MOLECULAR ECOLOGY

Molecular Ecology (2015) **24**, 3329–3344 doi: 10.1111/mec.13256

More than one way to evolve a weed: parallel evolution of US weedy rice through independent genetic mechanisms

XINSHUAI QI,*[1] YAN LIU,†[1] CYNTHIA C. VIGUEIRA,*‡ NELSON D. YOUNG,§ ANA L. CAICEDO,§ YULIN JIA,† DAVID R. GEALY† and KENNETH M. OLSEN*

*Department of Biology, Washington University, St. Louis, MO 63130, USA, †Dale Bumpers National Rice Research Center, USDA-ARS, Stuttgart, AR 72160, USA, ‡Department of Biology, High Point University, High Point, NC 27268, USA, §Department of Biology, University of Massachusetts, Amherst, MA 01003, USA

Abstract

Many different crop species were selected for a common suite of 'domestication traits', which facilitates their use for studies of parallel evolution. Within domesticated rice (*Oryza sativa*), there has also been independent evolution of weedy strains from different cultivated varieties. This makes it possible to examine the genetic basis of parallel weed evolution and the extent to which this process occurs through shared genetic mechanisms. We performed comparative QTL mapping of weediness traits using two recombinant inbred line populations derived from crosses between an *indica* crop variety and representatives of each of the two independently evolved weed strains found in US rice fields, strawhull (S) and blackhull awned (B). Genotyping-by-sequencing provided dense marker coverage for linkage map construction (average marker interval <0.25 cM), with 6016 and 13 730 SNPs mapped in F_5 lines of the S and B populations, respectively. For some weediness traits (awn length, hull pigmentation and pericarp pigmentation), QTL mapping and sequencing of underlying candidate genes confirmed that trait variation was largely attributable to individual loci. However, for more complex quantitative traits (including heading date, panicle length and seed shattering), we found multiple QTL, with little evidence of shared genetic bases between the S and B populations or across previous studies of weedy rice. Candidate gene sequencing revealed causal genetic bases for 8 of 27 total mapped QTL. Together these findings suggest that despite the genetic bottleneck that occurred during rice domestication, there is ample genetic variation in this crop to allow agricultural weed evolution through multiple genetic mechanisms.

Keywords: agricultural weeds, de-domestication, *Oryza sativa*, parallel evolution, QTL mapping, weedy rice

Received 20 March 2015; revision received 26 May 2015; accepted 27 May 2015

Introduction

Parallel evolution can be defined as independent origins of the same trait in closely related lineages (Conte *et al.* 2012). Studies of parallel evolution have emerged as an important approach for understanding the genetic mechanisms that underlie adaptation and the potential constraints on this process (Schluter *et al.* 2004; Stern 2013). Similar phenotypes may evolve independently through changes at orthologous genes, a phenomenon termed molecular convergence (Lenser & Theißen 2013), or potentially through altogether different genetic and developmental pathways (Arendt & Reznick 2008; Manceau *et al.* 2010; Losos 2011). For very closely related lineages, parallel evolution is generally expected to occur through shared genetic mechanisms simply as a

Correspondence: Kenneth M. Olsen, Fax: +1 314-935-4432; E-mail: kolsen@wustl.edu

[1]The first two authors contributed equally to this study.

function of their similar genetic make-up (Lenser & Theißen 2013). Studies examining the molecular basis of parallel evolution at specific candidate genes in crops have largely borne out this prediction (Lenser & Theißen 2013; Olsen & Wendel 2013b). However, such candidate gene studies may be inherently biased towards detecting evidence of shared genetic bases, and the role of shared genetic mechanisms is less clear when parallel evolution involves more complex quantitative trait variation. This is especially true for domesticated plant species, where, with the exception of flowering time (Olsen & Wendel 2013a), the parallel evolution of complex, polygenic traits remains largely uncharacterized (Fuller *et al.* 2014).

Domestication has long been recognized as a useful system for studying the process of evolution (Darwin 1859; Larson *et al.* 2014). Many crop species have undergone selection for a suite of shared 'domestication syndrome' traits (Hammer 1984; Harlan 1992), making them particularly appropriate for examining the genetics of parallel evolution. In cereal crops such as rice and maize, shared domestication traits include the loss of seed shattering and dormancy, erect plant architecture, uniform flowering and maturation in the field, and increased resource allocation to grains. Because of their economic importance, many cereal crop species have been a major focus of genomic characterizations and genome-enabled resource development (Morrell *et al.* 2011; Olsen & Wendel 2013b), which further facilitates their utility for studying the genetic mechanisms underlying parallel evolution.

Recently, the weedy relatives of crop species have been recognized as a valuable complement to domesticates for examining the genetics of adaptation in the agricultural context (Kane & Rieseberg 2008; Ellstrand *et al.* 2010; Vigueira *et al.* 2013b). Weedy relatives are close relatives of cultivated crop plants (in some cases, direct feral descendants) that aggressively outcompete their domesticated relatives. Unlike crops, where domestication traits are favoured that facilitate human cultivation, weedy relatives are characterized by traits that allow them to escape detection and eradication in crop fields and that confer a competitive growth advantage over crops. While they are adapted to crop fields and often restricted to such habitats, weedy relatives possess many traits that are more similar to wild species than domesticates; these include persistent seed dormancy, asynchronous reproduction and highly shattering seed, all features that promote their survival and proliferation as agricultural weeds (Vigueira *et al.* 2013b). As with domestication syndrome traits in cultivated crops, the repeated emergence of weed-adaptive traits in distinct lineages of weedy relatives provides a useful model for studying parallel evolution (Basu *et al.* 2004; Gross *et al.* 2010a; Thurber *et al.* 2011; Vigueira *et al.* 2013b). However, even less is known about the genetic basis of parallel complex trait evolution in weedy crop relatives than in crop species.

Weedy rice (*Oryza sativa* L.), also known as red rice, is a conspecific weedy relative of rice that is well suited for examining the parallel evolution of weed-adaptive traits. Weedy rice infests rice fields worldwide and has arisen multiple times, often through parallel evolution from different cultivated rice varieties (Cao *et al.* 2006; Akasaka *et al.* 2009; Chung & Park 2010; Reagon *et al.* 2010; Grimm *et al.* 2013; Song *et al.* 2014). As the conspecific weed of a genomic model species, the genetic basis of weediness traits can be examined in weedy rice with the benefits of an annotated reference genome and an abundance of functionally characterized domestication genes (Olsen *et al.* 2007; Reagon *et al.* 2010; Vigueira *et al.* 2013a). In addition, its selfing mating system and annual life history facilitate the development of advanced-generation recombinant inbred lines (RILs) for genetic mapping of weediness traits.

In the United States, the southern Mississippi valley is the primary region of rice production, and most cultivars grown there are *tropical japonica* varieties. Two major strains of weedy rice predominate in this region, which are largely distinguishable by grain morphology: blackhull awned (BHA) and strawhull awnless (SH). These appear to have originated through two independent 'de-domestications' from Asian *aus* and *indica* rice varieties, respectively (Londo & Schaal 2007; Reagon *et al.* 2010), which are genetically distinct from *tropical japonica* rice. As neither *aus* nor *indica* varieties were historically grown commercially in the United States, the presence of BHA and SH weeds in North America most likely reflects accidental introductions of weeds that evolved in Asia rather than weed origins in North America (Reagon *et al.* 2010). The independent evolution of these two strains provides a prime opportunity to study the parallel evolution of weediness.

