

RESEARCH ARTICLE

Open Access

Diversity and selective sweep in the *OsAMT1;1* genomic region of rice

Zehong Ding¹, Chongrong Wang¹, Sheng Chen², Sibin Yu^{1*}

Abstract

Background: Ammonium is one of the major forms in which nitrogen is available for plant growth. *OsAMT1;1* is a high-affinity ammonium transporter in rice (*Oryza sativa* L.), responsible for ammonium uptake at low nitrogen concentration. The expression pattern of the gene has been reported. However, variations in its nucleotides and the evolutionary pathway of its descent from wild progenitors are yet to be elucidated. In this study, nucleotide diversity of the gene *OsAMT1;1* and the diversity pattern of seven gene fragments spanning a genomic region approximately 150 kb long surrounding the gene were surveyed by sequencing a panel of 216 rice accessions including both cultivated rice and wild relatives.

Results: Nucleotide polymorphism (Pi) of *OsAMT1;1* was as low as 0.00004 in cultivated rice (*Oryza sativa*), only 2.3% of that in the common wild rice (*O. rufipogon*). A single dominant haplotype was fixed at the locus in *O. sativa*. The test values for neutrality were significantly negative in the entire region stretching 5' upstream and 3' downstream of the gene in all accessions. The value of linkage disequilibrium remained high across a 100 kb genomic region around *OsAMT1;1* in *O. sativa*, but fell rapidly in *O. rufipogon* on either side of the promoter of *OsAMT1;1*, demonstrating a strong natural selection within or nearby the ammonium transporter.

Conclusions: The severe reduction in nucleotide variation at *OsAMT1;1* in rice was caused by a selective sweep around *OsAMT1;1*, which may reflect the nitrogen uptake system under strong selection by the paddy soil during the domestication of rice. Purifying selection also occurred before the wild rice diverged into its two subspecies, namely *indica* and *japonica*. These findings would provide useful insights into the processes of evolution and domestication of nitrogen uptake genes in rice.

Background

Rice, one of the world's most important crops, is the staple food of over half of the world's population. Cultivated rice occurs in two forms, *Oryza sativa* and *Oryza glaberrima*. *Oryza sativa*, referred to as the Asian rice, is cultivated worldwide but distributed mainly in Asia and has two subspecies, *indica* and *japonica*, estimated to have diverged at least about 10,000 years ago [1,2]. The two subspecies are further divided into major subpopulations: *indica* into *aus* and *indica*, *japonica* into *temperate japonica*, *tropical japonica*, and the *aromatic* group [3]. It has been proposed that these subpopulations were

independently domesticated from divergent pools of *O. rufipogon* [3,4].

During the process of domestication, genetic diversity in rice has reduced dramatically as a result of natural and artificial selection [5-7]. With the rapid development of molecular genomics, the path of domestication has been traced in detail through characterization of several domestication genes associated with artificial selection. For example, a major gene for grain shattering in rice, *shattering 4* (*sh4*), has been characterized: nucleotide diversity of a ~50 kb region surrounding *sh4*, is nearly one-tenth of the average value across the entire chromosome 4 [8,9]. The non-shattering *sh4* allele was fixed in all the *O. sativa* varieties surveyed. A selective sweep nearby the *sh4* genomic region was probably due to strong artificial selection of the non-shattering allele, which confers a trait important to efficient grain harvest [10]. Similar findings have been reported in two other

* Correspondence: ysb@mail.hzau.edu.cn

¹National Key Laboratory of Crop Genetic Improvement, and the College of Plant Science and Technology, Huazhong Agricultural University, Wuhan 430070, PR China

Full list of author information is available at the end of the article

selection-related genes in rice, namely *Red pericarp (Rc)*, which is responsible for pericarp colour, and *waxy (wx)*, for glutinous grains. The two were under strong artificial selection given the human preference for white rice and low-amylose grains. As a result, the level of polymorphism in and around these two genes is markedly reduced [11,12]. Nucleotide diversity in the gene *teosinte branched 1 (tb1)*, which is involved in plant architecture, is much lower in maize than in *teosinte*, the wild relative of maize. Wang et al. [13] showed that maize has retained only 3% of the diversity found in *teosinte* in the 5' non-transcribed region of *tb1*, again as a result of human selection. This selection effect led to a selective sweep of a 60-90 kb long region across the 5' promoter to the *tb1*-transcribed sequence [14,15]. Similarly, nucleotide variation in the gene *phytoene synthase (Y1)*, which makes maize kernels yellow, was only 5.3% of that in the gene that makes them white, owing to the human preference for yellow grains, which have higher nutritional value [16]. Because of the effect of selection hitchhiking, nucleotide diversity was lower around the region of *Y1* from ~600 kb downstream to ~200 kb upstream [17]. These results provide significant insights into the selective sweeps brought about by artificial selection during the process of crop domestication.

Ammonium is one of the major forms in which nitrogen is available for plant growth and reproduction. Ammonium absorption from soil is mediated by ammonium transporters (AMT) through both high-affinity and low-affinity transport systems. The high-affinity transport system plays a major role when ammonium in soil is at micromolar ammonium concentrations, and the low-affinity transport system mainly at millimolar or higher concentrations [18,19]. Phylogenetic analysis shows that the two systems are represented by AMT subfamilies AMT1 and AMT2 respectively [20,21]. Kaiser et al. [22] reported that ammonium uptake decreased by about 30% in a T-DNA insertion line of *AtAMT1;1* compared to the wild-type, which encodes a high-affinity ammonium transporter in *Arabidopsis thaliana*. Other studies have also confirmed that the role of AMT1 in other species [23-26].

Rice carries at least twelve AMT genes [21], three of which, *OsAMT1;1* to *OsAMT1;3*, have been identified as members of AMT1, each showing a distinct expression pattern: *OsAMT1;1* shows constitutive expression in both shoots and roots; *OsAMT1;2* shows root-specific and ammonium-inducible expression; and *OsAMT1;3* shows root-specific and nitrogen-suppressible expression [27,28]. *OsAMT1;1* has only one open reading frame encoding 499 amino acid residues. Its transcript was abundant in roots of plants that were grown under low levels of NH_4^+ but decreased rapidly when the plants were transferred to a medium with high NH_4^+

concentration [27]. Moreover, *OsAMT1;1* expression was promoted by low ammonium in rice roots [28] and regulated by endogenous glutamine rather than by endogenous ammonium [29]. However, nucleotide variations in *OsAMT1;1* and in other key genes in the nitrogen uptake and assimilation pathway are yet to be elucidated, and their evolutionary path from wild rice to cultivated rice remains unclear.

Improving the uptake and assimilation efficiency of nitrogen has been a major breeding target in rice, and success depends largely on discovering allelic variations of nitrogen metabolism genes in rice germplasm and understanding their adaptive significance. The objectives of the present study were 1) to investigate allelic variations of *OsAMT1;1*, 2) to analyse its pattern of polymorphism and its evolution, and 3) to define the extent of selective sweep around the *OsAMT1;1* region in cultivated rice.

