Kobe University Repository : Kernel

PDF issue: 2026-11-20

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

Ishikawa, Ryo

Castillo, Cristina C.

Fuller, Dorian Q.

(Citation)
Archaeological and Anthropological Sciences,12:197

(Issue Date)
2020-08-01

(Resource Type)
journal article

(Version)
Accepted Manuscript

(Rights)
© Springer-Verlag GmbH Germany, part of Springer Nature 2020. This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature’s AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of…

(URL)
https://hdl.handle.net/20.500.14094/90009576

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

Ryo Ishikawa[1, 2] *, Cristina C. Castillo[2], Dorian Q. 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, 31-34 Gordon Square, London, WC1H 0PY, UK

* Corresponding Author:

Ryo Ishikawa e-mail: r-ishika@port.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.

Keywords

Palaeoethnobotany, *Oryza sativa*, Domestication, genome, seed shattering, awn

Subtitle

Domestication-related traits in rice

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 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 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 1,083 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 >1,400 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

91 material or to genotypes shared by phylogenetic linages. In this paper, we review some
92 domestication-related traits and genes or loci that determine them, with a particular focus
93 on two traits that likely played key roles in early rice domestication: closed panicles and
94 the loss of seed shattering. Furthermore, we discuss some of the results from studies using
95 wild rice introgression lines which provide information on different aspects of rice
96 domestication.
97
98 **Major traits selected in rice domestication**
99 Genetic variations in wild rice are important since mutations lead to
100 agronomically important traits (Ishii et al. 2011). Over the past decade, several important
101 genes involved in rice domestication have been identified and cloned. These traits are
102 often linked to ease of cultivation and harvest as well as an increase in edible portions of
103 rice (Table 1). These are also known to be involved in the domestication of maize, wheat,
104 barley, sorghum, oats, and other crops (Doebley et al. 2006). These recurrently selected
105 common traits between species represent convergent evolution and are known as the
106 domestication syndrome (Sakuma et al. 2011; Fuller 2007). The major traits selected in
107 rice domestication with causal mutation and molecular function are shown in Table 1.
108

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 (Figs. 1A and 1B). 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 (Figs. 1C and 1D). 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 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

145 study using the sequences from wild rice accessions showed that the one causal single
146 nucleotide polymorphism (SNP) at the *SPR3* region was associated with panicle shape
147 (Zhu et al. 2013), although the causal mutation at *SPR3* still remains to be elucidated.
148
149 ***Loss of seed shattering***
150 The loss of seed shattering is the most striking example of a trait that contributes
151 to easy harvesting with guaranteed yield, and its widely seen as the key trait that separated
152 domesticated cereals from wild cereals (e.g. Harlan 1975; Fuller 2007).. In wild rice, an
153 abscission layer formed in the basal part of the grain contributes to seed shattering (Figs.
154 1E and 1F). Variations in the degree of seed shattering are observed in cultivated rice
155 (Konishi et al. 2006), indicating that several genes or genetic changes are involved.
156 Two loci involved in the loss of seed shattering were previously identified by
157 quantitative trait loci (QTL) analyses (Table 1). Li et al. (2006a) reported *sh4* as a major
158 locus explaining the difference in the degree of seed shattering between *indica* cultivated
159 rice and annual wild rice, *O. nivara* (the annual form of *O. rufipogon sensu lato*).
160 Molecular genetic analyses showed that the causal mutation was found to be a SNP in a
161 gene encoding a Myb transcription factor (Li et al. 2006b, Lin 2007). Experiments
162 growing domesticated rice with two variants of the *sh4* gene indicated that seed-shattering

163 behaviour was restored in the wild allele at *sh4* as opposed to the cultivated allele. This
164 confirms that the wild *sh4* allele had a role in promoting seed shattering. The original
165 inference is that the mutation confers a non-shattering effect (Li et al. 2006b; Sang and
166 Ge 2007b).

167 Another locus controlling the degree of seed shattering is *qSH1*, which was
168 identified as controlling the degree of seed shattering between two rice cultivars,
169 'Kasalath' (representing subsp. *indica*) and Nipponbare (representing subsp. *japonica*)
170 (Konishi et al. 2006). This QTL explained approximately 70% of the phenotypic variance.
171 A causal mutation was found to be an SNP at the promoter region of the gene encoding a
172 homolog of *Arabidopsis REPLUMLESS* (*RPL*). The mutation at *qSH1* resulted in a
173 complete loss of shattering by eliminating abscission layer formation. Since mutation at
174 *qSH1* is only observed in '*temperate japonica*' cultivars (Konishi et al. 2006; Onishi et al.
175 2007b; Zhang et al. 2009), it is proposed that selection of the loss of the functional allele
176 at *qSH1* occurred after the *sh4* mutation was fixed, at a later stage of diversification of
177 breeding in domesticated rice. Because *sh4* and *qSH1* were identified in QTL analyses as
178 having large effects on seed shattering, they were believed to be the genes explaining the
179 loss of seed shattering during rice domestication. However, recent research through
180 longer term experiments using plant materials with wild rice introgression lines have

181 proved this view to be over-simplified (Ishii and Ishikawa 2018; see discussion below).

182

183 *Loss of awn formation*

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

197 Three genes involved in the loss of awn formation have been reported in Asian
198 cultivated rice (Table 1). *An-1*, encoding a basic helix-loop-helix (bHLH) protein

199 controlling cell division, was found on chromosome 4 (Luo et al. 2013). Molecular
200 analyses showed that the *O. rufipogon* allele at *An-1* was expressed at the distal end of
201 the lemma while the *O. sativa* allele was found to be down-regulated. *LABA1* (*Long And*
202 *Barbed Awn 1*) / *An-2*, a gene encoding a cytokinin-activating enzyme, was identified on
203 chromosome 4 from the segregating population between *O. sativa* subsp. *indica* and *O.*
204 *rufipogon* (Hua et al. 2015; Gu et al. 2015). Causal mutation was shown to be a 1 bp
205 deletion at the gene. Genome analyses showed that the mutation at *LABA1* / *An-2*
206 originated in some *japonica* rices and moved into the *indica* gene pool via introgression
207 (Hua et al. 2015). However, further studies are still required to understand the origin of
208 the mutation and directionality of introgression, because the mutation is frequently
209 observed in wild rice (Civáň and Brown 2017), and the possibility of introgression of
210 cultivated alleles into wild rice populations needs to be explored.

211 Recently, a QTL controlling awn development on chromosome 8, *RAE2*
212 (*Regulator of Awn Elongation 2*), was found to encode a gene of EPFL (Epidermal
213 Patterning Factor-Like) family member (Bessho-Uehara et al. 2016). A 2-bp insertion in
214 the second exon of the gene was found to give rise to a truncated protein due to a
215 frameshift mutation. Another research group also identified the same gene named *Grain*
216 *number, grain length and Awn Development 1* (*GAD1*), shared between wild rice *O.*

217 *rufipogon* and an *indica* rice cultivar (Jin et al. 2016). More recently, natural alleles at
218 *RAE2*/*GAD1* were reported as they have a distinct role on gene expression and protein
219 function (Zhang et al. 2019).

220 Interestingly, the loss of awn formation is also observed in African cultivated
221 rice, *O. glaberrima*. A recent comparative analysis of the loss of awn formation between
222 *O. sativa* and *O. glaberrima* reported that *O. sativa* lost its awn through mutations at two
223 loci, *RAE1*, which is identical to *An-1* on chromosome 4, and *RAE2* on chromosome 8,
224 while *O. glaberrima* lost it through mutation at *RAE3* on chromosome 6. These results
225 likely indicate that selection for awn loss in the two cultivated rice species is mediated by
226 different loci (Furuta et al., 2015). The selection for the same trait at different loci is an
227 example of convergent evolution, an important feature of domestication (Fuller et al.
228 2014).

229 The three genes involved in the awnless phenotypes in Asian cultivated rice, *O.*
230 *sativa*, are not enough to fully explain the complete loss of awn formation when these
231 alleles are evaluated in the genetic background of wild rice, *O. rufipogon*, indicating that
232 some additional mutations(s) may still be involved in the selection of shorter awn length
233 during rice domestication and diversification (Ikemoto et al. 2017; Amarasinghe et al.
234 2020).

235

236 *Seed and hull colour*

237 Colours of seeds and hulls were also selected for in rice domestication. Wild rice
238 has red-pigmented pericarps, while cultivated rice has predominantly white pericarps (Fig.
239 1I). However, these mutations are what are known as post-domestication genes since they
240 are not necessarily essential in improving harvests. They are related more to cultural
241 histories (Fuller and Castillo 2016). Many red pericarp rices are cultivated and possess
242 other domestication traits. While white rice is predominant in Asia, there are also dark
243 purple or "black" pericarp rices, which represent an alternative pathway of cultural
244 selection (Oikawa et al. 2015).