BHA and SH weeds share a number of weediness traits that distinguish them from domesticated rice and are characteristic of wild *Oryza* species. These include highly shattering seed; proanthocyanidin pigmentation of the pericarp (bran) (the source of the weed's common name, red rice); and persistent seed dormancy, a trait directly associated with pericarp pigmentation (Gu *et al.* 2011). The weeds are also characterized by rapid growth and competitive nutrient uptake (Burgos *et al.* 2006), aiding their persistence and proliferation in rice fields. The most noticeable differences between the two US weed strains are in grain characteristics, with SH weeds more closely resembling domesticated rice in that they lack awns and dark-hull pigmentation.

Biparental QTL (quantitative trait locus) mapping is a widely used approach in quantitative genetic studies, and it is highly applicable for studies of parallel evolution (Paterson 2002; Wood *et al.* 2005; Arendt & Reznick 2008). The effectiveness of this approach for fine-scale genetic mapping is dependent on marker density across the genome and the number of generations (and thus recombination events) following the initial parental cross (Beavis 1998; Takuno *et al.* 2012). Next-generation sequencing technology has greatly improved the ease with which dense, genomewide markers can be obtained for genetic mapping. Compared to more traditional molecular markers (e.g. SSRs, AFLPs), the recently developed genotyping-by-sequencing (GBS) strategy (Elshire *et al.* 2011) can yield thousands of genomewide SNPs at a reasonable cost and within a relatively short time frame. This dense marker coverage makes it possible to localize QTL within a small enough genomic region that candidate genes could in principle be recognized without the need for additional fine mapping. GBS-enabled QTL mapping should be particularly effective in rice and its relatives, given the relatively compact *O. sativa* genome (~430 Mb), the availability of an annotated reference genome and the abundance of functionally annotated genes for this species (Meyer & Purugganan 2013; Olsen & Wendel 2013b).

In this study, we used GBS-based QTL mapping in advanced-generation RILs derived from two weed × crop (*indica*) crosses to examine the genetic basis of weediness traits in SH and BHA weedy rice strains. The two biparental crosses analysed here were previously examined using an independently derived set of F_2 samples that were genotyped with SSR markers (54 and 72 loci for SH and BHA populations, respectively) (Thurber *et al.* 2013). This analysis is based on >6000 SNPs per mapping population and trait mapping in the F_5–F_7 generations, with the inclusion of several traits not examined in the earlier study; thus, it has the potential to provide highly increased resolution on the genetic architecture of weediness traits. This analysis includes quantitative traits that are likely under polygenic control (e.g. flowering time, inflorescence morphology, shattering), as well as discrete traits where individual genes may control a large proportion of observed phenotypic variation (e.g. grain and hull pigmentation). Molecular variation at functionally characterized candidate genes has previously been shown to underlie some traits in weedy rice, including hull colour (controlled by *Bh4*; Zhu *et al.* 2011; Vigueira *et al.* 2013a), pericarp pigmentation (controlled by *Rc*; Sweeney *et al.* 2006; Gross *et al.* 2010b) and flowering time (controlled by *Hd1*; Takahashi *et al.* 2009; Thurber *et al.* 2014). These candidate genes can thus serve as positive controls for testing the accuracy with which GBS-enabled QTL mapping can pinpoint the genomic location of causal genetic variation.

The objectives of this research were to (i) identify the QTL underlying weed-associated traits in SH and BHA US weedy rice; (ii) assess the extent to which the parallel evolution of weediness is determined by the same underlying QTL; and (iii) use known candidate gene variation to test the accuracy and precision of GBS-SNP markers for fine-scale genetic mapping. We find that while some genes/QTL have played a role in both weed evolution events, many QTL are not shared between the mapping populations, which suggests that multiple genetic routes to weediness are possible from domesticated rice.

Materials and methods

Plant materials

Two recombinant inbred line (RIL) mapping populations were generated through single-seed descent following the cross of a Taiwanese *indica* rice variety, Dee Geo Woo Gen (DGWG; PI 279131), with each of two weedy rice ecotypes, SH strain AR-2001-1135 (PI 653435) and BHA strain MS-1996-9 (PI 653419). All of the named cultivars and entries identified using the 'PI' prefix in this study can be accessed in GRIN (2015). Crosses and F_2 lines were developed in 2007–2009 under growth chamber conditions at the University of Massachusetts (Amherst, MA); further line advancement was performed under controlled greenhouse conditions in 2007–2012 at the USDA-ARS Dale Bumpers National Rice Research Center (Stuttgart, AR). The two mapping populations are referred to below as the S (for strawhull) and B (for blackhull) populations. A total of 175 and 224 RILs from the S and B mapping populations were used in genotyping (F_5 generation) and phenotyping (F_6 and F_7 generations).

Trait evaluation

Phenotypic data for seven traits were collected in the summers of 2012 (F_6 RILs) and 2013 (F_7 RILs), either in field plots at the Dale Bumpers National Rice Research Center (Stuttgart, AR) or, for grain characteristics, at Washington University (St. Louis, MO). The following traits were assessed: seedling emergence date, heading date (flowering time), panicle length, degree of shattering, hull colour, awn length and pericarp colour. Field plantings of RILs used a randomized complete block design with three replicate blocks. For each replication, several seeds per RIL were sown in one drop using a standard hill drop planting method. Fields were flooded when the seedlings reached the 4- to 7-leaf

stage. Five seedlings were separated and transplanted into a single row after flooding. Plants were spaced 30 cm apart within and 61 cm between rows. Each replicate block thus had five plants per row for each RIL, contained within a bay measuring ~40 m by ~16 m. The three parental lines were included in each replication in a randomly selected location, along with four standard cultivars (Zhe733, Purple Marker, Wells and Francis) and two weedy red rice strains (StgS [PI 653423] and StgB [PI 653422]) used as internal phenotyping controls. A hurricane-related storm in the summer of 2012 caused substantial damage to the field plots; therefore, data for only four traits (heading date, panicle length, shattering and awn length) were collected from the F_6 generation.

Emergence date was recorded as the number of days from sowing to 50% seedling emergence. Heading date was calculated as the number of days from sowing to ≥50% of plants within a line having one or more tillers with an inflorescence at anthesis. Panicle length was calculated as the average for two panicles per plant of the length from base (neck) to tip. Shattering was measured in grams of force required to remove individual seeds from the panicle using digital force gauge (Model DS2, IL, USA), based on the protocol described by Thurber *et al.* (2010); shattering measurements were calculated as the average from 10 mature seeds per panicle collected at 30 days after heading, with two panicles tested per RIL per field replication. Awn length was recorded on a scale ranging from absent (0), to some seeds with short awns (1), all seeds with short awns (5), some seeds with long awns (7) and all seeds with long awns (9); hull colour was scored on a scale ranging from straw (0), to gold (1), brown-spotted (2), brown-furrowed (3), solid brown (4), reddish-to-light purple (5), purple-spotted (6), purple-furrowed (7), solid purple (8) and solid black (9) (IRRI 1980). Pericarp colour was scored on a scale ranging from white (0), to light red (3), red (5), or dark red (7).

