Harvey, E. & Qin, L. (2007) Presumed domestication? Evidence for wild rice cultivation and domestication in the fifth millennium BC of the Lower Yangtze region. *Antiquity*, **81**, 316–331.

Fuller, D.Q. & Qin, L. (2009) Water management and labour in the origins and dispersal of Asian rice. *World Archaeology*, **41**, 88.

Fuller, D.Q., Qin, L., Zheng, Y., Zhao, Z., Chen, X., Hosoya, L.A. & Sun, G.-P. (2009) The domestication process and domestication rate in rice: spikelet bases from the lower Yangtze. *Science*, **323**, 1607–1610.

Fuller, D.Q., Willcox, G. & Allaby, R.G. (2012b) Early agricultural pathways: moving outside the "core area" hypothesis in Southwest Asia. *Journal of Experimental Botany*, **63**, 617 –633.

Galgaro, L., Valls, J.F.M. & Lopes, C.R. (1997) Study of the genetic variability and similarity among and within Arachis villosulicarpa, A. pietrarellii and A. hypogaea through isoenzyme analysis. *Genetic Resources and Crop Evolution*, **44**, 9–15.

Gan, Y.T., Miller, P.R. & McDonald, C.L. (2003) Response of kabuli chickpea to seed size and planting depth. *Canadian Journal of Plant Science*, **83**, 39–46.

Geritz, S.A.H., van der Meijden, E. & Metz, J.A.J. (1999) Evolutionary dynamics of seed size and seedling competitive ability. *Theoretical Population Biology*, **55**, 324–343.

Gifford, R.M. & Evans, L.T. (1981) Photosynthesis, carbon partitioning, and yield. *Annual Review of Plant Physiology*, **32**, 485–509.

Giles, R.J. & Brown, T.A. (2008) Improved methodology for extraction and amplification of DNA from single grains of charred wheat. *Journal of Archaeological Science*, **35**, 2585–2588.

Glover, J.D., Reganold, J.P., Bell, L.W., Borevitz, J., Brummer, E.C.,

Buckler, E.S., Cox, C.M., Cox, T.S., Crews, T.E., Culman, S.W., DeHaan, L.R., Eriksson, D., Gill, B.S., Holland, J., Hu, F., Hulke, B.S., Ibrahim, A.M.H., Jackson, W., Jones, S.S., Murray, S.C., Paterson, A.H., Ploschuk, E., Sacks, E.J., Snapp, S., Tao, D., Tassel, D.L.V., Wade, L.J., Wyse, D.L. & Xu, Y. (2010) Increased food and ecosystem security via perennial grains. *Science*, **328**, 1638–1639.

Gómez, J.M. (2004) Bigger Is not always better: conflicting selective pressures on seed size in Quercus ilex. *Evolution*, **58**, 71–80.

Gravano, L., Ipeirotis, P.G., Jagadish, H.V., Koudas, N., Muthukrishnan, S., Pietarinen, L. & Srivastava, D. (2001) Using q-grams in a DBMS for Approximate String Processing. *IEEE Data Engineering Bulletin*, **24**, 28–34.

Grime, J.P. (1974) Vegetation classification by reference to strategies. *Nature*, **250**, 26–31.

Grime, J.P. (1977) Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. *The American Naturalist*, **111**, 1169–1194.

Grime, J.P. (2001) *Plant Strategies, Vegetation Processes, and Ecosystem Properties*, 2nd ed. John Wiley & Sons, Chichester, UK.

Grubb, P.J., Coomes, D.A. & Metcalfe, D.J. (2005) Comment on "A brief history of seed size." *Science*, **310**, 783–783.

Gulliksen, S., Birks, H.H., Possnert, G. & Mangerud, J. (1998) A calendar age estimate of the Younger Dryas-Holocene boundary at Krakenes, western Norway. *The Holocene*, **8**, 249–259.

Guo, J., Wang, Y., Song, C., Zhou, J., Qiu, L., Huang, H. & Wang, Y. (2010) A single origin and moderate bottleneck during domestication of soybean (Glycine max): implications from microsatellites and nucleotide sequences. *Annals of Botany*, **106**, 505–514.

Gupta, Y.P. (1987) Anti-nutritional and toxic factors in food legumes: a review. *Plant Foods for Human Nutrition*, **37**, 201–228.

Gupta, A.K. (2004) Origin of agriculture and domestication of plants and animals linked to early Holocene climate amelioration. *Current Science*, **87**, 54–59.

Gutiérrez Salgado, A., Gepts, P. & Debouck, D.G. (1995) Evidence for two gene pools of the Lima bean, Phaseolus lunatus L., in the Americas. *Genetic Resources and Crop Evolution*, **42**, 15–28.

Hammer, K. (1984) Das Domestikationssyndrom. *Die Kulturpflanze*, **32**, 11–34.

Hammer, K. (2004) *Resolving the Challenge Posed by Agrobiodiversity and Plant Genetic Resources: An Attempt*. Kassel University Press, Kassel, Germany.

Hanson, V.D. (1999) *The Other Greeks: The Family Farm and the Agrarian Roots of Western Civilization*. University of California Press.

Harlan, J.R. (1965) The possible role of weed races in the evolution of cultivated plants. *Euphytica*, **14**, 173–176.

Harlan, J.R. (1967) A wild wheat harvest in Turkey. *Archaeology*, **20**, 197–201.

Harlan, J.R. (1971) Agricultural origins: centers and noncenters. *Science*, **174**, 468–474.

Harlan, J.R. (1976) *Origins of African Plant Domestication*. Walter de Gruyter.

Harlan, J.R. (1992) *Crops & Man*, 2nd ed. American Society of Agronomy & Crop Science Society of America, Madison, Wisconsin, USA.

Harlan, J.R., de Wet, J.M.J. & Price, E.G. (1973) Comparative evolution of cereals. *Evolution*, **27**, 311–325.

Hartley, M.L., Tshamekeng, E. & Thomas, S.M. (2002) Functional heterostyly in Tylosema esculentum (Caesalpinioideae). *Annals of Botany*, **89**, 67–76.

Hayden, B. (2009) The proof is in the pudding: feasting and the origins of domestication. *Current Anthropology*, **50**, 597–601.

Heiser, C.B. (1988) Aspects of unconscious selection and the evolution of domesticated plants. *Euphytica*, **37**, 77–81.

Hemsley, W.B. (1892) A Drift-seed (Ipomoea tuberosa, L.). *Annals of Botany*, 369–372.

Heun, M., Schäfer-Pregl, R., Klawan, D., Castagna, R., Accerbi, M., Borghi, B. & Salamini, F. (1997) Site of einkorn wheat domestication identified by DNA fingerprinting. *Science*, **278**, 1312–1314.

Hillman, G.C. & Davies, M.S. (1990a) Domestication rates in wild-type wheats and barley under primitive cultivation. *Biological Journal of the Linnean Society*, **39**, 39–78.

Hillman, G.C. & Davies, M.S. (1990b) Measured domestication rates in wild wheats and barley under primitive cultivation, and their archaeological implications. *Journal of World Prehistory*, **4**, 157–222.