245 In classical genetic studies, pericarp colour was known to be regulated by two
246 loci, *Rc* (brown pericarp and seed coat) and *Rd* (red pericarp and seed coat), and the
247 functional alleles at the two loci produce a red pigmented pericarp (Table 1; Sweeney et
248 al. 2006; Furukawa et al. 2006). Genetic analysis identified that *Rc* encoded a bHLH
249 protein. A 14-bp deletion within exon 6 that knocked out the bHLH domain of the protein
250 was observed in most cultivated rice. Some cultivars were found to encode loss of the
251 functional allele at *Rc* caused by an SNP, which was detected in some *aus* and aromatic
252 rice, implying an independent origin of the selection at the *Rc* locus (Sweeney et al. 2007).

253 A recent analysis by Civáň and Brown (2017) showed that some wild rice accessions were
254 found to carry a cultivated-type allele (a 14-bp deletion) for *Rc* gene, although it is not
255 known whether they express the white phenotype. It is also unclear whether this could
256 indicate some standing variation in wild rice or whether the presence of the allele in wild
257 rice might be caused by a gene flow from cultivated rice. The other locus, *Rd*, was found
258 to encode dihydroflavonol-4-reductase (DFR) involved in proanthocyanidin synthesis,
259 and two types of truncated DFR protein were caused by independent premature stop
260 codons in the first and second exons of the gene encoding DFR (Furukawa et al. 2006).
261 Changes in hull colour also took place with crop domestication, although these
262 would have occurred as post-domestication events (Purugganan and Fuller 2009). Black
263 seed hulls in wild rice and straw-white seed hulls in cultivated rice are commonly
264 observed (Fig. 1H). Genetic analysis of the causal gene for hull colour identified the *Bh4*
265 gene, encoding an amino acid transporter (Table 1). Two origins of the mutations causing
266 truncation of the BH4 protein were identified; one was a 22-bp deletion within exon 3 of
267 the *Bh4* gene and the other was an SNP causing a premature stop codon in the third exon
268 (Zhu et al. 2011). A survey of *Bh4* genotypes using a large panel of rice landraces showed
269 that most of them had the 22-bp deletion, while a few had the SNP. Interestingly, a few
270 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 plant 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 homolog of the *Rc* gene that is involved in grain tannins 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

307 plant domestication (Wright 2015; Denham et al. 2020). In rice, grain size increased from
308 wild to cultivated rice, especially grain breadth and thickness. Trends for increasing grain
309 size have been identified in time series of archaeological rice grain measurements in both
310 China and India (Fuller et al 2010; 2014), e.g. Fig. 2.

311 Loci possibly involved in an increase in grain size during rice domestication are
312 listed in Table 1. *GS3* was identified through a QTL analysis between two cultivated rice
313 species with small and large grains, and was found to encode a putative transmembrane
314 protein (Fan et al. 2006). *GS5* was shown to encode a putative serine carboxypeptidase.
315 A higher expression level of the *GS5* gene was found to be associated with increased grain
316 size (Li et al. 2011). A loss of a functional allele at *GW2* encoding a previously unknown
317 RING-type protein with E3 ubiquitin ligase activity was found to enhance grain width
318 and weight via an increase in cell numbers in the spikelet hull (Song et al. 2007). When
319 *qSW5* was cloned through a QTL analysis of grain size between *indica* and *japonica*
320 cultivated rice, 38.5% of the variation was explained in the F_2 population. A deletion at
321 *qSW5* was found to be a causal mutation leading to an increase in seed size, although the
322 molecular function of the encoded protein is still unknown (Shomura et al. 2008).
323 Genome-wide analyses of the four genes involved in the regulation of grain size suggested
324 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 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. 2011, 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

343 to a lesser extent loss of awns. Many of the domestication traits mentioned in the
344 preceding section are not visible archaeologically because of preservation issues. The
345 parts of the rice plant that preserve best in archaeological sites are the grain and spikelet
346 base, which shows the abscission scar. Hull colour and plant architecture do not preserve
347 as the archaeobotanical material is normally charred although theoretically, if one
348 encountered waterlogged material one would be able to evaluate these traits as well.
349 However, neither trait has been documented so far.

350 Spikelets or grains can be measured even when charred, although a shrinkage
351 factor needs to be taken into account in order to compare to modern examples. The
352 difference in grain size is evaluated using a dataset that has a temporal scale, and such
353 data indicate a gradual increase in the mean seed breadth and maximum seed breadth iver
354 thousand s of year (Fuller at at. 2010; 2014). This is also true for examining the loss of
355 seed shattering. Spikelet bases are identified as domesticated-type, wild-type, and
356 immature-type and compared across assemblages over time (Fig. 3). In the archaeological
357 study by Fuller et al (2009), spikelet bases of non-shattering rice were recovered in the
358 Lower Yangtze region of Zhejiang, China, showing archaeobotanical evidence of rice
359 domestication over a period of time (for additional data, see Fuller et al 2014; 2016; Deng
360 et. al. 2015; cf. Zheng et al 2016). Such data can be plotted as time series providing

various estimates of the extent of domestication overtime, 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 the reproductive success of cultivated rice, but also made increased 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

379 of grain from harvesting (Harlan et al. 1975; Fuller 2007). In modern cultivated rice,
380 mature seeds remain in the panicle due to the inhibition of the abscission layer formation.
381 Plants with inhibited abscission layer formation could have been present in wild rice
382 populations, in very low numbers, and may have been unintentionally gathered by hunter-
383 gatherers. The loss of seed shattering was also a key domestication trait in other crop
384 species such as barley (Pourkheirandish et al. 2015) and sorghum (Lin et al. 2011). The
385 increasing proportions of this trait in early cultivated population has also been
386 documented through time series of archaeobotanical data, including in barley and wheat
387 (Fuller et al. 2014; Allaby et al. 2017) and sorghum (Winchell et al. 2018).

388 In the initial stages of rice domestication, wild rice populations would have
389 contained some stands with reduced proportions of seed shattering. Therefore, it would
390 be more accurate to evaluate domestication-related mutated alleles in the wild rice genetic
391 background. Therefore, research carried out at Kobe University has focused on
392 maintaining a wild population of rice (the annual form of *O. rufipogon sensu lato*) and
393 introgressing alleles from domesticated lines into this population to assess the impact of
394 hypothesized domestication alleles on rice characters and through hand harvesting
395 experiments (Ishii and Ishikawa 2018). This is the opposite of most functional genetics
396 which introgresses targeted wild rice alleles into plants of domesticated rice background,

397 which means that interactions with other genes selected in domesticated populations can
398 go undetected..

399 Interestingly, generation of wild introgression lines by recurrent backcrossing
400 showed that the plants carrying domesticated alleles at either of the two seed-shattering
401 loci, *qSH1* or *sh4*, displayed complete wild-type shattering (Ishikawa et al. 2010).
402 Moreover, histological analysis showed complete abscission layer formation in both of
403 the introgression lines (Htun et al. 2014). These findings strongly indicated that a single
404 mutation at either *qSH1* or *sh4* was not sufficient to confer a loss of seed shattering under
405 natural conditions. Therefore, some other mutation(s) must have been required for an
406 initial loss of seed shattering. The seed shattering locus is *qSH3*, which will be described
407 in the next section, encodes a likely candidate gene (Htun et al. 2014; Ishikawa et al.
408 2017). However, the accumulation and fixation of such mutations took a long time. The
409 complicated process on initial loss of seed shattering raises the simple question whether
410 this trait can be considered a crucial step towards rice domestication (Ray and
411 Chakraborty 2017). Considering these results, there may have been other trait(s) that
412 promoted rice domestication prior to the selection of a loss of seed shattering.

