Genetic evaluation of domestication-related traits in rice: implications for the archaeobotany of rice origins

[Final manuscript] published in *Archaeological and Anthropological Sciences* 12, 197 (2020). https://doi.org/10.1007/s12520-020-01112-3]

Ryo Ishikawa[1, 2] *, Cristina Castillo[2], Dorian Fuller[2]

1. Laboratory of Plant Breeding, Graduate School of Agricultural Science, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe 657-8501, Japan

2. Institute of Archaeology, University College London, London, WC1H 0PY, United Kingdom

* Corresponding Author:

Ryo Ishikawa e-mail: r-ishika@pot.kobe-u.ac.jp

ORCID

Ryo Ishikawa 0000-0003-1848-2493

Cristina Castillo 0000-0001-8491-3314

Dorian Fuller 0000-0002-4859-080X

Abstract

Domestication is the process in which preferred genetic changes in wild plants and animals have been selected by humans. In other words, domesticated plants have become adapted to being part of human-managed ecosystems. Asian cultivated rice, Oryza sativa L., is one of the most important crops in the world and is known to have been domesticated from its wild ancestor, O. rufipogon. Many morphological changes in cultivated rice have been beneficial to humans in terms of increased efficiency of cultivation and yield that supported the development of human civilisations. The genetic mechanisms of these changes have been extensively studied since rice genome sequences were determined, and based on genome analyses, the origin of rice has been widely discussed. Most of the domestication-related traits and genes are, however, often evaluated based on the genetic background of cultivated rice, leading to misinterpretation of rice domestication. Here, we review several genetic changes and discuss the importance of evaluating these traits in the wild rice genetic background to understand the process of rice domestication. In this review, we also provide a phenotypic evaluation of domestication-related traits.

Introduction

In prehistoric times, humans obtained their food by hunting and gathering wild plants to survive. The several transitions from hunting and gathering to farming took place in the prehistoric period when wild plants and animals were domesticated and resulted in many advantages (Larson et al. 2014). By cultivating domesticated plants in fields, people successfully obtained more food, which could support a growing population. During the process of domestication, people's activities modified certain traits of plants and animals to meet human needs and increase their reliance on the species being domesticated. These modifications were mainly achieved by selecting natural variations that resulted in preferred traits or traits better suited to human cultivation, such as non-seed shattering or an increase in grain size.

Many crops have been domesticated all over the world (Doebley et al. 2006). Of these, rice is one of the most important crops that provides the daily calories for almost half of the world's population (Khush 2005). There are mainly two species of cultivated rice: Asian cultivated rice, Oryza sativa L., and African cultivated rice, O. glaberrima Steud; these two species were domesticated from the ancestral wild rice species, O. rufipogon Griff and O. barthii A. Chev., respectively (Oka 1988). Asian cultivated rice is mainly classified into two subspecies, indica and japonica (Oka 1988; Fuller 2007). Currently, O. sativa is widely cultivated in the world, whereas O. glaberrima is cultivated only in Africa. The domestication of O. sativa has been extensively studied both through plant genetics and archaeology.

A breakthrough in rice domestication studies was achieved with the completion of the genome sequencing of the cultivated rice O. sativa subsp. indica (Yu et al. 2002) and subsp. japonica (Goff et al. 2002; International Rice Genome Sequencing Project 2005). Subsequent genetic studies have identified the loci responsible for some domestication-related traits and further genetic mapping has uncovered key genes and causal mutations involved in changes in these traits. These are detailed below.

Phylogenetic studies have also been possible based on a growing body of sequence data. However, these have resulted in alternative hypotheses with several ‘origins’ debated, both single and multiple (Kovach et al. 2007; Sang and Ge 2007a; Izawa et al. 2009; Panaud 2009; Gross and Zhao 2014; Civáň et al. 2015; Gutaker et al. 2020). A study by Huang et al. (2012) sequenced 446 accessions of O. rufipogon from Asia and 1083 cultivated rice strains of O. sativa subsp. indica and subsp. japonica. It was suggested using map-based sequence data that a specific population of O. rufipogon around the middle area of the Pearl River in southern China was the origin of rice domestication and that indica rice was produced by subsequent introgression of japonica with wild rice (Huang et al. 2012). Later, Civáň et al. (2015) re-analysed the sequence data of Huang et al. (2012) and proposed three geographically separate events in Asian rice domestication. Studies of maternal markers (Tang et al. 2004; Takahashi et al. 2008) and coalescent analyses have tended to support multiple origins (Choi et al. 2017), as do recent population genetic inferences from > 1400 whole genomes (Gutaker et al. 2020). Also, data from archaeological finds of sub-fossil rice (charred grains and chaff) favours a model of multiple domestications (Silva et al. 2015) but supports regions of origins that differ from those inferred from many of the genomic phylogenetic studies. Therefore, the area(s) of origin of rice domestication still remains unresolved.

To better understand the origin and chronology of rice domestication, we believe that more information related to phenotypic changes is required. Understanding the evolution of these phenotypic traits can inform both archaeological and phylogenetic approaches to origins, as they can be related to morphological features in sub-fossil material or to genotypes shared by phylogenetic linages. In this paper, we review some domestication-related traits and genes or loci that determine them, with a particular focus on two traits that likely played key roles in early rice domestication: closed panicles and the loss of seed shattering. Furthermore, we discuss some of the results from studies using wild rice introgression lines which provide information on different aspects of rice domestication.

Major traits selected for in rice domestication

Genetic variations in wild rice are important since mutations lead to agronomically important traits (Ishii et al. 2011). Over the past decade, several important genes involved in rice domestication have been identified and cloned. These traits are often linked to ease of cultivation and harvest as well as an increase in edible portions of rice (Table 1). These are also known to be involved in the domestication of maize, wheat, barley, sorghum, oats and other crops (Doebley et al. 2006). These recurrently selected common traits between species represent convergent evolution and are known as the domestication syndrome (Sakuma et al. 2011; Fuller 2007). The major traits selected in rice domestication with causal mutation and molecular function are shown in Table 1.

Plant architecture

An obvious change between wild and cultivated rice is observed in the plant architecture; wild rice has prostrate tiller morphology, while cultivated rice shows erect growth (Fig. 1a, b). In wild rice, prostrate morphology helps to widely distribute seeds in the growing area. However, prostate morphology is not a favoured trait in paddy fields whereas erect growth habit is a preferred trait for cultivation.

Two research groups identified PROG1, a gene responsible for prostate morphology, encoding a zinc-finger protein through molecular cloning of the causal gene (Table 1; Tan et al. 2008; Jin et al. 2008). Furthermore, one of the causal gene sequences at PROG1 was found to be an amino acid substitution by transformation experiments (Jin et al. 2008). However, other possibilities for the underlying mutation are also reported (Tan et al. 2008). These results suggest that gene sequences leading to prostrate tiller morphology may have arisen independently. Since the cultivated-type allele of PROG1 is observed in wild rice, it might have been selected in cultivation from standing variation in wild populations (Civáň and Brown 2017) and the phenotypic effect might be caused by an interaction with other mutation(s). Recently, a deletion in the region called SEMI-PROSTRATE GROWTH 1 (SPROG1), linked to PROG1 was also shown to be important for erect growth and higher yield in cultivars (Wu et al. 2018). These genes indicate that early selection under cultivation favoured erect plants that could be grown together more compactly, which parallels selection for apical dominance in other cereals (Fuller 2007; Fuller et al. 2010).

Panicle shape differs between wild and cultivated rice; wild rice has open panicles, whereas cultivated rice, O. sativa, has closed panicles (Fig. 1c, d). A closed panicle has the advantage of promoting self-fertilisation and for humans gathering wild rice, the closed panicle helps to retain the spikelets longer on the panicle, increasing harvest returns (Ishii et al. 2013). Experimental data indicates that harvesting of closed panicle rice, with otherwise wild features, could increase harvested yields by ~ 50% (Ishii et al. 2013). QTL analysis identified a single locus on chromosome 4 with 80.1% phenotypic variance (Ishii et al. 2013). The position was similar to the previously identified locus called SPR3 detected by genetic analysis of panicle shape between Asian cultivated rice, O sativa and African cultivated rice, O. glaberrima (Luo et al. 2008). Molecular genetic analyses showed that the 9.3-kb region is responsible for panicle shape, but no annotated genes were located in the region. A transformation study introducing a wild chromosomal segment into cultivated rice showed that the region together with a downstream gene, OsLG1, was sufficient to change panicle shape (Table 1; Ishii et al. 2013). Therefore, the region may act as a transcriptional regulator of OsLG1. Another study using the sequences from wild rice accessions showed that the one causal single nucleotide polymorphism (SNP) at the SPR3 region was associated with panicle shape (Zhu et al. 2013), although the causal mutation at SPR3 still remains to be elucidated.

Review

Published: 01 August 2020

Genetic evaluation of domestication-related traits in rice: implications for the archaeobotany of rice origins

Ryo Ishikawa, Cristina C. Castillo & Dorian Q. Fuller

Archaeological and Anthropological Sciences volume 12, Article number: 197 (2020) Cite this article

Abstract

Domestication is the process in which preferred genetic changes in wild plants and animals have been selected by humans. In other words, domesticated plants have become adapted to being part of human-managed ecosystems. Asian cultivated rice, Oryza sativa L., is one of the most important crops in the world and is known to have been domesticated from its wild ancestor, O. rufipogon. Many morphological changes in cultivated rice have been beneficial to humans in terms of increased efficiency of cultivation and yield that supported the development of human civilisations. The genetic mechanisms of these changes have been extensively studied since rice genome sequences were determined, and based on genome analyses, the origin of rice has been widely discussed. Most of the domestication-related traits and genes are, however, often evaluated based on the genetic background of cultivated rice, leading to misinterpretation of rice domestication. Here, we review several genetic changes and discuss the importance of evaluating these traits in the wild rice genetic background to understand the process of rice domestication. In this review, we also provide a phenotypic evaluation of domestication-related traits.

Introduction

In prehistoric times, humans obtained their food by hunting and gathering wild plants to survive. The several transitions from hunting and gathering to farming took place in the prehistoric period when wild plants and animals were domesticated and resulted in many advantages (Larson et al. 2014). By cultivating domesticated plants in fields, people successfully obtained more food, which could support a growing population. During the process of domestication, people's activities modified certain traits of plants and animals to meet human needs and increase their reliance on the species being domesticated. These modifications were mainly achieved by selecting natural variations that resulted in preferred traits or traits better suited to human cultivation, such as non-seed shattering or an increase in grain size.

Many crops have been domesticated all over the world (Doebley et al. 2006). Of these, rice is one of the most important crops that provides the daily calories for almost half of the world's population (Khush 2005). There are mainly two species of cultivated rice: Asian cultivated rice, Oryza sativa L., and African cultivated rice, O. glaberrima Steud; these two species were domesticated from the ancestral wild rice species, O. rufipogon Griff and O. barthii A. Chev., respectively (Oka 1988). Asian cultivated rice is mainly classified into two subspecies, indica and japonica (Oka 1988; Fuller 2007). Currently, O. sativa is widely cultivated in the world, whereas O. glaberrima is cultivated only in Africa. The domestication of O. sativa has been extensively studied both through plant genetics and archaeology.

A breakthrough in rice domestication studies was achieved with the completion of the genome sequencing of the cultivated rice O. sativa subsp. indica (Yu et al. 2002) and subsp. japonica (Goff et al. 2002; International Rice Genome Sequencing Project 2005). Subsequent genetic studies have identified the loci responsible for some domestication-related traits and further genetic mapping has uncovered key genes and causal mutations involved in changes in these traits. These are detailed below.

Phylogenetic studies have also been possible based on a growing body of sequence data.