Genotyping

Total genomic DNA was extracted from F_5 RILs using QIAamp DNA Mini Kits with a modified plant tissue extraction protocol (Qiagen, MD, USA) and digested using *Ape*KI. Genotyping-by-sequencing (GBS) libraries were constructed at Cornell University's Genomic Diversity Facility using the protocol described by Burgos *et al.* (2014). Both libraries were run twice on an Illumina HiSeq platform. Raw data were handled using the standard *TASSEL* GBS pipeline. We filtered out reads with any of the following characteristics: (1) 'N' in the first 72 bases; (2) no match to any of the barcodes used in the study; or (3) tags with fewer than 5 reads. SNP calls were made based on the published *Oryza sativa* 'Nipponbare' reference *japonica* genome sequence, build MSU 6.0 (available online at ftp://ftp.plantbiology.msu.edu/pub/data/Eukaryotic_Projects/o_sativa/annotation_dbs/pseudomolecules/version_6.0/) using Burrows–Wheeler alignment (BWA). SAMConverter was then employed to convert SAM format to TagsOnPhysicalMap (TOPM) file. SNPs and small indels were identified from each tag alignment and recorded in a HapMap format. SNPs with >10% missing data, SNPs with >15% heterozygosity, individuals with >95% missing data and monomorphic sites with missing data were all removed before further analysis.

Genetic map construction and QTL mapping

Genetic positions of SNPs were calculated using cM Converter (http://mapdisto.free.fr/cMconverter) based on their physical positions in MSU 6.0. Linkage maps for each cross were generated using the *R/QTL* package (Broman *et al.* 2003). GBS data were transferred from HapMap format to *R/qtl* format using Perl scripts before the initial analysis. For each phenotype, QTL were analysed using both single-marker analysis (SMA) and composite interval mapping (CIM), and both mapping methods were performed using Haley–Knott (HK) regression (Haley & Knott 1992) and multiple imputation (IMP) (Sen & Churchill 2001). There is currently no consensus as to which combination of these analyses is best for QTL mapping, and we applied all four combinations (SMA/HK, SMA/IMP, CIM/HK and CIM/IMP) as a conservative approach (K. Broman, personal communication). Haley–Knott (HK) regression is a widely used approach for standard interval mapping, while multiple imputation (IMP) is a computationally more intensive method that deals with missing data by imputing genotypes (Broman & Sen 2009). The SMA method assesses each point along the chromosome individually as a potential QTL; the CIM method considers other nearby markers as covariates. CIM performs well in reducing residual variation, identifying separate linked QTL and identifying interactions among QTL (Broman & Sen 2009). For traits where data were included from both 2012 and 2013, the 2 years' data were analysed separately.

For each of the four QTL mapping analyses (SMA/HK, SMA/IMP, CIM/HK and CIM/IMP), conditional genotype probabilities were calculated using the *calc.genoprob* function with *Kosambi* map function, maximum distance 1 cM, and a genotyping error rate of 0.01. SMA and CIM were performed using *scanone* and *cim* functions, respectively. To obtain a genomewide significance threshold, a permutation test was carried out for each data set (n = 1000). The 95% confidence interval of

each detected QTL was calculated using the *bayesint* function. The nearest marker to the LOD peak of each QTL was selected using the *find.marker* function. Allele effect was determined using *plotPXG* function. QTL interaction effects were estimated using the *addint* function. For those phenotypes with more than one QTL detected, the percentage contribution (R^2) for each QTL and related additive QTL pairs were calculated in a multiple QTL model. Statistical significance of individual QTL and additive effects were also calculated. QTL and QTL-interaction combinations with nonsignificant *P*-values were excluded from the multiple QTL model using 'drop one QTL ANOVA test' in the *fitqtl* function. All other parameters in the above analyses were performed using default settings. All *R/qtl* analyses were run on the High Performance Computer platform in Biology Department at Washington University.

For QTL identified through each of the four methods, confidence intervals for corresponding genomic regions were obtained by combining the ranges of all four intervals. Physical positions of QTL were determined by mapping to the most recent rice genome annotation release (version 7.0: IRGSP-1.0, referred to as MSU 7 hereafter), using the Assembly converter tool on Gramene (http://www.gramene.org/). Where possible, we mapped the genomic locations of genetic markers used in previous QTL studies of weedy rice onto the rice reference genome (Kawahara *et al.* 2013) for comparison with our results.

Potential candidate genes located within 0.5 Mb of QTL map intervals (defined as the genomic region encompassed by the confidence intervals of the four combined mapping methods) were identified using the Rice Genome Annotation Project website (RGAP; http://rice.plantbiology.msu.edu/). These included the flowering time genes *Hd1*, *Hd3a*, *DTH7* and *DTH8* (Takahashi *et al.* 2009; Wei *et al.* 2010; Takahashi & Shimamoto 2011; Fujino *et al.* 2013; Gao *et al.* 2014; Thurber *et al.* 2014); *Bh4*, controlling hull colour (Zhu *et al.* 2011); *An-1*, controlling awn length (Luo *et al.* 2013); and *Rc*, controlling pericarp colour (Sweeney *et al.* 2006). Of these seven genes, three of them (*Bh4*, *Hd1* and *Rc*) have previously been found to contain molecular variation associated with weedy rice trait variation (Gross *et al.* 2010b; Vigueira *et al.* 2013a; Thurber *et al.* 2014), while the remaining four have not been previously examined in weedy rice. For each candidate gene, we tested whether molecular variation between the parental lines supported the conclusion that the gene controlled observed phenotypic variation. Where published sequences were not already available for the parental lines, we used PCR and direct Sanger sequencing to obtain DNA sequence haplotypes, following previously described protocols (Vigueira *et al.* 2013a).

Results

Phenotypic variation

Seven phenotypes were evaluated in the F_7 generation of the two weed × crop RIL populations; data for heading date, panicle length, shattering and awn length were also included from the F_6 generation (Table 1). Both of the weed strains differed significantly from the parental crop accession in emergence date, shattering and pericarp colour; only BHA differed from the crop in hull colour and awn length; and only SH differed significantly from the crop in heading date and panicle length (Fig. S1, Supporting information; Table 1).

Table 1 Phenotypic values for parental lines and mean values for F_6 (2012) and F_7 (2013) RILs in the S and B mapping populations. Asterisks indicate significant differences between crop and weed parents ($P \leq 0.05$)

	S population			B population			
	Parents			Parents			
	AR-2001-1135	DGWG	RIL Mean (SD)	MS-1996-9	DGWG	RIL Mean (SD)	Units
Emergence date 2013	9.7*	11.4	10.27 (0.77)	9.7*	11	10.98 (0.85)	Days
Heading date 2012	77*	83	82.23 (10.47)	91.1	87.3	86.07 (8.28)	Days
Heading date 2013	73.1*	84.7	78.62 (9.77)	85.9	84.7	83.88 (7.28)	Days
Panicle length 2012	24.5*	20.1	23.16 (1.83)	22.1	20.1	22.02 (2.72)	cm
Panicle length 2013	25.9*	23.5	24.77 (2.04)	22	23.5	23.37 (3.23)	cm
Shattering 2012	0*	80.94	39.21 (30.92)	33.2*	80.94	43.47 (32.37)	g
Shattering 2013	0*	82.3	36.05 (27.06)	22.17*	82.3	46.58 (29.98)	g
Hull colour 2013	0	0	0.25 (0.69)	9*	0	4.87 (3.47)	–
Awn length 2012	–	–	–	9	0	5.37 (3.7)	cm
Awn length 2013	–	–	–	9	0	4.51 (3.97)	cm
Pericarp colour 2013	5*	0	3.25 (2.55)	5*	0	1.86 (2.27)	–

Linkage map construction

After data filtering, 6016 and 13 730 SNPs were detected in the S and B populations, respectively. Given that *indica* rice is more closely related to SH weeds than to BHA weeds (Reagon *et al.* 2010), the lower SNP number for the S population is not unexpected. Total genetic map lengths are 1522.91 cM (S population) and 1528.66 cM (B population). These values are broadly consistent with those obtained previously for the S and B populations using independently generated F_2 lines and 126 SSR markers (Thurber *et al.* 2013; 1577 cM for the S population, 1687 cM for the B population), as well as for other mapping populations derived from different weed–crop crosses (e.g. 1423 cM, Mispan *et al.* 2013; 1410 and 1574 cM, Subudhi *et al.* 2014).