413 One candidate trait that may have been the target of selection before the selection
414 of loss of seed shattering is panicle shape. To understand the role of panicle shape in rice

415 domestication, a near isogenic line (NIL) having a chromosomal segment from *O. sativa*
416 Nipponbare covering at the *SPR3* region in the *O. rufipogon* acc. W630 genetic
417 background was produced by backcrossing. Since wild rice spikelets have long awns on
418 the lemma tips, a closed panicle was found to increase seed retention, since the lower
419 awns supported the upper seeds to stay longer in the panicle (Ishii et al. 2013). To confirm
420 the effect of seed retention by closed panicle on harvesting efficiency, two field
421 experiments were conducted. The time required for seed shattering was evaluated, and
422 closed panicles were found to retain seeds for approximately a day longer than open
423 panicles. This could have visibly aided seed collection in prehistory. A yield test by
424 gathering seeds was also performed in the experimental paddy field. The gathering rates
425 of wild rice and three NILs (*SPR3*, *qSH1*, and *sh4*) were calculated. Results indicated that
426 NIL(*SPR3*) with a closed panicle had a higher yield than W630 and the other NILs with
427 an open panicle, confirming the advantage of closed panicles on yield, increasing
428 gathered grain yields by around 50%. These field results suggest that panicle shape could
429 have been selected for at a relatively early stage of rice domestication (Ishii et al. 2013),
430 even before non-shattering spikelet bases. In addition to the yield, the closed panicle
431 provided an important advantage for rice as a “crop”. During flowering, long awns from
432 the lower flowers blocked the free exposure of the anther when a panicle was closed. This

433 increased the self-pollination rate; which is an important change since recessive mutations
434 are more easily fixed through self-pollination. Self-pollination has often been
435 hypothesized to be a trait that could have aided the domestication process in many cereals
436 (Harlan 1975; Allard 1988; Zohary 1969; 2004). The change in panicle structure from
437 open to closed may have had enormous effects on the selection of other domestication-
438 related traits. High self-pollination behaviour would increase the appearance of valuable
439 traits caused by recessive mutations in the segregating F_2 generation. Once people started
440 to cultivate what they harvested, these plants entered a cycle of cultivation. Therefore,
441 seed shattering and pollinating behaviours were greatly influenced by a simple
442 morphological change that may have emerged at the very beginning of cultivation or even
443 intensive harvesting from the wild (Ishii et al. 2013). Closed panicles would increase
444 return rates with typical hunter-gatherer harvesting methods such as hand stripping or
445 basket swinging.
446
447 **The processes of loss of seed shattering is more complicated than expected**
448 Loss of seed shattering occurs when the abscission layer, which is
449 developmentally programmed in the basal part of the spikelet, is disrupted. In wild rice,
450 after successful fertilization, the abscission layer begins to degrade as the seeds mature.

451 They easily dehisce, resulting in successful propagation. As mentioned in the previous
452 section, neither a single mutation at *sh4* nor *qSH1* was sufficient to confer a loss of seed
453 shattering in the wild rice genetic background (Ishikawa et al. 2010; Ishii et al. 2013).
454 Therefore, other mutation(s) together with *sh4* must have been required for the loss of
455 seed shattering in the initial stage of rice domestication. To search for gene(s) involved in
456 the loss of shattering, wild rice lines with chromosomal segments of cultivated rice at *sh4*
457 and *qSH1* that still showed a weak seed-shattering phenotype were further crossed with
458 Nipponbare. In the resulting F_2 population, the two loci were fixed and were not
459 responsible for the variation in the degree of seed shattering. QTL analysis identified a
460 locus on chromosome 3, and the wild allele had a role in enhancing seed shattering (Htun
461 et al. 2014). This locus was named *qSH3* (Table 1), as it was also detected in a previous
462 study (Onishi et al. 2007a); however, the effects of *qSH3* on seed shattering were
463 relatively lower than those of the two loci in the genetic background of cultivated rice.
464 The difference in the effect of *qSH3* suggested that a genetic interaction among seed-
465 shattering loci could have been important in the loss of seed shattering.

466 Since the candidate region at *qSH3* showed the signature of a selective sweep in
467 both *indica* and *japonica* cultivated rices (Xu et al. 2012), the mutation at *qSH3* may have
468 been selected for during rice domestication. Although the *sh4* domesticated allele is

469 commonly observed in all of the cultivated rice tested (Lin et al. 2007), it is not sufficient
470 on its own to confer a loss of seed shattering in wild rice background (Ishikawa et al.
471 2010; Ishii et al. 2013). This contradiction was also observed in some weedy rice that had
472 a non-functional allele at *sh4* but had seed-shattering behaviour (Thurber et al. 2010; Zhu
473 et al. 2012). Therefore, a combination of mutations at seed-shattering loci, like *sh4* and
474 *qSH3*, may have played an important role in the initial loss of seed shattering. To test this
475 idea, introgression lines having chromosomal segments of Nipponbare at sh4 and *qSH3*
476 were produced in a wild rice genetic background. The introgression line showed partial
477 disruption of the abscission layer around the vascular bundle (Inoue et al. 2015). Slight
478 inhibition of abscission layer formation may have had an advantage in that seeds could
479 stay longer in the panicle but could also be easily harvested without tools, i.e. by "forager"
480 methods such as basket swinging, hand paddling or hand stripping. It would be of interest
481 to understand the molecular nature that caused the partial disruption of abscission layer
482 formation.

483 Further research then explored how many genes or mutations were involved in
484 the complete loss of seed shattering in *japonica* rice. The degree of seed shattering of an
485 introgression line of wild rice having the domesticated alleles at the three loci was
486 compared with that of Nipponbare. The breaking tensile strength (BTS) value of the

487 introgression line was significantly lower than that of Nipponbare (Htun et al 2014),
488 suggesting that some other loci may still be involved in the loss of seed shattering in
489 *japonica* rice. To identify the rest of the gene(s), an approach that involves fixing the
490 known QTL(s) in the segregating population would be effective for exploring and
491 uncovering the novel QTL(s). Based on this approach, we recently detected *qCSS3* (*QTL*
492 *for the Control of Seed Shattering in rice on chromosome 3*), a novel locus involved in a
493 loss of seed shattering in *japonica* rice cultivar *O. sativa* Nipponbare (Tsujimura et al.
494 2019). Once all the mutations related to a loss of seed shattering are clarified, information
495 on the causal mutations would be very useful to understand the change in the degree of
496 seed shattering conferred by each locus and their additive effects in combination. This
497 will help to clarify how the mutation took place and the process of rice domestication.

498 It is important to understand how the various mutations in the seed-shattering
499 loci were accumulated and fixed. The domesticated alleles at the three seed-shattering
500 loci all act recessively (Li et al. 2006a; Konishi et al. 2006; Htun et al. 2014), and other
501 domestication-related mutations are often recessive. Therefore, to exhibit a reduction in
502 seed shattering, these recessive alleles must have been fixed through self-pollination. As
503 wild rice has a partial out-crossing behaviour (Phan et al. 2011), the selection of the closed
504 panicle trait might have been an important step allowing mutations to be more easily fixed

505 (Ishii et al. 2013). Therefore, closed panicles may have contributed to the accumulation
506 of these recessive alleles causing loss of seed shattering. Although the order of the
507 selections is unknown, it would be of great interest to test this idea in the future. Present
508 accumulated data suggest that the process of a loss of seed shattering may have been more
509 complicated than expected.
510
511 **Future aspects of the study of rice domestication: power of ancient DNA**
512 The origin of rice domestication has been discussed using data obtained from
513 population genetic analyses (Londo et al. 2006; Molina et al. 2011; Huang et al. 2012; He
514 et al. 2014; Civáň et al. 2015). These data are based on phylogenetic analyses that
515 compare DNA sequences from modern wild rice accessions and cultivars. When the DNA
516 of modern accessions is studied, their domestication-related phenotypes may derive from
517 underlying genetic redundancies, additive loci or alternative loci. Since wild and
518 cultivated rice have few reproductive barriers, gene flow between them is sometimes
519 observed in their natural habitat, where wild rice are grown near cultivated rice, which is
520 often the case through much of South and Southeast Asia, and therefore genes selected
521 under domestication may have been introduced into wild rice populations. Weedy rice is
522 an example of introgression (He et al. 2014; Thurber et al. 2014). For instance, in the case

523 of wild rice outcrossed with cultivated rice, some chromosomal fragments are maintained,
524 although most parts of the chromosome will be replaced in the wild genome by
525 backcrossing. These plants show cultivar-like phenotypes. Thus, we need to pay more
526 attention to these cultivar-like wild rice species recorded as “wild rice” in gene banks.
527 Moreover, most present genetic studies have evaluated the function of the genes
528 responsible for domestication-related traits in the genetic background of cultivated rice.
529 Since cultivated rice may still harbour unidentified mutation(s) involved in the traits,
530 inferences on the domestication-related phenotypes is misleading. The use of wild rice as
531 the genetic background has been shown to provide more accurate information on rice
532 domestication.

533 Despite the identification of a number of genes involved in rice domestication,
534 the limitation of the present genetic studies is the lack of information about when and
535 where the domestication-related mutations were selected. Ancient DNA (aDNA) would
536 help in determining the chronology of rice domestication. Analyses using fragments of
537 preserved DNA would provide more information relative antiquity than those using DNA
538 from modern rice. In barley, ancient DNA extraction from the entire range of occupancy
539 of the site at Qasr Ibrim was performed, and the *Vrs1* locus responsible for row number
540 in barley was studied. The results showed that reassertion of the two-row condition was

541 selected from six-row ancestry for the Qasr Ibrim barley (Palmer et al. 2009). There have
542 been very few aDNA studies conducted on rice (Castillo et al. 2016; Tanaka et al. 2015),
543 and these have used PCR instead of NGS, with limited potential to recover nuclear DNA
544 which provides details on domestication-related genes.