Methods

Rice materials

The nucleotide variation of the *OsAMT1;1* gene was assessed in a panel of 216 *Oryza* accessions comprising 190 accessions of *O. sativa* (102 *indica*, 85 *japonica*, 2 aromatic, and 1 *aus* varieties), and 25 of wild rice (19 *O. rufipogon*, 5 *O. nivara*, and 1 *O. barthii*), and 1 of the African cultivated rice *O. glaberrima* (Additional file 1, Table S1). Most of the *O. sativa* accessions - 103 of them were of landraces (<http://icgr.caas.net.cn/>, Additional file 1, Table S1) - came from a collection of Chinese rice varieties estimated to represent about 70% of the total diversity in Chinese rice germplasm [30]. The other accessions of *O. sativa*, including 8 landraces, were selected from the OryzaSNP project <http://www.oryzasnp.org/>. The 25 wild relatives with different original sources were obtained from the International Rice Research Institute (IRRI).

Nucleotide diversity across the genomic region around *OsAMT1;1* was also surveyed by sequencing seven gene segments from 94 accessions of *O. sativa* and 19 accessions of *O. rufipogon* (Additional file 1, Table S1).

PCR and DNA sequencing

Nucleotide diversity in the gene *OsAMT1;1* and the surrounding genomic region was assessed by PCR amplification and sequencing. Plant DNA was extracted from young leaves of each accession using cetyltrimethylammonium bromide (CTAB) method described by Rogers and Bendich [31]. All primers for the PCR were designed by Primer Premier 5.0 based on the Nipponbare genomic sequence available at the TIGR website <http://rice.plantbiology.msu.edu/cgi-bin/gbrowse/rice/>. For *OsAMT1;1*, four pairs of primers were used for PCR amplification to cover the entire coding sequence,

~1080 bp of the 5' promoter and ~110 bp of the 3' untranslated region. For each locus around *OsAMT1;1*, one pair of primer was designed to amplify fragments of the genomic DNA approximately 860-1060 bp long (Additional file 2, Table S2). To ensure that only the targeted genomic region is amplified, all PCR primers were confirmed by BLAST against the Nipponbare genomic sequence in the NCBI database <http://blast.ncbi.nlm.nih.gov/Blast.cgi>.

The PCR experiments were performed on 20 μ L amplification reaction consisting of 0.75 unit of *Taq* polymerase (Takara, Dalian, China), 10 mM Tris-HCl, 50 mM KCl, 1.5 mM MgCl₂, 0.2 mM dNTPs, 0.15 μ M of each primer, and 50 ng of genomic DNA. PCR reactions were carried out in GeneAmp PCR System 9700 (Applied Biosystems, USA) under the following conditions: 5 min at 94°C, followed by 35 cycles of 1 min at 94°C, 1 min at a primer-specific temperature (Additional file 2, Table S2), and 1 min at 72°C and, for a final extension, 5 min at 72°C.

PCR products were directly sequenced with the Big-Dye Terminator Cycle Sequencing v3.1 (Applied Biosystems) after digestion and purification according to the manufacturer's specifications. Briefly, the PCR products were digested with Exonuclease I (Biolabs Inc.) and Shrimp Alkaline Phosphatase (Takara), and purified with 95% ethanol and 3 M sodium acetate (pH 5.2). Sequencing reactions were run on an automated capillary sequencer ABI 3730, and sequence contigs were assembled and edited using Sequencher 4.1.2 (Gene Codes Corporation, USA). To ensure reliability of the sequencing data, PCR amplicons were sequenced in both orientations, and any remaining ambiguities in the sequences resolved by repeated PCR amplification and re-sequencing. Almost all fragments were homozygous except two fragments from the two *O. rufipogon* accessions (DWR35 and DWR38) with heterozygous sites. PCR products of these two fragments were cloned into pGEM-T easy vectors (Promega, USA) and five independent clones were sequenced for each fragment. If two different alleles were found, the predominant allele of the five clones was arbitrarily chosen for further analysis [11].

Data analysis

Sequences alignment was conducted with ClustalX 1.83 <ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/>. The number of segregating sites and levels of nucleotide diversity π (π), the average number of nucleotide differences per site between two sequences [32], and θ , an estimate of $4N_e\mu$, where N_e is the effective population size and μ is the mutation rate per nucleotide [33], were computed in DnaSP (version 5.10.00). Tajima's D test and Fu and Li's D* test were also performed in DnaSP for testing

selections deviating from neutrality [34]. Haplotype analysis was conducted using TASSEL (version 2.0.1) [35]. To visualize the impact of sweep selection across the *OsAMT1;1* genomic region, extended haplotype homozygosity (EHH) analysis was performed to measure the degree of linkage disequilibrium (LD), as sweep selection is expected to elevate the LD nearby the selected site. EHH was reported on a scale from 0 to 1, with 0 indicating that all extended haplotypes are different and 1 indicating that they are the same [36].

To investigate selection pressure on *OsAMT1;1*, *Ka/Ks* ratios (non-synonymous substitution rate / synonymous substitution rate) between *OsAMT1;1* in rice and its homologs in the relative species were calculated using the method described by Goldman and Yang [37]. A *Ka/Ks* ratio significantly greater than one implies positive selection, less than one implies purifying selection, and a ratio of one may indicate neutrality [38].