The average genetic distance between adjacent SNPs is 0.253 ± 0.067 cM and 0.111 ± 0.031 cM in the S and B populations, respectively. These values indicate a level of genetic resolution that is approximately two orders of magnitude greater than that of previous studies of US weedy rice mapping populations (Thurber *et al.* 2013: 29.8 cM for the S population, and 24.1 cM for B population; Mispan *et al.* 2013: 13 ± 7 cM). Linkage maps for the S and B populations are presented in Fig. 1. Markers are well distributed across the genome, yielding linkage groups that are generally proportional to chromosome sizes (MSU 7.0).

Trait mapping

Across the two mapping populations, a total of 27 statistically supported QTL (11 in the S population, 16 in the B population) were identified for the seven mapped phenotypes. Fifteen of these accounted for ≥10% of phenotypic variation for a given trait, and seven accounted for ≥30% of variation (see PVE values, Table 2). Four statistically significant QTL interactions were identified, three in the B population and one in the S population; those in the B population explain only a small proportion of the total phenotypic variation (PVE ≤1.6%), while the S population interaction is more substantial (*qHD6S* × *qHD7S*, affecting heading date; PVE = 6.4%). For traits scored in both 2012 and 2013, QTL mapping results were generally consistent between years; for those with inconsistent results (panicle length in S population and shattering in B population), we focus on the 2013 data, as hurricane-related storm damage may have led to unrepresentative measurements in 2012 (see Methods). Below we describe QTL for each mapped trait in order of developmental stage, with comparisons to previous mapping studies and to known candidate genes where possible.

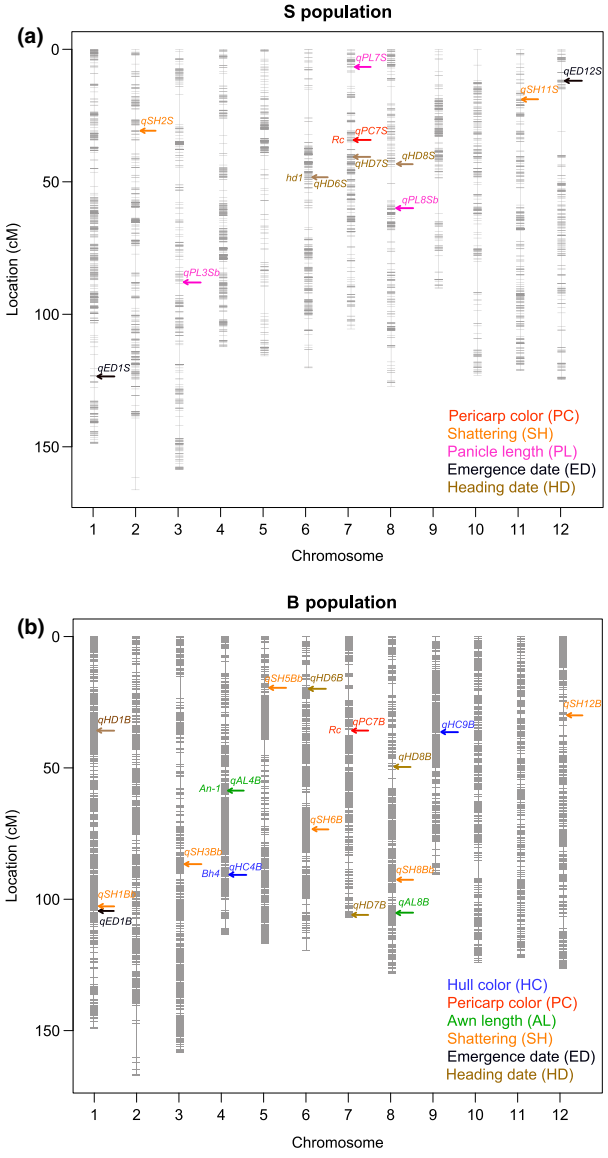

Fig. 1 Linkage map and QTL locations for (a) the S population and (b) the B population. Vertical lines represent the 12 rice chromosomes; short horizontal lines indicate SNP marker locations. Arrows indicate QTL positions. Map distances (cM) are derived from Kosambi's map function. Known rice candidate genes for weedy traits are indicated on the left side of the chromosome according to their physical positions in the Rice Genome Annotation Project database (MSU 7).

Emergence date. Both the SH and BHA weed strains showed earlier seedling emergence than the parental crop line, DGWG (Fig. S1, Supporting information, Table 1). Three statistically supported QTL were detected in the two mapping populations (Table 2). There is no overlap in QTL locations between the populations (Fig. 1, Table 2; see also Table S1, Supporting information), which suggests that early seedling emergence is controlled by different genetic mechanisms in the two weed strains. In

Table 2 QTL for weediness traits in the two mapping populations. Physical position is the genomic location of the nearest marker in the MSU 6.0 rice reference genome; PVE describes the per cent phenotypic variation explained. Analysis of variance (ANOVA) was implemented under a multiple QTL model, with statistical significance shown for post hoc Fisher's exact tests: ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$. LOD values larger than significance thresholds are indicated by lowercase letters, based on (a) HK SMA; (b) HK CIM; (c) IMP SMA; and (d) IMP CIM methods