545 One of the limitations of aDNA studies is the state of preservation, which directly
546 complicates DNA recovery. However, the technology for analysing aDNA has improved;
547 therefore, more information could be obtained once the aDNA analysis of rice becomes successful.
548 Therefore, information should be obtained from aDNA or even from some modifications
549 of DNA such as DNA methylation, an important modification for the regulation of gene
550 expression (Smith et al. 2014). A genotyping survey of the domesticated alleles at a
551 certain age and place will provide evidence of selections and the process of domestication.
552 In addition to aDNA analysis, the use of population genetic approaches such as McSwan,
553 a method to detect and date past and recent natural selection events, will be a great help
554 in understanding of the timeframe of selection of various loci (Tournebize et al. 2018).
555 These multidimensional perspectives will contribute to fully understanding the process
556 and chronology of rice domestication.

557 Sang and Ge (2007b) provided two models for rice domestication: snow-balling
558 and combination. In the snow-balling model, a single initial selection triggered rice

559 domestication, and then introgressions from wild rice were repeated in various areas,
560 creating subspecies such as *indica* and *japonica*. This model was essentially the same
561 as that posited by Huang et al (2012). In the combination model, independent selection
562 of useful traits occurred in different locations, and these partially domesticated plants
563 were followed by subsequent combination of the most useful variants. Subsequently,
564 coalescence-based modelling of demographic parameters on modern genomes of 7 rices
565 (two wild and five cultivated rice) confirmed an important level of gene-flow from
566 ancestral domesticated *japonica* rice into *indica* and *aus* rice (Choi et al. 2017). These
567 data are consistent with a form of the combination model, referred to as the "Proto-*indica*
568 hypothesis", which infers that ancient cultivation of ancestral *indica* in South Asia that
569 did not develop a population fixed in domestication features until after introgressions
570 from domesticated *japonica*, thought to have taken place after 4,000 years ago (Fuller and
571 Qin 2009; Fuller 2011; Silva et al 2015). Recently, Civáň and Brown (2018) argued
572 against introgression from *japonica* into early *indica*, and argue instead for three distinct
573 domestications of subsp. *indica*, subsp. *japonica*, and the "*aus*" group of rices with later
574 post-domestication introgression. The structure of genomic diversity in modern rice
575 populations makes it clear that *indica* and *japonica* (and "*aus*") rice had distinct histories
576 of geographical dispersal (Gutaker et al 2020), but a more detailed understanding in each

577 domestication sequence from only modern genetics remains difficult; it also cannot
578 account for any posited extinct or extirpated lineages (Fuller 2011). For testing these
579 alternative hypotheses about domestication, aDNA could prove essential, providing direct
580 integration of archaeobotanical and genomic approaches.
581
582 **Conclusions**
583 In prehistory, humans selected plants based on phenotypes, not based on
584 genomic information. The genetic changes were therefore unintended consequences. It is
585 necessary to pay more attention to the evaluation of phenotypic changes when discussing
586 early events in rice domestication. Recent research has demonstrated that key phenotypic
587 changes were often affected by multiple alleles, and therefore the evaluation of additive
588 effects and interactions of multiple alleles is important. Far from single recessive
589 mutations being the sole cause of domestication morphotypes, as long assumed for non-
590 shattering in cereals (Harlan 1975; Zohary 1969 ; 2004), multiple alleles may contribute
591 to additive effects and interact with other phenotypic changes selected during early
592 harvesting and cultivation, such as awns, closed versus open panicles, dormancy, and
593 other traits. Although the history of rice domestication is still a subject of debate, the
594 accumulation of data related to phenotypic effects from multiple alleles is essential as is

595 archaeological evidence for the appearance and spread of certain phenotypes. The
596 promise of aDNA is that the occurrence of alleles can be directly placed in past time and
597 past geography and connected to visible morphological features of archaeological rice.
598 We are therefore at the beginning of unravelling rice domestication processes and
599 achieving a deeper understanding how this occurred.
600
601 Acknowledgements
602 We thank Dr. T. Ishii for his valuable comments on this manuscript. The wild
603 rice accession, *O. rufipogon* W630, used in the study at Kobe University was provided
604 by the National Institute of Genetics, Mishima, Japan. This project was supported in part
605 by a JSPS bilateral open partnership joint research project (JPJSBP120189948) and
606 Grant-in-Aid from the Japanese Society for Promotion of Science (15KK0280 and
607 18K05594).
608
609 References
610 Allaby RG, Stevens C, Lucas L, Maeda O, Fuller DQ (2017) Geographic mosaics and
611 changing rates of cereal domestication. Philos Trans R Soc Lond B Biol Sci. 372.
612 Allard RW (1988) Genetic changes associated with the evolution of adaptedness in

613 cultivated plants and their wild progenitors. J Hered 79: 225-238.
614 Amarasinghe YPJ, Kuwata R, Nishimura A, Phan PDT, Ishikawa R, Ishii T (2020)
615 Evaluation of domestication loci associated with awnlessness in cultivated rice,
616 *Oryza sativa*. Rice 13: 26.
617 Bessho-Uehara K, Wang DR, Furuta T, Minami A, Nagai K, Gamuyao R, Asano K
618 Angeles-Shim RB, Shimizu Y Ayano M, Komeda N, Doi K, Miura K, Toda Y
619 Kinoshita T, Okuda S, Higashiyama T, Nomoto M, Tada Y, Shinohara H,
620 Matsubayashi Y, Greenberg A, Wu J, Yasui H, Yoshimura A, Mori H, McCouch
621 SR, Ashikari M (2016) Loss of function at *RAE2*, a previously unidentified EPFL,
622 is required for awnlessness in cultivated Asian rice. Proc Natl Acad Sci U S A.
623 113: 8969-8974.
624 Castillo CC, Tanaka K, Sato Y-I, Ishikawa R, Bellina B, Higham C, Chang N, Mohanty
625 R, Kajale M, Fuller DQ (2016) Archaeogenetic study of prehistoric rice remains
626 from Thailand and India: Evidence of early *japonica* in South and Southeast Asia.
627 Archaeological and Anthropological Sciences 8: 523-543.
628 Choi JY, Platts AE, Fuller DQ, Hsing YI, Wing RA, Purugganan MD (2017) The rice
629 paradox: multiple origins but single domestication in Asian rice. Mol Biol Evol. 34:
630 969-979.

631 Civáň P, Craig H, Cox CJ, Brown TA (2015) Three geographically separate
632 domestications of Asian rice. Nat Plants 1: 15164.
633 Civáň P, Brown TA (2017) Origin of rice (*Oryza sativa* L.) domestication genes. Genet
634 Resour Crop Evol 64:1125-1132.
635 Civáň P, Brown TA (2018) Role of genetic introgression during the evolution of cultivated
636 rice (*Oryza sativa* L.). BMC Evol Biol. 18: 57.
637 Civáň P (2019) When bitter is better. Nat Plants 5: 1205-1206.
638 Deng Z, Qin L, Gao Y, Weisskopf AR, Zhang C, Fuller DQ (2015) From early
639 domesticated rice of the middle Yangtze Basin to millet, rice and wheat agriculture:
640 Archaeobotanical macro-remains from Baligang, Nanyang Basin, Central China
641 (6700–500 BC). PLoS One 10: e0139885.
642 Denham T, Barton H, Castillo C, Crowther A, Dotte-Sarout E, Florin A, Pritchard J,
643 Barron A, Zhang Y, Fuller DQ (2020) The domestication syndrome in vegetatively-
644 propagated field crops. Ann Bot 125: 581-597.
645 Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication.
646 Cell 127: 1309-1321.
647 Elbaum R, Zaltzman L, Burgert I, Fratzl P (2007) The role of wheat awns in the seed
648 dispersal unit. Science 316: 884-886.