However, these have resulted in alternative hypotheses with several 'origins' debated, both single and multiple (Kovach et al. 2007; Sang and Ge 2007a; Izawa et al. 2009; Panaud 2009; Gross and Zhao 2014; Civáň et al. 2015; Gutaker et al. 2020). A study by Huang et al. (2012) sequenced 446 accessions of O. rufipogon from Asia and 1083 cultivated rice strains of O. sativa subsp. indica and subsp. japonica. It was suggested using map-based sequence data that a specific population of O. rufipogon around the middle area of the Pearl River in southern China was the origin of rice domestication and that indica rice was produced by subsequent introgression of japonica with wild rice (Huang et al. 2012). Later, Civáň et al. (2015) re-analysed the sequence data of Huang et al. (2012) and proposed three geographically separate events in Asian rice domestication. Studies of maternal markers (Tang et al. 2004; Takahashi et al. 2008) and coalescent analyses have tended to support multiple origins (Choi et al. 2017), as do recent population genetic inferences from > 1400 whole genomes (Gutaker et al. 2020). Also, data from archaeological finds of sub-fossil rice (charred grains and chaff) favours a model of multiple domestications (Silva et al. 2015) but supports regions of origins that differ from those inferred from many of the genomic phylogenetic studies. Therefore, the area(s) of origin of rice domestication still remains unresolved.

To better understand the origin and chronology of rice domestication, we believe that more information related to phenotypic changes is required. Understanding the evolution of these phenotypic traits can inform both archaeological and phylogenetic approaches to origins, as they can be related to morphological features in sub-fossil material or to genotypes shared by phylogenetic linages. In this paper, we review some domestication-related traits and genes or loci that determine them, with a particular focus on two traits that likely played key roles in early rice domestication: closed panicles and the loss of seed shattering. Furthermore, we discuss some of the results from studies using wild rice introgression lines which provide information on different aspects of rice domestication.

Major traits selected for in rice domestication

Genetic variations in wild rice are important since mutations lead to agronomically important traits (Ishii et al. 2011). Over the past decade, several important genes involved in rice domestication have been identified and cloned. These traits are often linked to ease of cultivation and harvest as well as an increase in edible portions of rice (Table 1). These are also known to be involved in the domestication of maize, wheat, barley, sorghum, oats and other crops (Doebley et al. 2006). These recurrently selected common traits between species represent convergent evolution and are known as the domestication syndrome (Sakuma et al. 2011; Fuller 2007). The major traits selected in rice domestication with causal mutation and molecular function are shown in Table 1.

Table 1 Major traits and genes/loci involved in rice domestication

Full size table

Plant architecture

An obvious change between wild and cultivated rice is observed in the plant architecture; wild rice has prostrate tiller morphology, while cultivated rice shows erect growth (Fig. 1a, b). In wild rice, prostrate morphology helps to widely distribute seeds in the growing area. However, prostate morphology is not a favoured trait in paddy fields whereas erect growth habit is a preferred trait for cultivation.

Fig. 1

figure1

Comparison of plant morphologies between wild rice, Oryza rufipogon acc. W630, and a japonica rice cultivar, O. sativa Nipponbare. Plant architecture (a, b), panicle shape (c, d) and abscission layer formation (e, f) of W630 (a, c and e) and Nipponbare (b, d and f). Awn formation (g), seed colour (h) and seed dormancy (i), of W630 (left) and Nipponbare (right). Scale bars, 20 cm (a, b), 5 cm (c, d), 50 µm (e, f) and 5 mm (g–i). In the photo of abscission layer formation (e, f), black arrowheads indicate the abscission layer. VB, vascular bundle. In the photo of seed dormancy (i), W630 and Nipponbare seeds were induced for germination by water soaking a month after harvest.

Full size image

Two research groups identified PROG1, a gene responsible for prostate morphology, encoding a zinc-finger protein through molecular cloning of the causal gene (Table 1; Tan et al. 2008; Jin et al. 2008). Furthermore, one of the causal gene sequences at PROG1 was found to be an amino acid substitution by transformation experiments (Jin et al. 2008). However, other possibilities for the underlying mutation are also reported (Tan et al. 2008). These results suggest that gene sequences leading to prostrate tiller morphology may have arisen independently. Since the cultivated-type allele of PROG1 is observed in wild rice, it might have been selected in cultivation from standing variation in wild populations (Civáň and Brown 2017) and the phenotypic effect might be caused by an interaction with other mutation(s). Recently, a deletion in the region called SEMI-PROSTRATE GROWTH 1 (SPROG1), linked to PROG1 was also shown to be important for erect growth and higher yield in cultivars (Wu et al. 2018). These genes indicate that early selection under cultivation favoured erect plants that could be grown together more compactly, which parallels selection for apical dominance in other cereals (Fuller 2007; Fuller et al. 2010).

Panicle shape differs between wild and cultivated rice; wild rice has open panicles, whereas cultivated rice, O. sativa, has closed panicles (Fig. 1c, d). A closed panicle has the advantage of promoting self-fertilisation and for humans gathering wild rice, the closed panicle helps to retain the spikelets longer on the panicle, increasing harvest returns (Ishii et al. 2013). Experimental data indicates that harvesting of closed panicle rice, with otherwise wild features, could increase harvested yields by ~ 50% (Ishii et al. 2013). QTL analysis identified a single locus on chromosome 4 with 80.1% phenotypic variance (Ishii et al. 2013). The position was similar to the previously identified locus called SPR3 detected by genetic analysis of panicle shape between Asian cultivated rice, O sativa and African cultivated rice, O. glaberrima (Luo et al. 2008). Molecular genetic analyses showed that the 9.3-kb region is responsible for panicle shape, but no annotated genes were located in the region. A transformation study introducing a wild chromosomal segment into cultivated rice showed that the region together with a downstream gene, OsLG1, was sufficient to change panicle shape (Table 1; Ishii et al. 2013). Therefore, the region may act as a transcriptional regulator of OsLG1. Another study using the sequences from wild rice accessions showed that the one causal single nucleotide polymorphism (SNP) at the SPR3 region was associated with panicle shape (Zhu et al. 2013), although the causal mutation at SPR3 still remains to be elucidated.

Comparison of archaeological time series of grain size increase (mean grain width and standard deviation, plotted against estimated median age), from a northern India and b eastern China (Lower Yangtze) (data from Fuller et al. 2010, 2014). In India, some sites are plotted as median and range and when only the maximum and minimum metrics were available in the primary sources. Eastern Harappan sites in the upper Indus valley tributaries are indicated as a potential distinct group

Loss of seed shattering

The loss of seed shattering is the most striking example of a trait that contributes to easy harvesting with guaranteed yield, and its widely seen as the key trait that separated domesticated cereals from wild cereals (e.g. Harlan 1975; Fuller 2007). In wild rice, an abscission layer formed in the basal part of the grain contributes to seed shattering (Fig. 1e, f). Variations in the degree of seed shattering are observed in cultivated rice (Konishi et al. 2006), indicating that several genes or genetic changes are involved.

Two loci involved in the loss of seed shattering were previously identified by quantitative trait loci (QTL) analyses (Table 1). Li et al. (2006a) reported sh4 as a major locus explaining the difference in the degree of seed shattering between indica cultivated rice and annual wild rice, O. nivara (the annual form of O. rufipogon sensu lato). Molecular genetic analyses showed that the causal mutation was found to be a SNP in a gene encoding a Myb transcription factor (Li et al. 2006b; Lin et al. 2007). Experiments growing domesticated rice with two variants of the sh4 gene indicated that seed-shattering behaviour was restored in the wild allele at sh4 as opposed to the cultivated allele. This confirms that the wild sh4 allele had a role in promoting seed shattering. The original inference is that the mutation confers a non-shattering effect (Li et al. 2006b; Sang and Ge 2007b).

Another locus controlling the degree of seed shattering is qSH1, which was identified as controlling the degree of seed shattering between two rice cultivars, ‘Kasalath’ (representing subsp. indica) and ‘Nipponbare’ (representing subsp. japonica) (Konishi et al. 2006). This QTL explained approximately 70% of the phenotypic variance. A causal mutation was found to be a SNP at the promoter region of the gene encoding a homologue of Arabidopsis REPLUMLESS (RPL). The mutation at qSH1 resulted in a complete loss of shattering by eliminating abscission layer formation. Since mutation at qSH1 is only observed in ‘temperate japonica’ cultivars

(Konishi et al. 2006; Onishi et al. 2007b; Zhang et al. 2009), it is proposed that selection of the loss of the functional allele at qSH1 occurred after the sh4 mutation was fixed, at a later stage of diversification of breeding in domesticated rice. Because sh4 and qSH1 were identified in QTL analyses as having large effects on seed shattering, they were believed to be the genes explaining the loss of seed shattering during rice domestication. However, recent research through longer term experiments using plant materials with wild rice introgression lines have proved this view to be over-simplified (Ishii and Ishikawa 2018; see discussion below).

Loss of awn formation

In wild rice, a long, hair-like structure forms on the top of the seed, while it is mostly absent in cultivated rice (Fig. 1h). This tissue is called an awn, and it plays important roles in the propagation of wild relatives of cereal crops. This appendage allows the dehisced spikelet to travel some distances in water. Awns also act as a propelling mechanism to embed the spikelet in the soil. This has been demonstrated in other cereals, such as wild wheat, in which awns bend by means of cellulose fibrils fuelled by the daily humidity cycles (Elbaum et al. 2007). Therefore, awns evolved as an essential tissue for seed propagation in wild plants. The loss of awn formation was an important change affecting the efficiency of harvesting. In addition, human harvesting and sowing would have removed natural selection in favour of well-developed awns, and it has been observed that many domesticated rices have fewer and shorter awns (Tang et al. 1996; Fuller and Allaby 2009), although some japonica cultivars have awns which help deter seed predation by animals (Hua et al. 2015).

Three genes involved in the loss of awn formation have been reported in Asian cultivated rice (Table 1). An-1, encoding a basic helix-loop-helix (bHLH) protein controlling cell division, was found on chromosome 4 (Luo et al. 2013). Molecular analyses showed that the O.

rufipogon allele at An-1 was expressed at the distal end of the lemma while the O. sativa allele was found to be down-regulated. Long And Barbed Awn 1 (LABA1)/An-2, a gene encoding a cytokinin-activating enzyme, was identified on chromosome 4 from the segregating population between O. sativa subsp. indica and O. rufipogon (Hua et al. 2015; Gu et al. 2015). Causal mutation was shown to be a 1-bp deletion at the gene. Genome analyses showed that the mutation at LABA1/An-2 originated in some japonica rices and moved into the indica gene pool via introgression (Hua et al. 2015). However, further studies are still required to understand the origin of the mutation and directionality of introgression, because the mutation is frequently observed in wild rice (Civáň and Brown 2017) and the possibility of introgression of cultivated alleles into wild-rice populations needs to be explored.

Recently, a QTL controlling awn development on chromosome 8, Regulator of Awn Elongation 2 (RAE2), was found to encode a gene of Epidermal Patterning Factor-Like (EPFL) family member (Bessho-Uehara et al. 2016). A 2-bp insertion in the second exon of the gene was found to give rise to a truncated protein due to a frameshift mutation. Another research group also identified the same gene named Grain number, grain length and Awn Development 1 (GAD1), shared between wild rice O. rufipogon and an indica rice cultivar (Jin et al. 2016). More recently, natural alleles at RAE2/GAD1 were reported as they have a distinct role on gene expression and protein function (Zhang et al. 2019).