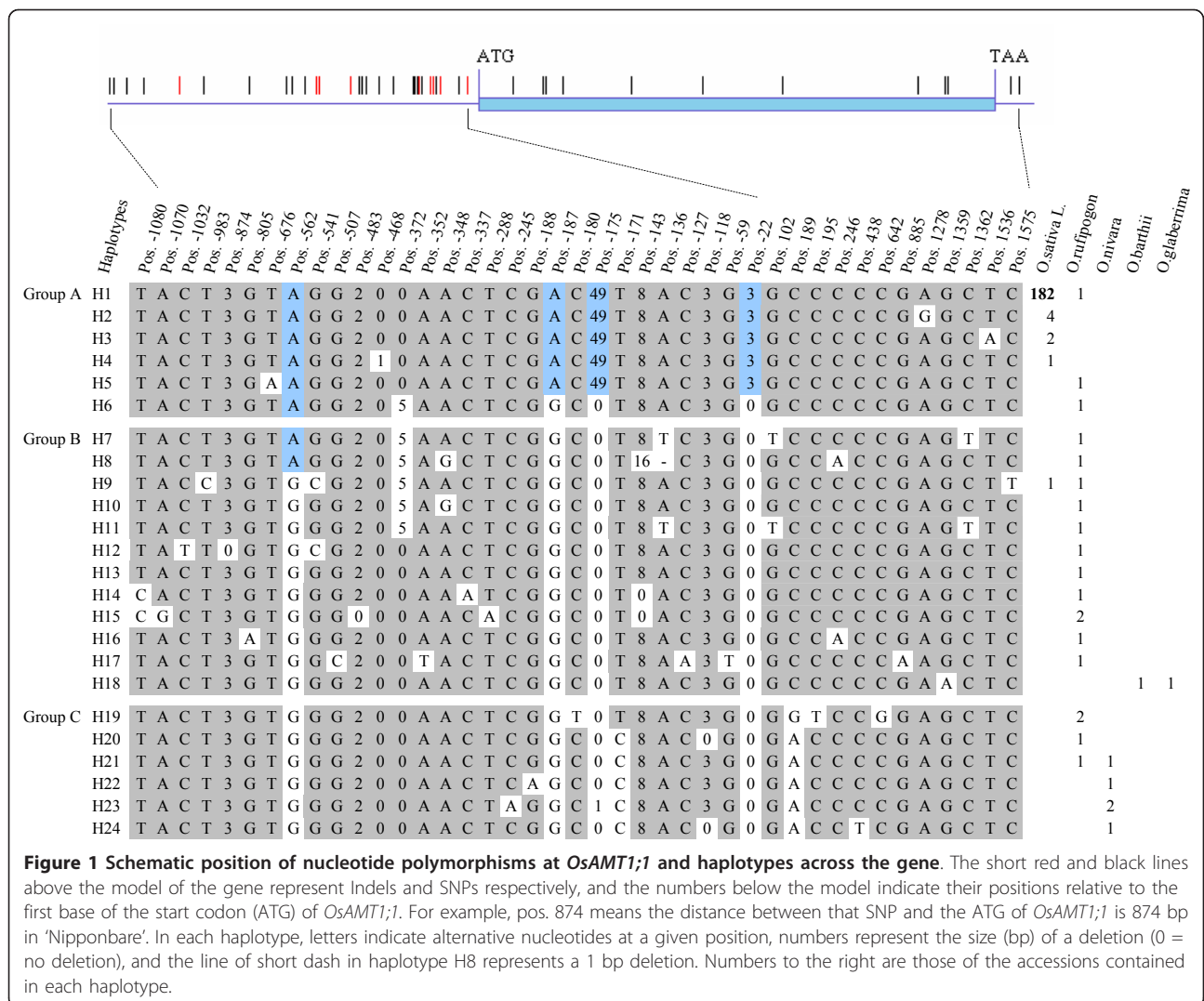
Results

Sequence diversity at *OsAMT1;1*

Nucleotide polymorphism was examined in 216 accessions across the entire 2.68 kb length of *OsAMT1;1* including a 1083 bp fragment of the 5' promoter region and a 111 bp fragment of the 3' untranslated region. The examination yielded 32 single nucleotide polymorphisms (SNP) and 9 insertions or deletions (Indels) in this 2.68-kb region (Figure 1), of which 30 polymorphism sites (25 SNPs and 5 Indels) were observed in the wild rice accessions and most were located in the non-coding region. Five polymorphism sites, at positions -562, -372, -187, -175, and -22 of the promoter region, occurred in high or intermediate frequencies in all the accessions (Figure 1), whereas at the coding region, ten SNPs - all synonymous variations - occurred rare frequency (well below 5%), indicating that the gene function had been conserved.

Oryza rufipogon was found to be at least 30 times as diverse as *O. sativa* across the entire sequenced region, including the coding and non-coding regions of *OsAMT1;1* (Table 1). Nucleotide diversity π of the entire gene sequence was 0.00172, and θ was 0.00278 for *O. rufipogon*, while estimates of π for *indica* and *japonica* were much lower ($\pi = 0.00006$ and 0.00002 and $\theta = 0.00043$ and 0.00007 , respectively). These results indicate that the diversity of *OsAMT1;1* was much less in *O. sativa* than in its wild relatives. In addition, nucleotide polymorphism in the coding region was much lower than that in the non-coding region for all accessions (Table 1). Nucleotide diversity in the coding region was about a third of that in the non-coding region in *O. rufipogon* and about half in cultivated rice (Table 1).

Such extreme reduction in diversity of *OsAMT1;1* was observed in both landraces of *O. sativa* and its elite



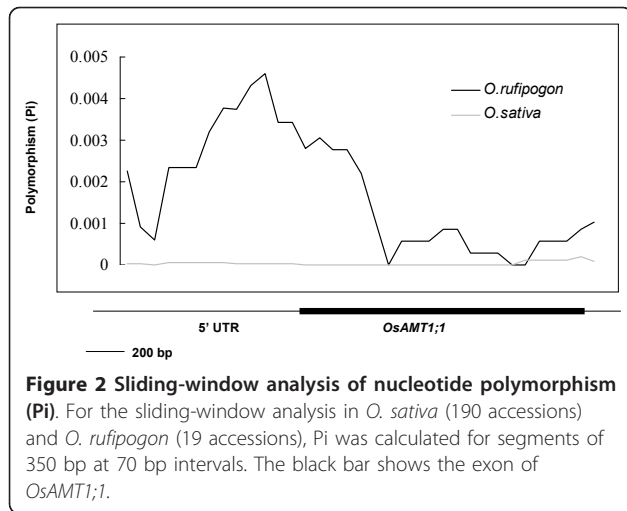
varieties (Additional file 5, Figure S1). On the other hand, in the case of two other nitrogen-related genes (Loc_Os02 g50240 and Loc_Os01 g48960), the reduction in diversity in landraces and elite varieties was more modest; in fact in the case of Loc_Os01 g48960, the extent of diversity in landraces was nearly the same as that in elite lines of *O. sativa* (Additional file 5, Figure S1).

Results of Tajima's D and Fu and Li's D* tests for the entire region and the non-coding region of *OsAMT1;1* were negative and significantly deviated from neutrality in all accessions, as well as in *O. sativa* samples. Fu and Li's D* values were not significant in the coding region in any of the accessions and subsamples (Table 1). Sliding window analysis of the pattern of polymorphism

Table 1 Nucleotide diversity and neutral test at *OsAMT1;1*

| Test set ^a | Entire sequenced region (2688 bp) | | | | | Coding region (1494 bp) | | | | | Non-coding region (1194 bp) | | | | |
|------------------------------|-----------------------------------|-------------------|----------------------|---------|---------|-------------------------|-------------------|----------------------|--------|-------|-----------------------------|-------------------|----------------------|---------|---------|
| | S | $\pi \times 10^3$ | $\theta \times 10^3$ | D1 | D2 | S | $\pi \times 10^3$ | $\theta \times 10^3$ | D1 | D2 | S | $\pi \times 10^3$ | $\theta \times 10^3$ | D1 | D2 |
| All (n = 216) | 32 | 0.37 | 2.01 | -2.33** | -2.01 | 10 | 0.16 | 1.12 | -2.04* | -0.10 | 22 | 0.63 | 3.13 | -2.15** | -2.57* |
| <i>O. rufipogon</i> (n = 19) | 26 | 1.72 | 2.78 | -1.59 | -0.57 | 7 | 0.99 | 1.34 | -1.22 | 0.79 | 19 | 2.65 | 4.60 | -1.62 | -1.16 |
| <i>O. sativa</i> (n = 190) | 7 | 0.04 | 0.45 | -1.93* | -3.58** | 1 | 0.03 | 0.11 | -0.77 | 0.45 | 6 | 0.06 | 0.87 | -1.89* | -4.08** |
| <i>Indica</i> (n = 102) | 6 | 0.06 | 0.43 | -1.90* | -3.67** | 1 | 0.05 | 0.13 | -0.69 | 0.49 | 5 | 0.08 | 0.81 | -1.90* | -4.29** |
| <i>Japonica</i> (n = 85) | 1 | 0.02 | 0.07 | -0.90 | 0.50 | 0 | 0 | 0 | NA | NA | 1 | 0.04 | 0.17 | -0.91 | 0.50 |

^an in parentheses is the number of accessions used; S, number of segregating sites; π , nucleotide diversity per site; θ , Watterson's estimator; D1, Tajima's D; D2, Fu and Li's D*; NA, not applied; *, ** significant at 0.05 and 0.01(or 0.02) level for Tajima's D (or Fu and Li's D*), respectively.



across the gene in *O. sativa* (190 accessions) and *O. rufipogon* (19 accessions) revealed a sharp peak at a point of about 200 bp upstream of the start codon of the gene (Figure 2). These results suggest a strong selection at *OsAMT1;1*, particularly in the 5' promoter region.