Hillman, G., Hedges, R., Moore, A., Colledge, S. & Pettitt, P. (2001) New evidence of Lateglacial cereal cultivation at Abu Hureyra on the Euphrates. *The Holocene*, **11**, 383–393.

Hillman, G.C., Mason, S., de Moulins, D. & Nesbitt, M. (1996) Identification of archaeological remains of wheat: the 1992 London workshop. *Circaea*, **12**, 195–209.

Hilu, K.W. (1993) Polyploidy and the evolution of domesticated plants. *American Journal of Botany*, **80**, 1494–1499.

Hilu, K.W. (1994) Evidence from RAPD markers in the evolution of Echinochloa millets (Poaceae). *Plant Systematics and Evolution*, **189**, 247–257.

Hodgson, J.G., Wilson, P.J., Hunt, R., Grime, J.P. & Thompson, K. (1999) Allocating C-S-R plant functional types: a soft approach to a hard problem. *Oikos*, **85**, 282–294.

Horton, P. (2000) Prospects for crop improvement through the genetic manipulation of photosynthesis: morphological and biochemical aspects of light capture. *Journal of Experimental Botany*, **51**, 475–485.

Howe, H.F. & Richter, W.M. (1982) Effects of seed size on seedling size in Virola surinamensis; a within and between tree analysis. *Oecologia*, **53**, 347–351.

Huamán, Z. & Spooner, D.M. (2002) Reclassification of landrace populations of cultivated potatoes (Solanum sect. Petota). *American Journal of Botany*, **89**, 947–965.

Hubbard, R.N.L.B. & al Azm, A. (1990) Quantifying preservation and distortion in carbonized seeds; and investigating the history of friké production. *Journal of Archaeological Science*, **17**, 103–106.

Hu, X., Li, T., Wang, J., Wang, Y., Baskin, C.C. & Baskin, J.M. (2013) Seed dormancy in four Tibetan Plateau Vicia species and characterization of physiological changes in response of seeds to environmental factors. *Seed Science Research*, **23**, 133–140.

Hunt, H.V., Campana, M.G., Lawes, M.C., Park, Y.-J., Bower, M.A., Howe, C.J. & Jones, M.K. (2011) Genetic diversity and phylogeography of broomcorn millet (Panicum miliaceum L.) across Eurasia. *Molecular Ecology*, **20**, 4756–4771.

Ingenhousz, J. (1779) *Experiments Upon Vegetables: Discovering Their Great Power of Purifying the Common Air in the Sun-shine, and of Injuring It in the Shade and at Night.* Printed for P. Elmsly and H. Payne, London.

International Commission on Zoological Nomenclature. (1999) *International Code of Zoological Nomenclature*, 4th Revised edition. International Trust for Zoological Nomenclature.

International Legume Database & Information Service. (2005) ILDIS World Database of Legumes. URL http://www.ildis.org/ [accessed 15 March 2012]

Iorizzo, M., Senalik, D.A., Ellison, S.L., Grzebelus, D., Cavagnaro, P.F., Allender, C., Brunet, J., Spooner, D.M., Deynze, A.V. & Simon, P.W. (2013) Genetic structure and domestication of carrot (Daucus carota subsp. sativus) (Apiaceae). *American Journal of Botany*.

Iqbal, A., Khalil, I.A., Ateeq, N. & Sayyar Khan, M. (2006) Nutritional quality of important food legumes. *Food Chemistry*, **97**, 331–335.

Islebe, G.A. & Hooghiemstra, H. (2006) Effects of the Younger Dryas cooling event on late quaternary montane oak forest in Costa Rica. *Ecology and Conservation of Neotropical Montane Oak Forests* Ecological Studies. (ed M. Kapelle), pp. 29–37. Spinger, Heidelberg, Germany.

Iwanaga, M., Freyre, R. & Watanabe, K. (1991) Breaking the crossability barriers between disomic tetraploid Solanum acaule and tetrasomic tetraploid S. tuberosum. *Euphytica*, **52**, 183–191.

Jackson, M.T., Hawkes, J.G. & Rowe, P.R. (1980) An ethnobotanical field study of promitive potato varieties in Peru. *Euphytica*, **29**, 107–113.

Jakobsson, A. & Eriksson, O. (2000) A comparative study of seed number, seed size, seedling size and recruitment in grassland plants. *Oikos*, **88**, 494–502.

Janzen, D.H. (1977) Variation in seed size within a crop of a Costa Rican Mucuna andreana (Leguminosae). *American Journal of Botany*, **64**, 347–349.

Johns, T. & Keen, S.L. (1986) Ongoing evolution of the potato on the Altiplano of Western Bolivia. *Economic Botany*, **40**, 409–424.

Johnson, R.C., Mornhinweg, D.W., Ferris, D.M. & Heitholt, J.J. (1987) Leaf photosynthesis and conductance of selected Triticum

species at different water potentials. *Plant Physiology*, **83**, 1014–1017.

Jones, G. (2005) Garden cultivation of staple crops and its implications for settlement location and continuity. *World Archaeology*, **37**, 164–176.

Jones, S.E., Hunt, C. & Reimer, P.J. (2013) A 2300 yr record of sago and rice use from the southern Kelabit Highlands of Sarawak, Malaysian Borneo. *The Holocene*, **23**, 708–720.

Jones, M.K. & Liu, X. (2009) Origins of agriculture in East Asia. *Science*, **324**, 730–731.

Jones, G., Valamoti, S. & Charles, M. (2000) Early crop diversity: A "new" glume wheat from northern Greece. *Vegetation History and Archaeobotany*, **9**, 133–146.

Judkins, G., Smith, M. & Keys, E. (2008) Determinism within human–environment research and the rediscovery of environmental causation. *Geographical Journal*, **174**, 17–29.

Kaplan, D. (2000) The darker side of the "original affluent society." *Journal of Anthropological Research*, **56**, 301–324.

Kerem, Z., Lev-Yadun, S., Gopher, A., Weinberg, P. & Abbo, S. (2007) Chickpea domestication in the Neolithic Levant through the nutritional perspective. *Journal of Archaeological Science*, **34**, 1289–1293.

Khan, M.A. & Tsunoda, S. (1970) Evolutionary trends in leaf photosynthesis and related leaf characters among cultivated wheat species and its wild relatives. *Japanese Journal of Breeding*, **20**, 133–140.

Khazaei, J., Jafari, S. & Noorolah, S. (2008) Lognormal vs. Normal and Weibull distributions for modeling the mass and size distributions of sunflower seeds and kernels. *Agricultural Information and IT* (eds T. Nagatsuka & S. Ninomiya), pp. 7–16. Tokyo University of Agriculture, Tokyo.

Khurana, E., Sagar, R. & Singh, J.S. (2006) Seed size: a key trait determining species distribution and diversity of dry tropical forest in northern India. *Acta Oecologica*, **29**, 196–204.

