nature genetics

Deletion in a gene associated with grain size increased yields during rice domestication

Ayahiko Shomura[1,5], Takeshi Izawa[2,5], Kaworu Ebana[3], Takeshi Ebitani[4], Hiromi Kanegae[3], Saeko Konishi[2] & Masahiro Yano[3]

The domestication of crops involves a complex process of selection in plant evolution and is associated with changes in the DNA regulating agronomically important traits. Here we report the cloning of a newly identified QTL, *qSW5* (*QTL for seed width on chromosome 5*), involved in the determination of grain width in rice. Through fine mapping, complementation testing and association analysis, we found that a deletion in *qSW5* resulted in a significant increase in sink size owing to an increase in cell number in the outer glume of the rice flower; this trait might have been selected by ancient humans to increase yield of rice grains. In addition, we mapped two other defective functional nucleotide polymorphisms of rice domestication-related genes with genome-wide RFLP polymorphisms of various rice landraces. These analyses show that the *qSW5* deletion had an important historical role in artificial selection, propagation of cultivation and natural crossings in rice domestication, and shed light on how the rice genome was domesticated.

As Charles Darwin said[1], domestication is a good example of evolution, as domesticated species and their corresponding wild species can be easily compared. This should be all the more true for domestication of crops with known genome information[2–5]. However, crop domestication is a complex process of selection in which multiple agronomically important traits have been involved, and few details are known about how it has proceeded[6,7]. Domestication-related genes for various traits have been cloned from several crops—especially cereals such as maize[8,9], wheat[10], barley[11] and rice[12,13]. The domestication traits are associated with grain size, grain number, panicle size, grain quality, flowering time, plant architecture and seed shattering. One important motive behind the domestication of rice was the increase of grain yield, as all accessions of the wild species of Asian rice, *Oryza rufipogon*, have thin panicles with thin grains and relatively low fertility.

To identify the genes involved in the increase in grain yield that occurred during rice domestication, we carried out a QTL (quantitative trait locus) analysis for grain size in an F_2 cross population between Nipponbare (*japonica*) and Kasalath (*indica*) cultivars, which we thought might have distinct domestication histories. We detected several QTLs for grain width and focused on a major one, termed *qSW5* (QTL for rice seed width on chromosome 5), which explained 38.5% of natural variation in the F_2 population (**Fig. 1a** and **Supplementary Fig. 1a** online). Using SL22, a line with substitution of Kasalath chromosome 5 in a Nipponbare genetic background (see URLs section in Methods and **Fig. 1b**), and NIL(*qSW5*), a nearly isogenic line (NIL) that contained around 90 kbp of Kasalath fragments of the *qSW5* region in a Nipponbare background (**Fig. 1c**), for comparison, we first observed the appearance of rice grains (**Fig. 1d,e**) and found that the number of rows of specialized cells with rigidified walls in the upper epidermis—and especially of the outer glume (lemma)—of the rice flower were increased in Nipponbare but not in SL22 (**Table 1**), indicating that the primary cause of the increase in grain width was the increase in size of the outer glumes (**Fig. 1f,g**). The size of the rice glume is one of the determinants of rice endosperm size or grain size[14,15]. We also counted the number of lower epidermis cells inside the rice glumes of SL22 (**Fig. 1f** and **Table 1**) and found that the number was higher in Nipponbare, indicating that the *qSW5* gene may control cell number of the outer glume of the rice flower. As the Nipponbare allele of *qSW5* behaves in a recessive manner in inheritance, it might have acquired a defect during domestication. Fine mapping of *qSW5* using F_3 and F_4 progeny of a F_2 plant, 94BC_3F_2–7 (**Supplementary Fig. 1b**), delimited the functional nucleotide polymorphisms (FNPs) for *qSW5* within a 2,263-bp fragment of Kasalath genomic region (**Fig. 2** and **Supplementary Fig. 1a**). Compared with the corresponding region of Kasalath, the Nipponbare region harbored a 1,212-bp deletion and several SNPs (**Fig. 2b–e**). It is likely that this deletion is the FNP for *qSW5* (**Fig. 2b–e**). We next carried out a complementation test by transforming several Kasalath fragments covering the FNP region into Nipponbare (**Fig. 2c**). Only transformation of an 11.2-kbp fragment covering the deletion region resulted in thin rice grains; thus, we succeeded in cloning the *qSW5* gene (**Fig. 2f** and **Supplementary**

[1]Institute of the Society for Techno-Innovation of Agriculture, Forestry, and Fisheries, 446-1 Ippaizuka, Kamiyokoba, Tsukuba, Ibaraki 305-0854, Japan. [2]Plant Genome Research Unit, National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan. [3]QTL Genomics Research Center, National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan. [4]Toyama Agricultural Research Center, 1124-1, Yoshioka, Toyama 939-8153, Japan. [5]These authors contributed equally to this work. Correspondence should be addressed to T.I. (tizawa@nias.affrc.go.jp).