Notably, nine SNPs in the coding region of *OsAMT1;1* that occurred in the wild relatives were all synonymous (Figure 1), suggesting that *OsAMT1;1* is under purifying selection in the wild. To further validate this suggestion, *Ka/Ks* ratios were also calculated between *OsAMT1;1* and its homologs in *O. barthii* (DWR50), *Sorghum bicolor* (Sb06 g022230), and *Zea mays* (GRMZM2G175140) from the TIGR website. The ratios ranged from 0.0010 to 0.0344, significantly less than 1.0 (Additional file 3, Table S3), suggesting strong purifying selection of *OsAMT1;1*.

Haplotype variations across *OsAMT1;1*

A total of 24 haplotypes were identified in the entire region of gene *OsAMT1;1* in the 216 accessions. Phylogenetic analysis showed that these distinct haplotypes fell into three groups (A, B and C). In group A, four haplotypes, H1 to H4, were detected in *O. sativa*. Most (182/190) accessions of the Asian cultivated rice (both *indica* and *japonica*) belonged to haplotype H1. Each of the haplotypes H2 to H5 in group A differed from H1 in only a single nucleotide site, whereas haplotype H6 differed from H1 in several sites (Figure 1). The 19 accessions of *O. rufipogon* fell into as many as 17 distinct haplotypes, indicating a high level of genetic variation in *O. rufipogon* (Figure 1). In group B, *O. glaberrima* and *O. barthii* had the same haplotype, H18, indicating their close genetic relationship. Five *O. nivara* and four *O. rufipogon* belonged to group C. As can be seen in Figure 1, almost all accessions of

O. sativa are of haplotype H1, which incorporates the alleles at positions -562, -187, -175 and -22 in the 5' promoter region, whereas practically every accession of the wild rice is of a different haplotype. Incidentally, the one exception from *O. sativa* was of haplotype H9.

Selective sweep surrounding *OsAMT1;1*

To assess the impact of selection at *OsAMT1;1*, the extent of nucleotide polymorphism was examined at seven loci within a region around *OsAMT1;1* approximately 150 kb long. The average nucleotide diversity (π) for *O. sativa* was low at five loci from -74.1 kb to 27.5 kb (Figure 3A and Table 2) but markedly high at the other two loci of -101.9 kb and 49.5 kb. The relative levels of nucleotide diversity between *O. sativa* and *O. rufipogon* were also calculated across the *OsAMT1;1* genomic region. A striking reduction in the relative diversity (the ratio ranged from 0.012 to 0.077) was observed along the same region mentioned above (-74.1 kb to 27.5 kb) (Figure 3B and Table 2). These results indicate a selective sweep around *OsAMT1;1* extending about 100 kb in length.

Similar patterns were drawn from Hd, a measure of haplotype diversity, the value of which for *O. sativa* was markedly lower for the five loci along the stretch from -74.1 kb to 27.5 kb than that for either of the loci at -101.9 kb or 49.5 kb. However, Hd for *O. rufipogon* was much higher across all six loci including the gene *OsAMT1;1* (Table 2). The values of Fu and Li's D^* tests were significantly negative for the five loci (-74.1 kb to 27.5 kb) in *O. sativa*, which is a signature of selection in *OsAMT1;1* genomic region from *O. rufipogon* to *O. sativa* (Table 2).

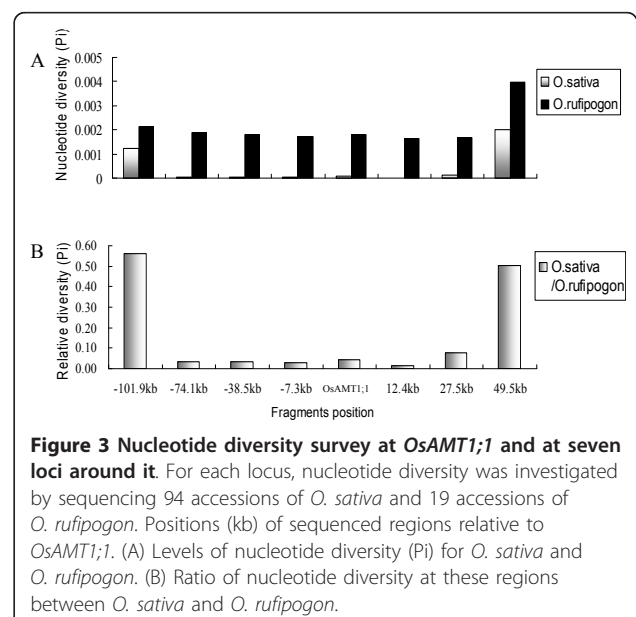


Table 2 Nucleotide diversity at *OsAMT1;1* and at seven loci around it in *O. sativa* and *O. rufipogon*

| Genes predicted | Position ^a | Length ^b | <i>O. sativa</i> (n = 94) ^c | | | | <i>O. rufipogon</i> (n = 19) | | | | The ratio of Pi ^d |
|------------------------------------|-----------------------|---------------------|--|-------|--------|---------|------------------------------|-------|--------|---------|------------------------------|
| | | | Pi×10 ³ | Hd | D1 | D2 | Pi×10 ³ | Hd | D1 | D2 | |
| Glutaredoxin subgroup I | -101.9 | 853 | 1.21 | 0.516 | 1.38 | -0.57 | 2.15 | 0.830 | -2.02* | -3.11** | 0.563 |
| C3HC4 type zinc finger family | -74.1 | 757 | 0.06 | 0.042 | -1.39 | -2.80* | 1.87 | 0.766 | -1.31 | -1.44 | 0.032 |
| Lipase class 3 family | -38.5 | 756 | 0.06 | 0.042 | -1.39 | -2.80* | 1.81 | 0.754 | -1.63 | -2.17 | 0.033 |
| Cupin superfamily | -7.3 | 863 | 0.05 | 0.021 | -1.39 | -2.80* | 1.72 | 0.865 | -1.20 | -1.44 | 0.029 |
| <i>OsAMT1;1</i> | 0 | 2688 | 0.08 | 0.123 | -2.00* | -3.13* | 1.72 | 0.982 | -1.59 | -0.57 | 0.047 |
| Hypothetical protein | 12.4 | 877 | 0.02 | 0.021 | -1.03 | -2.02 | 1.65 | 0.807 | -1.25 | -0.88 | 0.012 |
| Transcriptional corepressor LEUNIG | 27.5 | 638 | 0.13 | 0.042 | -1.79* | -3.84** | 1.69 | 0.678 | -1.21 | -1.51 | 0.077 |
| Expressed protein | 49.5 | 832 | 2.01 | 0.576 | 0.52 | -1.40 | 3.99 | 0.906 | -0.50 | -0.41 | 0.504 |

^aPosition of sampled loci as far to the first nucleotide of *OsAMT1;1* that was considered as 0 kb; ^bLength (bp) of sampled alignment regions is relative to Nipponbare genomic sequence; ^cPi (π), nucleotide diversity per site; Hd, haplotype diversity; D1, Tajama's D; D2, Fu and Li's D*; *, ** significant at 0.05 and 0.02 level, respectively; ^dThe relative ratio of Pi for *O. sativa* to *O. rufipogon*.

