Introduction

The birth of Asian cultivated rice has been a contentious issue which mostly dwells on two contrasting hypotheses, single and multiple origin. The single origin posits a prior origin of the key genes in *japonica*. Attainment of non-shattering, a marker; appeared least integral to the initial phases of domestication. The other archeological determinants were less discernible in specimens. Existence of the key domestication genes in the wild rice and absence of introgression signature in *indica* further weakened the hypothesis. Moreover, *japonica* introduction from China happened in a backdrop of a culture exploiting domesticated rice. Summarizing, we propose that *proto-indica* model has a little bearing on rice domestication.

We have evaluated the *proto-indica* model that is the proponent of multiple domestication of rice but a single origin of the key genes in *japonica*. Attainment of non-shattering, a marker; appeared least integral to the initial phases of domestication. The other archeological determinants were less discernible in specimens. Existence of the key domestication genes in the wild rice and absence of introgression signature in *indica* further weakened the hypothesis. Moreover, *japonica* introduction from China happened in a backdrop of a culture exploiting domesticated rice. Summarizing, we propose that *proto-indica* model has a little bearing on rice domestication.

**Figure 1a–b:** The proto-Silk routes for putative *japonica* introgression after 2000 B.C. (excluding the maritime Silk route and the southern Himalayan/Tibet route).
transfer of key domestication traits (Fuller and Qin 2009; Stevens et al. 2016).

The proto-indica model has a great implication on the origin of rice agriculture in the Indian subcontinent. The primary basis of the early rice agriculture in India is contingent on the findings from the cultural site of Lahuradewa of the upper Gangetic plains (Tewari et al. 2006, 2008, Kingwell-Banham et al. 2015). While researchers claimed it to be the oldest site thriving on domesticated rice, it has been severely criticized to be wild rice gathering rather than cultivation. This criticism formed the primary foundation of the proto-indica model that is currently reformulated into a multiple origin and single domestication model. It advocated for the multiple origin of indica and japonica but the single origin of domestication genes in japonica followed by an introgression into indica (Choi et al. 2017).

Against this backdrop, we attempted to critically evaluate the proto-indica model. In doing so, we invoked support from various disciplines and demonstrated that it has a little bearing on rice domestication in India.

**Shattering or not shattering in rice domestication**

Cereal domestication stimulated a radical transformation in the plant manifested through several underlying traits over a long timeframe (Harlan 1992; Ladizinsky 1998). The control of shattering was one such crucial step in cereal domestication that facilitated increased anthropo-dependence of for their survival (Fuller and Allaby 2009). Unlike barley or wheat, it has not been well documented in rice archaeologically (Fuller and Allaby 2009). Yet, an increased frequency of the non-shattering-type grains was a trend often observed in rice cultural sites of China and considered as one of the key markers of domestication (but see van Driem 2017). The method to distinguish two types, wild and domesticated, has drawn support from morphological features of the rice spikelet bases preserved in archaeological samples. While domesticated types exhibited rough scars due to non-shattering rachis, the wild-type shattering rachis produced smooth scars (Harvey 2006; Fuller and Qin 2009).

The shattering trait of rice has a crucial bearing on the proto-indica model since it is proposed to be the key trait introgressed from japonica to indica (Fuller et al. 2010; Fuller 2011). The findings at Lahuradewa, succeeded by Tokwa, Kunjhn, Chopani-Mando, Senuwar, Mahagara, Atranjikhera are some of the major cornerstones of ancient rice agriculture in India (Table 1). It is, however, a highly debated issue among the commentators whether the sites exhibited the marks of domesticated rice cultivation or wild rice gathering. The criticism surfaced from the morphometrics of excavated rice grains, i.e., L/W/Th ratios, and the nature of scars of excavated spikelet bases at Lahuradewa (Table 1). It exhibited a resemblance with wild rice (Oryza nivara S.D. Sharma & Shasray, O. rufipogon Griff, and Oryza officinalis Wall. ex Watt), thus refuting the notion of domestication (Fuller et al. 2010). Moreover, the non-shattering forms had not been witnessed until 2000 BC in the Indian archaeological sites, such as Mesudipur I, VII, Bahola (Bates et al. 2017). This formed the primary foundation of argument in favor of the proto-indica model and the shattering trait as a 'Chinese loan' seemed to be at the center of this discussion.