649 Fan C, Xing Y, Mao H, Lu T, Han B, Xu C, Li X, Zhang Q (2006) *GS3*, a major QTL for
650 grain length and weight and minor QTL for grain width and thickness in rice,
651 encodes a putative transmembrane protein. Theor Appl Genet 112: 1164-1171.
652 Finkelstein R, Reeves W, Ariizumi T, Steber C (2008) Molecular aspects of seed
653 dormancy. Annu Rev Plant Biol. 59: 387-415.
654 Fuller DQ (2007) Contrasting patterns in crop domestication and domestication rates,
655 recent archaeobotanical insights from the Old World. Ann. Bot. 100: 903-924.
656 Fuller DQ (2011) Pathways to Asian civilizations: Tracing the origins and spread of rice
657 and rice cultures. Rice 4: 78-92.
658 Fuller DQ, Allaby R (2009) Seed dispersal and crop domestication: shattering,
659 germination and seasonality in evolution under cultivation. In Annual Plant
660 Reviews, Vol. 38, Fruit Development and Seed Dispersal (ed.: Ostergaard L.). pp.
661 238-295. Wiley-Blackwell, Oxford.
662 Fuller DQ, Qin L (2009) Water management and labour in the origins and dispersal of
663 Asian rice. World Archaeology 41: 88-111.
664 Fuller DQ, Qin L, Zheng Y, Zhao Z, Chen X, Hosoya LA, Sun GP (2009) Science 323:
665 1607-1610.
666 Fuller DQ, Allaby RG, Stevens C (2010) Domestication as innovation: the entanglement

667 of techniques, technology and chance in the domestication of cereal crops. World
668 Archaeology 42: 13-28.
669 Fuller DQ, Denham T, Arroyo-Kalin M, Lucas L, Stevens CJ, Qin L, Allaby RG,
670 Purugganan MD (2014) Convergent evolution and parallelism in plant
671 domestication revealed by an expanding archaeological record. Proc Natl Acad Sci
672 U S A 111: 6147-6152.
673 Fuller DQ, Castillo CC (2016) Diversification and cultural construction of a crop: the
674 case of glutinous rice and waxy cereals in the food cultures of eastern Asia, In;
675 Thorp JL, Katzenberg MA (eds) The Oxford handbook of the archaeology of food
676 and diet. Oxford: Oxford University Press.
677 Fuller DQ, Weisskopf AR, Castillo C (2016) Pathways of rice diversification across Asia.
678 Archaeology International 19: 84-96.
679 Furukawa T, Maekawa M, Oki T, Suda I, Iida S, Shimada H, Takamure I, Kadowaki K
680 (2006) The *Rc* and *Rd* genes are involved in proanthocyanidin synthesis in rice
681 pericarp. Plant J 49: 91-102.
682 Furuta T, Komeda N, Asano K, Uehara K, Gamuyao R, Angeles-Shim RB, Nagai K, Doi
683 K, Wang DR, Yasui H, Yoshimura A, Wu J, McCouch SR, Ashikari M (2015)
684 Convergent loss of awn in two cultivated rice species *Oryza sativa* and *Oryza*

685 *glaberrima* is caused by mutations in different loci. G3 5: 2267-2274.
686 Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A,
687 Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM,
688 Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S,
689 Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing
690 R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R,
691 Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM,
692 Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T,
693 Oliphant A, Briggs S (2002) A draft sequence of the rice genome (*Oryza sativa* L.
694 ssp. *japonica*). Science 296: 92-100.
695 Gross BL, Zhao Z (2014) Archaeological and genetic insights into the origins of
696 domesticated rice. Proc Natl Acad Sci U S A 111: 6190-6197.
697 Gu B, Zhou T, Luo J, Liu H, Wang Y, Shangguan Y, Zhu J, Li Y, Sang T, Wang Z, Han B
698 (2015) *An-2* encodes a cytokinin synthesis enzyme that regulates awn length and
699 grain production in rice. Mol Plant 8: 1635-1650.
700 Gu XY, Kianian SF, Hareland GA, Hoffer BL, Foley ME (2005) Genetic analysis of
701 adaptive syndromes interrelated with seed dormancy in weedy rice (*Oryza sativa*
702 L.). Theor Appl Genet 110: 1108-1118.

703 Gu XY, Foley ME, Horvath DP, Anderson JV, Feng J, Zhang L, Mowry CR, Ye H, Suttle
704 JC, Kadowaki K, Chen Z (2011) Association between seed dormancy and pericarp
705 color is controlled by a pleiotropic gene that regulates abscisic acid and flavonoid
706 synthesis in weedy red rice. Genetics 189: 1515-1524.
707 Gutaker RM, Groen SC, Bellis ES, Choi JY, Pires IS, Bocinsky RK, Slayton ER, Wilkins
708 O, Castillo CC, Negrão S, Oliveira MM (2020) Genomic history and ecology of the
709 geographic spread of rice. Nature Plants 6: 492-502.
710 Harlan JR (1975) Crops and Man. American Society of Agronomy, Madison.
711 He Z, Jiang X, Ratnasekera D, Grassi F, Perera U, Lu BR (2014) Seed-mediated gene
712 flow promotes genetic diversity of weedy rice within populations: implications for
713 weed management. PLoS One 9: e112778.
714 He Z, Zhai W, Wen H, Tang T, Wang Y, Lu X, Greenberg AJ, Hudson RR, Wu CI, Shi S
715 (2011) Two evolutionary histories in the genome of rice: the roles of domestication
716 genes. PLoS Genet 7: e1002100.
717 Htun TM, Inoue C, Chhourn O, Ishii T, Ishikawa R (2014) Effect of quantitative trait loci
718 for seed shattering on abscission layer formation in Asian wild rice *Oryza rufipogon*.
719 Breed Sci 64: 199-205.
720 Hua L, Wang DR, Tan L, Fu Y, Liu F, Xiao L, Zhu Z, Fu Q, Sun X, Gu P, Cai H, McCouch SR,

721 Sun C (2015) *LABA1*, a domestication gene associated with long, barbed awns in wild rice.
722 Plant Cell 27: 1875-1888.
723 Huang X, Kurata N, Wei X, Wang ZX, Wang A, Zhao Q, Zhao Y, Liu K, Lu H, Li W, Guo
724 Y, Lu Y, Zhou C, Fan D, Weng Q, Zhu C, Huang T, Zhang L, Wang Y, Feng L,
725 Furuumi H, Kubo T, Miyabayashi T, Yuan X, Xu Q, Dong G, Zhan Q, Li C,
726 Fujiyama A, Toyoda A, Lu T, Feng Q, Qian Q, Li J, Han B (2012) A map of rice
727 genome variation reveals the origin of cultivated rice. Nature 490: 497-501.
728 International rice genome sequencing project 2005. The map-based sequence of the rice
729 genome. Nature 436: 793-780.
730 Inoue C, Htun TM, Inoue K, Ikeda K, Ishii T, Ishikawa R (2015) Inhibition of abscission
731 layer formation by an interaction of two seed-shattering loci, *sh4* and *qSH3*, in rice.
732 Genes Genet Syst 90: 1-9.
733 Ishii T, Hiraoka T, Kanzaki K, Akimoto M, Shishido R, Ishikawa R (2011) Evaluation of
734 genetic variation among wild populations and local varieties of rice. Rice 4: 170-
735 177.
736 Ishii T, Numaguchi K, Miura K, Yoshida K, Thanh PT, Htun TM, Yamasaki M, Komeda
737 N, Matsumoto T, Terauchi R, Ishikawa R, Ashikari M. (2013) *OsLG1* regulates a
738 closed panicle trait in domesticated rice. Nat Genet 45; 462-465.

739 Ishii T, Ishikawa R (2018) Domestication loci controlling panicle shape, seed shattering,
740 and seed awning. In: Sasaki T, Ashikari M (eds) Rice Genomics, Genetics and
741 Breeding. Springer, Singapore, pp. 207-221.
742 Ishikawa R, Thanh PT, Nimura N, Htun TM, Yamasaki M, Ishii T (2010) Allelic
743 interaction at seed-shattering loci in the genetic backgrounds of wild and cultivated
744 rice species. Genes Genet Syst 85: 265-271.
745 Ishikawa R, Nishimura A, Htun TM, Nishioka R, Oka Y, Tsujimura Y, Inoue C, Ishii T
746 (2017) Estimation of loci involved in non-shattering of seeds in early rice
747 domestication. Genetica 145: 201-207.
748 Izawa T, Konishi S, Shomura A, Yano M 2009. DNA changes tell us about rice
749 domestication. Curr Opin Plant Biol. 12: 185-192.
750 Jin J, Huang W, Gao JP, Yang J, Shi M, Zhu MZ, Luo D, Lin HX (2008) Genetic control
751 of rice plant architecture under domestication. Nat Genet 40: 1365-1369.
752 Jin J, Hua L, Zhu Z, Tan L, Zhao X, Zhang W, Liu F, Fu Y, Cai H, Sun X, Gu P, Xie D,
753 Sun C (2016) *GAD1* encodes a secreted peptide that regulates grain number, grain
754 length, and awn development in rice domestication. Plant Cell 28: 2453-2463.
755 Khush GS (2005) What it will take to feed 5.0 billion rice consumers in 2030. Plant Mol
756 Biol 59: 1-6.