Kimata, M., Ashok, E.G. & Seetharam, A. (2000) Domestication, cultivation and utilization of two small millets, Brachiaria ramosa and Setaria glauca (Poaceae), in South India. *Economic Botany*, **54**, 217–227.

Kimber, C.T. (2000) Origins of domesticated sorghum and its early

diffusion to India and China. *Sorghum: Origin, History, Technology, and Production* (eds C.W. Smith & R.A. Frederiksen), John Wiley & Sons.

Kislev, M.E., Hartmann, A. & Bar-Yosef, O. (2006a) Early domesticated fig in the Jordan Valley. *Science*, **312**, 1372–1374.

Kislev, M.E., Hartmann, A. & Bar-Yosef, O. (2006b) Response to comment on "Early domesticated fig in the Jordan Valley." *Science*, **314**, 1683b.

Kislev, M., Nadel, D. & Carmi, I. (1992) Epipalaeolithic (19,000 BP) cereal and fruit diet at Ohalo II, Sea of Galilee, Israel. *Review of Palaeobotany and Palynology*, **73**, 161–166.

Kitajima, K., Mulkey, S.S., Samaniego, M. & Wright, S.J. (2002) Decline of photosynthetic capacity with leaf age and position in two tropical pioneer tree species. *American Journal of Botany*, **89**, 1925–1932.

Klu, G.Y.P. (1996) *Efforts to Accelerate Domestication of Winged Bean (Psophocarpus Tetragonolobus (L.) DC.) by Means of Induced Mutations and Tissue Culture*. Landbouwuniversiteit Wageningen, Wageningen, Netherlands.

Kluyver, T.A. & Osborne, C.P. (2013) Taxonome: a software package for linking biological species data. *Ecology and Evolution*, **3**, 1262–1265.

Kochert, G., Stalker, H.T., Gimenes, M., Galgaro, L., Lopes, C.R. & Moore, K. (1996) RFLP and Cytogenetic Evidence on the Origin and Evolution of Allotetraploid Domesticated Peanut, Arachis hypogaea (Leguminosae). *American Journal of Botany*, **83**, 1282–1291.

Kochieva, E.Z., Goryunova, S.V. & Pomortsev, A.A. (2001) RAPD Analysis of the Genome in Species of the Genus Hordeum. *Russian Journal of Genetics*, **37**, 905–910.

Kokubun, M. & Wardlaw, I.F. (1988) Temperature adaptation of Glycine species as expressed by germination, photosynthesis, photosynthate accumulation and growth. *Japanese journal of crop science*, **57**, 211–219.

Koziol, L., Rieseberg, L.H., Kane, N. & Bever, J.D. (2012) Reduced drought tolerance during domestication and the evolution of weediness results from tolerance–growth trade-offs. *Evolution*, **66**, 3803–3814.

Kuroda, E. & Kumura, A. (1990) Difference in single leaf photosynthesis between old and new rice varieties : I.

Single-leaf photosynthesis and its dependence on stomatal conductance. *Japanese journal of crop science*, **59**, 283–292.

Ladizinsky, G. (1975) Collection of wild cereals in the upper Jordan Valley. *Economic Botany*, **29**, 264–267.

Ladizinsky, G. (1987) Pulse domestication before cultivation. *Economic Botany*, **41**, 60–65.

Ladizinsky, G. (1993) Lentil domestication: on the quality of evidence and arguments. *Economic Botany*, **47**, 60–64.

Lal, R., Reicosky, D.C. & Hanson, J.D. (2007) Evolution of the plow over 10,000 years and the rationale for no-till farming. *Soil and Tillage Research*, **93**, 1–12.

Larsen, C.S. (2006) The agricultural revolution as environmental catastrophe: Implications for health and lifestyle in the Holocene. *Quaternary International*, **150**, 12–20.

Lebot, V. (2009) *Tropical Root and Tuber Crops: Cassava, Sweet Potato, Yams and Aroids*. CABI.

Lee, R.B. (1968) What hunters do for a living, or, how to make out on scarce resources. *Man the Hunter* (eds R.B. Lee & I. DeVore), pp. 30–48. Aldine Transaction, New Jersey, USA.

Lee, G.-A., Crawford, G.W., Liu, L., Sasaki, Y. & Chen, X. (2011) Archaeological soybean (Glycine max) in East Asia: does size matter? *PLoS ONE*, **6**, e26720.

Leishman, M.R. & Westoby, M. (1994) The role of seed size in seedling establishment in dry soil conditions -- experimental evidence from semi-arid species. *Journal of Ecology*, **82**, 249–258.

Leishman, M.R., Westoby, M. & Jurado, E. (1995) Correlates of seed size variation: a comparison among five temperate floras. *Journal of Ecology*, **83**, 517–529.

Lempiäinen, T. (1989) Germination of the seeds of ulluco (Ullucus tuberosus, Basellaceae). *Economic Botany*, **43**, 456–463.

Linder, C.R., Taha, I., Seiler, G.J., Snow, A.A. & Rieseberg, L.H. (1998) Long-term introgression of crop genes into wild sunflower populations. *TAG Theoretical and Applied Genetics*, **96**, 339–347.

Liu, H., Edwards, E.J., Freckleton, R.P. & Osborne, C.P. (2012) Phylogenetic niche conservatism in C4 grasses. *Oecologia*, **170**, 835–845.

Liu, F. & Stützel, H. (2004) Biomass partitioning, specific leaf area, and

water use efficiency of vegetable amaranth (Amaranthus spp.) in response to drought stress. *Scientia Horticulturae*, **102**, 15–27.

Li, Q.Y., Zhao, W.Z. & Fang, H.Y. (2006) Effects of sand burial depth and seed mass on seedling emergence and growth of Nitraria sphaerocarpa. *Plant Ecology*, **185**, 191–198.

Lloyd, J., Bloomfield, K., Domingues, T.F. & Farquhar, G.D. (2013) Photosynthetically relevant foliar traits correlating better on a mass vs an area basis: of ecophysiological relevance or just a case of mathematical imperatives and statistical quicksand? *New Phytologist*, **199**, 311–321.

Londo, J.P., Chiang, Y.-C., Hung, K.-H., Chiang, T.-Y. & Schaal, B.A. (2006) Phylogeography of Asian wild rice, Oryza rufipogon, reveals multiple independent domestications of cultivated rice, Oryza sativa. *Proceedings of the National Academy of Sciences*, **103**, 9578–9583.

Long, S.P. & Hällgren, J.-E. (1993) Measurement of CO2 assimilation by plants in the field and the laboratory. *Photosynthesis and production in a changing environment* (eds D.O. Hall, J.M.. Scurlock, H.R. Bolhàr-Nordenkampf, R.C. Leegood & S.P. Long), pp. 129–167. Chapman and Hall, London.

Long, S.P., Zhu, X.-G., Naidu, S.L. & Ort, D.R. (2006) Can improvement in photosynthesis increase crop yields? *Plant, Cell & Environment*, **29**, 315–330.

Lonsdale, W.M. (1990) The self-thinning rule: dead or alive? *Ecology*, **71**, 1373–1388.