Interestingly, the loss of awn formation is also observed in African cultivated rice, O. glaberrima. A recent comparative analysis of the loss of awn formation between O. sativa and O. glaberrima reported that O. sativa lost its awn through mutations at two loci, RAE1, which is identical to An-1 on chromosome 4, and RAE2 on chromosome 8, while O. glaberrima lost it through mutation at RAE3 on chromosome 6. These results likely indicate that selection for awn loss in the two cultivated rice species is mediated by different loci (Furuta et al. 2015). The selection for the same trait at different loci is an example of convergent evolution, an important feature of domestication (Fuller et al. 2014).

The three genes involved in the awnless phenotypes in Asian cultivated rice, O. sativa, are not enough to fully explain the complete loss of awn formation when these alleles are evaluated in the genetic background of wild rice, O. rufipogon, indicating that some additional mutation(s) may still be involved in the selection of shorter awn length during rice domestication and diversification (Ikemoto et al. 2017; Amarasinghe et al. 2020).

Seed and hull colour

Colours of seeds and hulls were also selected for in rice domestication. Wild rice has red-pigmented pericarps, while cultivated rice has predominantly white pericarps (Fig. 1h). However, these mutations are what are known as post-domestication genes since they are not necessarily essential in improving harvests. They are related more to cultural histories (Fuller and Castillo 2016). Many red pericarp rices are cultivated and possess other domestication traits. While white rice is predominant in Asia, there are also dark purple or ‘black’ pericarp rices, which represent an alternative pathway of cultural selection (Oikawa et al. 2015).

In classical genetic studies, pericarp colour was known to be regulated by two loci, Rc (brown pericarp and seed coat) and Rd (red pericarp and seed coat), and the functional alleles at the two loci produce a red pigmented pericarp (Table 1; Sweeney et al. 2006; Furukawa et al. 2006). Genetic analysis identified that Rc encoded a bHLH protein. A 14-bp deletion within exon 6 that knocked out the bHLH domain of the protein was observed in most cultivated rice. Some cultivars were found to encode loss of the functional allele at Rc caused by a SNP, which was detected in some aus and aromatic rice, implying an independent origin of the selection at the Rc locus (Sweeney et al. 2006). A recent analysis by Civáň and Brown (2017) showed that some wild rice accessions were found to carry a cultivated-type allele (a 14-bp deletion) for Rc gene, although it is not known whether they express the white phenotype. It is also unclear whether this could indicate some standing variation in wild rice or whether the presence of the allele in wild rice might be caused by gene flow from cultivated rice. The other locus, Rd, was found to encode dihydroflavonol-4-reductase (DFR) involved in proanthocyanidin synthesis, and two types of truncated DFR protein were caused by independent premature stop codons in the first and second exons of the gene encoding DFR (Furukawa et al. 2006).

Changes in hull colour also took place with crop domestication, although these would have occurred as post-domestication events (Purugganan and Fuller 2009). Black seed hulls in wild rice and straw-white seed hulls in cultivated rice are commonly observed (Fig. 1h). Genetic analysis of the causal gene for hull colour identified the Bh4 gene, encoding an amino acid transporter (Table 1). Two origins of the mutations causing truncation of the BH4 protein were identified; one was a 22-bp deletion within exon 3 of the Bh4 gene and the other was a SNP causing a premature stop codon in the third exon (Zhu et al. 2011). A survey of Bh4 genotypes using a large panel of rice landraces showed that most of them had the 22-bp deletion, while a few had the SNP. Interestingly, a few landraces tested showed the same allele as wild rice at Bh4, even with a straw-white seed hull, suggesting that other genetic change(s) might have been involved in the change in hull colour. These findings suggest that several independent selections occurred for changes in seed hull colour.

Seed dormancy

When seeds are sown, some germinate well, while others do not. This character of dormancy is adaptive in many wild plants but is disadvantageous in plants that rely on human harvesting and sowing (Harlan 1975; Fuller and Allaby 2009). Seed dormancy is an evolutionary strategy that prevents early germination before the next season (Shu et al. 2015). Furthermore, extending the germination period contributes to a decrease in competition between individuals of the same species, which acts to increase chances of favourable conditions for some of the progeny (Finkelstein et al. 2008). In contrast, plants carrying a mutation to germinate soon after harvest should have an advantage in early farming, as they will get harvested and re-sown in larger numbers.

Wild rice has strong seed dormancy, while cultivated rice easily germinates (Fig. 1i). Several loci involved in differences in seed dormancy have been identified in rice (Table 1). Sdr4 was found as a QTL that contributed to the difference in seed dormancy between Nipponbare (subsp. japonica) and Kasalath (subsp. indica) cultivars (Sugimoto et al. 2010). Recently, OsG, an orthologue of classical stay-green G gene from soybean, was shown to be involved in the control of seed dormancy in rice. A genome-wide selection screening study using wild and cultivated rice reported that OsG is located in a selective sweep region and possibly selected in the early stages of rice domestication (Wang et al. 2018).

Seed dormancy and red pericarp colour have been found to be associated, and the responsible locus was detected in the same position in chromosome 7 in weedy red rice (Gu et al. 2005). Genetic analysis identified Rc, a gene encoding grain pericarp colour (as discussed above, Sweeney et al. 2006), and association between seed dormancy and pericarp colour is explained by a pleiotropic gene regulating abscissic acid and flavonoid levels (Gu et al. 2011). A recent genetic study in sorghum identified a homologue of the Rc gene that is involved in grain tannin content (Wu et al. 2019). Since condensed tannins inhibit alpha-amylase which is necessary for germination, the selection of the mutations at the gene in both rice and sorghum is an example of ‘domestication convergence’ (Civáň 2019).

Grain size

An increase in the edible portion is commonly observed during both animal and plant domestication (Wright 2015; Denham et al. 2020). In rice, grain size increased from wild to cultivated rice, especially grain breadth and thickness. Trends for increasing grain size have been identified in time series of archaeological rice grain measurements in both China and India (Fuller et al. 2010, 2014), e.g. Fig. 2.

Loci possibly involved in an increase in grain size during rice domestication are listed in Table 1. GS3 was identified through a QTL analysis between two cultivated rice species with small and large grains, and was found to encode a putative transmembrane protein (Fan et al. 2006). GS5 was shown to encode a putative serine carboxypeptidase. A higher expression level of the GS5 gene was found to be associated with increased grain size (Li et al. 2011). A loss of a functional allele at GW2 encoding a previously unknown RING-type protein with E3 ubiquitin ligase activity was found to enhance grain width and weight via an increase in cell numbers in the spikelet hull (Song et al. 2007). When qSW5 was cloned through a QTL analysis of grain size between indica and japonica cultivated rice, 38.5% of the variation was explained in the F2 population. A deletion at qSW5 was found to be a causal mutation leading to an increase in seed size, although the molecular function of the encoded protein is still unknown (Shomura et al. 2008). Genome-wide analyses of the four genes involved in the regulation of grain size suggested that qSW5 and GS3 had strong effects on grain size, while GW2 and GS5 had weaker effects (Lu et al. 2013). Since the increase in grain size occurred gradually, several genetic changes might have been involved. Therefore, the control of seed size seems to be more complicated than the control of the other domestication-related traits.

Selective sweep: a signature of domestication

Population genetic studies have been conducted using a large number of cultivated and wild rice species to evaluate the nature of selection for domestication-related genes. The evidence showing that a locus was selected during domestication, known as a selective sweep, is supported by the loss of nucleotide diversity around the chromosomal region carrying a domestication-related gene in cultivated rice. The selective sweeps were detected in most of the genes mentioned above (see references for the genes). In addition, whole genome sequence analysis identified unknown hot spots with a selective sweep in rice (Xu et al. 2012; Huang et al. 2012; Zhao et al. 2018), suggesting that some other unidentified genes were involved in rice domestication.

An archaeological view of rice domestication

Domestication of rice studied through the archaeobotanical lens focuses on a few traits which are visible, such as loss of seed shattering, increase in grain size (Fig. 2) and to a lesser extent loss of awns. Many of the domestication traits mentioned in the preceding section are not visible archaeologically because of preservation issues. The parts of the rice plant that preserve best in archaeological sites are the grain and spikelet base, which shows the abscission scar. Hull colour and plant architecture do not preserve as the archaeobotanical material is normally charred although theoretically, if one encountered waterlogged material, one would be able to evaluate these traits as well. However, neither trait has been documented so far.

Spikelets or grains can be measured even when charred, although a shrinkage factor needs to be taken into account in order to compare with modern examples. The difference in grain size is evaluated using a dataset that has a temporal scale, and such data indicate a gradual increase in the mean seed breadth and maximum seed breadth over thousands of years (Fig. 2; Fuller et al. 2010, 2014). This is also true for examining the loss of seed shattering. Spikelet bases are identified as domesticated type, wild type and immature type and compared across assemblages over time (Fig. 3). In the archaeological study by Fuller et al. (2009), spikelet bases of non-shattering rice were recovered in the Lower Yangtze region of Zhejiang, China, showing archaeobotanical evidence of rice domestication over a period of time (for additional data, see Fuller et al. 2014, 2016; Deng et al. 2015; cf. Zheng et al. 2016). Such data can be plotted as time series providing various estimates of the extent of domestication over time, based on whether or not 'immature' are classed with wild types or used to estimate uncertainty over the non-shattering proportion (Fig. 4). Lastly, lemma apices that preserve are examined for awns. The combination of mainly wild-type spikelet bases and lemma apices with awns would be highly suggestive that the rice being used in a particular site was still wild.

What traits were targeted for selection in early rice domestication?

The definition of domestication is focused on genetic changes affecting adaptations to cultivation that differ from the adaptations of wild rice. The genetic changes that benefited human needs were selected, albeit unconsciously, through morphological manifestations (Zohary 2004). These traits that made rice more productive, easier to harvest and plant under cultivation, increased not only the reproductive success of cultivated rice but also the interdependence of rice cultivators and rice plants (Fuller et al. 2010). A question is which traits were targeted in the initial domestication of rice, and in which order are traits likely to have been selected as a result of the changing human practices of early cultivation.

It is probable that one of the traits selected for in the very early part of crop domestication was the loss of seed shattering, which would have enabled greater returns of grain from harvesting (Harlan 1975; Fuller 2007). In modern cultivated rice, mature seeds remain in the panicle due to the inhibition of the abscission layer formation. Plants with inhibited abscission layer formation could have been present in wild rice populations, in very low numbers, and may have been unintentionally gathered by hunter-gatherers. The loss of seed shattering was also a key domestication trait in other crop species such as barley (Pourkheirandish et al. 2015) and sorghum (Lin et al. 2012). The increasing proportions of this trait in early cultivated population has also been documented through time series of archaeobotanical data, including in barley and wheat (Fuller et al. 2014; Allaby et al. 2017) and sorghum (Winchell et al. 2018).

In the initial stages of rice domestication, wild rice populations would have contained some stands with reduced proportions of seed shattering. Therefore, it would be more accurate to evaluate domestication-related mutated alleles in the wild rice genetic background. Therefore, research carried out at Kobe University has focused on maintaining a wild population of rice (the annual form of O. rufipogon sensu lato) and introgressing alleles from domesticated lines into this population to assess the impact of hypothesised domestication alleles on rice characters and through hand-harvesting experiments (Ishii and Ishikawa 2018). This is the opposite of most functional genetics which introgresses targeted wild rice alleles into plants of domesticated rice background, which means that interactions with other genes selected in domesticated populations can go undetected.

Interestingly, generation of wild introgression lines by recurrent backcrossing showed that the plants carrying domesticated alleles at either of the two seed-shattering loci, qSH1 or sh4, displayed complete wild-type shattering (Ishikawa et al. 2010). Moreover, histological analysis showed complete abscission layer formation in both the introgression lines (Htun et al. 2014).

