

Rice domestication occurred through single origin and multiple introgressions

To the Editor — Rice domestication has long been a hot topic. Civián *et al.*¹ recently suggested that the Asian rice varieties *indica*, *aus* and *japonica* were domesticated independently in different parts of Asia. They reanalysed previously published genomic data to ascertain the loci under selection in *indica*, *aus* and *japonica*, and identified 38 co-located low-diversity genomic regions (CLDGRs) by applying the $\pi_{(wild)}/\pi_{(domesticated)}$ threshold of 4 for *indica*, *aus* and *japonica* separately. In 25 CLDGR trees, *indica* and *japonica* branched separately, leading to the conclusion that the three groups were domesticated independently. We argue that the methodology Civián *et al.* used to identify domestication loci is technically flawed.

We have previously sequenced the genomes of 1,083 accessions of cultivated rice (*Oryza sativa*) and 446 accessions of wild rice (*Oryza rufipogon*)². Analysing these data, we showed that *indica* and *japonica* descend from different subpopulations of wild rice, namely Or-I and Or-IIIa, respectively. We identified the domestication loci, of which the validity was confirmed by the identification of quantitative trait loci (QTLs) for domestication traits. We further found that, unlike the genome-wide pattern, the two subspecies were often clustered together at the domestication loci. Integrating all the data, we proposed a demographic scenario in which *japonica* was first domesticated from Or-IIIa, whereas *indica* was subsequently developed from Or-I with the adoption of many domestication alleles from *japonica*.

We noticed that the domestication loci reported by Civián *et al.* were mostly different from the domestication sweeps we had previously identified. The well-characterized domestication genes (such as *An1*, *Bh4* and *OsC1*)^{3–5} were located in our previous 55 domestication sweeps, but were not included in the 38 CLDGRs of

Civián *et al.* We suggest that the procedure they used to identify domestication loci was flawed. Specifically, there was a strong genetic bottleneck in *japonica*, but Civián *et al.* still applied the $\pi_{(wild)}/\pi_{(domesticated)}$ cutoff of 4 in this population, which meant that approximately half of the rice genome was considered as ‘low-diversity genomic regions’ during CLDGR detection. There certainly existed *indica*-specific selective sweeps, as many natural and human selections must have occurred over the past several thousand years. Simply identifying the low-diversity regions in *indica* separately and then overlapping them with half of the rice genome, as Civián *et al.* did, would have recognized many footprints of other types of selection (such as natural selection or selection for improvement after domestication). As such, we suggest that building trees from these regions does not tell us about domestication.

Civián *et al.* also found evidence supporting a single domestication origin in six CLDGRs, one of which included the *PROG1* gene^{6,7}. They suggest that such selection could have taken place multiple times during independent domestication processes. However, the occurrence of identical domestication alleles and very similar adjacent haplotypes suggests that this is unlikely. Moreover, it wasn't entirely clear how they classified the CLDGR trees into single or multiple domestication origins. Particularly, we noticed that the domestication gene *sh4* (which controls seed shattering), which apparently originated only once⁸, was not classified into the single domestication origin group (see the position of CLDGR tree 13 in Table 1 of ref. 1).

Civián *et al.* were puzzled by the pattern of the phylogenetic tree for the 55 domestication loci in our work, where the *japonica* group is nested within *indica*. Careful analysis of our paper would find that for each individual domestication locus,

japonica and *indica* varieties were all mixed together (see Supplementary Fig. 21 in ref. 2). At some domestication loci (such as *OsC1*), the domestication haplotypes were almost fixed in the *japonica* population but not fixed in *indica* rice. In other words, the *indica* group has many allelic combinations at the 55 domestication loci but *japonica* has very few. The phylogenetic tree derived from the genomic data of all 55 domestication loci together thus understandably results in a pattern whereby *japonica* is nested within the *indica* group.

In summary, we suggest that our previous conclusion was well founded: that *japonica* was domesticated first, whereas *indica* subsequently emerged by adopting domestication alleles from *japonica*. Observations of a single domestication origin, where the alleles are fixed in both *indica* and *japonica*, are indeed infrequent — among ~40,000 rice genes, there are probably only a few tens of domestication genes. However, it was these genes that finally generated a new species, *Oryza sativa*. The role of the genes associated with key domestication traits has been explained well in many previous studies^{3–10}.

References

1. Civián, P., Craig, H., Cox, C. J. & Brown, T. A. *Nature Plants* **1**, 15164 (2015).
2. Huang, X. *et al.* *Nature* **490**, 497–501 (2012).
3. Zhu, B. F. *et al.* *Plant Physiol.* **155**, 1301–1311 (2011).
4. Saitoh, K., Onishi, K., Mikami, I., Thidar, K. & Sano, Y. *Genetics* **168**, 997–1007 (2004).
5. Luo, J. *et al.* *Plant Cell* **25**, 3360–3376 (2013).
6. Tan, L. *et al.* *Nature Genet.* **40**, 1360–1364 (2008).
7. Jin, J. *et al.* *Nature Genet.* **40**, 1365–1369 (2008).
8. Li, C., Zhou, A. & Sang, T. *Science* **311**, 1936–1939 (2006).
9. Hua, D. R. *et al.* *Plant Cell* **27**, 1875–1888 (2015).
10. Gu B. *et al.* *Mol. Plant* **8**, 1635–1650 (2015).

Xuehui Huang and Bin Han

National Center for Gene Research, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200233, China. e-mail: xhuang@ncgr.ac.cn; bhan@ncgr.ac.cn