Received 7 April; accepted 12 May; published online 6 July 2008; doi:10.1038/ng.169

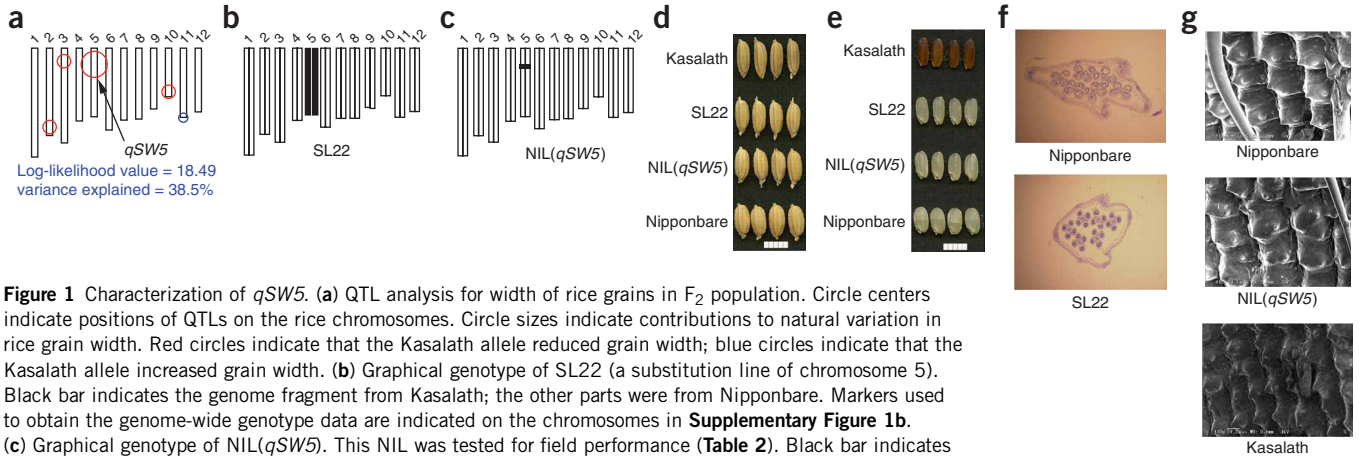

Figure 1 Characterization of *qSW5*. (**a**) QTL analysis for width of rice grains in F_2 population. Circle centers indicate positions of QTLs on the rice chromosomes. Circle sizes indicate contributions to natural variation in rice grain width. Red circles indicate that the Kasalath allele reduced grain width; blue circles indicate that the Kasalath allele increased grain width. (**b**) Graphical genotype of SL22 (a substitution line of chromosome 5). Black bar indicates the genome fragment from Kasalath; the other parts were from Nipponbare. Markers used to obtain the genome-wide genotype data are indicated on the chromosomes in **Supplementary Figure 1b**. (**c**) Graphical genotype of NIL(*qSW5*). This NIL was tested for field performance (**Table 2**). Black bar indicates the fragment from Kasalath; the other parts were from Nipponbare. Markers used to obtain the genome-wide genotype data are indicated on the chromosomes in **Supplementary Figure 1b**. (**d**) Photos of rough rice grains. (**e**) Photos of hulled rice grains without glumes. Grain widths were different because of the glume size restrictions apparent in **d**. (**f**) Transverse sections of rice florets before flowering. Sizes of the outer glumes were distinct. Outer glume of Nipponbare was bigger than that of SL22. (**g**) Scanning electron microscope photos of outer glume surfaces. Nipponbare and NIL(*qSW5*) have stripes of the same width. Kasalath has thinner stripes because of the influence of other, minor QTLs.

Fig. 2 online). We predicted three ORFs in the 11.2-kbp fragment. RT-PCR analysis of these ORFs identified a putative transcribed ORF, termed ORF1, for the *qSW5* gene product (**Fig. 2d**), although the sequence information of the ORF gave no clues as to the biochemical function of the *qSW5* gene (**Supplementary Fig. 3** online). To confirm that the transcript for the ORF was responsible for the phenotypic changes induced by *qSW5*, we transformed an RNAi construct for ORF1 (**Fig. 2d**) into Kasalath. The seed weight of T0 plants was increased in most of the RNAi transgenic lines (**Fig. 2h**), strongly suggesting that ORF1 is the *qSW5* gene product. We next sequenced the PCR-amplified *qSW5* regions of more than 100 rice landraces, including *japonica* and *indica*, and examined the grain width of each cultivar (**Supplementary Table 1** online). Several haplogroups were identified, and the deletion in the Nipponbare allele of *qSW5* was clearly associated with an increase in rice grain width (**Fig. 2g**), suggesting that the deletion was an FNP that might have been selected by ancient humans during rice domestication. From this finding, together with the evidence from the cloning of *qSW5*, we concluded that *qSW5* is a domestication-related gene in rice. Note that because of the complex population structure of rice, the association of DNA change with some trait changes alone was not enough to conclude that *qSW5* is a domestication gene.