Lush, W.M. & Rawson, H.M. (1979) Effects of domestication and region of origin on leaf gas exchange in cowpea (Vignua unguiculata (L.) Walp.). *Photosynthetica*, **13**, 419–427.

Lush, W.M. & Wien, H.C. (1980) The importance of seed size in early growth of wild and domesticated cowpeas. *The Journal of Agricultural Science*, **94**, 177–182.

Lu, H., Zhang, J., Liu, K., Wu, N., Li, Y., Zhou, K., Ye, M., Zhang, T., Zhang, H., Yang, X., Shen, L., Xu, D. & Li, Q. (2009) Earliest domestication of common millet (Panicum miliaceum) in East Asia extended to 10,000 years ago. *Proceedings of the National Academy of Sciences*, **106**, 7367–7372.

Maass, B.L. (2005) Changes in seed morphology, dormancy and germination from wild to cultivated hyacinth bean germplasm (Lablab purpureus: Papilionoideae). *Genetic Resources and Crop Evolution*, **53**, 1127–1135.

Madon, O. & Médail, F. (1997) The ecological significance of annuals on a Mediterranean grassland (Mt Ventoux, France). *Plant Ecology*, **129**, 189–199.

Van der Maesen, L.J.G. (1995) Pigeonpea. *Evolution of crop plants* (eds J. Smartt & N.W. Simmonds), pp. 251–255. Longman Scientific and Technical, London.

Maisels, C.K. (1993) *The Emergence of Civilisation: From Hunting and Gathering to Agriculture, Cities and the State of the Near East*. Routledge.

Malagamba, J.P. & Monares, A. (1988) *True Potato Seed: Past and Present Uses*. International Potato Center (CIP), Lima, Peru.

Malice, M. (2009) *Genetic Diversity and Structure of Three Andean Tubers: Oxalis Tuberosa Molina, Ullucus Tuberosus Caldas and Tropaeolum Tuberosum Ruiz & Pav.* Ph.D. Thesis, Gembloux Agricutural University, Belgium.

Malice, M., Vogt, C.L.V., Pissard, A., Arbizu, C. & Baudoin, J.-P. (2009) Genetic diversity of the Andean tuber crop species Ullucus tuberosus as revealed by molecular (ISSR) and morphological markers. *Belgian Journal of Botany*, **142**, 68–82.

Manga, V.K. & Yadav, O.P. (1995) Effect of seed size on developmental traits andability to tolerate drought in pearl millet. *Journal of Arid Environments*, **29**, 169–172.

Markison, S., Thompson, B.L., Smith, J.C. & Spector, A.C. (2000) Time course and pattern of compensatory ingestive behavioral adjustments to lysine deficiency in rats. *J. Nutr.*, **130**, 1320–1328.

Marshall, F. & Hildebrand, E. (2002) Cattle before crops: the beginnings of food production in Africa. *Journal of World Prehistory*, **16**, 99–143.

Martin, F.W. & Cabanillas, E. (1966) Post-pollen-germination barriers to seed set in sweet-potato. *Euphytica*, **15**, 404–411.

Martin, F.W. & Cabanillas, E. (1976) Leren (Calathea allouia), a little known tuberous root crop of the Caribbean. *Economic Botany*, **30**, 249–256.

Massad, T.J., Dyer, L.A. & Vega C., G. (2012) Costs of defense and a test of the carbon-nutrient balance and growth-differentiation balance hypotheses for two co-occurring classes of plant defense. *PLoS ONE*, **7**, e47554.

Massei, G. & Hartley, S.E. (2000) Disarmed by domestication? Induced

responses to browsing in wild and cultivated olive. *Oecologia*, **122**, 225–231.

Matsuoka, Y., Vigouroux, Y., Goodman, M.M., G, J.S., Buckler, E. & Doebley, J. (2002) A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences*, **99**, 6080–6084.

McNeill, J., Barrie, F.R., Burdet, H.M., Demoulin, V., Hawksworth, D.L., Marhold, K., Nicolson, D.H., Prado, J., Silva, P.C., Skog, J.E., Wiersema, J.H. & Turland, N.J. (eds). (2006) *International Code of Botanical Nomenclature (Vienna Code). Regnum Vegetabile 146.* Gantner Verlag KG, Ruggell, Liechtenstein.

Meadows, J. (2005) The Younger Dryas episode and the radiocarbon chronologies of the Lake Huleh and Ghab Valley pollen diagrams, Israel and Syria. *The Holocene*, **15**, 631–636.

Mendoza, E. & Dirzo, R. (2007) Seed-size variation determines interspecific differential predation by mammals in a neotropical rain forest. *Oikos*, **116**, 1841–1852.

Menezes, E.G., Franke, L.B. & Dall'agnol, M. (2004) Seed yield components and seed yield of Adesmia latifolia (Spreng.) in two physiographic regions of Rio Grande do Sul. *Revista Brasileira de Sementes*, **26**, 25–32.

Miller, J., Funk, V., Wagner, W., Barrie, F.R., Hoch, P.C. & Herendeen, P.S. (2011) Outcomes of the 2011 Botanical Nomenclature Section at the XVIII International Botanical Congress. *PhytoKeys*, **5**, 1.

Miralles, D.J. & Slafer, G.A. (2007) Sink limitations to yield in wheat. *The Journal of Agricultural Science*, **145**, 139–149.

Molina-Cano, J.-L., Russell, J.R., Moralejo, M.A., Escacena, J.L., Arias, G. & Powell, W. (2005) Chloroplast DNA microsatellite analysis supports a polyphyletic origin for barley. *TAG Theoretical and Applied Genetics*, **110**, 613–619.

Monnin, E., Indermuhle, A., Dallenbach, A., Fluckiger, J., Stauffer, B., Stocker, T.F., Raynaud, D. & Barnola, J.-M. (2001) Atmospheric CO2 Concentrations over the Last Glacial Termination. *Science*, **291**, 112–114.

Morrell, P.L. & Clegg, M.T. (2007) Genetic evidence for a second domestication of barley (Hordeum vulgare) east of the Fertile Crescent. *Proceedings of the National Academy of Sciences*, **104**, 3289–3294.

Mosse, J. (1990) Nitrogen-to-protein conversion factor for ten cereals

and six legumes or oilseeds. A reappraisal of its definition and determination. Variation according to species and to seed protein content. *Journal of Agricultural and Food Chemistry*, **38**, 18–24.

Murphy, D.J. (2007) *People, Plants and Genes: The Story of Crops and Humanity*. OUP Oxford.

Nabhan, G. & de Wet, J.M.J. (1984) Panicum sonorum in Sonoran Desert Agriculture. *Economic Botany*, **38**, 65–82.

Nabity, P.D., Zavala, J.A. & DeLucia, E.H. (2013) Herbivore induction of jasmonic acid and chemical defences reduce photosynthesis in Nicotiana attenuata. *Journal of Experimental Botany*, **64**, 685–694.

