Article

Chinese Science Bulletin

Crop Genetics

October 2011 Vol.56 No.28-29: 3108–3113
doi: 10.1007/s11434-011-4691-9

Ancient DNA sequences of rice from the low Yangtze reveal significant genotypic divergence

FAN LongJiang[1*], GUI YiJie[1], ZHENG YunFei[2], WANG Yu[1], CAI DaGuang[3] & YOU XiuLing[4]

[1] *Department of Agronomy, Zhejiang University, Hangzhou 310058, China;*
[2] *Zhejiang Provincial Institute of Archaeology and Cultural Relics, Hangzhou 310014, China;*
[3] *Department of Molekulare Phytopathologie, Christian-Albrechts-Universität zu Kiel, D-24118 Kiel, Germany;*
[4] *School of Humanities, Zhejiang University, Hangzhou 310028, China*

Received February 22, 2011; accepted July 11, 2011

Rice (*Oryza sativa*) was first domesticated in the lower and middle Yangtze regions of China, and rice remains have been found in many Chinese archaeological sites. Until now, only phenotypic archeobotanical evidence, such as the spikelet bases of ancient grains, has been used to speculate on the domestication process and domestication rate of rice. In this study, we sequenced 4 genomic segments from rice remains in Tianluoshan, a site of the local Hemudu Neolithic culture in the low Yangtze and two other archaeological sites (~2400 and 1200 BC, respectively). We compared our sequences with those of the current domesticated and wild rice (*O. rufipogon*) populations. At least two genotypes were found in the remains from each site, suggesting a heterozygotic state of the rice seeds. One ancient genotype was not found in the current domesticated population and might have been lost. The rice remains belonged to the *japonica* group, and most if not all were *japonica*-type, suggesting that the remains might be at an early stage of *indica-japonica* divergence or an *indica-japonica* mixture. We also identified sequences with significant similarity to those from species of Sapindales, Zygophyllales, and Brassicales, which is consistent with the identification of other plant remains in the Tianluoshan site and the common rice field weeds such as mustards in southern China.

***Oryza sativa*, ancient DNA sequence, Hemudu Neolithic culture, rice domestication**

Citation: Fan L J, Gui Y J, Zheng Y F, et al. Ancient DNA sequences of rice from the low Yangtze reveal significant genotypic divergence. Chinese Sci Bull, 2011, 56: 3108–3113, doi: 10.1007/s11434-011-4691-9

Rice (*Oryza sativa*) is one of the most important crops in Asia. Wild rice was first domesticated in the Yangtze region of China to 2 cultivated subspecies, *indica* and *japonica* [1]. Phylogeographic investigations have suggested that cultivated rice has experienced at least two independent domestications: *indica* rice was domesticated within a region south of the Himalaya mountains whereas *japonica* was domesticated from wild rice in southern China [2]. Many ancient rice remains have been found in China, such as in Hemudu in Zhejiang Province, Lixian in Hunan Province, and Jiahu in Henan Province [3]. In the 1970s, ancient rice grains were found at an archaeological site of the Hemudu Neolithic culture (about 7000 years ago) in the low Yangtze [4]. In the rice remains, domesticated seeds were observed according to their short rachillae [5,6]. Ancient rice is an important node in the process of domestication from wild to cultivated rice and can provide key evidence regarding rice domestication. Newly developed techniques for DNA extraction and sequencing provide great promise for molecular archeology research. For example, the whole genomes of several ancient humans and animals from archeological sites have been sequenced and have provided important insights into their evolution. Partial ancient DNA sequences have been obtained and analyzed from plant remains, mostly from maize [7–12], but not yet from rice.

In this study, we determined four genomic segments of

*Corresponding author (email: fanlj@zju.edu.cn)

ancient DNA from rice remains in 3 archaeological sites, including a Neolithic site of the Hemudu Neolithic culture in the lower Yangtze region of China, and compared them with those from wild and cultivated rice.