Using the NIL(*qSW5*), we further carried out a field test in a paddy field in Japan (**Table 2**). We found that NIL(*qSW5*) showed more than 10% reduced grain yield (**Table 2**), possibly as a result of reduced grain width (**Fig. 1e**). On the other hand, most of the RNAi lines of ORF1 in Kasalath (**Fig. 2d**) produced seeds with increased grain weight (**Fig. 2h**), suggesting the possible use of the defective *qSW5* allele for a breeding program of new *indica* cultivars.

Two more FNPs of domestication-related genes of rice were identified previously. One FNP was found at the junction of the first exon and intron of the *Wx* (*Waxy*) gene, which encodes a granule-bound starch synthase and controls taste and texture of cooked rice grains[16]; natural variations in *Wx* genes have been extensively analyzed in landraces of rice[17,18]. The other one was an SNP found at the RY-repeat *cis*-element in the promoter region of *qSH1* (*QTL for seed shattering in chromosome 1*), which played a critical role in the loss of seed shattering trait in temperate *japonica* group[13]. As defective alleles were somehow selected during domestication in these cases, we could judge in current rice landraces whether the genotypes of these three FNPs were still original types or whether they were the defective types resulting from selection. Therefore, we mapped the genotypes of these defective FNPs in various rice landraces (in total 142 cultivars, including a few modern cultivars; **Supplementary Table 2** online) with the local origins (**Fig. 3a**). We further matched genome-wide RFLP data on the 142 rice landraces in order to elucidate the changes in genome structure during domestication (**Supplementary Figs. 4** and **5** online); each dataset contained one of three RFLP genotypes—Nipponbare, Kasalath or 'other'—for 179 loci distributed genome-wide over the 12 chromosomes[19] (**Supplementary Fig. 4**).

Examination of this genome distance map (**Supplementary Fig. 4**) led us to identify 'heritage' *japonica* landraces of rice—that is, landraces that have all the original alleles for the three genes (**Fig. 3b**)—as these FNPs were defective and had occurred as mutations in the past. Because the same mutation rarely occurred several times within the 10,000 years over which rice was domesticated[6,7,13], we can consider these mutations as single events in rice domestication. First, most tested *indica* landraces carried Kasalath-biased genome structures and all the original alleles, suggesting that the FNPs were not facilitated

Table 1 Phenotypes of SL22

Traits		Nipponbare (%)	SL22 (%)
Grain width (mm)		3.3 ± 0.1 (100)	2.8 ± 0.2 (85)
Grain length (mm)		7.2 ± 0.2 (100)	7.6 ± 0.2 (95)
Circumference (mm)	Lemma	5.7 ± 0.2 (100)	4.9 ± 0.1 (86)
	Palea	2.6 ± 0.2 (100)	2.5 ± 0.1 (96)
Cell row number of upper	Lemma	82.7 ± 2.5 (100)	61.5 ± 2.6 (74)
epidermis	Palea	32.9 ± 2.4 (100)	30.1 ± 2.1 (91)
Cell number in lemma	Upper	106.0 ± 3.6 (100)	79.3 ± 0.9 (75)
	Lower	157.7 ± 5.6 (100)	126.0 ± 2.4 (80)
Cell number in palea	Upper	44.0 ± 3.3 (100)	42.7 ± 3.1 (97)
	Lower	72.3 ± 2.6 (100)	61.3 ± 2.6 (85)