Contradicting this, Ray and Chakraborty (2018) have contended that the control of shattering may not be imperative to rice domestication process, at least in the initial phases of domestication. They cited various examples, such as a natural variation in threshability in a large fraction of the landraces and cultivars and variable shattering ability, i.e., indica shatters easily than japonica (Lin et al. 2007). So, it was suggested that there could have been a geographic mosaic of the degree of shattering, and from where non-shattering phenotypes were preferred much later in history (Ray and Chakraborty 2018). Furthermore, they also presented examples of wild rice gathering ethnic tribal groups who harvest either using swinging or sweeping basket or bundling or canoe and flail technique. Pertinently, archaeologists have long attempted to relate the method of harvesting to the origin of non-shattering in cereals. In wheat and barley, the sickles had exerted a selective pressure that preferred non-shattering over shattering (Hillman and Davies, 1999; Willcox, 1999). But, the evolution of non-shattering in response to the use sickles has been a topic of conjecture (Fuller and Allaby 2009, but see Fuller 2007; Maeda et al. 2016). In rice, the use of stone harvesting hand-knives or sickles from Chinese sites (Liangzhu period ca. 3,000–2,000 BC, Lower Yangtze) cropped up after domestication (Fuller and Qin 2009). Similar evidence from the Indian subcontinent was reported from Semthan, Burzahom, Kashmir (after 2000 BC), adjacent Swat, northern Pakistan (around 1800 BC) (Fuller 2011; Stevens et al. 2016). The occurrence of Chinese stone harvesting knife forms and square stone artifacts in these areas were argued as proofs of ‘Chinese horizon’. Probably, the places were also a part of the proto-Silk route labyrinth through which japonica introgression occurred after 2000 BC (Fuller 2011; Figure 1).

On the contrary, we have instances of long-term sustenance of rice agriculture based on shattering-type rice throughout Southeast Asia, Australia, South America, and West Africa. West African wild rice (Oryza barthii A. Chev., O. breviligulata A. Chev. & Roehrl.) harvest mostly counted on sweeping basket technique to collect shattering grains since the specific technique exploited shattering grains non-shattering was never selected for (National Research Council 1996; Zaroug et al. 2002). Similarly, Hilbert et al. (2017) have discovered a mid-Holocene domestication of rice in South America focussing on an equivalent canoe and flail strategy to exploit shattering wild rice. Their prehistoric cultivation used shattering species of Oryza glumaepatula Steud, and Oryza latifolia Desv. using beating and flailing of wild rice in canoes. Altogether, both the examples convey that the harvesting technique could be a driver of non-shattering trait determining its fixation. The trait could have been selected while maintaining grain abscission such that the yield increment was not undermined by making threshing problematic. This was opined to be selected as a subtle balance in the regulatory region of the gene for achieving fine adjustment between shattering or threshing (Li et al. 2006). The other cases of wild rice cultivation is resonated in a plethora of prehistoric records in China, e.g., Pengtoushan (7000–6000
Ray et al: A Critical Evaluation Revealed the Proto-indica Model Rests on a Weaker Foundation and Has a Minimal Bearing on Rice Domestication

Art. 8, page 3 of 8

while the support for non-shattering as domestication marker is not so compelling (van Driem 2017), the genetic data largely relied on a mutation\(^4\) in \(sh4\) (Li et al. 2006; Lin et al. 2007, Fuller 2011), and qSH1 loci (Konishi et al. 2006). It has been observed that a substitution\(^5\) (‘G’ to ‘T’) in \(sh4\) locus underlying the morphological change turned the fruits non-shedding on maturation. The fact also helped the premise of rice domestication by reduced shattering to gain momentum since a loss of shattering has been a common trend in cereals.

However, many recent discoveries have detected a suite of cultivated rice and wild accessions possessing this mutation tend to shatter ripened seeds, implying the causal mutation may not be indispensable to the domestication by reduced shattering (Izawa et al. 2008; Thurber et al. 2010; Zhu et al. 2012).

It also revealed that a significant portion of shattering phenotypes in wild rice with \(sh4\) mutation originated from India (22.5%) and Bangladesh.