757 Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, Yano M (2006) An SNP caused
758 loss of seed shattering during rice domestication. Science 312: 1392-1396.
759 Kovach MJ, Sweeney MT, McCouch SR (2007) New insights into the history of rice
760 domestication Trends Genet 23: 578-587.
761 Larson G, Piperno DR, Allaby RG, Purugganan MD, Andersson L, Arroyo-Kalin M,
762 Barton L, Vigueira CC, Denham T, Dobney K, Doust AN, Geots P, Gilbert MPT,
763 Gremillion KJ, Lucas L, Lukens L, Marshall FB, Olsen KM, Pires JC, Richerson
764 PJ, Rubio de Casa R, Sanjur OI, Thomas MG, Fuller DQ (2014) Current
765 perspectives and the future of domestication studies. Proc Natl Acad Sci U S A 111:
766 6139-46.
767 Lee SJ, Oh CS, Suh JP, McCouch SR, Ahn SN (2005) Identification of QTLs for
768 domestication-related and agronomic traits in an *Oryza sativa* × *O. rufipogon* BC_1F_7
769 population. Plant Breed 124: 209-219.
770 Li C, Zhou A, Sang T (2006a) Genetic analysis of rice domestication syndrome with the
771 wild annual species, *Oryza nivara*. New Phytol 170: 185-193.
772 Li C, Zhou A, Sang T (2006b) Rice domestication by reducing shattering. Science 311:
773 1936-1939.
774 Li Y, Fan C, Xing Y, Jiang Y, Luo L, Sun L, Shao D, Xu C, Li X, Xiao J, He Y, Zhang Q

775 (2011) Natural variation in *GS5* plays an important role in regulating grain size and
776 yield in rice. Nat Genet 43: 1266-1269.
777 Lin Z, Griffith ME, Li X, Zhu Z, Tan L, FuY, Zhang W, Wang X, Xie D, Sun C (2007)
778 Origin of seed shattering in rice (*Oryza sativa* L.). Planta 226: 11-20.
779 Lin Z, Li X, Shannon LM, Yeh CT, Wang ML, Bai G, Peng Z, Li J, Trick HN, Clemente
780 TE, Doebley J, Schnable PS, Tuinstra MR, Tesso TT, White F, Yu J (2012) Parallel
781 domestication of the *Shattering1* genes in cereals. Nat Genet 44: 720-724.
782 Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA (2006) Phylogeography of Asian
783 wild rice, *Oryza rufipogon*, reveals multiple independent domestications of
784 cultivated rice, *Oryza sativa*. Proc Natl Acad Sci U S A 103: 9578-9583.
785 Lu L, Shao D, Qiu X, Sun L, Yan W, Zhou X, Yang L, He Y, Yu S, Xing Y (2013) Natural
786 variation and artificial selection in four genes determine grain shape in rice. New
787 Phytol 200: 1269-1280.
788 Luo JJ, Hao W, Jin J, Gao JP, Lin HX (2008) Fine mapping of *Spr3*, a locus for spreading
789 panicle from African cultivated rice (*Oryza glaberrima* Steud.). Mol Plant 1: 830-
790 838.
791 Luo J, Liu H, Zhou T, Gu B, Huang X, Shangguan Y, Zhu J, Li Y, Zhao Y, Wang Y, Zhao
792 Q, Wang A, Wang Z, Sang T, Wang Z, Han B (2013) *An-1* encodes a basic helix-

793 loop-helix protein that regulates awn development, grain size, and grain number in
794 rice. Plant Cell 25: 3360-3376.
795 Molina J, Sikora M, Garud N, Flowers JM, Rubinstein S, Reynolds A, Huang P, Jackson
796 S, Schaal BA, Bustamante CD, Boyko AR, Purugganan MD (2011) Molecular
797 evidence for a single evolutionary origin of domesticated rice. Proc Natl Acad Sci
798 U S A 108: 8351-8356.
799 Oka HI (1988) Origin of Cultivated Rice. Elsevier, Tokyo, p. 254.
800 Oikawa T, Maeda H, Oguchi T, Yamaguchi T, Tanabe N, Ebana K, Yano M, Ebitani T,
801 Izawa T (2015) The birth of a black rice gene and its local spread by introgression.
802 The Plant Cell 27: 2401-2414.
803 Onishi K, Horiuchi Y, Ishigoh-Oka N, Takagi K, Ichikawa N, Maruoka M, Sano Y (2007a)
804 A QTL cluster for plant architecture and its ecological significance in Asian wild
805 rice. Breed. Sci. 57: 7-16.
806 Onishi K, Takagi K, Kontani M, Tanaka T, Sano Y (2007b) Different patterns of
807 genealogical relationships found in the two major QTLs causing reduction of seed
808 shattering during rice domestication. Genome 50: 757-766.
809 Panaud O (2009) The molecular bases of cereal domestication and the history of rice.
810 C R Biol. 332: 267-272.

811 Palmer SA, Moore JD, Clapham AJ, Rose P, Allaby RG (2009) Archaeogenetic evidence
812 of ancient nubian barley evolution from six to two-row indicates local adaptation.
813 PLoS One 4: e6301.
814 Phan PD, Kageyama H, Ishikawa R, Ishii T (2012) Estimation of the outcrossing rate for
815 annual Asian wild rice under field conditions. Breed Sci. 62: 256-262.
816 Pourkheirandish M, Hensel G, Kilian B, Senthil N, Chen G, Sameri M, Azhaguvel P,
817 Sakuma S, Dhanagond S, Sharma R, Mascher M, Himmelbach A, Gottwald S, Nair
818 SK, Tagiri A, Yukuhiro F, Nagamura Y, Kanamori H, Matsumoto T, Willcox G,
819 Middleton CP, Wicker T, Walther A, Waugh R, Fincher GB, Stein N, Kumlehn J,
820 Sato K, Komatsuda T (2015) Evolution of the grain dispersal system in barley. Cell
821 162: 527-539.
822 Purugganan MD, Fuller DQ (2009) The nature of selection during plant domestication.
823 Nature 457: 843-848.
824 Ray A, Chakraborty D (2017) Shattering or not shattering: that is the question in
825 domestication of rice (*Oryza sativa* L.). Genet Resour Crop Evol 65: 391- 395.
826 Sakuma S, Salomon B, Komatsuda T (2011) The domestication syndrome genes
827 responsible for the major changes in plant form in the Triticeae crops. Plant Cell
828 Physiol 52: 738-749.

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

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

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

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

837 Silva F, Stevens CJ, Weisskopf A, Castillo C, Qin L, Bevan A, Fuller DQ (2015)
838 Modelling the geographical origin of rice cultivation in Asia using the rice
839 archaeological database. PLoS One 10: e0137024.

840 Smith O, Momber G, Bates R, Garwood P, Fitch S, Pallen M, Gaffney V, Allaby RG
841 (2014) Sedimentary DNA from a submerged site reveals wheat in the British Isles
842 8000 years ago. Science 347: 998-1001.

843 Song XJ, Huang W, Shi M, Zhu MZ, Lin HX (2007) A QTL for rice grain width and
844 weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nat Genet
845 39: 623-630.

846 Sugimoto K, Takeuchi Y, Ebana K, Miyao A, Hirochika H, Hara N, Ishiyama K,

847 Kobayashi M, Ban Y, Hattori T, Yano M (2010) Proc Natl Acad Sci U S A 107:
848 5792-5797.
849 Sweeney MT, Thomson MJ, Pfeil BE, McCouch S (2006) Caught red-handed: Rc encodes
850 a basic helix-loop-helix protein conditioning red pericarp in rice. Plant Cell 18: 283-
851 294.
852 Tan L, Li X, Liu F, Sun X, Li C, Zhu Z, Fu Y, Cai H, Wang X, Xie D, Sun C (2008)
853 Control of a key transition from prostrate to erect growth in rice domestication. Nat
854 Genet 40: 1360-1364.
855 Tang J, Xia H, Cao M, Zhang X, Zeng W, Hu S, Tong W,Wang J,Wang J, Yu J, Yang H,
856 Zhu L (2004) A comparison of rice chloroplast genomes. Plant Physiol 135: 412–
857 20.
858 Tang S, Min S, Sato YI (1996) Exploration on origin of keng rice (*japonica*) in China. In:
859 Wang X and Sun C (eds) Origin and differentiation of Chinese cultivated rice.
860 Chinese Agricultural University Press, Beijing, pp 72–80.
861 Takahashi H, Sato YI, Nakamura I (2008) Evolutionary analysis for two plastid DNA
862 sequences in cultivated and wild species of *Oryza*. Breeding Science 58: 225–233.
863 Thurber CS, Reagon M, Gross BL, Olsen KM, Jia Y, Caicedo AL (2010) Molecular
864 evolution of shattering loci in US weedy rice. Mol Ecol 19: 3271-3284.