1 Materials and methods

1.1 Materials

Rice remains used in this study were excavated from 3 archaeological sites in the lower Yangtze region, including Tianluoshan, a Neolithic site of the Hemudu Neolithic culture in Yuyao, Zhejiang Province, China (Table 1, Figure 1). We also used 51 Asian rice (*O. sativa*) cultivars, including 29 *japonica* (6 landraces from Zhejiang Province) and 22 *indica* strains; 15 Asian wild rice (*O. rufipogon*), 2 African rice cultivars (*O. glaberrima*), *Leersia oryzoides* (kindly provided by GE Song, Institute of Botany, Chinese Academy of Sciences), and *Zizania latifolia* (Table 1).

1.2 DNA extraction, sequencing, and analysis

Because of the limited DNA content in a single piece of ancient grain, whole rice remains or environmental samples from an archaeological site were used to extract DNA. We employed a published magnetic bead approach for nucleic acid purification [14], with slight modification. During the DNA extraction, nonionic and cationic detergents were used to protect ancient DNA chains and decrease the amounts of inhibitors such as humic acid and/or phenolic components in the environmental samples.

Two degenerate primer pairs (Angio_1F/Angio_2R and HrbcL252F/HrbcL320R [9]) and two to amplify the rice simple sequence repeat (SSR) RM211 (http://www.gramene.org) and a rice coding gene, respectively, were used (Table 2). PCR was performed in 25 μL volumes containing 1× PCR buffer (50 mmol L^{-1} KCl, 10 mmol L^{-1} Tris-HCl pH 8.8), 2.5 mmol L^{-1} $MgCl_2$, 1.0 μmol L^{-1} each primer, 0.2 mmol L^{-1} dNTPs, and 1 U of *Taq* polymerase (Sangon, Shanghai, China). The thermocycling conditions comprised an initial denaturation step for 5 min at 95°C, 35 cycles of 30 s at 95°C, 30 s at 55°C, and 20 s at 72°C, and a final extension step of 72°C for 10 min. Then, secondary 25-μL PCR reactions were carried out using 2 μL of 1:10-diluted primary PCR product as the template. The amplification products were checked by electrophoresis on 2.0% (w/v) agarose, then were purified and directly cloned and sequenced. ClustalW (www.ebi.ac.uk/clustalw/) was used for multiple sequence alignment. All sequences generated in this study have been deposited in GenBank with the accession numbers JN169832–JN169947.

2 Results and discussion

2.1 DNA extraction, amplification, and contamination control

Rice remains (including seeds and chaff) were cleaned and stored in distilled water. DNA extraction and amplification were carried out at two independent laboratories, including one (for the DNA extraction) that never deals with rice samples. An example gel showing the target bands after 2 rounds of amplification is shown in Figure 2.

Contamination control and detection are key steps for ancient DNA investigation [15]. To verify our results, we sent our DNA samples to Prof. YU XuPing at the Laboratory of Animal Immunology, College of Animal Science, Zhejiang University. Here, the DNA was amplified following our protocol and the sequences were determined independently. Prof. Yu's results were identical to our own. Moreover, several lines of evidence indicated that our results were not derived from DNA contaminants: (i) one genotype is expected to dominate in a contaminated sample. Each of our 3 samples displayed several genotypes (Figure 3); (ii) 2 genotypes of the ancient rice remains were not found in our current modern rice collection (Table 1), whereas some sequences were likely to be derived from ancient rice field weeds (next section) which have never

Table 1 Plant materials used in this study

Species	Tissue	Number[a)]	Culture/type	Estimated time	Site/origin[b)]
O. sativa	Chaff	/	The Hemudu Neolithic Culture	~7000 years ago	Tianluoshan, Zhejiang
O. sativa	Seed	/	The Warring States Period	~2400 years ago	Xiguan, Jiangxi
O. sativa	Chaff	/	Tang Dynasty	~1200 years ago	Huzhou, Zhejiang
O. sativa	Leaf	51	Asia cultivar	Present	IRRI and CNRRI
O. sativa	Leaf	374	Asia cultivar	Present	[13]
O. rufipogon	Leaf	15	Asia wild	Present	IRRI
Leersia oryzoides	Leaf	1	Wild	Present	Institute of Botany, Chinese Academy of Sciences
Zizania latifolia	Leaf	1	Wild	Present	Zhejiang University

a) Ancient rice remains are environmental samples and their genotype numbers are unknown; b) IRRI: International Rice Research Institute; CNRRI: China National Rice Research Institute.