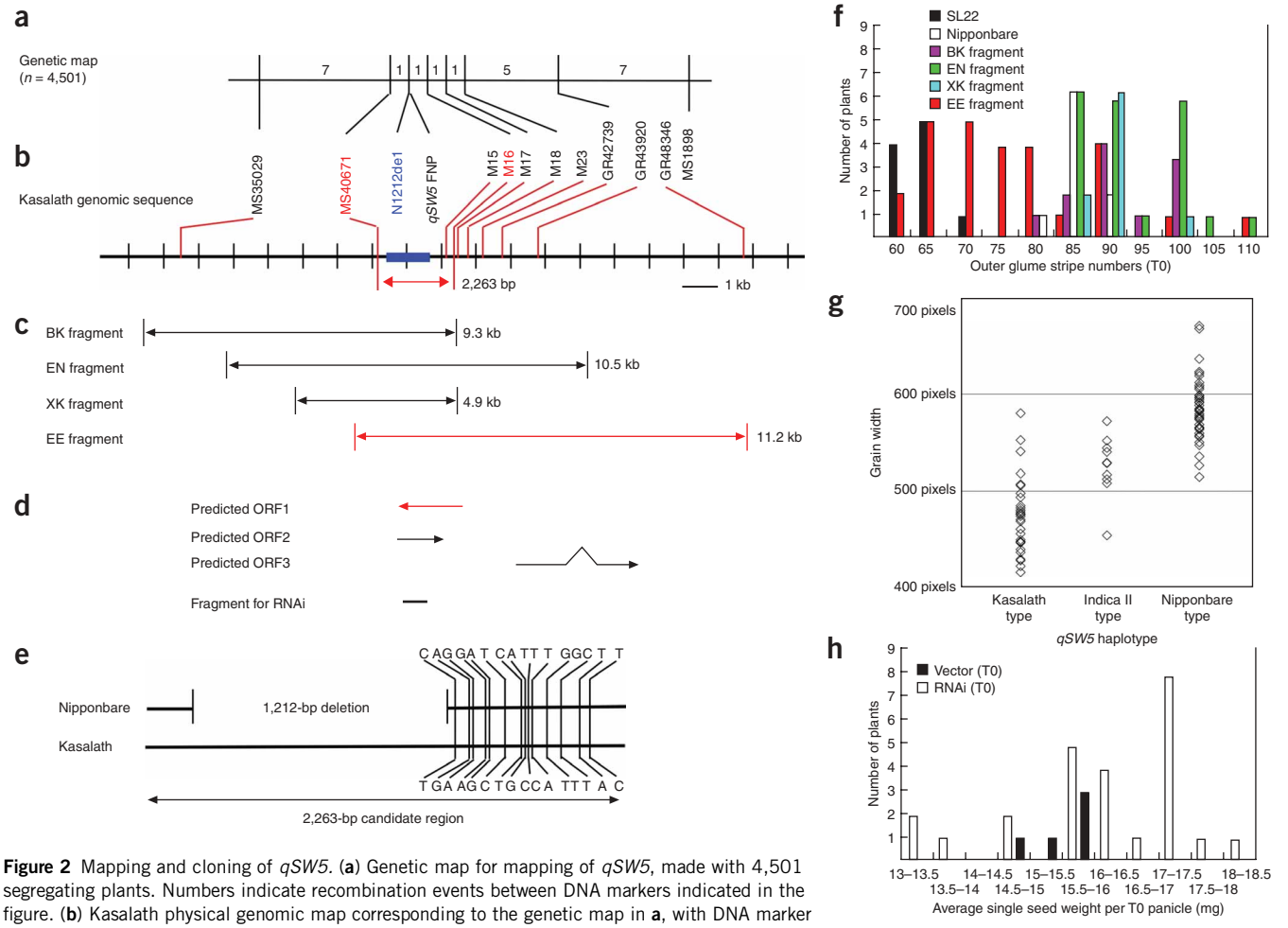

Figure 2 Mapping and cloning of *qSW5*. (**a**) Genetic map for mapping of *qSW5*, made with 4,501 segregating plants. Numbers indicate recombination events between DNA markers indicated in the figure. (**b**) Kasalath physical genomic map corresponding to the genetic map in **a**, with DNA marker positions. *qSW5* FNP was delimited to a 2,263-bp region. (**c**) Four Kasalath genome fragments used for the complementation test. (**d**) Predicted ORFs in the *qSW5* candidate regions. Of the three ORFs, two contained no intron but were long ORFs in the opposite orientation in the same region. ORF1 is the *qSW5* gene product candidate based on the expression analysis (**Supplementary Fig. 3**). (**e**) Polymorphisms between Kasalath and Nipponbare in the candidate region. A major difference was a 1,212-bp deletion. (**f**) Results of the complementation test. Average numbers of stripes in the outer glumes of five florets from each T_0 transgenic plant were measured. Only the 11.2-kb Kasalath EE fragment showed clear complementation. (**g**) Association analysis of more than 100 rice cultivars. Cultivars were categorized into three groups by allele polymorphisms in the *qSW5* FNP. Width of ten seed grains was measured from charge-coupled device (CCD)-captured images, and the average values are plotted on the graph. Landraces and raw data are listed in **Supplementary Table 1**. In total, 44 Nipponbare-type landraces, 30 Kasalath-type landraces, and 10 *indica* type II landraces were examined for this association analysis. (**h**) An RNAi construct was introduced into Kasalath, and most RNAi lines showed increased grain weight in the T_0 generation. The position of the genomic fragment used for the RNAi construct is indicated in **d**.

during the domestication of *indica* rice (**Supplementary Figs. 4** and **5**). In contrast, mapping with all functional alleles of the three domestication-related genes highlighted a local group of *japonica* landraces (**Fig. 3b**). This finding is consistent with the fact that at least two independent domestication processes occurred to form Asian cultivated rice (*O. sativa*): one for *indica*, the other for *japonica*[20–22]. The local origins of the *japonica* landrace group with original alleles of all three genes were mainly distributed around the Philippines, Indonesia and partly Indochina, suggesting that these were the locations where *japonica* domesticated rice originated (**Fig. 3a**,**b**). The 'heritage' landraces identified in this work had similar RFLP patterns in that they were mixtures between Nipponbare and Kasalath genome structures. Most tested loci were biased to either Nipponbare- or Kasalath-type genotypes (**Supplementary Fig. 4**). The Nipponbare-type allele was a majority at many Nipponbare-biased loci of these heritage landraces, whereas the Kasalath-biased loci were somehow mixtures between Kasalath- and Nipponbare-type alleles (**Supplementary Fig. 4**).