Nakano, H., Makino, A. & Mae, T. (1995) Effects of panicle removal on the photosynthetic characteristics of the flag leaf of rice plants during the ripening stage. *Plant and Cell Physiology*, **36**, 653–659.

Nakano, H., Muramatsu, S., Makino, A. & Mae, T. (2000) Relationship between the suppression of photosynthesis and starch accumulation in the pod-removed bean. *Functional Plant Biology*, **27**, 167–173.

Nautiyal, P.., Rachaputi, N.R. & Joshi, Y.. (2002) Moisture-deficit-induced changes in leaf-water content, leaf carbon exchange rate and biomass production in groundnut cultivars differing in specific leaf area. *Field Crops Research*, **74**, 67–79.

Osnas, J.L.D., Lichstein, J.W., Reich, P.B. & Pacala, S.W. (2013) Global leaf trait relationships: mass, area, and the leaf economics spectrum. *Science*, **340**, 741–744.

Ovchinnikova, A., Krylova, E., Gavrilenko, T., Smekalova, T., Zhuk, M., Knapp, S. & Spooner, D.M. (2011) Taxonomy of cultivated potatoes (Solanum section Petota: Solanaceae). *Botanical Journal of the Linnean Society*, **165**, 107–155.

Pallais, N., Villagarcia, S., Fong, N., Tapia, J. & Garcia, R. (1987) Effect of supplemental nitrogen on true potato seed weight. *American Potato Journal*, **64**, 483–491.

Parera, C.A. & Ruiz, M. (2003) Adesmia subterranea Clos germination physiology and presowing treatments. *Journal of Range Management*, **56**, 273–276.

Paul, M.J. & Foyer, C.H. (2001) Sink regulation of photosynthesis. *Journal of Experimental Botany*, **52**, 1383–1400.

Pearson, T.R.H., Burslem, D.F.R.P., Mullins, C.E. & Dalling, J.W. (2002) Germination ecology of neotropical pioneers: interacting effects of environmental conditions and seed size. *Ecology*, **83**, 2798–2807.

Peleg, Z., Fahima, T., Korol, A.B., Abbo, S. & Saranga, Y. (2011) Genetic analysis of wheat domestication and evolution under domestication. *Journal of Experimental Botany*, **62**, 5051–5061.

Peng, S., Krieg, D.R. & Girma, F.S. (1991) Leaf photosynthetic rate is correlated with biomass and grain production in grain sorghum lines. *Photosynthesis Research*, **28**, 1–7.

Peng, J., Ronin, Y., Fahima, T., Röder, M.S., Li, Y., Nevo, E. & Korol, A. (2003) Domestication quantitative trait loci in Triticum dicoccoides, the progenitor of wheat. *Proceedings of the National Academy of Sciences*, **100**, 2489–2494.

Perry, L., Dickau, R., Zarrillo, S., Holst, I., Pearsall, D.M., Piperno, D.R., Berman, M.J., Cooke, R.G., Rademaker, K., Ranere, A.J., Raymond, J.S., Sandweiss, D.H., Scaramelli, F., Tarble, K. & Zeidler, J.A. (2007) Starch fossils and the domestication and dispersal of chili peppers (Capsicum spp. L.) in the Americas. *Science*, **315**, 986–988.

Petersen, G., Seberg, O., Yde, M. & Berthelsen, K. (2006) Phylogenetic relationships of Triticum and Aegilops and evidence for the origin of the A, B, and D genomes of common wheat (Triticum aestivum). *Molecular Phylogenetics and Evolution*, **39**, 70–82.

Pichare, M.M. & Kachole, M.S. (1996) Protease inhibitors of pigeonpea (Cajanus cajan) and its wild relatives. *Physiologia Plantarum*, **98**, 845–851.

Pickersgill, B. (2007) Domestication of plants in the Americas: insights from Mendelian and molecular genetics. *Annals of Botany*, **100**, 925–940.

Piperno, D.R. & Dillehay, T.D. (2008) Starch grains on human teeth reveal early broad crop diet in northern Peru. *Proceedings of the National Academy of Sciences*, **105**, 19622–19627.

Piperno, D.R., Ranere, A.J., Holst, I. & Hansell, P. (2000) Starch grains reveal early root crop horticulture in the Panamanian tropical forest. *Nature*, **407**, 894–897.

Piperno, D.R., Ranere, A.J., Holst, I., Iriarte, J. & Dickau, R. (2009) Starch grain and phytolith evidence for early ninth millennium B.P. maize from the Central Balsas River Valley, Mexico. *Proceedings of the National Academy of Sciences*, **106**, 5019–5024.

Piperno, D.R., Weiss, E., Holst, I. & Nadel, D. (2004) Processing of wild cereal grains in the Upper Palaeolithic revealed by starch grain analysis. *Nature*, **430**, 670–673.

Portères, R. (1976) African cereals: Eleusine, Fonio, Black Fonio, Teff, Brachiaria, Paspalum, Pennisetum and African rice. *Origins of African Plant Domestication* (ed J.R. Harlan), Walter de Gruyter.

Prendergast, M.E., Yuan, J. & Bar-Yosef, O. (2009) Resource intensification in the Late Upper Paleolithic: a view from southern China. *Journal of Archaeological Science*, **36**, 1027–1037.

Pujol, B., David, P. & McKey, D. (2005a) Microevolution in agricultural environments: how a traditional Amerindian farming practice favours heterozygosity in cassava (Manihot esculenta Crantz, Euphorbiaceae). *Ecology Letters*, **8**, 138–147.

Pujol, B., Gigot, G., Laurent, G., Pinheiro-Kluppel, M., Elias, M., Hossaert-McKey, M. & McKey, D. (2002) Germination ecology of cassava (Manihot esculenta Crantz, Euphorbiaceae) in traditional agroecosystems: seed and seedling biology of a vegetatively propagated domesticated plant. *Economic Botany*, **56**, 366–379.

Pujol, B., Mühlen, G., Garwood, N., Horoszowski, Y., Douzery, E.J.P. & McKey, D. (2005b) Evolution under domestication: contrasting functional morphology of seedlings in domesticated cassava and its closest wild relatives. *New Phytologist*, **166**, 305–318.

Pujol, B., Salager, J., Beltran, M., Bousquet, S. & McKey, D. (2008) Photosynthesis and leaf structure in domesticated cassava (Euphorbiaceae) and a close wild relative: have leaf photosynthetic parameters evolved under domestication? *Biotropica*, **40**, 305–312.

Pulse Western Australia. (2013) About Lupins. URL http://www.lupins.org/lupins/ [accessed 17 May 2013]

Purugganan, M.D. & Fuller, D.Q. (2009) The nature of selection during plant domestication. *Nature*, **457**, 843–848.

Putnam, D.H., Oplinger, E.S., Teynor, T.M., Oelke, E.A., Kelling, K.A. & Doll, J.D. (1991) Peanut. URL http://www.hort.purdue.edu/newcrop/afcm/peanut.html [accessed 15 June 2013]

Du Puy, D.J., Labat, J.N., Rabevohitra, R. & Villiers, J.F. (2002) *The Leguminosae of Madagascar*. Royal Botanic Gardens, Kew, London.