Phenotype	Mapping population	QTL	Chromosome	Genetic Position (cM)	Physical position	PVE (R^2)	LOD	Increased effect†
Emergence date	S	*qED1S*	1	124.3	40556649	10.9***	abcd	WEED-SH
		qED12S	12	12.8	2701924	7.2***	c	CROP
	B	*qED1B*	1	103.8	38361220	15.1***	abcd	CROP
Heading date	S	*qHD6S*	6	50.4	10107767	32.1***	abcd	CROP
		qHD7S	7	41.7	9487660	23.6***	abcd	WEED-SH
		qHD8S	8	44.2	4211713	22.8***	abcd	WEED-SH
		qHD6S × *qHD7S*				6.4***		
	B	*qHD1B*	1	35.90	6590103	3.5***	cd	CROP
		qHD6B	6	22.50	2532300	11.9***	abcd	CROP
		qHD7B	7	106.01	29474756	15.1***	bcd	WEED-BHA
		qHD8B	8	50.02	4603005	37.8***	abcd	WEED-BHA
		qHD6B × *qHD8B*				1.0*		
Panicle length	S	*qPL3Sb*	3	88.02	25234852	8.4***	d	WEED-SH
		qPL7S	7	7.48	1754679	8.0**	d	CROP
		qPL8Sb	8	60.57	5315626	9.3***	abcd	WEED-SH
Shattering	S	*qSH2S*	2	31.4	4352643	51.7***	abcd	CROP
		qSH11S	11	19.4	4320500	2.0***	a	CROP
	B	*qSH1Bb*	1	103.75	38361220	7.2***	abcd	CROP
		qSH3Bb	3	86.93	25000637	14.1***	abcd	CROP
		qSH5Bb	5	19.82	1786217	6.9***	cd	CROP
		qSH6B	6	76.02	20058587	1.9**	abcd	CROP
		qSH8Bb	8	92.96	20478068	4.6***	d	CROP
		qSH12B	12	30.63	4712192	5.5***	abd	WEED-BHA
		qSH3Bb × *qSH5Bb*				1.6*		
Hull colour	B	*qHC4B*	4	76.32	22365441	32.6***	abcd	WEED-BHA
		qHC9B	9	21.6	7044615	5.1***	d	WEED-BHA
		qHC4B × *qHC9B*				1.4*		
Awn length	B	*qAL4B*	4	58.32	16884832	35.4***	abcd	WEED-BHA
		qAL8B	8	104.51	24140854	17.6***	abcd	WEED-BHA
Pericarp colour	S	*qPC7S*	7	35.83	6372531	59.6***	abcd	WEED-SH
	B	*qPC7B*	7	34.00	6068228	38.3***	abcd	WEED-BHA

†Increased effect is the parental source of the allele causing an increase in the phenotypic value (see Fig. S1, Supporting information), based on the plots of the phenotype data against the genotypes at the specified marker.

addition, the direction of the effect is not the same in the two populations for chromosome 1 weed alleles: BHA allele *qED1B* is associated with earlier emergence than the crop allele, while SH allele *qED1S* is associated with later emergence; the latter pattern points to genetic background-specific effects in the earlier emergence phenotype of SH weeds. As seedling emergence date has not been mapped in earlier studies of weedy rice, no comparison across studies is possible for this trait.

Heading date. The SH weed parent showed earlier flowering than DGWG, while the BHA strain flowered later than the crop (albeit not significantly) (Table 1, Fig. S1, Supporting information). This pattern is consistent with previous observations of the US weed strains showing divergence in flowering time in the field and in common garden experiments (Shivrain *et al.* 2010; Reagon *et al.* 2011; Thurber *et al.* 2014). Three statistically supported QTL for heading date were detected in the S mapping population, and four were detected in the B population; there is one QTL per mapping population on chromosomes 6, 7 and 8, plus one QTL on chromosome 1 in the B population (Fig. 1, Table 2). The two chromosome 8 QTL (*qHD8S*, *qHD8B*) occur in the same approximate genomic location and could potentially correspond to the same underlying locus. The direction

of the weed allele effect is the same in both mapping populations, potentially consistent with allele sharing at this locus in the two weed strains. These may also correspond to QTL detected in a previous study of the S and B populations (*qHD8.1s*, *qHD8.1b*; Thurber *et al.* 2013; Table 3; see also Table S2, Supporting information). The genomic locations of the other B population QTL suggest that these too may correspond to heading date QTL detected in earlier studies (shaded cells in Table 3 and Table S2, Supporting information).

The molecular basis of flowering time variation has been extensively studied in rice, and five of the seven heading date QTL occur in genomic regions that overlap or are near (<0.1 Mb) well-characterized flowering time genes. In four of these five cases, we found causal genetic variation in the candidate gene sequences that can explain the associated QTL (Table 4). We describe these heading date candidate gene results below.

The flowering time gene *Hd1* has been implicated as a key target of selection for a loss of photoperiod sensitivity during rice domestication (Takahashi *et al.* 2009; Fujino *et al.* 2010; Takahashi & Shimamoto 2011). One of the three QTL detected in the S population, *qHD6S*, maps to a genomic interval that contains the *Hd1* locus (Table 4), with the weed allele conferring earlier flowering (Table 2). Consistent with this finding, the crop parent DGWG carries a common *Hd1* loss-of-function mutation found in many cultivated varieties, while the SH weed carries an ancestral functional allele that would be expected to confer earlier flowering under the photoperiod experienced by the mapping populations (Table 4; Thurber *et al.* 2014). The *qHD6S*/*Hd1* QTL accounts for >30% of observed flowering time variation in the S population (Table 2). Interestingly, this major-effect QTL was not detected in the earlier study of the S population cross (Table 3; Thurber *et al.* 2013), a difference that may be attributable to photoperiod differences between the studies. Unlike SH, the BHA weed strain carries a croplike nonfunctional *Hd1* allele (Thurber *et al.* 2014), which accounts for the absence of any *Hd1*-associated QTL in the B population.

The *qHD8S* and *qHD8B* QTL occur in close proximity to *DTH8* (also referred to as *Ghd8* or *Hd5*), another major flowering time regulatory gene (Wei *et al.* 2010; Yan *et al.* 2011; Fujino *et al.* 2013) (Table 4). Frameshift indels in the *DTH8* coding region are associated with reduced photoperiod sensitivity, and rice varieties that carry these loss-of-function mutations show earlier flowering at a range of latitudes. As the weed alleles at both *qHD8S* and *qHD8B* show later flowering than the crop allele (Table 2), the presence of a *DTH8* loss-of-function allele in the *indica* crop parent would account for the observed heading date variation. Consistent with this prediction, we found that DGWG carries a 1-bp frameshift deletion that occurs commonly in *indica* cultivars (Wei *et al.* 2010), while no loss-of-function mutations were detected in the SH and BHA *DTH8* sequences (Table 4).

The *qHD7B* QTL occurs near another photoperiod response regulatory gene, *DTH7*, which was recently cloned and functionally characterized (Gao *et al.* 2014). Like *DTH8*, loss-of-function mutations in this gene result in earlier flowering under long-day photoperiod; an extensive survey of rice germplasm by Gao and colleagues revealed four common loss-of-function alleles. We found that *DTH7* sequences of both DGWG and the SH weed carry an 8-bp frameshift deletion that is common in *indica* rice (Gao *et al.* 2014), with no corresponding loss-of-function allele detected in the BHA weed (Table 4). This pattern can account for the *qHD7B* QTL, where the BHA weed allele is associated with later flowering (Table 2); it also explains the absence of any corresponding QTL in the S population, as both parents carry *DTH7* loss-of-function alleles. The occurrence of the *DTH7* frameshift deletion in the SH weed is also likely a contributing factor in the early heading date of this weed strain (Table 1; Fig. S1, Supporting information).

For *qHD6B*, where the crop allele shows earlier flowering, the nearest candidate gene is *Hd3a*, a flowering time regulator that functions downstream of *Hd1* (Table 4). Two *Hd3a* promoter types, distinguishable by multiple SNPs and a 12-bp indel, are associated with differences in gene expression and flowering time variation under short-day photoperiod; promoter type A is associated with lower expression and later flowering, while type B shows higher expression and earlier flowering (Takahashi *et al.* 2009). Sequencing the *Hd3a* promoter in the B population parents revealed that the *indica* crop parent and BHA weed both carry haplotypes that are identical or very similar to previously described B promoter haplotypes (Table 4; see also Thurber *et al.* 2014). Thus, unlike the candidate genes associated with *qHD6S*, *qHD7B* and *qHD8S*/*qHD8B*, there is no obvious molecular variation in this flowering time gene that would easily account for the *qHD6B* QTL.