These findings strongly indicated that a single mutation at either qSH1 or sh4 was not sufficient to confer a loss of seed shattering under natural conditions. Therefore, some other mutation(s) must have been required for an initial loss of seed shattering. The seed-shattering locus is qSH3, which will be described in the next section, encodes a likely candidate gene (Htun et al. 2014; Ishikawa et al. 2017). However, the accumulation and fixation of such mutations took a long time. The complicated process on initial loss of seed shattering raises the simple question whether this trait can be considered a crucial step towards rice domestication (Ray and Chakraborty 2017). Considering these results, there may have been other trait(s) that promoted rice domestication prior to the selection of a loss of seed shattering.

One candidate trait that may have been the target of selection before the selection of loss of seed shattering is panicle shape. To understand the role of panicle shape in rice domestication, a near isogenic line (NIL) having a chromosomal segment from O. sativa Nipponbare covering the SPR3 region in the O. rufipogon acc. W630 genetic background was produced by backcrossing. Since wild rice spikelets have long awns on the lemma tips, a closed panicle was found to increase seed retention, since the lower awns supported the upper seeds to stay longer in the panicle (Ishii et al. 2013). To confirm the effect of seed retention by closed panicle on harvesting efficiency, two field experiments were conducted. The time required for seed shattering was evaluated, and closed panicles were found to retain seeds for approximately a day longer than open panicles. This could have visibly aided seed collection in prehistory. A yield test by gathering seeds was also performed in the experimental paddy field. The gathering rates of wild rice and three NILs (SPR3, qSH1 and sh4) were calculated. Results indicated that NIL(SPR3) with a closed panicle had a higher yield than W630 and the other NILs with an open panicle, confirming the advantage of closed panicles on yield, increasing gathered grain yields by around 50%. These field results suggest that panicle shape could have been selected for at a relatively early stage of rice domestication (Ishii et al. 2013), even before non-shattering spikelet bases. In addition to the yield, the closed panicle provided an important advantage for rice as a ‘crop’. During flowering, long awns from the lower flowers blocked the free exposure of the anther when a panicle was closed. This increased the self-pollination rate, which is an important change since recessive mutations are more easily fixed through self-pollination. Self-pollination has often been hypothesised to be a trait that could have aided the domestication process in many cereals (Harlan 1975; Allard 1988; Zohary 1969, 2004). The change in panicle structure from open to closed may have had enormous effects on the selection of other domestication-related traits. High self-pollination behaviour would increase the appearance of valuable traits caused by recessive mutations in the segregating F2 generation. Once people started to cultivate what they harvested, these plants entered a cycle of cultivation. Therefore, seed-shattering and pollinating behaviours were greatly influenced by a simple morphological change that may have emerged at the very beginning of cultivation or even intensive harvesting from the wild (Ishii et al. 2013). Closed panicles would increase return rates with typical hunter-gatherer harvesting methods such as hand stripping or basket swinging.

The processes of loss of seed shattering is more complicated than expected

Loss of seed shattering occurs when the abscission layer, which is developmentally programmed in the basal part of the spikelet, is disrupted. In wild rice, after successful fertilisation, the abscission layer begins to degrade as the seeds mature. They easily dehisce, resulting in successful propagation. As mentioned in the previous section, neither a single mutation at sh4 nor qSH1 was sufficient to confer a loss of seed shattering in the wild rice genetic background (Ishikawa et al. 2010; Ishii et al. 2013). Therefore, other mutation(s) together with sh4 must have been required for the loss of seed shattering in the initial stage of rice domestication. To search for gene(s) involved in the loss of shattering, wild rice lines with chromosomal segments of cultivated rice at sh4 and qSH1 that still showed a weak seed-shattering phenotype were further crossed with Nipponbare. In the resulting F2 population, the two loci were fixed and were not responsible for the variation in the degree of seed shattering. QTL analysis identified a locus on chromosome 3, and the wild allele had a role in enhancing seed shattering (Htun et al. 2014). This locus was named qSH3 (Table 1), as it was also detected in a previous study (Onishi et al. 2007a); however, the effects of qSH3 on seed shattering were relatively lower than those of the two loci in the genetic background of cultivated rice. The difference in the effect of qSH3 suggested that a genetic interaction among seed-shattering loci could have been important in the loss of seed shattering.

Since the candidate region at qSH3 showed the signature of a selective sweep in both indica and japonica cultivated rices (Xu et al. 2012), the mutation at qSH3 may have been selected for during rice domestication. Although the sh4 domesticated allele is commonly observed in all of the cultivated rice tested (Lin et al. 2007), it is not sufficient on its own to confer a loss of seed shattering in wild rice background (Ishikawa et al. 2010; Ishii et al. 2013). This contradiction was also observed in some weedy rice that had a non-functional allele at sh4 but had seed-shattering behaviour (Thurber et al. 2010; Zhu et al. 2012). Therefore, a combination of mutations at seed-shattering loci, like sh4 and qSH3, may have played an important role in the initial loss of seed shattering. To test this idea, introgression lines having a chromosomal segments of Nipponbare at sh4 and qSH3 were produced in a wild rice genetic background. The introgression line showed partial disruption of the abscission layer around the vascular bundle (Inoue et al. 2015). Slight inhibition of abscission layer formation may have had an advantage in that seeds could stay longer in the panicle but could also be easily harvested without tools, i.e. by ‘forager’ methods such as basket swinging, hand paddling or hand stripping. It would be of interest to understand the molecular nature that caused the partial disruption of abscission layer formation.

Further research then explored how many genes or mutations were involved in the complete loss of seed shattering in japonica rice. The degree of seed shattering of an introgression line of wild rice having the domesticated alleles at the three loci was compared with that of Nipponbare. The breaking tensile strength (BTS) value of the introgression line was significantly lower than that of Nipponbare (Htun et al. 2014), suggesting that some other loci may still be involved in the loss of seed shattering in japonica rice. To identify the rest of the gene(s), an approach that involves fixing the known QTL(s) in the segregating population would be effective for exploring and uncovering the novel QTL(s). Based on this approach, we recently detected QTL for the Control of Seed Shattering in rice on chromosome 3 (qCSS3), a novel locus involved in a loss of seed shattering in japonica rice cultivar O. sativa Nipponbare (Tsujimura et al. 2019). Once all the mutations related to a loss of seed shattering are clarified, information on the causal mutations would be very useful to understand the change in the degree of seed shattering conferred by each locus and their additive effects in combination. This will help to clarify how the mutation took place and the process of rice domestication.

It is important to understand how the various mutations in the seed-shattering loci were accumulated and fixed. The domesticated alleles at the three seed-shattering loci all act recessively (Li et al. 2006a; Konishi et al. 2006; Htun et al. 2014), and other domestication-related mutations are often recessive. Therefore, to exhibit a reduction in seed shattering, these recessive alleles must have been fixed through self-pollination. As wild rice has a partial out-crossing behaviour (Phan et al. 2012), the selection of the closed panicle trait might have been an important step allowing mutations to be more easily fixed (Ishii et al. 2013). Therefore, closed panicles may have contributed to the accumulation of these recessive alleles causing loss of seed shattering. Although the order of the selections is unknown, it would be of great interest to test this idea in the future. Present accumulated data suggest that the process of a loss of seed shattering may have been more complicated than expected.

Future aspects of the study of rice domestication: power of ancient DNA

The origin of rice domestication has been discussed using data obtained from population genetic analyses (Londo et al. 2006; Molina et al. 2011; Huang et al. 2012; He et al. 2014; Civáň et al. 2015). These data are based on phylogenetic analyses that compare DNA sequences from modern wild rice accessions and cultivars. When the DNA of modern accessions is studied, their domestication-related phenotypes may derive from underlying genetic redundancies, additive loci or alternative loci. Since wild and cultivated rice have few reproductive barriers, gene flow between them is sometimes observed in their natural habitat, where wild rice are grown near cultivated rice, which is often the case through much of South and Southeast Asia, and therefore genes selected under domestication may have been introduced into wild rice populations. Weedy rice is an example of introgression (He et al. 2014; Thurber et al. 2014). For instance, in the case of wild rice outcrossed with cultivated rice, some chromosomal fragments are maintained, although most parts of the chromosome will be replaced in the wild genome by backcrossing. These plants show cultivar-like phenotypes. Thus, we need to pay more attention to these cultivar-like wild rice species recorded as 'wild rice' in gene banks. Moreover, most present genetic studies have evaluated the function of the genes responsible for domestication-related traits in the genetic background of cultivated rice. Since cultivated rice may still harbour unidentified mutation(s) involved in the traits, inferences on the domestication-related phenotypes is misleading. The use of wild rice as the genetic background has been shown to provide more accurate information on rice domestication.

Despite the identification of a number of genes involved in rice domestication, the limitation of the present genetic studies is the lack of information about when and where the domestication-related mutations were selected. Ancient DNA (aDNA) would help in determining the chronology of rice domestication. Analyses using fragments of preserved DNA would provide more information relative antiquity than those using DNA from modern rice. In barley, ancient DNA extraction from the entire range of occupancy of the site at Qasr Ibrim was performed, and the Vrs1 locus responsible for row number in barley was studied. The results showed that reassertion of the two-row condition was selected from six-row ancestry for the Qasr Ibrim barley (Palmer et al. 2009). There have been very few aDNA studies conducted on rice (Castillo et al. 2016; Tanaka et al. 2015), and these have used PCR instead of NGS, with limited potential to recover nuclear DNA which provides details on domestication-related genes.

One of the limitations of aDNA studies is the state of preservation, which directly complicates DNA recovery. However, the technology for analysing aDNA has improved; therefore, more information could be obtained once the aDNA analysis of rice becomes successful. Therefore, information should be obtained from aDNA or even from some modifications of DNA such as DNA methylation, an important modification for the regulation of gene expression (Smith et al. 2014). A genotyping survey of the domesticated alleles at a certain age and place will provide evidence of selections and the process of domestication. In addition to aDNA analysis, the use of population genetic approaches such as McSwan, a method to detect and date past and recent natural selection events, will be a great help in understanding of the timeframe of selection of various loci (Tournebize et al. 2019). These multidimensional perspectives will contribute to fully understanding the process and chronology of rice domestication.

Sang and Ge (2007b) provided two models for rice domestication: snow-balling and combination. In the snow-balling model, a single initial selection triggered rice domestication, and then introgressions from wild rice were repeated in various areas, creating subspecies such as indica and japonica. This model was essentially the same as that posited by Huang et al. (2012). In the combination model, independent selection of useful traits occurred in different locations, and these partially domesticated plants were followed by subsequent combination of the most useful variants. Subsequently, coalescence-based modelling of demographic parameters on modern genomes of seven rices (two wild and five cultivated rice) confirmed an important level of gene flow from ancestral domesticated japonica rice into indica and aus rice (Choi et al. 2017). These data are consistent with a form of the combination model, referred to as the 'Proto-indica hypothesis', which infers that ancient cultivation of ancestral indica in South Asia that did not develop a population fixed in domestication features until after introgressions from domesticated japonica, thought to have taken place after 4000 years ago (Fuller and Qin 2009; Fuller 2011; Silva et al. 2015). Recently, Civáň and Brown (2018) argued against introgression from japonica into early indica, and argue instead for three distinct domestications of subsp. indica, subsp. japonica and the 'aus' group of rices with later post-domestication introgression. The structure of genomic diversity in modern rice populations makes it clear that indica and japonica (and aus) rice had distinct histories of geographical dispersal (Gutaker et al. 2020), but a more detailed understanding in each domestication sequence from only modern genetics remains difficult; it also cannot account for any posited extinct or extirpated lineages (Fuller 2011). For testing these alternative hypotheses about domestication, aDNA could prove essential, providing direct integration of archaeobotanical and genomic approaches.