Figure 1 Ancient rice samples used in this study from the Tianluoshan archaeological site (the Hemudu culture).

been studied in our lab; (iii) our DNA samples were short in length (<500 bp) and PCR amplicons of over 500 bp expected length could not be amplified from these samples; this is in accordance with the known characteristics of ancient DNA samples, in which their long exposure to the environment causes the DNA to fragment into segments of less than 500 bp [15]. To avoid errors originating from PCR amplification and sequencing, at least two independent clone sequences were used to confirm a genotype.

2.2 Ancient DNA sequences

Four genomic segments from the rice remains were cloned and sequenced, and the sequence of 142 clones was determined successfully. In the 142 sequences, most (130, 91.5%) matched known sequences from rice or other plants, suggesting that they were the target sequences. Of the 130 sequences, 105 (80.8%) were highly similar to that of rice (Table 3).

Four primer pairs were used in this study (Table 2): 2 (A and C) are highly degenerate and can amplify target sequences from almost all plants [9]. Figure 3 illustrates an alignment of sequences from our ancient samples amplified using the C primers. The other two primer pairs (Y and R) are based on the modern rice genomic sequence (c.v. Nipponbare) — one in a coding gene and the other in an intergenic region (containing an SSR marker); these are therefore rice-specific.

2.3 The ancient DNA was not from *Leerxia* or *Zizania*

Leerxia and *Zizania* are the closest 2 neighboring genera to *Oryza* [16] (www.ncbi.nlm.nih.gov/Taxonomy/). Seeds of *L. oryzoides* are highly similar to those of rice; *Z. latifolia* is an ancient cereal reported in Chinese history that was replaced by rice approximately 1000 years ago. To test the possibility that our remains were seeds from either of these 2 genera, we amplified DNA from the above 2 species using our 4 primer sets. The rice-specific R primer pair failed to amplify any product, but PCR products were observed for the other 3 primer pairs (Y primer pair which was designed on coding regions of a rice gene is but not highly degenerate to the 2 rice neighbors) (Figure 4). Sequencing of the PCR products indicated that the A primer genotypes (T-A-A and T-G-A) and the sequences amplified by the Y primer pair from the 2 species (JN169842 and JN169843) were different from those of rice. This confirmed that our ancient DNA was not derived from either *L. oryzoides* or *Z. latifolia*.

Table 2 Primer pairs used in this study

Name	Code	Targets	Forward (5′→3′)	Reverse (5′→3′)	Expected length (bp)	Reference
Angio	A	18S rDNA	TGCAGTTAAAAAGCTCGTAG	GCACTCTAATTTCTTCAAAG	159	[9]
HbcL	C	rbcL	TAGCGGCGGAATCTTCTACT	TATGATAGCATCGTCGTTTG	89	[9]
RM211	R	Intergenic	CCGATCTCATCAACCAACTG	CTTCACGAGGATCTCAAAGG	161	www.gramene.org
YANG4	Y	CDK inhibitor	AGAGCTGGAAGCGTTCTTCG	GGCAGTCATTCACAGGATCAAAG	230	This study