865 Thurber CS, Reagon M, Olsen KM, Jia Y, Caicedo AL (2014) The evolution of flowering
866 strategies in US weedy rice. Am J Bot 101: 1737-1747.
867 Tournebize R, Poncet V, Jakobsson M, Vigouroux Y, Manel S (2019) McSwan: A joint
868 site frequency spectrum method to detect and date selective sweeps across multiple
869 population genomes. Mol Ecol Resour. 19: 283-295.
870 Tsujimura Y, Sugiyama S, Otsuka K, Htun TM, Numaguchi K, Castillo C, Akagi T, Ishii
871 T, Ishikawa R (2019) Detection of a novel locus involved in non-seed-shattering
872 behaviour of Japonica rice cultivar, *Oryza sativa* 'Nipponbare' Theor Appl Genet.
873 132: 2615-2623.
874 Wang M, Li W, Fang C, Xu F, Liu Y, Wang Z, Yang R, Zhang M, Liu S, Lu S, Lin T, Tang
875 J, Wang Y, Wang H, Lin H, Zhu B, Chen M, Kong F, Liu B, Zeng D, Jackson SA,
876 Chu C, Tian Z (2018) Parallel selection on a dormancy gene during domestication
877 of crops from multiple families. Nat Genet 50: 1435-1441.
878 Weng J, Gu S, Wan X, Gao H, Guo T, Su N, Lei C, Zhang X, Cheng Z, Guo X, Wang J,
879 Jiang L, Zhai H, Wan J (2008) Isolation and initial characterization of *GW5*, a major
880 QTL associated with rice grain width and weight. Cell Res 18: 1199-1209.
881 Winchell F, Brass M, Manzo A, Beldados A, Perna V, Murphy C, Stevens C, Fuller DQ
882 (2018) On the origins and dissemination of domesticated sorghum and pearl millet

883 across Africa and into India: a view from the Butana Group of the Far Eastern Sahel.
884 Afr Archaeol Rev 35: 483-505.
885 Wright D (2015) The genetic architecture of domestication in animals. Bioinform Biol
886 Insights 9: 11-20.
887 Wu Y, Zhao S, Li X, Zhang B, Jiang L, Tang Y, Zhao J, Ma X, Cai H, Sun C, Tan L (2018)
888 Deletions linked to *PROG1* gene participate in plant architecture domestication in
889 Asian and African rice. Nat Commun 9: 4157.
890 Wu Y, Guo T, Mu Q, Wang J, Li X, Wu Y, Tian B, Wang ML, Bai G, Perumal R, Trick
891 HN, Bean SR, Dweikat IM, Tuinstra MR, Morris G, Tesso TT, Yu J, Li X (2019)
892 Allelochemicals targeted to balance competing selections in African
893 agroecosystems. Nat Plants 5: 1229-1236.
894 Xu X, Liu X, Ge S, Jensen JD, Hu F, Li X, Dong Y, Gutenkunst RN, Fang L, Huang L,
895 Li J, He W, Zhang G, Zheng X, Zhang F, Li Y, Yu C, Kristiansen K, Zhang X, Wang
896 J, Wright M, McCouch S, Nielsen R, Wang J, Wang W (2012) Resequencing 50
897 accessions of cultivated and wild rice yields markers for identifying agronomically
898 important genes. Nat Biotechnol 30: 105-111.
899 Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M,
900 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

919 genomic variation in cultivated and wild rice. Nat Genet 50: 278-284.
920 Zheng Y, Crawford GW, Jiang L, Chen X (2016) Rice domestication revealed by reduced
921 shattering of archaeological rice from the Lower Yangtze valley. Scientific Reports
922 6: 28136.
923 Zhu Y, Ellstrand NC, Lu BR (2012) Sequence polymorphisms in wild, weedy, and
924 cultivated rice suggest seed-shattering locus *sh4* played a minor role in Asian rice
925 domestication. Ecol Evol 2: 2106-2113.
926 Zhu BF, Si L, Wang Z, Zhou Y, Zhu J, Shangguan Y, Lu D, Fan D, Li C, Lin H, Qian Q,
927 Sang T, Zhou B, Minobe Y, Han B (2011) Genetic control of a transition from black
928 to straw-white seed hull in rice domestication. Plant Physiol 155: 1301-1311.
929 Zhu Z, Tan L, Fu Y, Liu F, Cai H, Xie D, Wu F, Wu J, Matsumoto T, Sun C (2013) Genetic
930 control of inflorescence architecture during rice domestication. Nat Commun 4:
931 2200.
932 Zohary D (1969) The progenitors of wheat and barley in relation to domestication and
933 agricultural dispersal in the Old World. In: Ucko PJ, Dimbleby GW (eds) The
934 domestication and exploitation of plants and animals. Duckworth, London, pp. 47-
935 66.
936 Zohary D (2004) Unconscious selection and the evolution of domesticated plants. Econ

937 Bot 58: 5–10.

938

939 Figure legends

940 Fig. 1. Comparison of plant morphologies between wild rice, *Oryza rufipogon* acc. W630,
941 and a *japonica* rice cultivar, *O. sativa* 'Nipponbare'. Plant architecture (A and B), panicle
942 shape (C and D), and abscission layer formation (E and F) of W630 (A, C, and E) and
943 Nipponbare (B, D, and F). Awn formation (G), seed colour (H) and seed dormancy (I), of
944 W630 (left) and Nipponbare (right). Scale bars = 20 cm (A and B), 5 cm (C and D), 50
945 μm (E and F), and 5 mm (G, H, and I). In the photo of abscission layer formation (E and
946 F), black arrowheads indicate the abscission layer. VB, vascular bundle. In the photo of
947 seed dormancy (i), W630 and Nipponbare seeds were induced for germination by water
948 soaking a month after harvest.

949

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

955 tributaries are indicated as a potential distinct group.

956

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

961

962 Fig. 4. Time series of archaeological spikelet base data from eastern China (Lower
963 Yangtze), comparing different estimated of the domesticated proportion: a maximum
964 estimate in which "immature" forms are counted as domesticated (after Zheng et al 2007;
965 2016), and minimum estimate which counted "immature" forms with wild types (as per
966 Fuller et al 2009) and a mean domesticated fraction based on the ratio of domesticated
967 and wild forms only for which "immature" types provide an estimate of error (standard
968 deviation) (as per Fuller et al 2014). Archaeological site names for data points indicated.