Panicle length. Whereas the BHA weed differs only negligibly from DGWG in panicle length (with a slightly larger mean value in 2012 measurements and a slightly smaller mean in 2013), the panicle length of SH is consistently larger than that of the crop (Table 1, Fig. S1, Supporting information). Three QTL were detected for panicle length in the S population based on 2013 phenotype data (Table 2); two of these, *qPL3S* and *qPL8S*, were also detected based on 2012 data (Table S3,

Table 3 Comparison of QTL identified in this study with the earlier studies of weedy rice. Grey-shaded cells indicate QTL that are potentially shared with those detected in the present study for the B population (dark shading) and S population (light shading) based on chromosomal location. Confirmed candidate genes are indicated in bold font

Phenotype	Study*	Chr. 1	Chr. 2	Chr. 3	Chr. 4	Chr. 5	Chr. 6	Chr. 7	Chr. 8	Chr. 9	Chr. 10	Chr. 11	Chr. 12
Heading date	Bres-Patry *et al.* 2001	–	–	*qHD3*	–	*qHD5*	–	*qHD7*	–	–	–	–	–
	Mispan *et al.* (2013)	*qFT1*	–	–	–	–	*qFT6*	*qFT7*	–	–	–	–	–
	Thurber *et al.* (2013) (S)	–	–	–	*qHD4s*	–	–	*qHD7s*	*qHD8.1s* *qHD8.2s*	–	–	–	–
	Thurber *et al.* (2013) (B)	–	–	–	*qHD4b*	–	–	–	*qHD8.1b* *qHD8.2b*	–	–	–	–
	This study (S)	–	–	–	–	–	*qHD6S*/ ***Hd1***	*qHD7S*	*qHD8S*/ ***DTH8***	–	–	–	–
	This study (B)	*qHD1B*	–	–	–	–	*qHD6B*	*qHD7B*/ ***DTH7***	*qHD8B*/ ***DTH8***	–	–	–	–
Shattering	Bres-Patry *et al.* (2001)	*qSHT1*/*Sh-2*	–	–	–	–	–	–	–	–	–	–	–
	Gu *et al.* (2005)	–	–	*qSH3*	*qSH4*	–	–	*qSH7*	*qSH8*	–	–	–	–
	Mispan *et al.* (2013)	–	–	*qSH3*	–	–	–	*qSH7*	–	*qSH9*	–	–	–
	Subudhi *et al.* (2014) (BR)	–	*qSH2*	*qSH3.1*	*qSH4*	–	–	*qSH7*	–	–	*qSH10*	–	–
	Subudhi *et al.* (2014) (CR)	–	–	–	*qSH4*	–	*qSH6*	–	–	–	*qSH10*	–	–
	Thurber *et al.* (2013) (S)	–	*qSS2s*	–	–	–	–	–	–	–	–	*qSS11s*	*qSS12s*
	Thurber *et al.* (2013) (B)	*qSS1b*	–	–	–	–	–	–	–	–	–	–	–
	This study (S)	–	*qSH2S*	–	–	–	–	–	–	–	–	*qSH11S*	–
	This study (B)	*qSH1Bb*	–	*qSH3Bb*	–	*qSH5Bb*	*qSH6B*	–	*qSH8Bb*	–	–	–	*qSH12B*
Hull colour	Gu *et al.* (2005)	–	–	–	*qHC4*	–	–	*qHC7*	–	–	–	–	–
	Mispan *et al.* (2013)	*qHC1*	–	–	*qHC4*	–	–	–	*qHC8*	–	–	–	–
	Thurber *et al.* (2013) (B)	1	–	–	4	–	–	–	–	–	–	–	–
	This study (B)	–	–	–	*qHC4B*/***Bh4***	–	–	–	–	*qHC9B*	–	–	–
Awn length	Bres-Patry *et al.* (2001)	–	–	*qAWN3*/*An-3*	*qAWN4*/*An-1*	–	–	–	–	–	–	–	–
	Gu *et al.* (2005)	–	–	–	*qAL4-1*; *qAL4-2*	–	–	–	*qAL8*	–	–	–	–
	Mispan *et al.* (2013)	–	*qAN2*	*qAN3*	*qAN4*	–	*qAN6.1*; *qAN6.2*	–	*qAN8*	–	–	–	–
	Thurber *et al.* (2013) (B)	–	–	3	–	–	–	–	–	9	–	11	–
	This study (B)	–	–	–	*qAL4B*/***An-1***	–	–	–	*qAL8B*	–	–	–	–
Pericarp colour	Bres-Patry *et al.* (2001)	*qPC1*/*Rd*	–	–	–	–	–	*qPC7*/*Rc*	–	–	–	–	–
	Gu *et al.* (2005)	–	–	–	–	–	–	*qPC7*	–	–	–	–	–
	Mispan *et al.* (2013)	–	–	–	–	–	*qPC6*	*qPC7*	–	–	–	–	*qPC12*
	This study (S)	–	–	–	–	–	–	*qPC7S*/***Rc***	–	–	–	–	–
	This study (B)	–	–	–	–	–	–	*qPC7B*/***Rc***	–	–	–	–	–

*Mapping population details for cited studies:

1. Bres-Patry *et al.* (2001): C6 (*temperate japonica*-type weed) × Miara (*japonica* cultivar), DH population (N = 151), 68 SSRs + 31 AFLPs.
2. Gu *et al.* (2005): SS18-2 (*indica*-type Thai weed) × EM93-1 (*indica* cultivar), BC1 population (N = 204), 175 SSRs.
3. Mispan *et al.* (2013): US1 (*indica*-type US weed, BHA) × EM93-1 (*indica* cultivar), F2 population (N = 188), 123 SSRs.
4. Subudhi *et al.* (2014) (BR population): Bengal (US *tropical japonica* cultivar) × PSRR-1 (US weed), $F_{7:8}$ population (N = 198), 212 SSRs.
5. Subudhi *et al.* (2014) (CR population): Cypress (US *tropical japonica* cultivar) × PSRR-1 (weed), $F_{8:9}$ population (N = 174), 189 SSRs.
6. Thurber *et al.* (2013) (S population): AR-2001-1135 (SH weed) × Dee Geo Woo Gen (*indica* cultivar); F_2 population (N = 184), 54 SSRs.
7. Thurber *et al.* (2013) (B population): MS-1996-9 (BHA weed) × Dee Geo Woo Gen (*indica* cultivar), F_2 population (N = 159), 72 SSRs.
8. This study (S population): AR-2001-1135 (SH weed) × Dee Geo Woo Gen (*indica* cultivar); F_5/F_6 population (N = 175), 6016 SNPs.
9. This study (B population): MS-1996-9 (BHA weed) × Dee Geo Woo Gen (*indica* cultivar), F_5/F_6 population (N = 224), 13 730 SNPs.