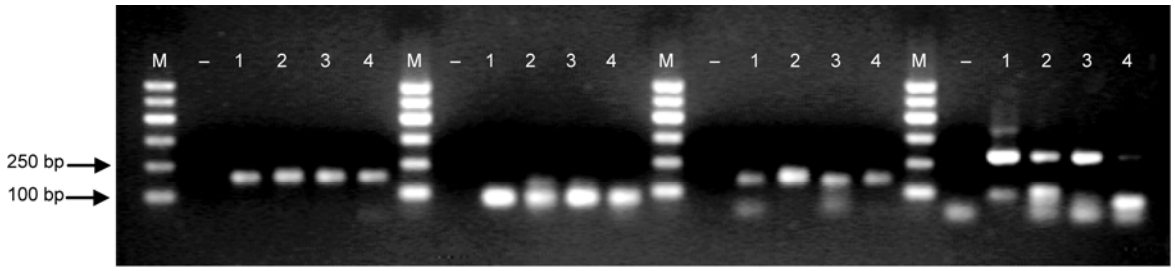

Figure 2 PCR amplification of ancient rice DNA. The results from four pairs of primers (A, C, R, and Y) are arranged from left to right, respectively. 1, Huzhou; 2 and 3, Tianluoshan; 4, Xingan. M, marker; "–", negative control.

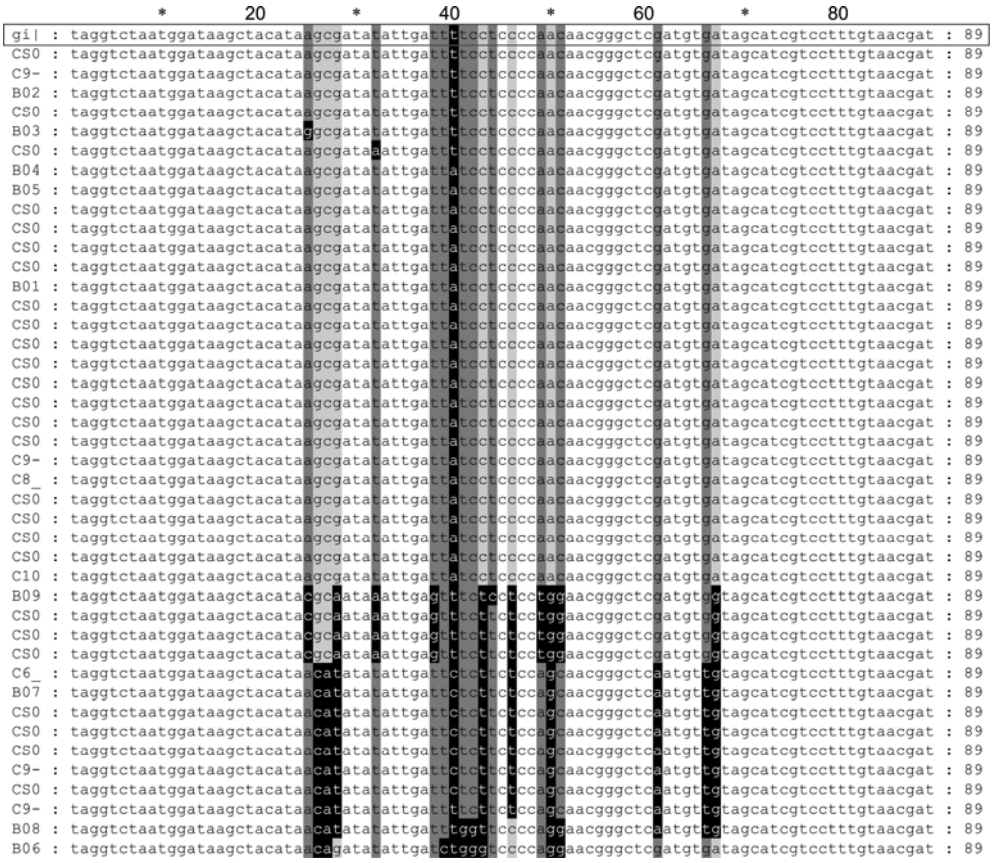

Figure 3 Multiple alignment of sequences from the 3 archaeological sites amplified with the C primer pair. The sequence (AK242631) in the box is from the modern rice cultivar Nipponbare. The columns with different nucleotides were labled.