Rice landraces with the single mutation of *qSW5* and original alleles in the other domestication-related genes had genome structures slightly different from those of the heritage landraces and were

Table 2 A yield test in a paddy between Nipponbare and NILs

Traits[a]	Nipponbare	NIL(*qSW5-Nip*)[b]	NIL(*qSW5-Kas*)
100 seeds weight (g)	2.19 (100%)	–	1.94 (89%)
Dry plant weight (kg)	2.04	2.34 (100%)	2.18 (93%)
Total grain weight (kg)	1.71	1.78 (100%)	1.56 (87%)
Rough seed weight (kg)	1.38	1.44 (100%)	1.24 (86%)

[a]From a field test in Toyama, Japan in 2004. We tested 40 plants for Nipponbare and each NIL; plant density was 25 cm × 25 cm. [b]A segregated-out line fixed with Nipponbare allele of *qSW5* when the NIL(*qSW5-Kas*) plant was selected from the progeny of a heterologous parent plant. All the genotypes of tested markers were for the Nipponbare allele with NIL(*qSW5-Nip*).

distributed not only in the original areas but also more broadly into areas such as India, China and Japan (**Fig. 3b** and **Supplementary Fig. 5**). Landraces with the *wx* mutation showed genome structures that were similar to, although distinct from, those of the heritage and *qsw5*-defective landraces (**Fig. 3b** and **Supplementary Fig. 5**). The *wx* landraces were also distributed into the broader areas. As we could find only a landrace with the single *wx* mutation in the original local area, the *wx* mutation might have occurred, or have been preferred, outside the original area (**Fig. 3b**). These results indicate that these two mutations occurred independently, being propagated in adaptation to broader local areas in accordance with changes in the corresponding traits. All landraces with defective FNPs of both *qSW5* and *qSH1* were of Chinese origin and had closely related genomic structures (**Fig. 3b**). However, no cultivar had only a defective FNP of *qSH1*, indicating that the *qSH1* FNP has a more recent origin than the *qSW5* FNP and might have occurred as a second mutation in a *qsw5* background (**Fig. 3b**). When we searched for landraces with defective FNPs of both *qSW5* and *Waxy*, we found a marked expansion in both numbers and cultivation areas of the corresponding landraces (**Fig. 3b**). Among the *qsw5 wx* double mutant landraces, the genome structures were largely classified into three local groups: tropical *japonica*, temperate *japonica* and Japanese upland rice, suggesting that these local groups were derived from at least three independent crossings between ancestors of *qsw5*- and *wx*-defective landraces (**Fig. 3c**). For instance, one of these three groups, the *qsw5 wx* double mutant landraces in temperate *japonica*, grow mainly in China and Japan. Furthermore, the genetic combination from a subsequent crossing between *qsh1 qsw5* and *wx qsw5* temperate *japonica* landraces produced landraces with all three defective mutations (*wx qsw5 qsh1* triple mutants), which were favored to produce a major group of the temperate *japonica*