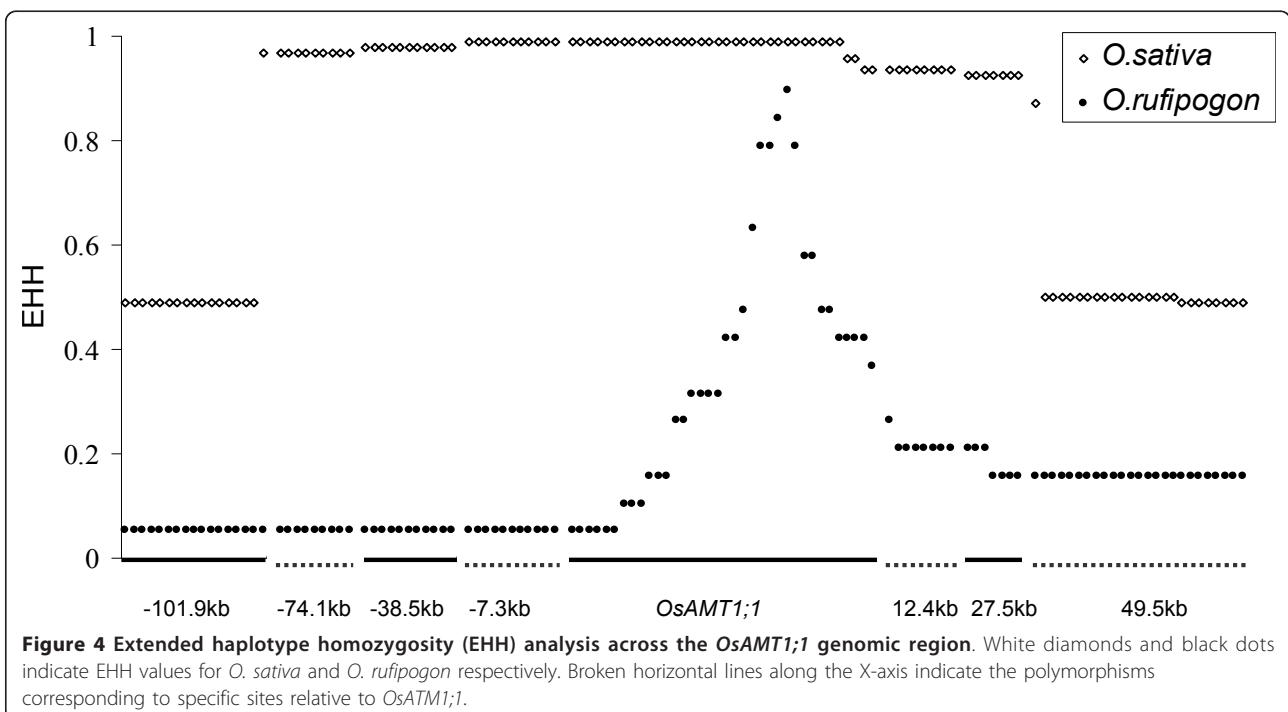
Lastly, the extent of linkage disequilibrium (LD) was measured by calculating the extended haplotype homozygosity (EHH) for *O. sativa* and *O. rufipogon*. The average EHH value was 0.926 in *O. sativa* across the region from -74.1 kb to 27.5 kb but abruptly dipped to 0.489 at the other two loci (-101.9 kb and 49.5 kb). The high value of LD along a distance of ~100 kb (from -74.1 kb to 27.5 kb) in *O. sativa* and the marked reduction in LD in *O. rufipogon* on either side of the core location in the promoter of *OsAMT1;1* (Figure 4) support the idea of a strong selection acting on *OsAMT1;1*.

Discussion

Diversity and selection of *OsAMT1;1*

Artificial selection is believed to have played an important role in the severe reduction of genetic diversity in

cultivated rice during its domestication from its wild ancestors. *OsAMT1;1* plays a key role in ammonium uptake and assimilation in rice but whether *OsAMT1;1* and other nitrogen-uptake genes were subjected to selection pressure during the domestication process is an open question. In the present study, we sequenced the entire gene *OsAMT1;1* in a panel of 216 rice accessions, drawn mostly from Chinese cultivated rice (*O. sativa*), and observed the predominance of one major haplotype (H1) and extreme reduction in genetic diversity of *OsAMT1;1* in cultivated rice compared to that in its wild progenitor *O. rufipogon*. The level of nucleotide polymorphism ($\pi = 0.00006$) in *OsAMT1;1* in *O. sativa* was also markedly lower than that in 111 randomly chosen gene fragments ($\pi = 0.0023$) [6] and in 10 reference genes from *O. sativa* ($\pi = 0.0037$) [7]. This



severe reduction of genetic diversity in cultivated rice and *Ka/Ks* ratios far below 1.0 (Additional file 3, Table S3) are probably characteristic of very strong selection pressure acting on *OsATM1;1*.

The severe reduction of genetic diversity and predominance of a single common haplotype in both subspecies may be a consequence of strong selection for the promoter region that potentially affects transcription of *OsAMT1;1*. This strong selection for the non-coding regions of *OsATM1;1* in *O. sativa* is supported by the significantly negative values of Tajima's D and Fu and Li's D* tests (Table 1). We found that *O. sativa* specifically carried three polymorphic sites differing from the most wild rice accessions at positions of -22, -175, and -187 that were adjacent to several core domains such as TATA box in the promoter, which is a *cis*-regulatory DNA region that facilitates transcription of a particular gene for functional adaptation. We observed different expression levels of *OsAMT1;1* in several near-isogenic lines that carried the alleles of *OsAMT1;1* from *O. rufipogon* and *O. sativa* in the similar genetic background (Additional file 6, Figure S2). The expression of the *O. rufipogon* alleles in shoot of the near-isogenic line was significantly higher than that of ZS97. However, further investigation is necessary to explain the advantages conferred by the selected allelic variations in *OsAMT1;1* during domestication. Several studies have shown that sequence variations in the promoter region result in differential regulation of transcription. For example, Crawford et al. [39] revealed that the *lactate dehydrogenase-B* (*Ldh-B*) proximal promoter sequence affected transcriptional processes and constituted a phenotypic change that was subject to natural selection by temperature differences. Konishi et al. [40] demonstrated that a single-nucleotide polymorphism in the 5' regulatory region of the QTL for seed shattering in chromosome 1 (*qSH1*) causes loss of seed shattering and was therefore the target of artificial selection during rice domestication. Similarly, the promoter and proximal intergenic region of *tb1* was the target of selection, and a selective sweep ~90 kb in length across this 5' region occurred during maize domestication [14,15].

Furthermore, the LD value was high along a ~100 kb region in *O. sativa*, but fell sharply in *O. rufipogon*. Fu and Li's D* tests were significant for five loci located in a region from -74.1 kb to 27.5 kb. These observations indicate a selective sweep within and nearby *OsAMT1;1* (Figure 4 and Table 2). Many agronomically important genes such as *sh4* and *waxy* were under strong artificial selection leading to a selective sweep around the site of those genes. We infer a similar selective sweep around *OsAMT1;1* caused largely by natural selection for an ammonium transporter.