Table 3 Number of clones sequenced from rice remains in this study

Primer code	A	C	R	Y	Total
Total	21	47	45	29	142
Target: rice	11	28	43	23	105
Target: others	10	15	0	0	25
Non-targets	0	4	2	6	12

2.4 Comparison of sequences from ancient and modern rice

Based on the multiple alignment of sequences generated using the 2 rice-specific primers (R and Y), at least two genotypes were found in each sample of rice remains from the three archaeological sites, suggesting a heterozygotic state in the rice seeds at that time. For example, two genotypes were found in the rice remains from the Tianluoshan site (Table 4). These results were consistent with the phenotypic observations of the ancient rice chaffs (glumes) by us and other studies (e.g., [6]) in that their shapes were not identical and some of them were novel.

For a comparison between ancient and modern rice, we sampled and sequenced wild and cultivated rice collected Asia-wide, including *japonica* landraces from Zhejiang Province; the results of a recent investigation of over 500

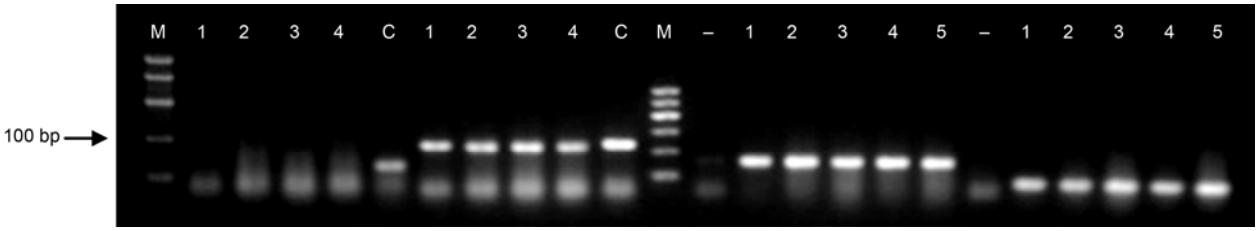

Figure 4 PCR amplification results for *L. oryzoides* and *Z. latifolia* using the R, Y, A, and C primers (from left to right). No PCR products were observed using the rice-specific primer pair R. 1, *L. oryzoides*; 2–5, *Z. latifolia*; C, rice (Nipponbare); "–", negative control; M, marker.

Table 4 Genotyping of ancient and modern rice[a)]

Primer	Genotype	Archaeological site				Asia cultivar (1)[b)]			Asia wild	Asia cultivar (2)[b)]
		Tianluoshan	Xingan	Huzhou	Total	*indica*	*japonica*	Total		*indica*
R	A-T-A-$(TC)_4$	12	7	9	28	9	18	27	1	17
	C-T-A-$(TC)_4$	0	0	0	0	0	0	0	0	210
	C-C-A$(TC)_4$	0	0	0	0	0	0	0	0	121
	A-C-T-$(TC)_8$	0	6	2	8	0	0	0	0	0
	A-C-T-$(TC)_9$	5	0	0	5	2	1	3	1	0
	Others	0	0	0	0	1	0	1	13	26
	Total	17	13	11	41	12	19	31	15	374
Y	T-G	12	5	1	18	1	15	16	2	0
	C-G	0	0	0	0	11	4	15	4	345
	C-C	2	1	0	3	0	0	0	0	0
	C-A	0	0	0	0	0	0	0	0	28
	Total	14	6	1	21	12	19	31	6	373

a) The table shows the distribution of clone/accession numbers with different genotypes from rice remains and modern rice amplified with the R and Y primer pairs. b) Asia cultivar (1) and (2) according to this study and Huang et al. [13], respectively.