Table 4 Comparison of physical positions of weed-associated QTL and known candidate genes. All physical positions are based on the MSU 7 reference genome. GenBank accession numbers are for parents used in generating mapping populations

Phenotype	QTL	QTL physical map position (bp)	Candidate gene	Locus name*	Candidate gene location (bp)	GenBank accessions	Candidate gene distance from QTL	Causal variation detected in candidate gene
Heading date	*qHD6S*	9289670–10554031	*Hd1*	Os06g16370	9336359–9338643	KM063570 (crop) KM063448 (weed)	Direct overlap	Yes
Heading date	*qHD6B*	2472896–2536344	*Hd3a* promoter	Os06g06320	2940004–2942452	KR815347 (crop) KR815348 (weed)	0.40 Mb	No
Heading date	*qHD7B*	29285345–29521503	*DTH7*	Os07g49460	29616705–29629223	KR815352 (crop) KR815353 (weed)	0.10 Mb	Yes
Heading date	*qHD8S*	4102290–4212710	*DTH8*	Os08g07740	4335434–4333717	KR815349 (crop) KR815351 (weed)	0.12 Mb	Yes
Heading date	*qHD8B*	4316506–4604002	*DTH8*	Os08g07740	4335434–4333717	KR815349 (crop) KR815350 (weed)	Direct overlap	Yes
Hull colour	*qHC4B*	22530074–22835077	*Bh4*	Os04g38660–Os04g38670	22969845–22972312	KR815345 (crop) KR815346 (weed)	0.13 Mb	Yes
Awn length	*qAL4B*	16537059–17101340	*An-1*	Os04g28280	16734806–16732393	KR815343 (crop) KR815344 (weed)	Direct overlap	Yes
Pericarp colour	*qPC7S*	6241588–6373526	*Rc*	Os07g11020	6062889–6069317	DQ885818 (crop) GU261594 (weed)	0.17 Mb	Yes
Pericarp colour	*qPC7B*	6062910–6132908	*Rc*	Os07g11020	6062889–6069317	DQ885818 (crop) GU261605 (weed)	Direct overlap	Yes

*Rice Genome Annotation Project, MSU 7.

Supporting information). As panicle length has not been assessed in previous QTL studies of weedy rice, no comparison across studies is possible for this trait.

Shattering. Highly shattering seed is a hallmark of weedy rice, and both the BHA and SH weed parents show much higher levels of shattering than DGWG (Table 1, Fig. S1, Supporting information). Two statistically significant QTL were identified in the S population; these explain 51.7% and 2.0% of the total variation (Table 2). Both of these map to the general vicinity of shattering QTL identified in the same cross by Thurber *et al.* (2013) (Table 3; Table S2, Supporting information). Six QTL were identified in the B population and individually account for 1.6–14.1% of the total variation. One of these, *qSH1Bb*, may correspond to the sole shattering QTL identified in the earlier study of this cross (Thurber *et al.* 2013), as well as to a shattering QTL in the same genomic region that was identified by Bres-Patry *et al.* (2001) (Table 3; Table S2, Supporting information). Four of the other five B population QTL also occur in chromosomal regions broadly shared with previously reported shattering QTL, although a lack of genetic resolution limits direct comparability across studies (see Table S2, Supporting information).

Like heading date, the molecular basis of the shattering phenotype has been studied extensively in rice. Previous studies have determined that differences in the shattering phenotype between domesticated and US weedy rice are not attributable to genetic variation at the well-characterized *sh4* and *qsh1* shattering loci, where reduced-function mutations were favoured during rice domestication (Konishi *et al.* 2006; Li *et al.* 2006; Zhang *et al.* 2009; Thurber *et al.* 2010; Zhu *et al.* 2012). Consistent with this finding, we observed no colocalization of shattering QTL in the S or B populations with *sh4* or *qsh1*. Moreover, we found no obvious sharing of shattering QTL between the two mapping populations, which suggests different genetic mechanisms in the emergence of this trait in the two weed strains.

Hull colour. During rice domestication, most cultivated varieties underwent selection for a loss of the dark-hull pigmentation that characterizes wild *Oryza* species (Zhu *et al.* 2011). Among US weed strains, BHA weeds have the ancestral, dark-hull phenotype, while SH strains have the domestication trait and closely resemble straw-hull crop varieties. Correspondingly, no hull colour QTL were identified in the S population, while two significant QTL were detected in the B population (Table 2). One of these, *qHC4B*, explains approximately one-third of the total variation (32.6%) and maps near the *Bh4* gene, which has previously been shown to control hull colour variation in US weedy rice (Vigueira

et al. 2013a) (Fig. 1b, Table 4). Consistent with this result, we found that the *Bh4* sequence of the BHA weed parent carries a putatively functional *Bh4* haplotype that would be expected to generate the blackhull phenotype, while DGWG carries a haplotype with a 22-bp loss-of-function deletion found in most cultivated rice (Table 4; Zhu *et al.* 2011; Vigueira *et al.* 2013a). This gene also likely underlies the chromosome 4 hull colour QTL that has been detected in several previous studies (Table 3; Gu *et al.* 2005; Mispan *et al.* 2013; Thurber *et al.* 2013).

Awn length. During domestication and subsequent crop improvement, most cultivated rice underwent selection for a reduction in awns (bristles) on the grain. Awn variation in US weeds parallels hull colour variation, with BHA strains possessing the ancestral long-awned form and SH strains having the croplike awnless phenotype. Thus, as with hull colour, no awn length QTL was mapped in the S population. In the B population, two QTL were identified: *qAL4B* and *qAL8B*. These explain 35.4% and 17.6% of the total variation, respectively. Both QTL map to regions associated with awn length variation in previous weedy rice studies (Bres-Patry *et al.* 2001; Gu *et al.* 2005; Mispan *et al.* 2013), although not in the study of Thurber *et al.* (2013) (Table 3).

A recent study has shown that the awnless phenotype is associated with loss-of-function mutations at the *An-1* locus on chromosome 4 (Luo *et al.* 2013). This gene colocalizes with the mapped location of the largest effect QTL, *qAL4B* (Tables 2 and 3). Sequencing *An-1* confirmed that DGWG carries a 1-bp putative loss-of-function deletion in exon 2 that is present in most *indica* varieties, while the BHA weed caries a haplotype with no obvious loss-of-function mutations (Table 4). These findings are consistent with *An-1* playing a major role in awn length differences between BHA weeds and cultivated rice, and with this candidate gene underlying the previously reported QTL for awn length on chromosome 4 (Table 3; Bres-Patry *et al.* 2001; Gu *et al.* 2005; Mispan *et al.* 2013).

Pericarp colour. The alternate name of weedy rice, red rice, derives from the proanthocyanidin pericarp pigmentation that characterizes most weed strains, including both of the strains used in this study. Domesticated rice underwent selection for nonpigmented pericarps, and this was achieved largely through selection for loss-of-function mutations in the *Rc* gene, which encodes a MYB transcription factor that regulates the proanthocyanidin synthesis pathway (Sweeney *et al.* 2006). Mapping pericarp colour in both the S and B populations revealed major-effect QTL in the general region of the *Rc* locus on chromosome 7 (*qPC7S* and *qPC7B*, explaining 59.6% and 38.3% of the total variation, respectively) (Table 2). The *Rc* locus does not fall directly within the boundaries of the *qPC7S* map interval. Nonetheless, *Rc* gene sequences confirm that DGWG carries a common 14-bp loss-of-function deletion in exon 7 that is present in most cultivated varieties, and that both weeds strains carry putatively functional alleles that would be expected to generate the pigmented-pericarp phenotype (Table 4; Gross *et al.* 2010b). Thus, *Rc* is the causal gene for pericarp pigmentation variation in both mapping populations, and the lack of direct overlap with the S population QTL is most likely an artefact of imperfect mapping resolution. *Rc* also probably underlies previously reported pericarp colour QTL that map broadly within the genomic region of this locus on chromosome 7 (Table 3; Bres-Patry *et al.* 2001; Gu *et al.* 2005).