Reasons for *OsAMT1;1* selection

The common wild rice *O. rufipogon* is found in ditches, pools, and other sites with shallow, standing, or slow-running water. In these natural habitats, the level of nitrogen is extremely low [41]. Nitrogen, either as ammonia or as nitrite, is present only in traces both in water and in the sediment around such habitats. Therefore, a nitrogen uptake system that can manage to obtain nitrogen from such infertile soil was most likely a target of selection during the transition from the wild progenitor to cultivated rice. In addition, until the arrival of nitrogenous fertilizers, the major sources of nitrogen were animal excreta and plant and animal residues, the ultimately usable forms being ammonia, acid amides, and urea. As a result, ammonium was the main form of nitrogen in anaerobic rice paddy environments [42] - these habitats remained unchanged for thousands of years until the 20th century, which saw large-scale application of chemical fertilizers. Even for current rice cultivation in paddy fields, ammonium remains the major form of inorganic nitrogen, making rice almost dependent on ammonium nutrition during a large part of the cropping season. In terms of the efficiency of fertilizer use, ammonium is superior to nitrate in paddy soils [43]. Therefore, a high-affinity ammonia transporter such as *OsAMT1;1* for ammonium uptake in rice [27,28], would have proved advantageous in adapting to domestication and remained under continual selection pressure in low-nitrogen soils.

We note that of 21 annotated genes located in the ~150 kb genomic region around *OsAMT1;1* based on the TIGR website (Additional file 4, Table S4), nine are reported to be involved in plant responses to such biotic and/or abiotic stresses as extreme temperatures, drought, excess salt, and infectious pathogens [44-52]. Rice has adapted to changes in several environmental features (e.g. soil, water, photoperiod, and pathogens). Selection for these genes nearby *OsAMT1;1* rather than *OsAMT1;1* alone could not be excluded as the cause of the selective sweep around the *OsAMT1;1* genomic region. However, inadequate soil nitrogen should be the most widespread selection pressure in operation for both subspecies, irrespective of their locations. We therefore suggest that fixation of the alleles of *OsAMT1;1* common to the two subspecies *indica* and *japonica* is the result of a combined force of natural selection for more than one trait. Unlike the domestication genes, *rc*, *sh4*, and *wx* that were involved in traits strongly preferred by humans [9], the gene *OsAMT1;1* may have been selected as a result of soils deficient in nitrogen.

Evolutionary process of *OsAMT1;1*

In the present study, haplotype H1 is predominant and fixed at *OsAMT1;1* in both subspecies (*indica* and

japonica). This dominant haplotype was also found in *O. rufipogon* albeit in very low frequency - only one *O. rufipogon* accession, namely DWR25 from Malaysia, carried it - suggesting a possibility that haplotype H1 first occurred in one of the subspecies of *O. sativa* and introgressed into the wild rice. We also observed that one *O. rufipogon* accession (DWR27 from China) had only one SNP (T/A) at position -676 compared to H1 (Figure 1). This instance of high diversity of the gene in the wild relatives but uniform outcome because all mutations were synonymous suggests another possibility, namely that the haplotype H1 of *OsAMT1;1* occurred first in one of the accessions of *O. rufipogon* (e.g. DWR27), and underwent strong selection pressure during domestication. A similar pattern was reported in the non-shattering allele (*sh4*) fixed in 208 rice cultivars [53]. Actually, two models, snowballing model and combination model, were previously proposed considering single and multiple origins of cultivated rice, respectively [9]. These two models reconciled the data for *sh4*. These models could also be applied in *OsAMT1;1* as *OsAMT1;1* and *sh4* shared the same features in terms of the nucleotide diversity. However, current sequence data from *OsAMT1;1* and *sh4* might be not strong enough to distinctly support which model is correct. With the rapid development of sequencing technology, more genes similar to *OsAMT1;1* and *sh4* would be discovered in the rice genome. Further cloning and functional characterization of these genes will help test the models of rice domestication.

Additionally, although genetic diversity in *indica* was double or triple of that in *japonica*, genetic diversity in common wild rice *O. rufipogon* around *OsAMT1;1* was more than 30 times higher than that both *indica* and *japonica* rice (Table 1). It is remarkable that genetic diversity in two other nitrogen-related genes decreased gradually from the wild relatives to the cultivated rice, whereas the reduction was extreme in the case of *OsAMT1;1* (Additional file 5, Figure S1). Therefore the selection around *OsAMT1;1* may have occurred before the divergence of *indica* and *japonica*. Several domestication genes such as *rc* and *waxy* found in domesticated forms with their favorable alleles occurred first in the *japonica* cultivars and experienced the selective sweep under strong artificial selection during domestication, causing the alleles to become fixed in all subpopulations of rice [12,54]. *OsAMT1;1* with only one predominant haplotype in both *indica* and *japonica* cultivars even in landraces supports the contention that the gene was first under purifying selection in the wild and then under positive selection during domestication (Additional file 3, Table S3, and Figure 4). However, further investigations using many more samples drawn from

locations across the world are required for more definitive answers.

Conclusions

The ammonium transporter *OsAMT1;1* showed abundant genetic diversity in *O. rufipogon* - diversity that has been dramatically lost in *O. sativa* owing to natural and artificial selection during the domestication of rice. This severe reduction in the extent of variation in nucleotides of *OsAMT1;1* in the cultivated rice (*O. sativa*) extends to a genomic region about 100 kb long, indicating a selective sweep around *OsAMT1;1*. This study demonstrates that the selective sweep was caused by strong selection within or nearby the ammonium transporter during the process of domestication of rice from its wild progenitors to cultivated rice. Understanding the diversity pattern of the ammonium transporter gene in rice not only provides valuable information on the allelic distribution of the nitrogen-uptake gene, but also gives important insights into the process of domestication of cultivated rice. For example, given the alleles fixed at *OsAMT1;1* in *O. sativa*, it is possible to discover novel alleles in wild relatives to broaden the genetic variation for improving the efficiency of nitrogen uptake in plants. It will also be worthwhile to elucidate the pattern of diversity of homologs of AMT1 in other crops or of other nitrogen genes in rice.

Additional material

Additional file 1: Table S1: Details of sources, classification, and haplotypes of accessions of cultivated rice and wild rice used in this study.

Additional file 2: Table S2: Primers used for PCR amplification and sequencing. The targeted genes with positions relative to *OsAMT1;1*, annealing temperature, and predicted product size for amplification are listed.

Additional file 3: Table S3: Ka/Ks test for *OsAMT1;1*. Ka, Ks represent non-synonymous and synonymous substitutions rate respectively.

Additional file 4: Table S4: Summary of 21 annotated genes surrounding *OsAMT1;1*.

Additional file 5: Figure S1: Comparison of nucleotide diversity in *OsAMT1;1* and two other genes related to nitrogen metabolism in *O. rufipogon*, landraces, and elite rice. (n) = no. of the samples assayed in each subgroup.