Conclusions

In prehistory, humans selected plants based on phenotypes and not on genomic information. The genetic changes were therefore unintended consequences. It is necessary to pay more attention to the evaluation of phenotypic changes when discussing early events in rice domestication. Recent research has demonstrated that key phenotypic changes were often affected by multiple alleles, and therefore the evaluation of additive effects and interactions of multiple alleles is important. Far from single recessive mutations being the sole cause of domestication morphotypes, as long assumed for non-shattering in cereals (Harlan 1975; Zohary 1969; Zohary 2004), multiple alleles may contribute to additive effects and interact with other phenotypic changes selected during early harvesting and cultivation, such as awns, closed versus open panicles, dormancy and other traits. Although the history of rice domestication is still a subject of debate, the accumulation of data related to phenotypic effects from multiple alleles is essential as is archaeological evidence for the appearance and spread of certain phenotypes. The promise of aDNA is that the occurrence of alleles can be directly placed in past time and past geography, and connected to visible morphological features of archaeological rice. We are therefore at the beginning of unravelling rice domestication processes and achieving a deeper understanding how this occurred.

Acknowledgements

We thank Prof. T. Ishii for his valuable comments on this manuscript. This project was supported in part by a JSPS bilateral open partnership joint research project (JPJSBP120189948) and Grant-in-Aid from the Japanese Society for Promotion of Science (15KK0280 and 18K05594).

References

Allaby RG, Stevens C, Lucas L, Maeda O, Fuller DQ (2017) Geographic mosaics and changing rates of cereal domestication. Philos Trans R Soc Lond B Biol Sci. 372.

Allard RW (1988) Genetic changes associated with the evolution of adaptedness in cultivated plants and their wild progenitors. J Hered 79: 225-238.

Amarasinghe YPJ, Kuwata R, Nishimura A, Phan PDT, Ishikawa R, Ishii T (2020) Evaluation of domestication loci associated with awnlessness in cultivated rice, *Oryza sativa*. Rice 13: 26.

Bessho-Uehara K, Wang DR, Furuta T, Minami A, Nagai K, Gamuyao R, Asano K Angeles-Shim RB, Shimizu Y Ayano M, Komeda N, Doi K, Miura K, Toda Y Kinoshita T, Okuda S, Higashiyama T, Nomoto M, Tada Y, Shinohara H, Matsubayashi Y, Greenberg A, Wu J, Yasui H, Yoshimura A, Mori H, McCouch SR, Ashikari M (2016) Loss of function at *RAE2*, a previously unidentified EPFL, is required for awnlessness in cultivated Asian rice. Proc Natl Acad Sci U S A. 113: 8969-8974.

Castillo CC, Tanaka K, Sato Y-I, Ishikawa R, Bellina B, Higham C, Chang N, Mohanty R, Kajale M, Fuller DQ (2016) Archaeogenetic study of prehistoric rice remains from Thailand and India: Evidence of early *japonica* in South and Southeast Asia. Archaeological and Anthropological Sciences 8: 523-543.

Choi JY, Platts AE, Fuller DQ, Hsing YI, Wing RA, Purugganan MD (2017) The rice paradox:

multiple origins but single domestication in Asian rice. Mol Biol Evol. 34: 969-979.

Civáň, P., & Brown, T. A. (2018). Role of genetic introgression during the evolution of cultivated rice (Oryza sativa L.). BMC evolutionary biology 18(1), 57.

Civáň P, Craig H, Cox CJ, Brown TA (2015) Three geographically separate domestications of Asian rice. Nat Plants 1: 15164.

Civáň O. Civáň, Brown TA (2017) Origin of rice (*Oryza sativa* L.) domestication genes. Genet Resour Crop Evol 64:1125-1132.

Civáň P (2019) When bitter is better. Nat Plants 5: 1205-1206.

Deng Z, Qin L, Gao Y, Weisskopf AR, Zhang C, Fuller DQ (2015) From early domesticated rice of the middle Yangtze Basin to millet, rice and wheat agriculture: Archaeobotanical macro-remains from Baligang, Nanyang Basin, Central China (6700–500 BC). PLoS One 10: e0139885.

Denham T, Barton H, Castillo C, Crowther A, Dotte-Sarout E, Florin A, Pritchard J, Barron A, Zhang Y, Fuller DQ (2020) The domestication syndrome in vegetatively-propagated field crops. Annals of Botany 125(4): 581-597

Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. Cell 127: 1309-1321.

Elbaum R, Zaltzman L, Burgert I, Fratzl P (2007) The role of wheat awns in the seed dispersal unit. Science 316: 884-886.

Fan C, Xing Y, Mao H, Lu T, Han B, Xu C, Li X, Zhang Q (2006) *GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. Theor Appl Genet 112: 1164-1171.

Finkelstein R, Reeves W, Ariizumi T, Steber C (2008) Molecular aspects of seed dormancy. Annu Rev Plant Biol. 59: 387-415.

Fuller DQ (2007) Contrasting patterns in crop domestication and domestication rates, recent archaeobotanical insights from the Old World. Ann. Bot. 100: 903-924.

Fuller, D. Q. (2011). Pathways to Asian civilizations: Tracing the origins and spread of rice and rice cultures. Rice 4: 78-92.

Fuller DQ, Allaby R (2009) Seed dispersal and crop domestication: shattering, germination and seasonality in evolution under cultivation. In Annual Plant Reviews, Vol. 38, Fruit Development and Seed Dispersal (ed.: Ostergaard L.). pp. 238-295. Wiley-Blackwell, Oxford.

Fuller, D. Q., & Qin, L. (2009). Water management and labour in the origins and dispersal of Asian rice. World Archaeology 41(1): 88-111.

Fuller DQ, Qin L, Zheng Y, Zhao Z, Chen X, Hosoya LA, Sun GP (2009) Science 323: 1607-1610.

Fuller DQ, Allaby RG, Stevens C (2010) Domestication as innovation: the entanglement of techniques, technology and chance in the domestication of cereal crops. World Archaeology 42: 13-28.

Fuller DQ, Denham T, Arroyo-Kalin M, Lucas L, Stevens CJ, Qin L, Allaby RG, Purugganan MD (2014) Convergent evolution and parallelism in plant domestication revealed by an expanding archaeological record. Proc Natl Acad Sci U S A 111: 6147-6152.

Fuller DQ, Castillo CC (2016) Diversification and cultural construction of a crop: the case of glutinous rice and waxy cereals in the food cultures of eastern Asia, In; Thorp JL, Katzenberg MA (eds) The Oxford handbook of the archaeology of food and diet. Oxford: Oxford University Press.

Fuller DQ, Weisskopf AR, Castillo C (2016) Pathways of rice diversification across Asia. Archaeology International 19: 84-96.

Furukawa T, Maekawa M, Oki T, Suda I, Iida S, Shimada H, Takamure I, Kadowaki K (2006) The *Rc* and *Rd* genes are involved in proanthocyanidin synthesis in rice pericarp. Plant J

49: 91-102.

Furuta T, Komeda N, Asano K, Uehara K, Gamuyao R, Angeles-Shim RB, Nagai K, Doi K, Wang DR, Yasui H, Yoshimura A, Wu J, McCouch SR, Ashikari M (2015) Convergent loss of awn in two cultivated rice species *Oryza sativa* and *Oryza glaberrima* is caused by mutations in different loci. G3 5: 2267-2274.

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A, Briggs S (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science 296: 92-100.

Gross BL, Zhao Z (2014) Archaeological and genetic insights into the origins of domesticated rice. Proc Natl Acad Sci U S A 111: 6190-6197.

Gu B, Zhou T, Luo J, Liu H, Wang Y, Shangguan Y, Zhu J, Li Y, Sang T, Wang Z, Han B (2015) *An-2* encodes a cytokinin synthesis enzyme that regulates awn length and grain production in rice. Mol Plant 8: 1635-1650.

Gu XY, Kianian SF, Hareland GA, Hoffer BL, Foley ME (2005) Genetic analysis of adaptive syndromes interrelated with seed dormancy in weedy rice (*Oryza sativa* L.). Theor Appl Genet 110: 1108-1118.

Gu XY, Foley ME, Horvath DP, Anderson JV, Feng J, Zhang L, Mowry CR, Ye H, Suttle JC, Kadowaki K, Chen Z (2011) Association between seed dormancy and pericarp color is controlled by a pleiotropic gene that regulates abscisic acid and flavonoid synthesis in weedy red rice. Genetics 189: 1515-1524.

Gutaker RM, Groen SC, Bellis ES, Choi JY, Pires IS, Bocinsky RK, Slayton ER, Wilkins O, Castillo CC, Negrão S, Oliveira MM. (2020) Genomic history and ecology of the geographic spread of rice. Nature Plants 6(5): 492-502.

Harlan JR (1975) Crops and Man. American Society of Agronomy, Madison.

He Z, Jiang X, Ratnasekera D, Grassi F, Perera U, Lu BR (2014) Seed-mediated gene flow promotes genetic diversity of weedy rice within populations: implications for weed management. PLoS One 9: e112778.

He Z, Zhai W, Wen H, Tang T, Wang Y, Lu X, Greenberg AJ, Hudson RR, Wu CI, Shi S (2011) Two evolutionary histories in the genome of rice: the roles of domestication genes. PLoS Genet 7: e1002100.

Htun TM, Inoue C, Chhourn O, Ishii T, Ishikawa R (2014) Effect of quantitative trait loci for seed shattering on abscission layer formation in Asian wild rice *Oryza rufipogon*. Breed Sci 64: 199-205.

Hua L, Wang DR, Tan L, Fu Y, Liu F, Xiao L, Zhu Z, Fu Q, Sun X, Gu P, Cai H, McCouch SR, Sun C (2015) *LABA1*, a domestication gene associated with long, barbed awns in wild rice. Plant Cell 27: 1875-1888.

Huang X, Kurata N, Wei X, Wang ZX, Wang A, Zhao Q, Zhao Y, Liu K, Lu H, Li W, Guo Y, Lu Y, Zhou C, Fan D, Weng Q, Zhu C, Huang T, Zhang L, Wang Y, Feng L, Furuumi H, Kubo T, Miyabayashi T, Yuan X, Xu Q, Dong G, Zhan Q, Li C, Fujiyama A, Toyoda A, Lu T, Feng Q, Qian Q, Li J, Han B (2012) A map of rice genome variation reveals the origin of cultivated rice. Nature 490: 497-501.

International rice genome sequencing project 2005. The map-based sequence of the rice genome. Nature 436: 793-780.

Inoue C, Htun TM, Inoue K, Ikeda K, Ishii T, Ishikawa R (2015) Inhibition of abscission layer formation by an interaction of two seed-shattering loci, *sh4* and *qSH3*, in rice. Genes Genet Syst 90: 1-9.

Ishii T, Hiraoka T, Kanzaki K, Akimoto M, Shishido R, Ishikawa R (2011) Evaluation of genetic variation among wild populations and local varieties of rice. Rice 4: 170-177.

Ishii T, Numaguchi K, Miura K, Yoshida K, Thanh PT, Htun TM, Yamasaki M, Komeda N, Matsumoto T, Terauchi R, Ishikawa R, Ashikari M. (2013) *OsLG1* regulates a closed panicle trait in domesticated rice. Nat Genet 45; 462-465.

Ishii T, Ishikawa R (2018) Domestication loci controlling panicle shape, seed shattering, and seed awning. In: Sasaki T, Ashikari M (eds) Rice Genomics, Genetics and Breeding. Springer, Singapore, pp. 207-221.

Ishikawa R, Thanh PT, Nimura N, Htun TM, Yamasaki M, Ishii T (2010) Allelic interaction at seed-shattering loci in the genetic backgrounds of wild and cultivated rice species. Genes Genet Syst 85: 265-271.