Qiu, J. & Mosjidis, J.A. (1993) Influence of seed weight and planting depth on common vetch establishment and growth. *Field Crops Research*, **33**, 353–366.

Quiros, C.F., Ortega, R., Raamsdonk, L. van, Herrera-Montoya, M., Cisneros, P., Schmidt, E. & Brush, S.B. (1992) Increase of potato genetic resources in their center of diversity: the role of natural outcrossing and selection by the Andean farmer. *Genetic Resources and Crop Evolution*, **39**, 107–113.

Reader, R.J. (1993) Control of seedling emergence by ground cover and seed predation in relation to seed size for some old-field species. *Journal of Ecology*, **81**, 169–175.

Rees, M. (1995) Community structure in sand dune annuals: is seed weight a key quantity? *Journal of Ecology*, **83**, 857–863.

Rees, M. & Venable, D.L. (2007) Why do big plants make big seeds? *Journal of Ecology*, **95**, 926–936.

Richards, R.A. (2000) Selectable traits to increase crop photosynthesis and yield of grain crops. *Journal of Experimental Botany*, **51**, 447–458.

Richerson, P.J., Boyd, R. & Bettinger, R.L. (2001) Was agriculture impossible during the Pleistocene but mandatory during the Holocene? A climate change hypothesis. *American Antiquity*, **66**, 387–411.

Ries, S.K. & Everson, E.H. (1973) Protein content and seed size relationships with seedling vigor of wheat cultivars. *Agronomy Journal*, **65**, 884–886.

Rindos, D. (1984) *The Origins of Agriculture: An Evolutionary Perspective*. Academic Press, New York.

Robinson, R.G. (1975) *Pulse or Grain Legume Crops for Minnesota*. Agricultural Experiment Station, University of Minnesota.

Rodbell, D.T. (2000) The Younger Dryas: cold, cold everywhere? *Science*, **290**, 285–286.

Rosenthal, J.P. & Dirzo, R. (1997) Effects of life history, domestication and agronomic selection on plant defence against insects: Evidence from maizes and wild relatives. *Evolutionary Ecology*, **11**, 337–355.

Rousi, A., Jokela, P., Kalliola, R., Pietilä, L., Salo, J. & Yli-Rekola, M. (1989) Morphological variation among clones of ulluco (Ullucus tuberosus, Basellaceae) collected in Southern Peru. *Economic Botany*, **43**, 58–72.

Rowland, M. (2008) Lathyrus Species. URL http://www.lathyrus.info/species/species.html [accessed 17 May 2013]

Royal Botanic Gardens, Kew. (2008) Seed Information Database. URL http://data.kew.org/sid/ [accessed 10 April 2012]

Roy, T.S., Nishizawa, T. & Ali, M.H. (2007) Flower, berry and true potato seed productions in potato mother plants (Solanum tuberosum L.). 1. Effects of nitrogen and phosphorus fertilizers. *Journal of Agronomy*, **6**, 106–112.

Sage, R.F. (1995) Was low atmospheric CO2 during the Pleistocene a limiting factor for the origin of agriculture? *Global Change Biology*, **1**, 93–106.

Salick, J., Cellinese, N. & Knapp, S. (1997) Indigenous diversity of cassava: generation, maintenance, use and loss among the Amuesha, Peruvian Upper Amazon. *Economic Botany*, **51**, 6–19.

Sallon, S., Solowey, E., Cohen, Y., Korchinsky, R., Egli, M., Woodhatch, I., Simchoni, O. & Kislev, M. (2008) Germination, genetics, and growth of an ancient date seed. *Science*, **320**, 1464.

Sammour, R.H., Mustafa, A.E.-Z., Badr, S. & Tahr, W. (2007) Genetic variability of some quality traits in Lathyrus spp. germplasm. *Acta agriculturae Slovenica*, **90**, 33–43.

Schimpf, D.J. (1977) Seed Weight of Amaranthus Retroflexus in Relation to Moisture and Length of Growing Season. *Ecology*, **58**, 450–453.

Schloen, M., Peters, M. & Schultze-Kraft, R. (2013) Canavalia brasiliensis. URL http://www.fao.org/ag/AGP/AGPC/doc/Gbase/data/canbras.htm [accessed 17 May 2013]

Schlumbaum, A., Neuhaus, J.-M. & Jacomet, S. (1998) Coexistence of tetraploid and hexaploid naked wheat in a Neolithic lake dwelling of Central Europe: evidence from morphology and ancient DNA. *Journal of Archaeological Science*, **25**, 1111–1118.

Schnee, B.K. & Waller, D.M. (1986) Reproductive behavior of Amphicarpaea bracteata (Leguminosae), an amphicarpic annual. *American Journal of Botany*, **73**, 376–386.

Seijo, G., Lavia, G.I., Fernandez, A., Krapovickas, A., Ducasse, D.A., Bertioli, D.J. & Moscone, E.A. (2007) Genomic relationships between the cultivated peanut (Arachis hypogaea, Leguminosae) and its close relatives revealed by double GISH. *Am. J. Bot.*, **94**, 1963–1971.

Seiwa, K., Watanabe, A., Saitoh, T., Kannu, H. & Akasaka, S. (2002) Effects of burying depth and seed size on seedling establishment of Japanese chestnuts, Castanea crenata. *Forest Ecology and Management*, **164**, 149–156.

Sencer, H.A. & Hawkes, J.G. (1980) On the origin of cultivated rye. *Biological Journal of the Linnean Society*, **13**, 299–313.

Sharma, S.R., Dwivedi, S.K. & Swarup, D. (1997) Hypoglycaemic, antihyperglycaemic and hypolipidemic activities of Caesalpinia bonducella seeds in rats. *Journal of Ethnopharmacology*, **58**, 39–44.

Shigeta, M. (1990) Folk in-situ conservation of ensete [Ensete ventricosum (Welw.) E.E. Cheesman]: toward the interpretation of indigenous agricultural science of the Ari, Southwestern Ethiopia. *African Study Monographs*, **10**, 93–107.

Shomura, A., Izawa, T., Ebana, K., Ebitani, T., Kanegae, H., Konishi, S. & Yano, M. (2008) Deletion in a gene associated with grain size increased yields during rice domestication. *Nat Genet*, **40**, 1023–1028.

Siddhuraju, P. & Becker, K. (2005) Nutritional and antinutritional composition, in vitro amino acid availability, starch digestibility and predicted glycemic index of differentially processed mucuna beans (Mucuna pruriens var. utilis): an under-utilised legume. *Food Chemistry*, **91**, 275–286.

Siddhuraju, P., Becker, K. & Makkar, H.P.S. (2002) Chemical composition, protein fractionation, essential amino acid potential and antimetabolic constituents of an unconventional legume, Gila bean (Entada phaseoloides Merrill) seed kernel. *Journal of the Science of Food and Agriculture*, **82**, 192–202.