Chinese *indica* landraces by genome re-sequencing [13] were also included (Tables 2 and 4). We found that the main genotypes (with the highest number of clone sequences) in ancient rice remains [R: A-C-T-$(TC)_4$ and Y: T-G] were also the main genotypes of modern *japonica* strains (all 6 *japonica* landraces from Zhejiang Province had the same genotype), suggesting that the genotype has been maintained in the subsequent long process of genetic improvement. However, several differences between the ancient and modern rice were observed: (i) the ancient genotypes [R: A-C-T-$(TC)_8$ and Y: C-C] were not found in our rice cultivars, suggesting that they might have been lost in the current gene pool. (ii) Further investigation of more cultivated strains, particular *japonica*, is needed to confirm this. Crops, including rice, have experienced strong artificial selection during domestication and their genetic diversity has significantly decreased — the so-called bottleneck effect [17,18]. Our results provide new evidence for this observation. (iii) several genotypes of modern rice [R: C-T-A-$(TC)_4$, C-C-A$(TC)_4$, and Y: C-G] were not observed in our ancient rice remains, which might be because of our limited archaeological sites, or indicate that new mutations have occurred late or that new rice germplasts have been introduced from abroad. For example, many crops or new crop cultivars have been introduced from abroad throughout Chinese history. For rice, for example, an early mature cultivar introduced from Vietnam was widely planted in the south of China in the Song Dynasty [19,20].

The divergence of the *indica* and *japonica* subspecies is an important evolutionary event in rice domestication. We wonder which of these types our rice remains belonged. Based on the sequence alignments, the sequences from the rice remains were the same as those of most *japonica* strains (including the 6 *japonica* landraces from Zhejiang Province),

Table 5 The best sequence matches (hits) for the 25 sequences putatively derived from other ancient plants amplified from the rice remains by the degenerate primer pairs A and C

Primer	Genotype	Number of clones	Site[a)]	Best hit (*E*-value)[b)]	Taxonomy
A	A96	2	T/X	*Citrus trifoliate* etc. (1e-74)	Sapindales
	A105	6	T/X	*Sinapis alba* etc. (4e-69)	Brassicales
	Y8-C	1	T	*Cercomonas media* etc.(2e-79)	Rhizaria
	A99	1	T	*Pichia kluyveri* etc.(7e-62)	Fungi
C	C6-C	7	T/X/H	*Tribulus terrestris* etc.(2e-34)	Zygophyllales
	C9-5	1	T	*Guarea glabra* etc.(2e-34)	Sapindales
	C62	1	H	*Peganum harmala* etc.(2e-31)	Sapindales
	C61	4	T/X/H	*Arabidopsis thaliana* etc.(2e-34)	Brassicales
	C64	1	H	*Coix lacryma-jobi* etc.(3e-38)	Poales

a) T, Tianluoshan; X, Xingan; H, Huzhou. b) BLASTN.

but were different from those of most *indica* landraces. Therefore, we believe that the ancient rice remains are *japonica*-type, or most if not all are *japonica*; i.e., they might be at an early stage of the *indica-japonica* divergence or an *indica-japonica* mixture.

2.5 Sequences from other ancient plants

In the sequences identified from rice remains by the two degenerate primer pairs (A and C), 25 matched other plants perfectly or with only one mismatch, but did not match rice. This suggested that these sequences might be derived from other ancient plants existing in the same environment as the ancient rice. According to the best hit to known sequences by a BLAST search of GenBank, these non-rice sequences were mainly derived from Sapindales, Zygophyllales, and Brassicales (Table 5). The Sapindale species, including the Chinaberry tree (*Melia azedarach*) and citrus trees, are commonly found in the Yangtze regions of China. We unearthed many Chinaberry fruits from the Tianluoshan archeological site (Hemudu), which is consistent with our sequencing results. Among the best-hit Zygophyllales species was *Tribulus terrestris*, a common plant in south China that is also used in traditional Chinese medicine. The Brassicale species included well-known wet-field rice weeds, such as *Sinapis alba* and *Arabidopsis thaliana.* Moreover, a sequence from a traditional Chinese cereal *Coix lacryma-jobi* was also obtained from our samples.