### Table 1: Indian and Chinese sites showing evidences for the Neolithic rice culture.

<table>
<thead>
<tr>
<th>India</th>
<th>Time</th>
<th>Key findings</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokwa, Adwa Valley in the Mirzapur district, U.P.</td>
<td>4797 B.C.</td>
<td>Complete and broken rice (wild/domesticated) charred grains, rice husk impressed potsherds and burnt-mudclods, use of rice husk as degraisant on the surface and core of the potsherds</td>
<td>Misra et al. 2000, Pokharia 2008</td>
</tr>
<tr>
<td>Kunjhn, Sidhi district, M.P</td>
<td>3530 B.C.</td>
<td>Wild rice grain</td>
<td>Possehl and Rissman 1992, Misra et al. 2000</td>
</tr>
<tr>
<td>Senuwar, Vindhyan region, Rohtas district, Bihar</td>
<td>2125 B.C.</td>
<td>Carbonised rice grains (wild/domesticated)</td>
<td>Saraswat 2004</td>
</tr>
<tr>
<td>Bahola, Karnal district, Haryana</td>
<td>Before 2000 B.C.</td>
<td>Domesticated rice cultivation &amp; management</td>
<td>Bates et al. 2017</td>
</tr>
<tr>
<td>China</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pengtoushan culture site at the middle Yangtze region</td>
<td>7000–6000 B.C.</td>
<td>Wild rice cultivation</td>
<td>Fuller et al. 2007; Zheng et al. 2016</td>
</tr>
</tbody>
</table>

BC), the Shangshan of Huxi (9000–8400 BP), Kuahuqiao (8000–7700 BP), Tianluoshan and Luojiajiao (ca. 7000–6500 BP), Majiabang (6300–6600 BP) (Fuller et al. 2007; Zheng et al. 2016) (**Figure 2a–c**). The recent findings at three Harappan sites have also found a certain percentage of wild rice samples from various time periods (Bates et al. 2017). Altogether, these facts emphasized on rice cultures that were firmly contingent on tending wild forms.

**Genetic data support for shattering**

While the support for non-shattering as domestication marker is not so compelling (van Driem 2017), the genetic data largely relied on a mutation\(^4\) in \(sh4\) (Li et al. 2006; Lin et al. 2007, Fuller 2011), and qSH1 loci (Konishi et al. 2006). It has been observed that a substitution\(^5\) (‘G’ to ‘T’) in \(sh4\) locus underlying the morphological change turned the fruits non-shedding on maturation. The fact also helped the premise of rice domestication by reduced shattering to gain momentum since a loss of shattering has been a common trend in cereals.

However, many recent discoveries have detected a suite of cultivated rice and wild accessions possessing this mutation tend to shatter ripened seeds, implying the causal mutation may not be indispensable to the domestication by reduced shattering (Izawa et al. 2008; Thurber et al. 2010; Zhu et al. 2012). It also revealed that a significant portion of shattering phenotypes in wild rice with \(sh4\) mutation originated from India (22.5%) and Bangladesh.
Ray et al: A Critical Evaluation Revealed the Proto-indica Model Rests on a Weaker Foundation and Has a Minimal Bearing on Rice Domestication

(42.8%). On this line, a very recent origin of functional mutation in sh4 conferring non-shattering was proposed by Zhang et al. (2009). They have claimed the generation of this mutation is not an older event, but it could be only a few centuries old.

Another major QTL, \( qSH1 \) responsible for differential shattering between indica and japonica, was reported only in temperate japonica (Konishi et al. 2006; Thurber et al. 2010). Its selection was probably driven by the ease of mechanical threshing during the crop improvement phase of temperate japonica, i.e., long time after domestication (Thurber et al. 2010; Zhu et al. 2012). The current understanding of attainment of shattering in rice is enhanced after identification of several major and minor QTL (Zhu et al. 2012). Based on these facts, it seems that the prehistoric process to achieve non-shattering may not be simple to involve a single locus, but i) may have demanded an interplay of multiple loci (Zhu et al. 2012; Ishikawa et al. 2017), ii) different mutations independently arisen in geographically disjunct wild populations that gradually contributed towards a reduction in shattering but without completely eliminating abscission layer formation facilitating threshing (Zhu et al. 2012), and iii) both the non-shattering and shattering phenotype could have co-occurred in overlapping, adjacent, or in isolated regions before non-shattering selected unanimously and become fixed (Civan and Brown 2017; Ray and Chakraborty 2018). Therefore, non-shattering in rice may be inferred as a trait not so crucial for domestication (van Driem 2017) but proved useful in later improvement phase (Abbo et al. 2014).

Near-absence of other reliable markers in archaeological samples

Archaeological records of rice mainly hinged on three phenotypic traits which are recognized as the markers of domestication, namely, grain size increase, a loss of natural dispersal mode or inability to shatter, and a reduction or a complete loss of natural dispersal aids, e.g., long awn, dense, upwardly angled awn barbs, bristles, hairs, trichomes density and length (Fuller and Allaby 2009; Hua et al. 2015). These markers have often played a critical role in identifying domestication events across Chinese and Indian prehistoric sites, e.g., (Tang et al.1996; Zheng et al. 2007; Bates et al. 2017). Of these, segregation of grains from archaeological finds of India into wild or domesticated-type mostly relied on only one trait, i.e., shattering spikelet base. No other trait, such as grain size increase, was clearly discernible from the archaeo-botanical remains of the Indian subcontinent (Fuller 2011).