Siddique, K.H.M. & Loss, S.P. (1999) Studies on sowing depth for chickpea (Cicer arietinum L.), faba bean (Vicia faba L.) and lentil (Lens culinaris Medik) in a Mediterranean-type environment of south-western Australia. *Journal of Agronomy and Crop Science*, **182**, 105–112.

Siddique, K.H.M., Loss, S.P., French, B. & Veitch, C. (1997) *Sowing Depth for Chickpea, Faba Bean, Lentil and Field Pea*. Department of Agriculture and Food, Western Australia.

Simmonds, N.W. (1962) *The Evolution of the Bananas*. Longmans, London.

Singh, H. & Arora, R. (1972) Raishan (Digitaria sp.)—a minor millet of the Khasi Hills, India. *Economic Botany*, **26**, 376–380.

Singh, A.K. & Smartt, J. (1998) The genome donors of the groundnut/peanut (Arachis hypogaea L.) revisited. *Genetic Resources and Crop Evolution*, **45**, 113–116.

Slewinski, T.L. (2012) Non-structural carbohydrate partitioning in grass stems: a target to increase yield stability, stress tolerance, and biofuel production. *Journal of Experimental Botany*, **63**, 4647–4670.

Sluyter, A. (2003) Neo-environmental determinism, intellectual damage control, and nature/society science. *Antipode*, **35**, 813–817.

Smalley, J. & Blake, M. (2003) Sweet beginnings: stalk sugar and the domestication of maize. *Current Anthropology*, **44**, 675–703.

Smartt, J. (1985) Evolution of grain legumes. II. Old and New World pulses of lesser economic importance. *Experimental Agriculture*, **21**, 1–18.

Smartt, J. & Simmonds, N.W. (1995) *Evolution of Crop Plants*. Longman Scientific and Technical, London.

Smith, B.D. (1997) The initial domestication of Cucurbita pepo in the Americas 10,000 years ago. *Science*, **276**, 932–934.

Smith, B.D. (2006) Eastern North America as an independent center of plant domestication. *Proceedings of the National Academy of Sciences*, **103**, 12223–12228.

Smith, C.C. & Fretwell, S.D. (1974) The optimal balance between size and number of offspring. *The American Naturalist*, **108**, 499–506.

Sobrado, M.A. & Turner, N.C. (1986) Photosynthesis, dry matter accumulation and distribution in the wild sunflower Helianthus petiolaris and the cultivated sunflower Helianthus annuus as influenced by water deficits. *Oecologia*, **69**, 181–187.

Spooner, D.M., McLean, K., Ramsay, G., Waugh, R. & Bryan, G.J. (2005) A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 14694–14699.

Starling, A.P. & Stock, J.T. (2007) Dental indicators of health and stress in early Egyptian and Nubian agriculturalists: A difficult transition and gradual recovery. *American Journal of Physical Anthropology*, **134**, 520–528.

Steckel, R.H., Rose, J.C., Larsen, C.S. & Walker, P.L. (2002) Skeletal

health in the Western Hemisphere from 4000 B.C. to the present. *Evolutionary Anthropology: Issues, News, and Reviews*, **11**, 142–155.

Steele, K.P. & Wojciechowski, M.F. (2003) Phylogenetic analyses of tribes Trifolieae and Vicieae, based on sequences of the plastid gene matK (Papilionoideae: Leguminosae). *Advances in legume systematics, part*, **10**, 355–370.

Stiner, M.C. (2001) Thirty years on the "Broad Spectrum Revolution" and paleolithic demography. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 6993–6996.

Strauss, M.S., Michaud, J.D. & Arditti, J. (1979) Seed storage and germination and seedling proliferation in taro, Colocasia esculenta (L) Schott. *Annals of Botany*, **43**, 603–612.

Takahashi, R. (1955) The origin and evolution of cultivated barley. *Advances in Genetics* (ed M. Demerec), pp. 227–266. Academic Press.

Takano, Y. & Tsunoda, S. (1971) Curvilinear regression of the leaf photosynthetic rate on leaf nitrogen content among strains of Oryza species. *Japanese Journal of Breeding*, **21**, 69–76.

Tanno, K. & Willcox, G. (2006) How fast was wild wheat domesticated? *Science*, **311**, 1886.

Tanno, K. & Willcox, G. (2012) Distinguishing wild and domestic wheat and barley spikelets from early Holocene sites in the Near East. *Vegetation History and Archaeobotany*, **21**, 107–115.

Teixeira-Sá, D.M.A., Reicher, F., Braga, R.C., Beltramini, L.M. & de Azevedo Moreira, R. (2009) Isolation of a lectin and a galactoxyloglucan from Mucuna sloanei seeds. *Phytochemistry*, **70**, 1965–1972.

Tomooka, N. (2002) *The Asian Vigna: Genus Vigna Subgenus Ceratotropis Genetic Resources*. Springer.

Townsend, C.E. (1992) Seedling emergence of yellow-flowered alfalfa as influenced by seed weight and planting depth. *Agronomy Journal*, **84**, 821–826.

Turnbull, L.A., Paul-Victor, C., Schmid, B. & Purves, D.W. (2008) Growth rates, seed size and physiology: do small-seeded species really grow faster. *Ecology*, **89**, 1352–1363.

Turnbull, L.A., Philipson, C.D., Purves, D.W., Atkinson, R.L., Cunniff, J., Goodenough, A., Hautier, Y., Houghton, J., Marthews, T.R.,

Osborne, C.P., Paul-Victor, C., Rose, K.E., Saner, P., Taylor, S.H., Woodward, F.I., Hector, A. & Rees, M. (2012) Plant growth rates and seed size: a re-evaluation. *Ecology*, **93**, 1283–1289.

Turnbull, L.A., Rees, M. & Crawley, M.J. (1999) Seed mass and the competition/colonization trade-off: a sowing experiment. *Journal of Ecology*, **87**, 899–912.

Twiss, K. (2011) The complexities of home cooking: public feasts and private meals Inside the Çatalhöyük house. *eTopoi. Journal for Ancient Studies*, **Special volume 2**, 53–73.

Twiss, K.C., Bogaard, A., Bogdan, D., Carter, T., Charles, M.P., Farid, S., Russell, N., Stevanović, M., Nurcan Yalman, E. & Yeomans, L. (2008) Arson or Accident? The Burning of a Neolithic House at Çatalhöyük, Turkey. *Journal of Field Archaeology*, **33**, 41–57.

Tyagi, A.P., Taylor, M. & Deo, P.C. (2004) Seed germination and seedling development in Taro (Colocasia esculenta). *The South Pacific Journal of Natural and Applied Sciences*, **22**, 62–66.

Uauy, C., Distelfeld, A., Fahima, T., Blechl, A. & Dubcovsky, J. (2006) A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. *Science*, **314**, 1298–1301.

Ugent, D., Pozorski, S. & Pozorski, T. (1982) Archaeological potato tuber remains from the Casma Valley of Peru. *Economic Botany*, **36**, 182–192.

Upadhya, M.D. & Cabello, R. (2000) *Influence of Seed Size and Density on the Performance of Direct Seedling Transplants from Hybrid True Potato Seed*. CIP.