The authors thank Karsten Hofmann (PLANTON GmbH) for his help with the collection of ancient DNA. This work was supported by the National Basic Research Program of China (2011CB109306) and the National High Technology Research and Development Program of China (2006AA10A102); and partially by a pilot program (2007DKA20Z90) and a DAAD Fellowship.

1 Vaughan D A, Lu B, Tomooka N. The evolving story of rice evolution. Plant Sci, 2008, 174: 394–408
2 Londo J P, Chiang Y C, Hung K H, et al. Phylogeography of Asian wild rice, *Oryza rufipogon,* reveals multiple independent domestications of cultivated rice, *Oryza sativa*. Proc Natl Acad Sci USA, 2006, 103: 9578–9583
3 You X L. A Global Chinese History (the primitive society volume) (in Chinese). Beijing: China Agricultural Press, 2008. 175–198
4 You X L. Several view about rice and bone spade excavated from the 4th cultural layer in the Hemudu site (in Chinese). Cult Relics, 1976, (8): 20–23
5 Zheng Y F, Shun D G, Chen X G. Characteristics of the short rachillae of rice from archaeological sites dating to 7000 years ago. Chinese Sci Bull, 2007, 52: 1037–1041
6 Fuller D Q, Qin L, Zheng Y F, et al. The domestication process and domestication rate in rice: Spikelet bases from the lower Yangtze. Science, 2009, 323: 1607–1610
7 Goloubinoff P, Pääbo S, Wilson A. Evolution of maize inferred from sequence diversity of an *Adh2* gene segment from archaeological specimens. Proc Natl Acad Sci USA, 1993, 90: 1997–2001
8 Deakin W J, Rowley-Conwy P J, Shaw C H. Amplification and sequencing of DNA from preserved sorghum of up to 2800 years antiquity found at Qasr Ibrim. Ancient Biomol, 1998, 2: 27–41
9 Rollo F, Ubaldi M, Ermini L, et al. Otzi's last meals: DNA analysis of the intestinal content of the Neolithic glacier mummy from the Apls. Proc Natl Acad Sci USA, 2002, 99: 12594–12599
10 Jaenicke-Despres V, Buckler E S, Smith B D, et al. Early allelic selection in maize as revealed by ancient DNA. Science, 2003, 302: 1206–1208
11 Erickson D L, Smith B D, Clarke A C, et al. An Asian origin for a 10000-year-old domesticated plant in the Americas. Proc Natl Acad Sci USA, 2005, 102: 18315–18320
12 Nakamura I, Sato Y I. Amplification of DNA fragments isolated form a single seed of ancient rice (AD 800) by polymerase chain reaction. Chinese J Rice Sci, 1991, 5: 175–179
13 Huang X H, Wei X H, Sang T, et al. Genome-wide association studies of 14 agronomic traits in rice landraces. Nat Genet, 2010, 42: 961–976
14 Levison P R, Badger S E, Dennis J, et al. Recent developments of magnetic beads for use in nucleic acid purification. J Chromatography A, 1998, 816: 107–111
15 Yang D Y. Contamination controls and detection in ancient DNA studies. Acta Anthropol Sin, 2003, 22: 163–173
16 Ge S, Sang T, Lu B R, et al. Phylogeny of rice genomes with emphasis on origins of allotetraploid species. Proc Natl Acad Sci USA, 1999, 96: 14400–14405
17 Doebley J F, Gaut B S, Smith B D. The molecular genetics of crop domestication. Cell, 2006, 127: 1309–1321
18 Zhu Q, Zheng X, Luo J, et al. Multilocus analysis of nucleotide variation of *Oryza sativa* and its wild relatives: Severe bottleneck during domestication of rice. Mol Biol Evol, 2007, 24: 875–888
19 Zeng X S. Double-harvest rice in the Song Dynasty (in Chinese). Studies History Nat Sci, 2002, 21: 255–268
20 You X L. On Champa rice (in Chinese). Agri Archaeol, 1983, (1): 25–31