Ishikawa R, Nishimura A, Htun TM, Nishioka R, Oka Y, Tsujimura Y, Inoue C, Ishii T (2017) Estimation of loci involved in non-shattering of seeds in early rice domestication. Genetica 145: 201-207.

Izawa T, Konishi S, Shomura A, Yano M 2009. DNA changes tell us about rice domestication. Curr Opin Plant Biol. 12: 185-192.

Jin J, Huang W, Gao JP, Yang J, Shi M, Zhu MZ, Luo D, Lin HX (2008) Genetic control of rice plant architecture under domestication. Nat Genet 40: 1365-1369.

Jin J, Hua L, Zhu Z, Tan L, Zhao X, Zhang W, Liu F, Fu Y, Cai H, Sun X, Gu P, Xie D, Sun C (2016) *GAD1* encodes a secreted peptide that regulates grain number, grain length, and awn development in rice domestication. Plant Cell 28: 2453-2463.

Khush GS (2005) What it will take to feed 5.0 billion rice consumers in 2030. Plant Mol Biol 59: 1-6.

Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, Yano M (2006) An SNP caused loss of seed shattering during rice domestication. Science 312: 1392-1396.

Kovach MJ, Sweeney MT, McCouch SR (2007) New insights into the history of rice domestication Trends Genet 23: 578-587.

Larson G, Piperno DR, Allaby RG, Purugganan MD, Andersson L, Arroyo-Kalin M, Barton L, Vigueira CC, Denham T, Dobney K, Doust AN, Geots P, Gilbert MPT, Gremillion KJ, Lucas L, Lukens L, Marshall FB, Olsen KM, Pires JC, Richerson PJ, Rubio de Casa R, Sanjur OI, Thomas MG, Fuller DQ (2014) Current perspectives and the future of domestication studies. Proc Natl Acad Sci U S A 111(17): 6139-46.

Lee SJ, Oh CS, Suh JP, McCouch SR, Ahn SN (2005) Identification of QTLs for domestication-related and agronomic traits in an *Oryza sativa* × *O. rufipogon* BC_1F_7 population. Plant Breed 124: 209-219.

Li C, Zhou A, Sang T (2006a) Genetic analysis of rice domestication syndrome with the wild annual species, *Oryza nivara*. New Phytol 170: 185-193.

Li C, Zhou A, Sang T (2006b) Rice domestication by reducing shattering. Science 311: 1936-1939.

Li Y, Fan C, Xing Y, Jiang Y, Luo L, Sun L, Shao D, Xu C, Li X, Xiao J, He Y, Zhang Q (2011) Natural variation in *GS5* plays an important role in regulating grain size and yield in rice. Nat Genet 43: 1266-1269.

Lin Z, Griffith ME, Li X, Zhu Z, Tan L, FuY, Zhang W, Wang X, Xie D, Sun C (2007) Origin of seed shattering in rice (*Oryza sativa* L.). Planta 226: 11-20.

Lin Z, Li X, Shannon LM, Yeh CT, Wang ML, Bai G, Peng Z, Li J, Trick HN, Clemente TE, Doebley J, Schnable PS, Tuinstra MR, Tesso TT, White F, Yu J (2012) Parallel domestication of the *Shattering1* genes in cereals. Nat Genet 44: 720-724.

Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA (2006) Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. Proc Natl Acad Sci U S A 103: 9578-9583.

Lu L, Shao D, Qiu X, Sun L, Yan W, Zhou X, Yang L, He Y, Yu S, Xing Y (2013) Natural variation and artificial selection in four genes determine grain shape in rice. New Phytol 200: 1269-1280.

Luo JJ, Hao W, Jin J, Gao JP, Lin HX (2008) Fine mapping of *Spr3*, a locus for spreading panicle from African cultivated rice (*Oryza glaberrima* Steud.). Mol Plant 1: 830-838.

Luo J, Liu H, Zhou T, Gu B, Huang X, Shangguan Y, Zhu J, Li Y, Zhao Y, Wang Y, Zhao Q, Wang A, Wang Z, Sang T, Wang Z, Han B (2013) *An-1* encodes a basic helix-loop-helix protein that regulates awn development, grain size, and grain number in rice. Plant Cell 25: 3360-3376.

Molina J, Sikora M, Garud N, Flowers JM, Rubinstein S, Reynolds A, Huang P, Jackson S, Schaal BA, Bustamante CD, Boyko AR, Purugganan MD (2011) Molecular evidence for a single evolutionary origin of domesticated rice. Proc Natl Acad Sci U S A 108: 8351-8356.

Oikawa T, Maeda H, Oguchi T, Yamaguchi T, Tanabe N, Ebana K, Yano M, Ebitani T, Izawa T. (2015) The birth of a black rice gene and its local spread by introgression. The Plant Cell 27(9): 2401-2414.

Oka HI (1988) Origin of Cultivated Rice. Elsevier, Tokyo, p. 254.

Onishi K, Horiuchi Y, Ishigoh-Oka N, Takagi K, Ichikawa N, Maruoka M, Sano Y (2007a) A QTL cluster for plant architecture and its ecological significance in Asian wild rice. Breed. Sci. 57: 7-16.

Onishi K, Takagi K, Kontani M, Tanaka T, Sano Y (2007b) Different patterns of genealogical relationships found in the two major QTLs causing reduction of seed shattering during rice domestication. Genome 50: 757-766.

Panaud O (2009) The molecular bases of cereal domestication and the history of rice. C R Biol. 332: 267-272.

Palmer SA, Moore JD, Clapham AJ, Rose P, Allaby RG (2009) Archaeogenetic evidence of ancient nubian barley evolution from six to two-row indicates local adaptation. PLoS One 4: e6301.

Phan PD, Kageyama H, Ishikawa R, Ishii T (2012) Estimation of the outcrossing rate for annual Asian wild rice under field conditions. Breed Sci. 62: 256-262.

Pourkheirandish M, Hensel G, Kilian B, Senthil N, Chen G, Sameri M, Azhaguvel P, Sakuma S, Dhanagond S, Sharma R, Mascher M, Himmelbach A, Gottwald S, Nair SK, Tagiri A, Yukuhiro F, Nagamura Y, Kanamori H, Matsumoto T, Willcox G, Middleton CP, Wicker T, Walther A, Waugh R, Fincher GB, Stein N, Kumlehn J, Sato K, Komatsuda T (2015) Evolution of the grain dispersal system in barley. Cell 162: 527-539.

Ray A, Chakraborty D (2017) Shattering or not shattering: that is the question in domestication of rice (*Oryza sativa* L.). Genet Resour Crop Evol 65: 391- 395.

Sakuma S, Salomon B, Komatsuda T (2011) The domestication syndrome genes responsible for the major changes in plant form in the Triticeae crops. Plant Cell Physiol 52: 738-749.

Sang T, Ge S (2007a) The puzzle of rice domestication. J Integr Plant Biol 49: 760–768.

Sang T, Ge S (2007b) Genetics and phylogenetics of rice domestication. Curr Opin Genet Dev 17: 533-538.

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.

Shu K, Meng YJ, Shuai HW, Liu WG, Du JB, Liu J, Yang WY (2015) Dormancy and germination: How does the crop seed decide? Plant Biol 17: 1104-1112.

Silva F, Stevens CJ, Weisskopf A, Castillo C, Qin L, Bevan A, Fuller DQ (2015) Modelling the geographical origin of rice cultivation in Asia using the rice archaeological database.

PLoS One 10: e0137024.

Smith O, Momber G, Bates R, Garwood P, Fitch S, Pallen M, Gaffney V, Allaby RG (2014) Sedimentary DNA from a submerged site reveals wheat in the British Isles 8000 years ago. Science 347: 998-1001.

Song XJ, Huang W, Shi M, Zhu MZ, Lin HX (2007) A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nat Genet 39: 623-630.

Sugimoto K, Takeuchi Y, Ebana K, Miyao A, Hirochika H, Hara N, Ishiyama K, Kobayashi M, Ban Y, Hattori T, Yano M (2010) Proc Natl Acad Sci U S A 107: 5792-5797.

Sweeney MT, Thomson MJ, Pfeil BE, McCouch S (2006) Caught red-handed: Rc encodes a basic helix-loop-helix protein conditioning red pericarp in rice. Plant Cell 18: 283-294.

Tan L, Li X, Liu F, Sun X, Li C, Zhu Z, Fu Y, Cai H, Wang X, Xie D, Sun C (2008) Control of a key transition from prostrate to erect growth in rice domestication. Nat Genet 40: 1360-1364.

Tang J, Xia H, Cao M, Zhang X, Zeng W, Hu S, Tong W,Wang J,Wang J, Yu J, Yang H, Zhu L (2004) A comparison of rice chloroplast genomes. Plant Physiol 135: 412–20.

Tang S, Min S, Sato YI (1996) Exploration on origin of keng rice (*japonica*) in China. In: Wang X and Sun C (eds) Origin and differentiation of Chinese cultivated rice. Chinese Agricultural University Press, Beijing, pp 72–80.

Takahashi H, Sato YI, Nakamura I (2008) Evolutionary analysis for two plastid DNA sequences in cultivated and wild species of *Oryza*. Breeding Science 58: 225–233.

Thurber CS, Reagon M, Gross BL, Olsen KM, Jia Y, Caicedo AL (2010) Molecular evolution of shattering loci in US weedy rice. Mol Ecol 19: 3271-3284.

Thurber CS, Reagon M, Olsen KM, Jia Y, Caicedo AL (2014) The evolution of flowering strategies in US weedy rice. Am J Bot 101: 1737-1747.

Tournebize R, Poncet V, Jakobsson M, Vigouroux Y, Manel S (2019) McSwan: A joint site frequency spectrum method to detect and date selective sweeps across multiple population genomes. Mol Ecol Resour. 19: 283-295.

Tsujimura Y, Sugiyama S, Otsuka K, Htun TM, Numaguchi K, Castillo C, Akagi T, Ishii T, Ishikawa R (2019) Detection of a novel locus involved in non-seed-shattering behaviour of Japonica rice cultivar, *Oryza sativa* 'Nipponbare' Theor Appl Genet. 132: 2615-2623.

Wang M, Li W, Fang C, Xu F, Liu Y, Wang Z, Yang R, Zhang M, Liu S, Lu S, Lin T, Tang J, Wang Y, Wang H, Lin H, Zhu B, Chen M, Kong F, Liu B, Zeng D, Jackson SA, Chu C, Tian Z (2018) Parallel selection on a dormancy gene during domestication of crops from multiple families. Nat Genet 50: 1435-1441.

Weng J, Gu S, Wan X, Gao H, Guo T, Su N, Lei C, Zhang X, Cheng Z, Guo X, Wang J, Jiang L, Zhai H, Wan J (2008) Isolation and initial characterization of *GW5*, a major QTL associated with rice grain width and weight. Cell Res 18: 1199-1209.

Winchell F, Brass M, Manzo A, Beldados A, Perna V, Murphy C, Stevens C, Fuller DQ (2018) On the origins and dissemination of domesticated sorghum and pearl millet across Africa and into India: a view from the Butana Group of the Far Eastern Sahel. Afr Archaeol Rev 35: 483-505.

Wright D (2015) The genetic architecture of domestication in animals. Bioinform Biol Insights 9: 11-20.

Wu Y, Zhao S, Li X, Zhang B, Jiang L, Tang Y, Zhao J, Ma X, Cai H, Sun C, Tan L (2018) Deletions linked to *PROG1* gene participate in plant architecture domestication in Asian and African rice. Nat Commun 9: 4157.

Wu Y, Guo T, Mu Q, Wang J, Li X, Wu Y, Tian B, Wang ML, Bai G, Perumal R, Trick HN, Bean SR, Dweikat IM, Tuinstra MR, Morris G, Tesso TT, Yu J, Li X (2019) Allelochemicals targeted to balance competing selections in African agroecosystems. Nat Plants 5: 1229-1236.