In general, from the perspective of genetics, much-discussed domestication syndrome of rice is founded on a set of traits, e.g., \( Prog1 \) (gene conferring erect habit), \( rc \) (red pericarp), \( wxy \) (amylose content), etc. Their description and validity entirely stemmed from the comparison of the phenotypes and the underlying genotypes between wild extant gene pool and landraces or cultivated forms; this phenotypic difference is barely discoverable in the archaeobotanical finds. So, it is often hard to complement the actual time of origin and spread of such functional mutations computed from genetic data with that of archaeological records. In this context, the evidence for presumed gene flow from japonica to indica is impossible to obtain from archeological records, given the wear and tear of preservation (Fuller 2011; Choi et al. 2017). More often, the uncertainty in interpretation from genetic analyses makes it hard to supplement the archaeological finds at a finer scale, thus failing to form a cohesive intellectual framework towards a robust conclusion. In addition, in most of the studies, use of extant populations of the wild progenitors as surrogates of the ancestral gene pool is another impediment to error-free interpretation. The trajectory of the extant populations has not been static throughout and inevitably went through various evolutionary events, such as, bottleneck and local extinction, hybridization between wild and cultivated rice promoting back gene flow (Fuller 2011; Wang et al. 2017).
It might have erased the actual genomic signature to a varying extent prompting an erroneous inference.

**The key genes in wild gene pool and minimal gene flow**

The advocates of the *proto-indica* and multiple origin single domestication model relied on putative gene flow\(^1\) between *proto-indica* and *japonica* facilitating the transfer of key domestication genes leading to the formation of *indic*a (Fuller et al. 2010, Fuller 2011, Choi et al. 2017). The suggested vehicle for the gene movement was likely to be recurrent trade but demic movement\(^2\) through the proto-Silk routes via north-west India and southeast Asia established during the 2000 BC that have facilitated the transfer of key genes (Chakraborty and Ray, 2019; Fuller 2011). But, recent findings proposed an introduction of *japonica* around 2000 BC through the proto-Silk route have occurred in a backdrop of already domesticated rice culture (Petrie et al. 2016; Bates et al. 2017).

It also assumed a single origin of genes underlying domestication syndromes in *japonica* and subsequent transfer through admixture.\(^3\) The claim of *japonica* origin of the key traits suffered from serious weaknesses in light of the recent findings. Civan and Brown (2017) have demonstrated an origin of major domestication genes in wild gene pool, e.g., white pericarp, barbless awns, erect growth and non-shattering ear emerged in the ancestral pool prior to their domestication. They proposed, i) an occurrence of *sh4* in wild populations, ii) three haplotypes\(^4\) of *prog-1* namely TC-prog1, CC-prog1, CA-prog1 reported in both *indic*a and *japonica*, of which CC-prog1 is common haplotype; an absence of unique haplotypes of *prog1* can be best explained by its occurrence in wild rice followed by an introgression, iii) moderate frequencies (*sh4* – 26%, *rc* – 13%, *labal* – 15%) and high diversity of *rc*, *sh4*, *labal* alleles in wild population than in the cultivated types (Zhu et al. 2012; Civan and Brown 2017). Likewise, in a follow-up study by Civan and Brown (2018) have shown that several key loci in *indic*a have no signature of introgression from *japonica*.

Secondly, a few key phenotypes showed different underlying mutation and portrayed independent evolutionary trajectory. One such example is the mutation for red pericarp independently originated in *aus* though not much disseminated from its region of origin and remained regionally fixed to that specific subpopulation, whereas a separate mutation is responsible for red pericarp color in *indic*a and *japonica* that exhibited wider spatial spread and cultural acceptance (Sweeney et al. 2007).

Hence, the course of functional mutations may not sufficiently be explained by single or *japonica* origin and subsequent hybridization model. In order to decipher this, natural variation and its spatial correlates in wild rice population as well as in the landraces are to be characterized in great detail. With the advent of high-throughput genomic data, better understanding of natural variation became possible; it has also opened a new vista to offer further insights into the general trends of evolutionary changes underlying domestication, their nature, origin, and temporal and spatial flow of mutations.

**Concluding remarks**

Concluding, the *proto-indica* model seems to be too premature and lacks sufficient support. The origin and evolution of domestication traits may not be explained sufficiently by single mutation, while many are polygenic; it is quite likely that an interaction of multiple QTL underlies such traits. Their predicted *japonica* origin reflects the fact that wild rice gene pool and indigenous landraces have not been thoroughly characterized, especially in regard to the origin of domestication genes. Other factors operate over time and space, such as cultural attributes and population history of the agriculturists, local extinction, climate, an interaction between wild and cultivated rice, have left marks on the genome.