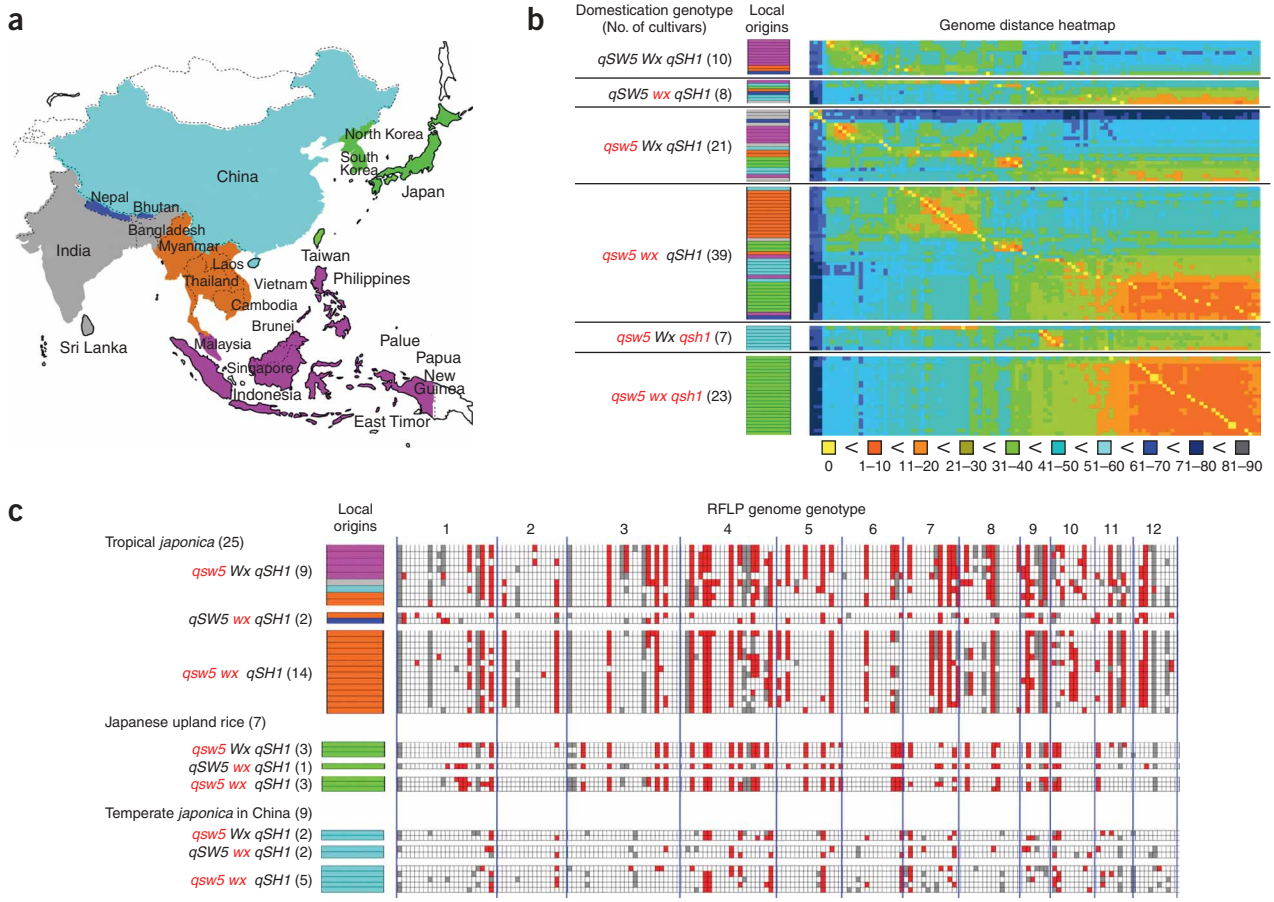

Figure 3 Genome dynamics during rice domestication. (**a**) Geographical origin of rice landraces used in this analysis. Six groups (indicated by different colors) are categorized. (**b**) Heatmap of genomic relationship among tested *japonica* rice landraces according to the increase in the occurrence of defects in three domestication-related genes: *qSW5, Waxy* and *qSH1*. Local origins of each landrace are indicated by color bars, which correspond to the colors on the map in **a**. Numbers in parentheses are numbers of landraces or cultivars with the corresponding genotypes. The original heatmap was clustered by pairwise genome distance calculated from RFLP patterns (**Supplementary Fig. 4**) and is shown in **Supplementary Figure 5**. Colors in the heatmap indicate genome distances, as indicated at the bottom, from other landraces tested. A genome distance of 0 means that the RFLP patterns are identical at 179 loci. The order of genome distances with 142 reference landraces for each landrace is shown at the top line in **Supplementary Figure 5**, in which the reference landraces are aligned according to the relative genome similarities calculated from the RFLP patterns shown in **Supplementary Figure 4**. (**c**) Evidence of multiple crossings to combine *qsw5* and *wx* mutations in several local areas. RFLP patterns were compared among three genotypes of domestication-related genes in three rice local groups: tropical *japonica*, Japanese upland rice and temperate *japonica* in China. Local origins of each landrace are indicated by color bars as in **a**. Each small square indicates the genotype of an RFLP marker of a landrace. RFLP markers are aligned in the physical order of the chromosome numbers shown on the top. In the RFLP analysis red squares indicate Kasalath alleles and white squares indicate Nipponbare alleles. Gray bars indicate other types. Genome patterns were more similar within groups than between groups, suggesting that there had been independent internal crossings in several local areas.

genome, cultivated only in Japan (**Fig. 3b**). These findings show that these three FNPs in landraces could highlight several key events in *japonica* rice domestication (**Supplementary Note** online).

We have demonstrated here that the three domestication-related genes were key genes involved in *japonica* rice domestication and that the FNPs might have been selected according to the propagation of rice cultivation areas. Both expansion of growing areas owing to local adaptation by either *qsw5* or *waxy* mutations and the creation of genetic combinations by independent natural crossing of *qsw5* and *waxy* mutants played critical roles in *japonica* rice domestication (**Supplementary Fig. 6** online).

By using the rice domestication process as a model, we could relatively easily follow how DNA changes were selected and adapted in local areas, as discussed here (**Supplementary Figs. 6** and **7** online). Elucidation of the rice domestication process may provide a unique chance to address some critical questions in evolution, including how the process of evolution linked molecular-level neutral selection and Darwinian selection with phenotypic changes in traits[1,4,23].

METHODS

QTL analysis. We measured the grain widths of 186 F_2 progeny from a cross between Nipponbare and Kasalath for QTL analysis with MAPMAKER/QTL. Five loci were scored with lod scores of more than 2.0. *qSW5* had the highest score and explained 38.5% of the natural variation in this F_2 population.