969

970 Table 1. Traits and genes / loci involved in rice domestication

971

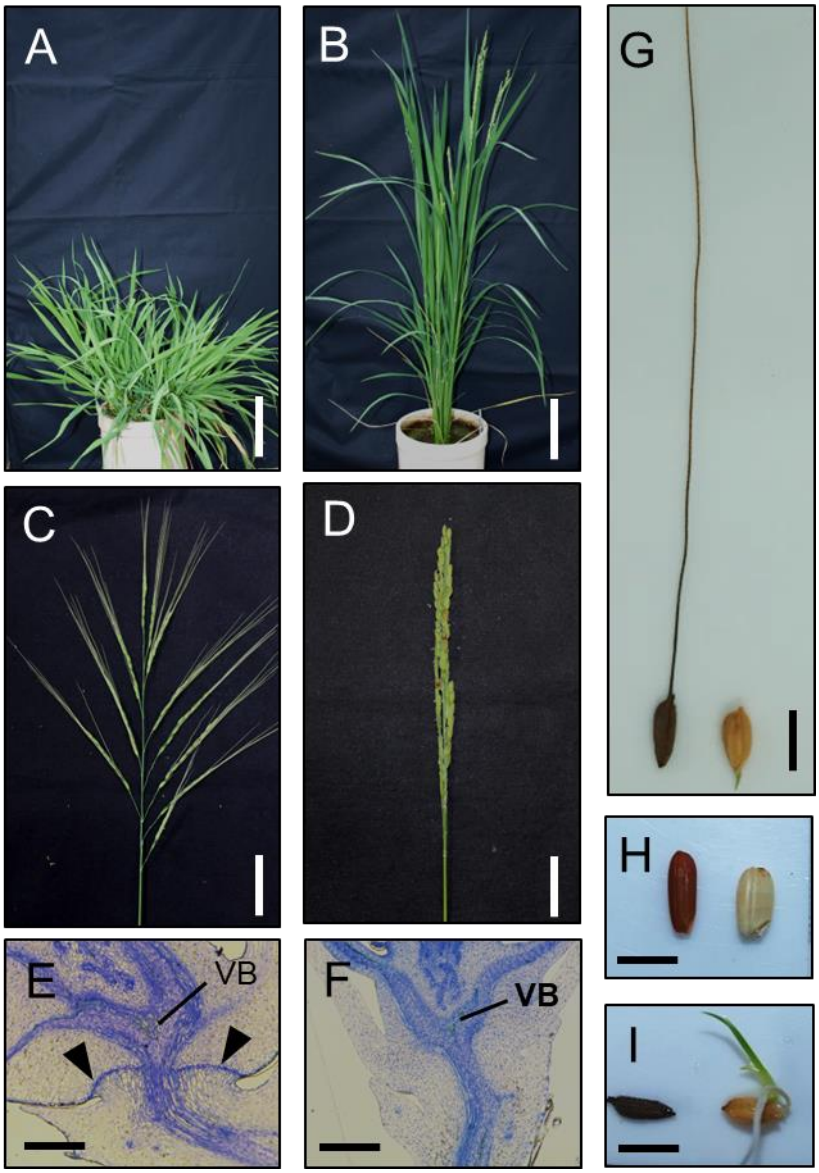
972

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 a. (2016
	qAWNL2	N.A.	N.A.	Amarashinghe et al. (2020)
Seed and hull colour	*Bh4*	Deletion	Amino acid transporter	Zhu et al. (2011)
	Rc[1)]	Deletion	Basic helix-loop-helix transcriptional regulator	Sweeney et al. (2006)
	Rd	SNP	Dihydroflavonol-4-reductase	Furukawa et al. (2007)
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)] *Rc* might be involved in the control of seed dormancy (Gu et al. 2011)

991 Ishikawa et al. Fig. 1.



992

993

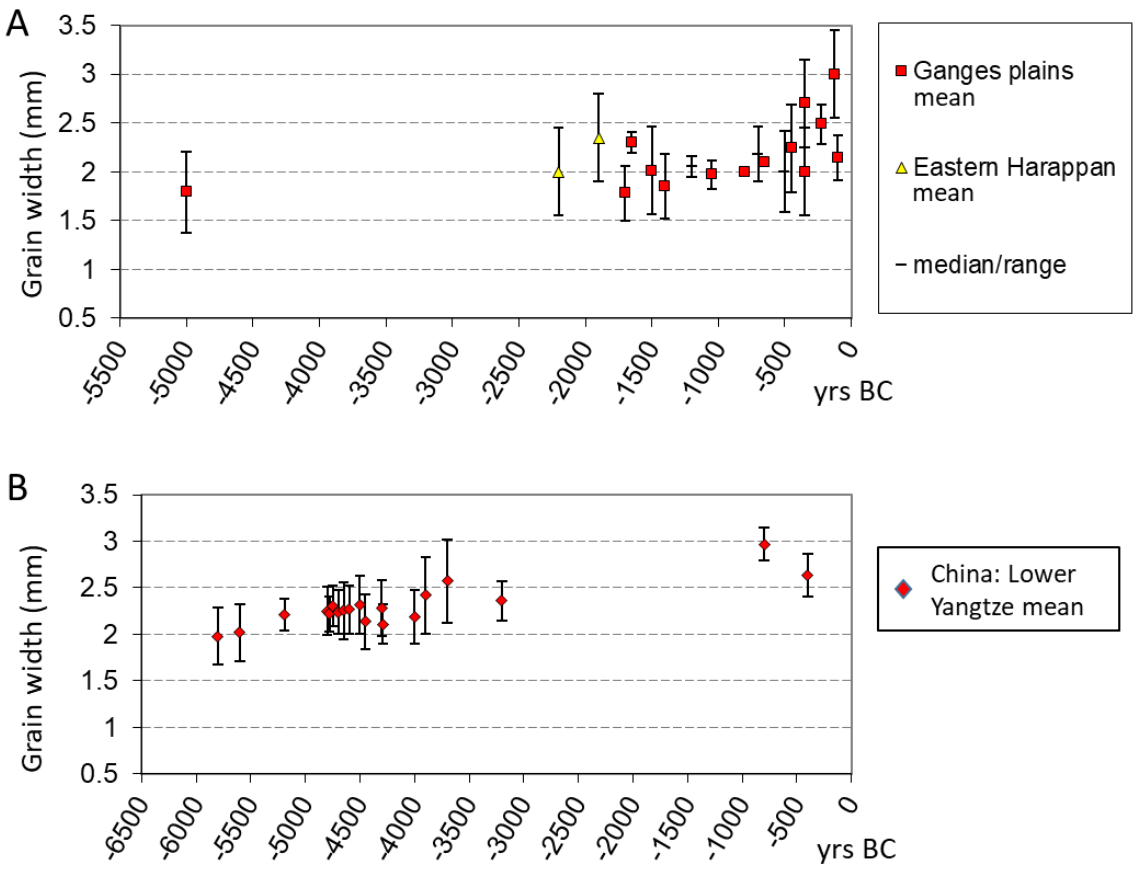
994

995

996

997

998 Ishikawa et al. Fig. 2.



999

1000

1001

1002

1003

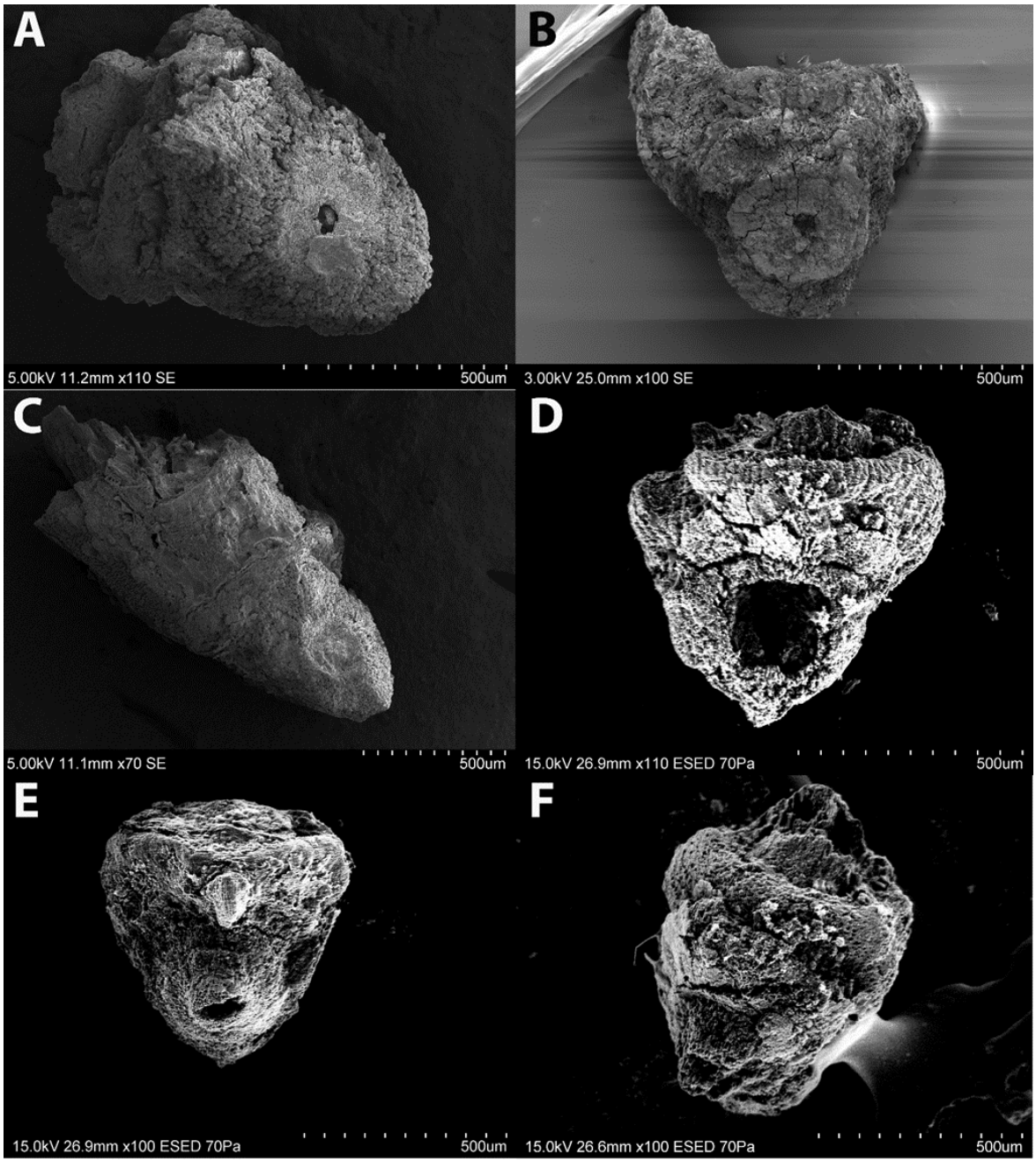
1004

1005

1006

1007

1008 Ishikawa et al. Fig. 3.



1009

1010