In light of above discussion, the following aspects deserve in-depth exploration to gain finer understandings, i) a reconciliation of the phases, i.e., wild rice cultivation in Lahuradewa and the fully domesticated forms in Mahagara, the elapsed time does not offer any clue to transition in rice culture, whether tending wild rice or cultivation, a point echoed in Bates et al. (2017); ii) further on, the recent discovery of domesticated type rice spikelet bases from Harappan sites of Masudpur VII (Early-Mature-Late Harappan), Masupdur I (Mature Harappan) and Bahola (Late Harappan-Painted GreyWare) proved that domesticated rice cultivation and various techniques like multi-cropping, irrigation had been already in practice before 2000 BC (Petrie et al. 2016; Bates et al. 2017). That implied *japonica* arrived from north-west via the proto-Silk routes in a cultural context of already domesticated rice (Bates et al. 2017); and iii) finally, the *proto-indica* model also calls attention to a re-examination of the key traits in light of domestication syndrome proposed by Abbo et al. (2012), (2014) with an aim to segregate into classic domestication traits and later diversification traits during crop evolution; whether they were essential to domestication from very early on, or arose lately while improvement or diversification stage (Abbo et al. 2012, 2014).

**Notes**

1. **Introgression** – It describes the transfer of genetic information (alleles) from one species (usually via hybridization and backcrossing) of from one entity (species or sub-species) into the gene pool of a second, divergent entity (species or sub-species).

2. **Selective pressure or natural selection** is imparted on organisms with certain phenotypes to have either a survival benefit or disadvantage. It is one of the causes of evolutionary change. For example, harvesting tools have long been regarded to exert selective pressure on the shattering traits of cereals.

3. **Fixation** – it is process through which an allele or genetic variant becomes a fixed or retained allele or genetic variant within a population. A fixed allele (e.g., an allele for non-shattering) is an allele that is the only variant that exists for that gene in all the population.

4. **Mutation** – It is the alteration in the nucleotide sequence (composed of strings of bases A, T, G, C) of the genome of an organism.
5 Substitution – It is a mutation that exchanges one base for another (i.e., a change in a single base such as switching a G to a T or vice versa).

6 Functional mutation – It is the type of mutation in which the altered gene product possesses a new molecular function and may cause a phenotypic change (e.g., shattering to non-shattering).

7 QTL or A quantitative trait locus – It is a locus or a region of DNA that is associated with a particular phenotypic trait and explains its underlying variation.

8 Bottleneck or population bottleneck – It is a drastic reduction in the population size due to environmental changes like earthquakes, famines, floods, fires, disease or due to anthropogenic activities like overhunting.

9 Local extinction – It is the condition of a species (or a population) that ceases to exist in a certain geographic area, though it may still exist elsewhere in its distribution range.

10 Hybridization – It is the process where two different species can lead to the formation of a new species through sexual reproduction.

11 Gene flow – It is any movement of individuals, and/or the genetic material or DNA they carry, from one population to another. Gene flow can happen in various ways, such as inter-population transfer of pollen from one flower to a flower of a distant location or people moving to new cities or countries.

12 Demic movement or demic diffusion – It is the population diffusion into an area that had not been previously inhabited by that group.

13 Admixture – It is the presence of genetic material or DNA in an individual(s) from a distantly-related population or species as a result of interbreeding between genetically divergent populations or species.

14 Haplotype – It is a set of DNA variants or polymorphisms that tend to be inherited together. It can also refer to a combination of various alleles found on the same chromosome.

Acknowledgements
Authors received no funding for this study.

Competing Interests
The authors have no competing interests to declare.

References


Civán, P and Brown, TA. 2017. Origin of rice (Oryza sativa L) domestication genes. Genetic resources and crop evolution, 64(6): 1125–32. DOI: https://doi.org/10.1007/s10722-017-0518-0


Fuller, DQ and Qin, L. 2009. Water management and labour in the origins and dispersal of Asian rice. World Archaeology, 41: 88e111. DOI: https://doi.org/10.1080/00438240802668321


Harvey, EL. 2006. Early Agricultural Communities in Northern and Eastern India: an Archaeobotanical
Ray et al: A Critical Evaluation Revealed the Proto-indica Model Rests on a Weaker Foundation and Has a Minimal Bearing on Rice Domestication


Ray et al: A Critical Evaluation Revealed the *Proto-indica* Model Rests on a Weaker Foundation and Has a Minimal Bearing on Rice Domestication