USDA Agricultural Research Service. (2012) GRIN National Genetic Resources Program. URL http://www.ars-grin.gov/ [accessed 5 March 2012]

USDA, NRCS. (2013) The PLANTS Database. Database. URL http://plants.usda.gov [accessed 17 May 2013]

Vamosi, S.M., Mazer, S.J. & Cornejo, F. (2008) Breeding systems and seed size in a neotropical flora: testing evolutionary hypotheses. *Ecology*, **89**, 2461–2472.

Vaughan, D.A., Lu, B.-R. & Tomooka, N. (2008) The evolving story of rice evolution. *Plant Science*, **174**, 394–408.

Vavilov, N.I. (1931) The problem concerning the origin of agriculture in the light of recent research [English translation]. *Socialist reconstruction and Science*, **1**, 34–43.

Vega, C.R.C., Andrade, F.H., Sadras, V.O., Uhart, S.A. & Valentinuz, O.R. (2001) Seed number as a function of growth. A comparative study in soybean, sunflower, and maize. *Crop Science*, **41**, 748.

Venable, D.L. (1992) Size-number trade-offs and the variation of seed size with plant resource status. *The American Naturalist*, **140**, 287–304.

Venable, D.L. & Rees, M. (2009) The scaling of seed size. *Journal of Ecology*, **97**, 27–31.

De Vries, I.M. (1997) Origin and domestication of Lactuca sativa L. *Genetic Resources and Crop Evolution*, **44**, 165–174.

Waines, J.G. (1978) Protein contents, grain weights, and breeding potential of wild and domesticated tepary beans. *Crop Science*, **18**, 587–589.

Walters, M.B. & Reich, P.B. (2000) Seed size, nitrogen supply, and growth rate affect tree seedling survival in deep shade. *Ecology*, **81**, 1887–1901.

Wang, J.H., Baskin, C.C., Cui, X.L. & Du, G.Z. (2009) Effect of phylogeny, life history and habitat correlates on seed germination of 69 arid and semi-arid zone species from northwest China. *Evolutionary Ecology*, **23**, 827–846.

Wang, J., Chen, W., Baskin, C.C., Baskin, J.M., Cui, X., Zhang, Y., Qiang, W. & Du, G. (2012) Variation in seed germination of 86 subalpine forest species from the eastern Tibetan Plateau: phylogeny and life-history correlates. *Ecological Research*, **27**, 453–465.

Weber, S.A. & Fuller, D.Q. (2008) Millets and their role in early agriculture. *Pragdhara*, **18**, 69–90.

Weiss, E., Kislev, M.E. & Hartmann, A. (2006) Autonomous cultivation before domestication. *Science*, **312**, 1608–1610.

Weiss, E., Kislev, M.E., Simchoni, O. & Nadel, D. (2004a) Small-grained wild grasses as staple food at the 23 000-year-old site of Ohalo II, Israel. *Economic Botany*, **58**, S125–S134.

Weiss, E., Wetterstrom, W., Nadel, D. & Bar-Yosef, O. (2004b) The broad spectrum revisited: Evidence from plant remains. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 9551–9555.

Weiss, E. & Zohary, D. (2011) The Neolithic Southwest Asian Founder Crops: their biology and archaeobotany. *Current Anthropology*, **52**, S237–S254.

Westoby, M. (1998) A leaf-height-seed (LHS) plant ecology strategy scheme. *Plant and Soil*, **199**, 213–227.

Westoby, M., Leishman, M., Lord, J., Poorter, H. & Schoen, D.J. (1996) Comparative ecology of seed size and dispersal. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **351**, 1309–1318.

White, J.W. & González, A. (1990) Characterization of the negative association between seed yield and seed size among genotypes of common bean. *Field Crops Research*, **23**, 159–175.

Whitlock, M.C. (2005) Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. *Journal of Evolutionary Biology*, **18**, 1368–1373.

Willcox, G. (2005) The distribution, natural habitats and availability of wild cereals in relation to their domestication in the Near East: multiple events, multiple centres. *Vegetation History and Archaeobotany*, **14**, 534–541.

Willcox, G., Buxo, R. & Herveux, L. (2009) Late Pleistocene and early Holocene climate and the beginnings of cultivation in northern Syria. *The Holocene*, **19**, 151–158.

Willcox, G., Fornite, S. & Herveux, L. (2008) Early Holocene cultivation before domestication in northern Syria. *Vegetation History and Archaeobotany*, **17**, 313–325.

Williams, P.R., Congdon, R.A., Grice, A.C. & Clarke, P.J. (2003) Fire-related cues break seed dormancy of six legumes of tropical eucalypt savannas in north-eastern Australia. *Austral Ecology*, **28**, 507–514.

Wu, G. & Du, G. (2007) Germination is related to seed mass in grasses (Poaceae) of the eastern Qinghai-Tibetan Plateau, China. *Nordic Journal of Botany*, **25**, 361–365.

Lev-Yadun, S., Ne'eman, G., Abbo, S. & Flaishman, M.A. (2006) Comment on "Early domesticated fig in the Jordan Valley." *Science*, **314**, 1683a.

Yang, X., Barton, H.J., Wan, Z., Li, Q., Ma, Z., Li, M., Zhang, D. & Wei, J. (2013) Sago-Type Palms Were an Important Plant Food Prior to Rice in Southern Subtropical China. *PLoS ONE*, **8**, e63148.

Yang, C., Yang, F., Ma, H. & Liu, P. (2010) Distribution and morphological variation of germplasm resource of Sophora alopecuroides. *Zhongguo Zhong yao za zhi (China journal of Chinese materia medica)*, **35**, 817–820.

Yen, D.E. (1960) The sweet potato in the Pacific: the propagation of the plant in relation to its distribution. *The Journal of the Polynesian Society*, **69**, 368–375.

Zeder, M.A. (2006) Central questions in the domestication of plants and animals. *Evolutionary Anthropology: Issues, News, and Reviews*, **15**, 105–117.

Zelitch, I. (1982) The close relationship between net photosynthesis and crop yield. *BioScience*, **32**, 796–802.

Zhang, S.T., Du, G.Z. & Chen, J.K. (2004) Seed size in relation to phylogeny, growth form and longevity in a subalpine meadow on the east of the Tibetan plateau. *Folia Geobotanica*, **39**, 129–142.

Zhao, Z. (2010) New data and new issues for the study of origin of rice agriculture in China. *Archaeological and Anthropological Sciences*, **2**, 99–105.

Zhou, X., Jellen, E.N. & Murphy, J.P. (1999) Progenitor germplasm of domesticated hexaploid oat. *Crop Science*, **39**, 1208–1214.

Zohary, D. (1989) Pulse domestication and cereal domestication: how different are they? *Economic Botany*, **43**, 31–34.

Zohary, D. (2004) Unconscious selection and the evolution of domesticated Plants. *Economic Botany*, **58**, 5–10.

Zohary, D. & Hopf, M. (2000) *Domestication of Plants in the Old World*. Oxford University Press.