Xu X, Liu X, Ge S, Jensen JD, Hu F, Li X, Dong Y, Gutenkunst RN, Fang L, Huang L, Li J, He W, Zhang G, Zheng X, Zhang F, Li Y, Yu C, Kristiansen K, Zhang X, Wang J, Wright M, McCouch S, Nielsen R, Wang J, Wang W (2012) Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. Nat Biotechnol 30: 105-111.

Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M, Liu J, Sun J, Tang J, Chen Y, Huang X, Lin W, Ye C, Tong W, Cong L, Geng J, Han Y, Li L, Li W, Hu G, Huang X, Li W, Li J, Liu Z, Li L, Liu J, Qi Q, Liu J, Li L, Li T, Wang X, Lu H, Wu T, Zhu M, Ni P, Han H, Dong W, Ren X, Feng X, Cui P, Li X, Wang H, Xu X, Zhai W, Xu Z, Zhang J, He S, Zhang J, Xu J, Zhang K, Zheng X, Dong J, Zeng W, Tao L, Ye J, Tan J, Ren X, Chen X, He J, Liu D, Tian W, Tian C, Xia H, Bao Q, Li G, Gao H, Cao T, Wang J, Zhao W, Li P, Chen W, Wang X, Zhang Y, Hu J, Wang J, Liu S, Yang J, Zhang G, Xiong Y, Li Z, Mao L, Zhou C, Zhu Z, Chen R, Hao B, Zheng W, Chen S, Guo W, Li G, Liu S, Tao M, Wang J, Zhu L, Yuan L, Yang H (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science 296: 79-92.

Zhang LB, Zhu Q, Wu ZQ, Ross-Ibarra J, Gaut BS, Ge S, Sang T (2009) Selection on grain shattering genes and rates of rice domestication. New Phytol 184: 708-720.

Zhang Y, Zhang Z, Sun X, Zhu X, Li B, Li J, Guo H, Chen C, Pan Y, Liang Y, Xu Z, Zhang H, Li Z (2019) Natural alleles of *GLA* for grain length and awn development were differently domesticated in rice subspecies *japonica* and *indica*. Plant Biotechnol J. 17: 1547-1559.

Zhao Q, Feng Q, Lu H, Li Y, Wang A, Tian Q, Zhan Q, Lu Y, Zhang L, Huang T, Wang Y, Fan D, Zhao Y, Wang Z, Zhou C, Chen J, Zhu C, Li W, Weng Q, Xu Q, Wang ZX, Wei X, Han B, Huang X (2018) Pan-genome analysis highlights the extent of genomic variation in cultivated and wild rice. Nat Genet 50: 278-284.

Zheng Y, Crawford GW, Jiang L, Chen X (2016) Rice domestication revealed by reduced shattering of archaeological rice from the Lower Yangtze valley. Scientific Reports 6: 28136.

Zhu Y, Ellstrand NC, Lu BR (2012) Sequence polymorphisms in wild, weedy, and cultivated rice suggest seed-shattering locus *sh4* played a minor role in Asian rice domestication. Ecol Evol 2: 2106-2113.

Zhu BF, Si L, Wang Z, Zhou Y, Zhu J, Shangguan Y, Lu D, Fan D, Li C, Lin H, Qian Q, Sang T, Zhou B, Minobe Y, Han B (2011) Genetic control of a transition from black to straw-white seed hull in rice domestication. Plant Physiol 155: 1301-1311.

Zhu Z, Tan L, Fu Y, Liu F, Cai H, Xie D, Wu F, Wu J, Matsumoto T, Sun C (2013) Genetic control of inflorescence architecture during rice domestication. Nat Commun 4: 2200.

Zohary D (1969) The progenitors of wheat and barley in relation to domestication and agricultural dispersal in the Old World. In: Ucko PJ, Dimbleby GW (eds) The domestication and exploitation of plants and animals. Duckworth, London, pp. 47-66.

Zohary D (2004) Unconscious selection and the evolution of domesticated plants. Econ Bot 58: 5–10.

Figures

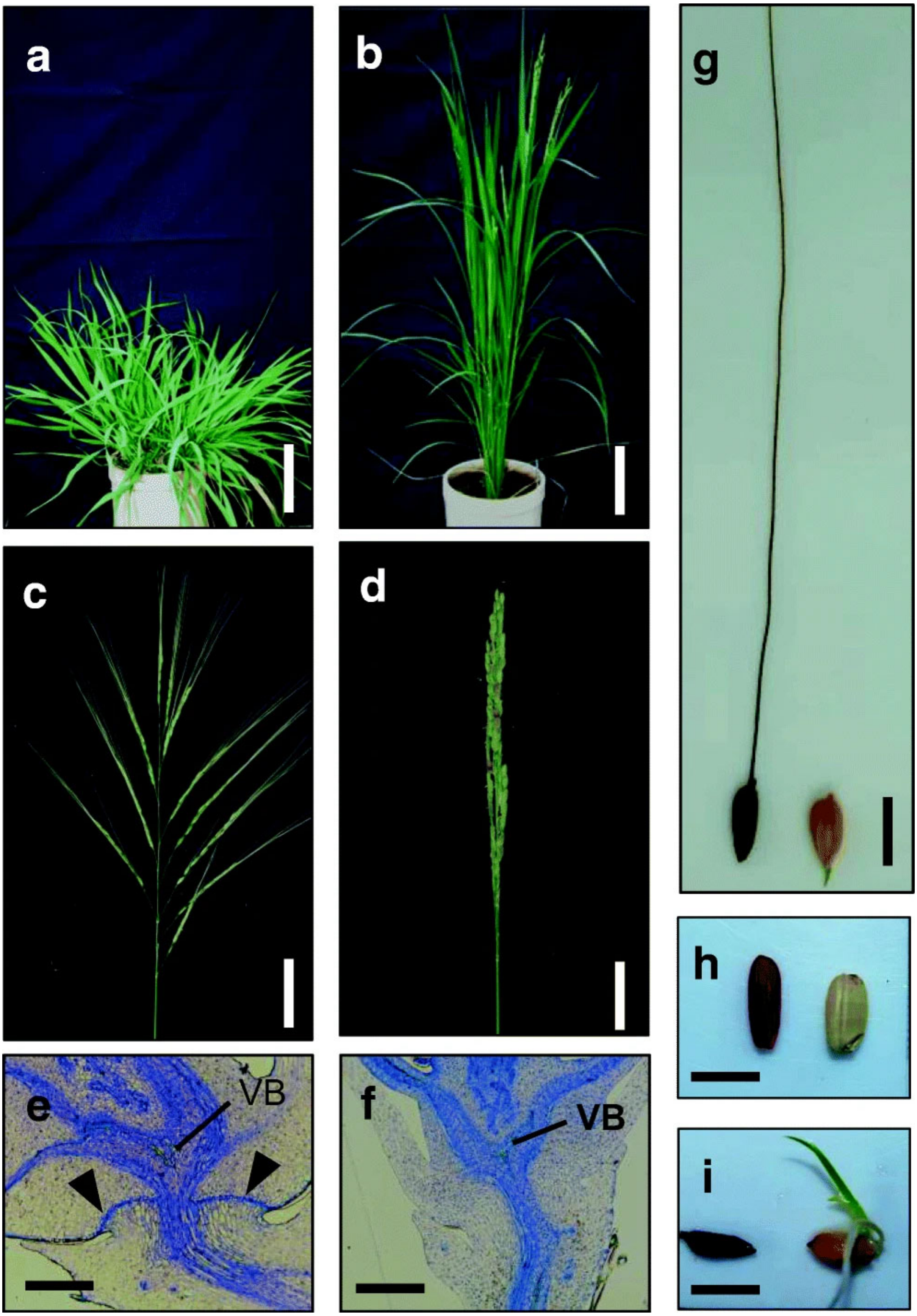

Fig. 1. Comparison of plant morphologies between wild rice, *Oryza rufipogon* acc. W630, and

a *japonica* rice cultivar, *O. sativa* ‘Nipponbare’. Plant architecture (A and B), panicle shape (C and D), and abscission layer formation (E and F) of W630 (A, C, and E) and Nipponbare (B, D, and F). Seed dormancy (G), awn formation (H), seed colour (I) of W630 (left) and ‘Nipponbare’ (right). Scale bars = 20 cm (A and B), 5 cm (C and D), 50 µm (E and F), and 5 mm (G, H, and I). In the photo of seed dormancy (G), W630 and Nipponbare seeds were induced for germination by water soaking a month after harvest. In the photo of abscission layer formation (E and F), black arrowheads indicate the abscission layer. VB, vascular bundle

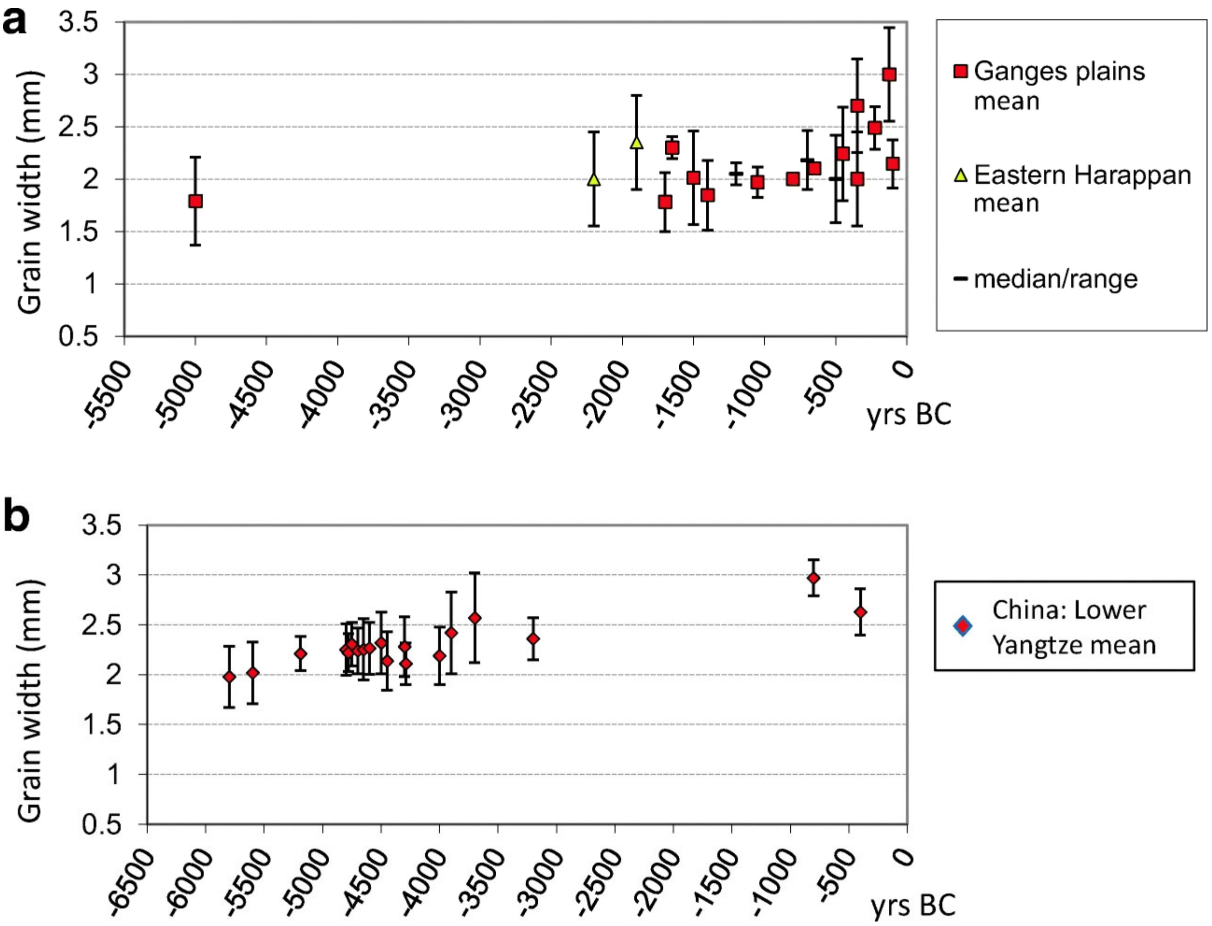

Fig. 2. Comparison of archaeological time series of grain size increase (mean grain width and standard deviation, plotted against estimated median age), from (A) northern India and (B) eastern China (Lower Yangtze) (data from Fuller et al 2010; 2014). In India some sites are plots as range and when only the maximum and minimum metrics were available in the primary

sources. Eastern Harappan sites in the upper Indus valley tributaries are indicated as a potential distinct group.