Mapping of *qSW5*. For the mapping of *qSW5*, we used the advanced backcross progeny BC_3F_1 population to remove other QTL effects. In this population, plants fixed with Nipponbare alleles at the other QTL loci on all chromosomes (except the *qSW5* region on chromosome 5) were selected. Fifty BC_3F_2 plants from the selected BC_3F_1 plants were finally used to map *qSW5* around several markers surrounding *qSW5*, such as W168A, Y1060L and Y1060R. For fine mapping, we used about 4,500 plants from the BC_3F_3 population from a F_2 plant, 94BC_3F_2–7. Seventy plants carrying a recombination in the interval between two flanking PCR markers, E1022 and P433SHC, were phenotyped. We judged the genotype as Kasalath or heterozygous on the basis of results using the progeny in the BC_3F_4 generation. Refer to **Supplementary Table 3** online for the PCR primer information.

Complementation test. Four Kasalath genomic fragments flanking the *qSW5* FNP were subcloned from the BAC clone KBM131B11 into a binary vector, pPZP2H-lac, and transformed into Nipponbare.

NIL(*qSW5*). We used 500 plants of 00GW-145 to select the 02F2-33 plant that contained Kasalath fragments only in the *qSW5* region. We backcrossed the plant with Nipponbare, and then we used the backcrossed progeny for selection for the NIL. In NIL(*qSW5*), the positions for MS40671 and GR42739 were the Kasalath alleles, and MS1898 was the Nipponbare allele. The other marker upstream of MS40671 is not shown in **Figure 2**. Refer to **Figure 2** and **Supplementary Table 3** for further information on PCR primers. After this process of backcrossing and marker selection, at most only a 90-kb Kasalath genome fragment was introgressed into Nipponbare in NIL(*qSW5*). We were able to judge the phenotype of *qSW5* at first glance, as we harvested at least ten panicles with hundreds of seeds per rice plant, and the difference was apparent from the appearance of the bunch of panicles (our patience in deciding to grow the progeny of some key plants the year after we had attempted to fine map *qSW5* and had realized the difficulty in phenotyping the key plants was rewarded). We also carried out detailed observations under a microscope for the fine mapping as needed.

Predictions of *qSW5* gene. We first searched for long ORFs without any introns in the FNP regions, and we identified two candidates (predicted ORF1 and ORF2). Several gene prediction software programs and a tblastn search were further applied to the 11.2-kb Kasalath fragment, and one ORF with an intron (predicted ORF3) was predicted.

Expression analysis. Several specific primer pairs for three predicted ORF (**Supplementary Table 3**) were used to amplify *qSW5* cDNA. We analyzed the PCR products by DNA blot hybridization using the corresponding genomic fragments as probes with an ECL hybridization kit. After the identification of *qSW5*, we carried out quantitative RT-PCR by Taq-Man PCR with an ABI7900 machine. The direction of transcription was estimated by Taq-Man PCR after cycle amplification with either of the single Taq-Man primers.

Making the RNAi lines. A complementary fragment (**Fig. 2d**) of the predicted candidate genes were amplified by PCR, subcloned into pDONR201 and transferred into the RNAi vector pHellsgate8. We weighed dried grains of each RNAi plant to determine the average weight; five seeds were measured for each independent transgenic line.

Clustering of rice landraces on the basis of RFLP patterns. The numbers of identical RFLP genotypes among 179 tested loci distributed over all 12 chromosomes in a pair of rice landraces (or cultivars) were determined as the genome distance between the landraces. We analyzed these pairwise genome distances by the pvclust clustering software in an R package for hierarchical clustering. The dendrogram obtained was arranged by eye to reduce inconsistency.

First, rice cultivars in the WRC rice core collection[19] were mapped with domestication FNPs. The results led us to focus on *japonica* landraces. In total, 142 landraces (including more than 100 *japonica* landraces from among a previously tested group of rice landraces[19]) were analyzed. One cultivar, Danyu1, was not assigned properly because its genome structure was a hybrid between *japonica* and *indica*.

Haplotype analysis. For *qSH1*, published primers were again used[13]. For the other genes, we sequenced PCR products amplified with the primers listed in the **Supplementary Table 3**. For *qSW5*, haplotypes in the WRC core collection and the FNP among 142 landraces were examined. For *Waxy*, the FNP was examined by PCR primers. For *Gn1a*, the exon 1 and exon 3 regions were sequenced among the 142 landraces (or cultivars). For *sh4*, FNP regions in several lines of *O. rufipogon* were sequenced.