Additional file 6: Figure S2: Expression of *OsAMT1;1* in paired near-isogenic lines under low nitrogen. NIL1 and NIL2 represent that near-isogenic line carried the *OsAMT1;1* allele from 'ACC10' (*O. rufipogon*) and from 'Nipponbare' (*japonica*), respectively within the same genetic background of 'Zhenshan97' (ZS97). Young seedlings at the two-leaf stage were transferred to a Yoshida nutrient solution with one modification representing low nitrogen (0.15 mM (NH₄)₂SO₄). The nutrient solution was replaced every three days. Seedlings were grown in the nutrient solution for seven days, after which their roots and shoots were harvested separately, frozen in liquid nitrogen, and stored at -70°C until required for RNA isolation. Total RNA was isolated using Trizol reagent (Invitrogen). qRT-PCR (quantitative real-time PCR) was performed using the forward primer 5'-CTGGGGTTGGTGGGTTCA-3' and reverse

primer 5'-CACTTGGTTGTTGCTGTTGGAG-3' for *OsAMT1;1* and the primers 5'-AACCGAGCTGAGGCCCAAGA-3' and 5'-ACGATTGATTAACCGAGTCATGA-3' for rice *ubiquitin* gene, which served as the internal control. Relative expression values are given as means \pm standard error from three biological replications each with three technical repeats, and the *p* value next to each bar represents the results of *t* test between a given NIL and ZS97. Error bars indicate standard error.

Acknowledgements

We thank Drs. X. Wei and Z. Li for providing most of the Chinese samples, and the International Rice Research Institutes for the IRGC and IRIS accessions. Thank the two anonymous reviewers for helpful comments on the manuscript. This study was supported by grants from National Program on Key Basic Research Project, National Special Program for Research of Transgenic Plant of China, and the National Natural Science Foundation of China.

Author details

¹National Key Laboratory of Crop Genetic Improvement, and the College of Plant Science and Technology, Huazhong Agricultural University, Wuhan 430070, PR China. ²School of Plant Biology, and International Centre for Plant Breeding Education and Research, The University of Western Australia, Crawley, WA 6009, Australia.

Authors' contributions

ZD conducted DNA sequencing, analyzed data and drafted the manuscript; CW conducted accessions collection; SC participated in development of the manuscript; SY designed the study, analyzed data, and wrote the paper; all authors read and approval the manuscript.

Received: 28 September 2010 Accepted: 8 March 2011

Published: 8 March 2011

References

1. Ma J, Bennetzen JL: **Rapid recent growth and divergence of rice nuclear genomes.** *Proc Natl Acad Sci USA* 2004, **101**:12404-12410.
2. Panaud CR: **The molecular bases of cereal domestication and the history of rice.** *Biologies* 2009, **332**:267-272.
3. Garris AJ, Tai TH, Coburn J, Kresovich S, McCouch SR: **Genetic structure and diversity in *Oryza sativa* L.** *Genetics* 2005, **169**:1631-1638.
4. McCouch SR, Sweeney M, Li JM, Jiang H, Thomson M, Septiningsih E, Edwards J, Moncada P, Xiao JH, Garris A, Tai T, Martinez C, Tohme J, Sugiono M, McClung A, Yuan LP, Ahn S: **Through the genetic bottleneck: *O. rufipogon* as a source of trait-enhancing alleles for *O. sativa*.** *Euphytica* 2006, **154**:317-339.
5. Tanksley SD, McCouch SR: **Seed banks and molecular maps: unlocking genetic potential from the wild.** *Science* 1997, **277**:1063-1066.
6. Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fledel-Alon A, York TL, Polato NR, Olsen KM, Nielsen R, McCouch SR, Bustamante CD, Purugganan MD: **Genome-wide patterns of nucleotide polymorphism in domesticated rice.** *PLOS Genetics* 2007, **3**(9):1745-1756.
7. Zhu QH, Zheng XM, Luo JC, Gaut BS, Ge S: **Multilocus analysis of nucleotide variation of *Oryza sativa* and its wild relatives: severe bottleneck during domestication of rice.** *Mol Biol Evol* 2007, **24**(3):875-888.
8. Li C, Zhou A, Sang T: **Rice domestication by reducing shattering.** *Science* 2006, **311**:1936-1939.
9. Sang T, Ge S: **Genetics and phylogenetics of rice domestication.** *Current Opinion in Genetics and Development* 2007, **17**:533-538.
10. Zhang LB, Zhu QH, Wu ZQ, Ibarra JR, Gaut BS, Ge S, Sang T: **Selection on grain shattering genes and rates of rice domestication.** *New Phytologist* 2009, **184**(3):708-720.
11. Olsen KM, Caicedo AL, Polato N, McClung A, McCouch S, Purugganan MD: **Selection under domestication: evidence for a sweep in the rice waxy genomic region.** *Genetics* 2006, **173**:975-983.
12. Sweeney MT, Thomson MJ, Cho YG, Park YJ, Williamson SH, Bustamante CD, McCouch SR: **Global dissemination of a single mutation conferring white pericarp in rice.** *PLOS Genet* 2007, **3**(8):1418-1424.
13. Wang RL, Stec A, Hey J, Lukens L, Doebley J: **The limits of selection during maize domestication.** *Nature* 1999, **398**:236-239.
14. Doebley J, Stec A, Hubbard L: **The evolution of apical dominance in maize.** *Nature* 1997, **386**:485-488.
15. Clark RM, Linton E, Messing J, Doebley JF: **Pattern of diversity in the genomic region near the maize domestication gene *tb1*.** *Proc Natl Acad Sci USA* 2002, **101**(3):700-707.
16. Palaisa KA, Morgante M, Williams M, Rafalski A: **Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci.** *The Plant Cell* 2003, **15**:1795-1806.
17. Palaisa K, Morgante M, Tingey S, Rafalski A: **Long-range patterns of diversity and linkage disequilibrium surrounding the maize *Y1* gene are indicative of an asymmetric selective sweep.** *Proc Natl Acad Sci USA* 2004, **101**(26):9885-9890.
18. Wang MY, Siddiqi MY, Ruth TJ, Glass ADM: **Ammonium uptake by rice roots (II. Kinetics of $^{13}\text{NH}_4^+$ influx across the plasmalemma).** *Plant Physiology* 1993, **103**(4):1259-1267.
19. Kronzucker HJ, Glass ADM, Siddiqi MY: **Kinetics of NH_4^+ influx in spruce.** *Plant Physiology* 1996, **110**(3):773-779.
20. Loque D, von Wieren N: **Regulatory levels for the transport of ammonium in plant roots.** *Journal of Experimental Botany* 2004, **55**(401):1293-1305.
21. Li BZ, Mike M, Li SM, Li HY, Zhu SW, Shi WM, Su YH: **Molecular basis and regulation of ammonium transporter in rice.** *Rice Science* 2009, **16**(4):314-322.
22. Kaiser BN, Rawat SR, Siddiqi MY, Masle J, Glass AD: **Functional analysis of an Arabidopsis T-DNA 'knockout' of the high-affinity NH_4^+ transporter *AtAMT1;1*.** *Plant Physiology* 2002, **130**:1263-1275.
23. Salvemini F, Marini AM, Riccio A, Patriarca EJ, Chiurazzi M: **Functional characterization of an ammonium transporter gene from *Lotus japonicus*.** *Gene* 2001, **270**:237-243.
24. Ludewig U, Wilken S, Wu B, Jost W, Obrdlik P, Ei Bakkoury M, Marini AM, Andre B, Hamacher T, Boles E, von Wieren N, Frommer WB: **Homo- and hetero-oligomerization of ammonium transporter-1 NH_4^+ uniporters.** *The Journal of Biological Chemistry* 2003, **278**(46):45603-45610.
25. Loque D, Yuan LX, Kojima S, Gojon A, Wirth J, Gazzarrini S, Ishiyama K, Takahashi H, von Wieren N: **Additive contribution of *AMT1;1* and *AMT1;3* to high-affinity ammonium uptake across the plasma membrane of nitrogen-deficient Arabidopsis root.** *The Plant Journal* 2006, **48**:522-534.
26. Yuan LX, Loque D, Kojima S, Rauch S, Ishiyama K, Inoue E, Takahashi H, von Wieren N: **The organization of high-affinity ammonium uptake in Arabidopsis roots depends on the spatial arrangement and biochemical properties of AMT1-type transporters.** *The Plant Cell* 2007, **19**:2636-2652.
27. Kumar A, Silim SN, Okamoto M, Siddiqi MY, Glass ADM: **Differential expression of three members of the AMT1 gene family encoding putative high-affinity NH_4^+ transporters in roots of *Oryza sativa* subspecies *indica*.** *Plant, Cell and Environment* 2003, **26**:907-914.
28. Sonoda Y, Ikeda A, Saiki S, von Wieren N, Yamaya T, Yamaguchi J: **Distinct expression and function of three ammonium transporter genes (*OsAMT1;1-1;3*) in rice.** *Plant Cell Physiology* 2003, **44**(7):726-734.
29. Sonoda Y, Ikeda A, Saiki S, Yamaya T, Yamaguchi J: **Feedback regulation of the ammonium transporter gene family AMT1 by glutamine in rice.** *Plant Cell Physiology* 2003, **44**(12):1396-1402.
30. Wen WW, Mei HW, Feng FJ, Yu SB, Huang ZC, Wu JH, Chen L, Xu XY, Luo LJ: **Population structure and association mapping on chromosome 7 using a diverse panel of Chinese germplasm of rice (*Oryza sativa* L.).** *Theor Appl Genet* 2009, **119**:459-470.
31. Rogers SO, Bendich AJ: **Extraction of DNA from plant tissue.** *Plant Mol Biol Manual* 1988, **A6**:1-10.
32. Nei M: **Molecular Evolutionary Genetics.** New York: Columbia Univ. Press; 1987.
33. Watterson GA: **On the number of segregating sites in genetical models without recombination.** *Theor Popul Biol* 1975, **7**:256-276.
34. Librado P, Rozas J: **DnaSP v5: A software for comprehensive analysis of DNA polymorphism data.** *Bioinformatics* 2009, **25**(11):1451-1452.
35. Bradbury PJ, Zhang ZW, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES: **TASSEL: Software for association mapping of complex traits in diverse samples.** *Bioinformatics* 2007, **23**(19):2633-2645.
36. Sabeti PC, Reich DE, Higgins JM, Levine HZ, Richter DJ, Schaffner SF, Gabriel SB, Platko JV, Patterson NJ, McDonald GJ, Ackerman HC, Campbell SJ, Altshuler D, Copper R, Kwiatkowski D, Ward R, Lander ES:

- Detecting recent positive selection in the human genome from haplotype structure. *Nature* 2002, **419**:832-837.
37. Goldman N, Yang Z: A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Molecular Biology and Evolution* 1994, **11**:725-736.
 38. Zhang Z, Li J, Zhao XQ, Wang J, Wong GK, Yu J: **KaKs_Calculator: Calculating Ka and Ks through model selection and model averaging.** *Genomics Proteomics Bioinformatics* 2006, **4**:259-263.
 39. Crawford DL, Segal JA, Barnett JL: **Evolutionary analysis of TATA-less proximal promoter function.** *Mol Biol Evol* 1999, **16**(2):194-207.
 40. Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, Yano M: **An SNP caused loss of seed shattering during rice domestication.** *Science* 2006, **312**:1392-1395.
 41. Kariiali E, Kuanar SR, Mohapatra PK: **Individual tiller dynamics of two wild *Oryza* species in contrasting habitats.** *Plant Production Science* 2008, **11**(3):355-360.
 42. Sasakawa H, Yamamoto Y: **Comparison of the uptake of nitrate and ammonium by rice seedlings.** *Plant Physiology* 1978, **62**(4):665-669.
 43. Yoshida S: **Fundamentals of rice crop science.** Philippines: The International Rice Research Institute; 1981.
 44. Keith B, Dong XN, Ausubel FM, Fink GR: **Differential induction of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase genes in *Arabidopsis thaliana* by wounding and pathogenic attack.** *Proc Nati Acad Sci USA* 1991, **88**(19):8821-8825.
 45. Guo Y, Halfter U, Ishitani M, Zhu JK: **Molecular characterization of functional domains in the protein kinase SOS2 that is required for plant salt tolerance.** *The Plant Cell* 2001, **13**:1383-1400.
 46. Vaillau F, Daniel X, Tronchet M, Montillet JL, Triantaphylides C, Roby D: **A R2R3-MYB gene, *AtMYB30*, acts as a positive regulator of the hypersensitive cell death program in plant in response to pathogen attack.** *Proc Nati Acad Sci USA* 2002, **99**(15):10179-10184.
 47. Nagaoka S, Takano T: **Salt tolerance-related protein STO binds to a Myb transcription factor homologue and confers salt tolerance in *Arabidopsis*.** *Journal of Experimental Botany* 2003, **54**(391):2231-2237.
 48. Dunwell JM, Purvis A, Khuri S: **Cupins: the most functionally diverse protein superfamily?** *Phytochemistry* 2004, **65**:7-17.
 49. Gong ZZ, Dong CH, Lee H, Zhu JH, Xiong LM, Gong DM, Stevenson B, Zhu JK: **A dead box RNA helicase is essential for mRNA export and important for development and stress responses in *Arabidopsis*.** *The Plant Cell* 2005, **17**:256-267.
 50. Gonzalez D, Bowen AJ, Carroll TS, Conlan RS: **The transcription corepressor LEUNIG interacts with the histone deacetylase HDA19 and mediator components MED14 (SWP) and CDK8 (HEN3) to repress transcription.** *Molecular and Cellular Biology* 2007, **27**(15):5306-5315.
 51. Liu B, Chen ZY, Song XW, Liu CY, Cui X, Zhao XF, Fang J, Xu WY, Zhang HY, Wang XJ, Chu CC, Deng XW, Xue YB, Cao XF: ***Oryza sativa* Dicer-like4 reveals a key role for small interfering RNA silencing in plant development.** *The Plant Cell* 2007, **19**(9):2705-2718.
 52. Ma K, Xiao JH, Li XH, Zhang QF, Lian XM: **Sequence and expression analysis of the C3HC4-type Ring finger gene family in rice.** *Gene* 2009, **444**:33-45.
 53. Lin ZW, Griffith ME, Li XR, Zhu ZF, Tan LB, Fu YC, Zhang WX, Wang XK, Xie DX, Sun CQ: **Origin of seed shattering in rice (*Oryza sativa* L.).** *Planta* 2007, **226**:11-20.
 54. Yamanaka S, Nakamura I, Watanabe KN, Sato YI: **Identification of SNPs in the waxy gene among glutinous rice cultivars and their evolutionary significance during the domestication process of rice.** *Theor Appl Genet* 2004, **108**:1200-1204.

doi:10.1186/1471-2148-11-61

Cite this article as: Ding et al.: Diversity and selective sweep in the *OsAMT1;1* genomic region of rice. *BMC Evolutionary Biology* 2011 **11**:61.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