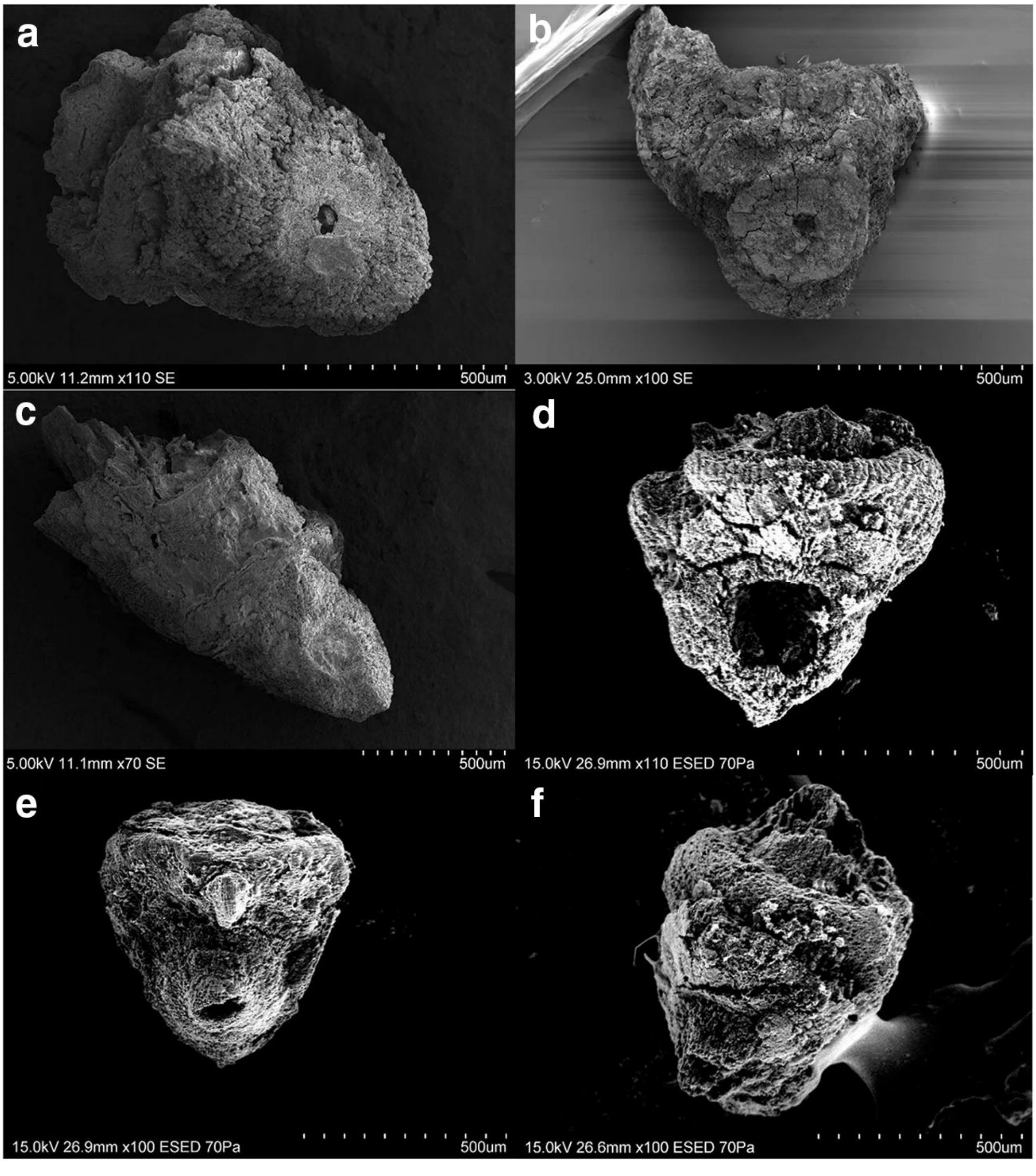

Fig. 3. Scanning electron micrographs of spikelet bases identified as wild-type from (A) Ban Non Wat (Thailand), (B) Tian Lou Shan (China); domesticated-type from (C) Wari Bateshwar (Bangladesh), (D) Tian Lou Shan (China); immature-type from (E) and (F) Tian Lou Shan

(China).

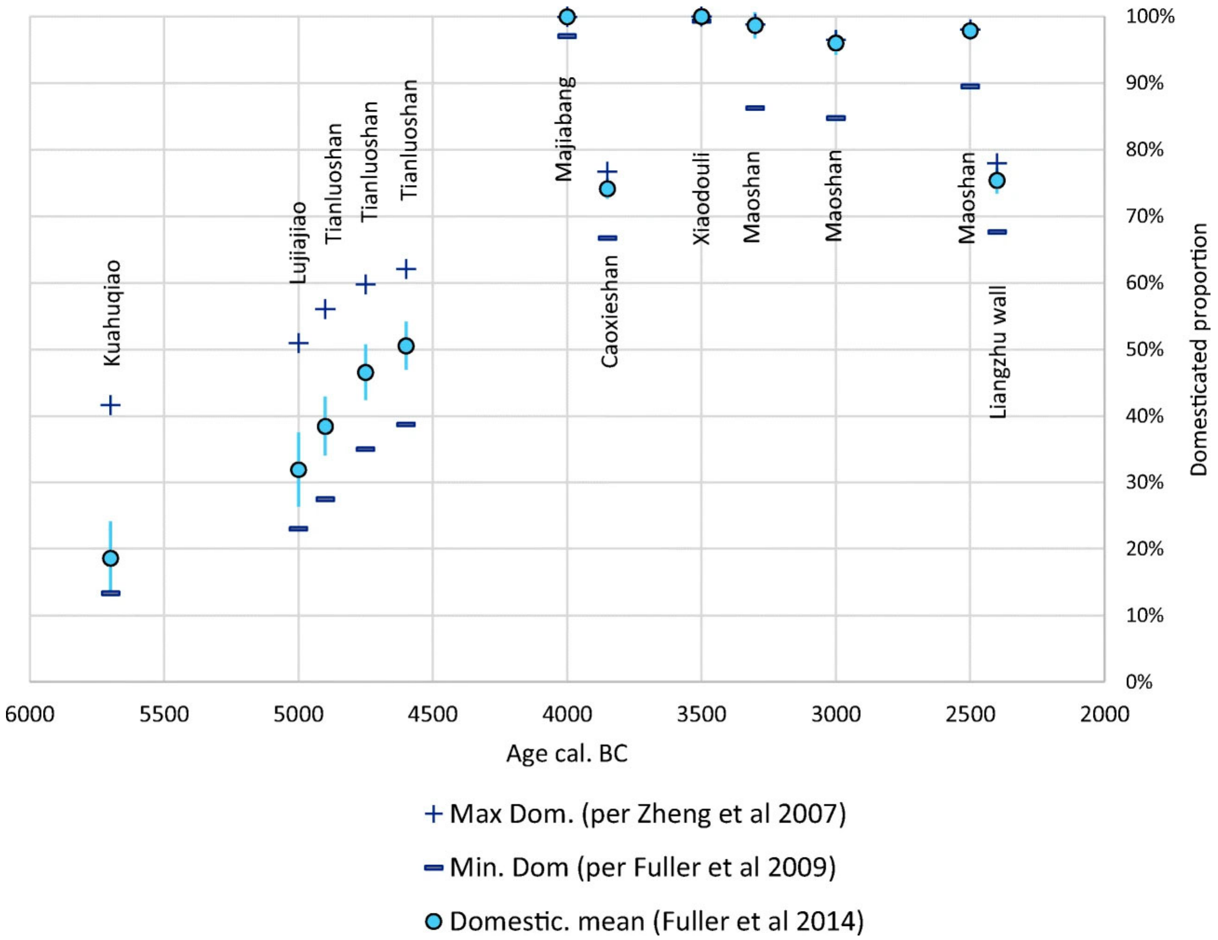

Fig. 4. Time series of archaeological spikelet base data from eastern China (Lower Yangtze), comparing different estimated of the domesticated proportion: a maximum estimate in which "immature" forms are counted as domesticated (after Zheng et al 2007; 2016), and minimum estimate which counted "immature" forms with wild types (as per Fuller et al 2009) and a mean domesticated fraction based on the ratio of domesticated and wild forms only for which "immature" types provide an estimate of error (standard deviation) (as per Fuller et al 2014). Archaeological site names for data points indicated.

Table 1 Major traits and genes/loci involved in rice domestication

Trait	Gene/locus	Mutation	Molecular function	References
Plant architecture	*PROG1*	SNP and other	C2H2 zinc-finger transcriptional regulator	Tan et al. (2008), Jin et al. (2008)
	SPROG1	Deletion	C2H2 zinc-finger transcriptional regulator	Wu et al. (2018)
	OsLG1	N.A. (SNP?)	*SQUAMOSA* promoter-binding protein	Ishii et al. (2013), Zhu et al. (2013)
Seed shattering	*sh4*	SNP	Myb3 transcriptional regulator	Li et al. (2006b)
	qSH1	SNP	BEL1-type homeobox transcriptional regulator	Konishi et al. (2006)
	qSH3	N.A.	N.A.	Htun et al. (2014)
	qCSS3	N.A.	N.A.	Tsujimura et al. (2019)
Awn	*An-1*	Deletion	Basic helix-loop-helix transcriptional regulator	Luo et al. (2013)
	LABA1 / *An-2*	Deletion	Cytokinin-activating enzyme	Hua et al. (2015), Gu et al. (2015)
	RAE2 / *GAD1*	Insertion	Epidermal pattering factor-like	Bessho-Uehara et al. (2016), Jin et al. (2016)
	qAWNL2	N.A.	N.A.	Amarasinghe et al. (2020)
Seed and hull colour	*Bh4*	Deletion	Amino acid transporter	Zhu et al. (2011)
	Rc[a]	Deletion	Basic helix-loop-helix transcriptional regulator	Sweeney et al. (2006)
	Rd	SNP	Dihydroflavonol-4-reductase	Furukawa et al. (2006)
Seed dormancy	*Sdr4*	SNP	Unknown	Sugimoto et al. (2010)
	OsG	SNP	CAAX amino-terminal protease protein	Wang et al. (2018)
Grain size	*qSW5*/*GW5*	Deletion	Novel nuclear protein	Shomura et al. (2008), Weng et al. (2008)
	GS3	SNP	Putative transmembrane protein	Fan et al. (2006)
	GS5	N.A. (promoter)	Serine carboxypeptidase	Li et al. (2011)
	GW2	Deletion	RING-type E3 ubiquitin ligase	Song et al. (2007)

1. [a] *Rc* might be involved in the control of seed dormancy (Gu et al. 2011)