URLs. Rice Genome Resource Center Seed stock, http://www.rgrc.dna.affrc.go.jp/stock.html

Accession codes. GenBank: Kasalath *qSW5* gene, AB433345.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank H. Kanamori of Institute of the Society for Techno-Innovation of Agriculture, Forestry and Fisheries and T. Matsumoto of National Institute of Agrobiological Sciences for genomic sequencing of the *qSW5* region of Kasalath, Y. Kojima for RFLP data production, and K. Ono for Kasalath transformation. M.Y. was supported by MP1113 (Integrated research project for plant, insect and animal using genome technology) and T.I. has been supported by GD2008 (Integrated research project for plant, insect and animal using genome technology) and QTL5001 (Genomics for Agricultural Innovation) of the Ministry of Agriculture, Forestry and Fisheries of Japan.

AUTHOR CONTRIBUTIONS

A.S. performed most of the experiments. S.K. helped A.S. with the experiments and carried out qRT-PCR expression analysis. H.K. performed the original QTL analysis with the F_2 population. T.E. field-tested NIL(*qSW5*). K.E. provided genome-wide RFLP data on rice landraces. M.Y. directed the QTL analysis, material production and fine mapping of *qSW5*. T.I. directed the research, designed the experiments for all the other parts and analyzed the FNPs with genome data, and wrote the manuscript. All authors contributed to improve the manuscript.

Published online at http://www.nature.com/naturegenetics/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Darwin, C. *The Variations of Animals and Plants under Domestication* (D. Appleton, New York, 1857).
2. Wright, S.I. *et al.* The effects of artificial selection on the maize genome. *Science* **308**, 1310–1314 (2005).

3. Doebley, J.F., Gaut, B.S. & Smith, B.D. The molecular genetics of crop domestication. *Cell* **127**, 1309–1321 (2006).
4. Roos-Ibarra, J., Morell, P.L. & Gaut, B.S. Plant domestication, a unique opportunity to identify the genetic basis of adaptation. *Proc. Natl. Acad. Sci. USA* **104**, 8641–8648 (2007).
5. Dubcovsky, J. & Dvorak, J. Genome plasticity: a key factor in the success of polyploidy wheat under domestication. *Science* **316**, 1862–1866 (2007).
6. Sang, T. & Ge, S. The puzzle of rice domestication. *J. Integr. Plant Biol.* **49**, 760–768 (2007).
7. Sweeney, M. & McCouch, S. The complex history of the domestication of rice. *Ann. Bot. (Lond.)* **100**, 951–957 (2007).
8. Doebley, J., Stec, A. & Hubbard, L. The evolution of apical dominance in maize. *Nature* **386**, 485–488 (1997).
9. Wang, H. *et al.* The origin of the naked grains of maize. *Nature* **436**, 714–719 (2005).
10. Simons, K.J. *et al.* Molecular characterization of the major wheat domestication gene Q. *Genetics* **172**, 547–555 (2006).
11. Komatsuda, T. *et al.* Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proc. Natl. Acad. Sci. USA* **104**, 1424–1429 (2007).
12. Li, C., Zhou, A. & Sang, T. Rice domestication by reducing shattering. *Science* **311**, 1936–1939 (2006).
13. Konishi, S. *et al.* An SNP caused loss of seed shattering during rice domestication. *Science* **312**, 1392–1396 (2006).
14. Hong, S.K., Kitano, H., Satoh, H. & Nagato, Y. How is embryo size genetically regulated in rice? *Development* **122**, 2051–2058 (1996).
15. Song, X.-J., Huang, W., Shi, M., Zhu, M.-Z. & Lin, H.-X.A. QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat. Genet.* **39**, 623–630 (2007).
16. Isshiki, M. *et al.* Naturally occurring functional allele of the rice waxy locus has a GT to TT mutation at the 5′ splice site of the first intron. *Plant J.* **15**, 133–138 (1998).
17. Olesen, K.M. & Purugganan, M.D. Molecular evidence on the origin and evolution of glutinous rice. *Genetics* **162**, 941–950 (2002).
18. Olsen, K.M. *et al.* Selection under domestication: evidence for a sweep in the rice *Waxy* genomic region. *Genetics* **173**, 975–983 (2006).
19. Kojima, Y., Ebana, K., Fukuoka, S., Nagamine, T. & Kawase, M. Development of an RFLP-based rice diversity research set of germplasm. *Breed. Sci.* **55**, 431–440 (2005).
20. Cheng, C. *et al.* Polyphyletic origin of cultivated rice: based on the interspersion pattern of SINEs. *Mol. Biol. Evol.* **20**, 67–75 (2003).
21. Ma, J. & Bennetzen, J.L. Rapid recent growth and divergence of rice nuclear genomes. *Proc. Natl. Acad. Sci. USA* **101**, 12404–12410 (2004).
22. Vitte, C., Ishii, T., Lamy, F., Brar, D. & Panaud, O. Genomic paleontology provides evidence for two distinct origins of Asian rice (*Oryza sativa* L.). *Mol. Gen. Genet.* **272**, 504–511 (2004).
23. Ehrenreich, I.M. & Purugganan, M.D. The molecular genetic basis of plant adaptation. *Am. J. Bot.* **93**, 953–962 (2006).