Discussion

Independent evolution of weediness traits

Studies of parallel evolution in crops and their relatives have primarily focused on traits that were selected upon during domestication and subsequent crop improvement (Paterson 2002; Hovav *et al.* 2008; Lenser & Theißen 2013; Fuller *et al.* 2014). Weedy crop relatives provide a complementary opportunity to study agroecological traits whose evolution is not directly determined by intentional human selection (Vigueira *et al.* 2013b). In the case of weedy rice, the independent origins of SH and BHA strains from domesticated *O. sativa* (Reagon *et al.* 2010) make this an especially valuable system for examining the roles of shared vs. different genetic mechanisms in the parallel emergence of weediness traits. At least two factors would seem to favour shared genetic mechanisms in this process. First, as members of the same species and subspecies (*O. sativa* ssp. *indica*), the similar genetic composition of BHA and SH weeds would be expected to predispose them to parallel evolution through shared genetic mechanisms (e.g. Lenser & Theißen 2013). Indeed, comparative QTL mapping in crops has often suggested shared genetic bases for domestication traits, not only within species but also between related species and genera (Fatokun *et al.* 1992; Paterson *et al.* 1995; Paterson 2002; Lin *et al.* 2012; Lenser & Theißen 2013; Olsen & Wendel 2013a). Second, rice, like other crop species, underwent a genomewide reduction in genetic diversity during domestication (Caicedo *et al.* 2007); this domestication bottleneck could further constrain the genetic avenues for weed evolution (but see also Moose *et al.* (2004) for examples of evolution in severely bottlenecked crop lines).

Nonetheless, similar to Thurber *et al.* (2013), who examined a subset of the traits studied here using F_2 lines, we find little evidence for shared genetic bases in the SH and BHA weedy rice strains. For the traits where both strains differed significantly from the crop parent (seedling emergence date, heading date, shattering and pericarp colour), only three QTL pairs out of 20 QTL are potentially shared between the two populations (*qED1S*/*qED1B*, controlling emergence date; *qHD8S*/*qHD8B*, controlling flowering time; and *qPC7S*/*qPC7B*, controlling pericarp colour; Tables 2 and 3). The QTL differences between mapping populations cannot be attributed to environmental effects, as the RILs were phenotyped simultaneously. Thus, it appears that many of the phenotypic changes associated with weediness have arisen through different underlying genetic mechanisms in the two major US weed strains, in many cases involving multiple QTL of relatively minor effect. Consistent with this hypothesis, comparison of the present study with previous QTL studies of weedy rice reveals relatively little obvious QTL sharing across studies (shaded cells, Table 3 and Table S2, Supporting information).

One factor that likely contributes to these findings is the complex developmental basis of many of the traits examined. With the exception of awn length, hull colour and pericarp pigmentation, where observed phenotypic variation is largely attributable to known major-effect genes (and where, as expected, the associated major-effect QTL are shared across mapping populations and studies; Tables 2 and 3), the remaining traits are determined through multiple QTL of small to moderate effect. This would maximize the opportunities for weedy phenotypes to emerge through different genetic mechanisms, potentially involving epistatic interactions that are unique to each weed strain's genomic background (e.g. *qHD6S* × *qHD7S*; Table 2). QTL mapping of complex traits in crosses of domesticated rice and its wild ancestor (*O. rufipogon*) provides support for this hypothesis; as with weedy rice studies, different crop–wild crosses have revealed a multitude of different major- and minor-effect QTL for traits such as heading date, panicle length and seed shattering, with little evidence of shared QTL across studies (Xiao *et al.* 1998; Cai & Morishima 2000; Moncada *et al.* 2001; Thomson *et al.* 2003). While some of these differences are likely due to environmental differences in the mapping experiments, the lack of shared QTL across studies is consistent with multiple underlying loci for these complex traits.

Genetic differences between the two putative crop ancestors of SH and BHA weeds (*indica* and *aus* rice, respectively) could further promote the evolution of weediness through independent genetic mechanisms. While both *indica* and *aus* rice are varieties of *O. sativa* ssp. *indica*, they are genetically distinct subgroups of the crop (Garris *et al.* 2005; Caicedo *et al.* 2007), and multiple hybrid sterility factors have been detected between them (Wan & Ikehashi 1997). In addition, molecular studies of some domestication genes, including *Rc* and *Bh4*, have revealed *aus*-specific mutations that were apparently selected independently of domestication alleles found in other rice varieties (Sweeney *et al.* 2007; Zhu *et al.* 2011). Thus, to the extent that *indica* and *aus* rice represent independent crop lineages, any associated genomic divergence would be expected to enhance possibilities for genetically independent mechanisms of weed evolution. As a corollary, it should be noted that because we used a single *indica* crop parent to create both the S and B RILs, the B population is likely segregating for genetic variation that differs not only between weedy and cultivated rice but also between *indica* and *aus* (see also Thurber *et al.* 2013). This is not a confounding effect for our mapping analyses, however, as the weed strains differ from both *aus* and *indica* rice in the weediness traits that are the focus of this study.

Directions of allele effects

For some of the traits controlled by multiple QTL, alleles from a given parent showed phenotypic effects in the same direction, with either the crop or weed alleles increasing the phenotype (e.g. awn length; Table 2). For others (emergence, heading date, panicle length and shattering), allele effects were not unidirectional in one or both mapping populations, so that the maximum and minimum ends of the RIL phenotypic range are associated with alternate combinations of crop and weed alleles (Fig. S4, Supporting information). Interestingly, this includes the well-studied *Hd1* flowering time gene (*qHD6S*), where the crop parent carries a loss-of-function allele that confers later flowering under long-day photoperiod. The effect of this crop allele is significantly augmented when present in combination with the weed allele of *qHD7S*, a QTL with which *qHD6S*/*Hd1* shows a significant epistatic interaction (Table 2); conversely, the earliest flowering phenotype is achieved by the presence of a functional (weed) *Hd1* allele in combination with the crop allele of the interacting QTL (Fig. S4b, Supporting information). Similarly, opposite combinations of weed and crop alleles confer the extremes of the observed phenotypic ranges for emergence date and shattering (Fig. S4a, c, d, Supporting information). Given that heading date and shattering are key traits of agronomic importance in cultivated rice, these patterns

suggest that weedy rice populations could provide a useful source of genetic variation for crop improvement.

Accuracy and precision of GBS-enabled QTL mapping

The GBS-generated SNPs employed in the present study provided marker coverage with approximately 100 times the genetic resolution of previous weedy rice QTL studies. With the benefits of dense marker coverage, advanced-generation RILs and a well-annotated reference genome, these data can serve as a useful test for whether candidate genes and functional nucleotide variation can be identified through a single round of QTL mapping, without additional fine mapping focused on specific genomic regions. Five of the 27 statistically supported QTL have confidence intervals of <1 cM (Table S1, Supporting information); this is typical resolution for secondary fine mapping in near-isogenic lines (Price 2006). Moreover, for the eight QTL that mapped near candidate genes where we confirmed causal molecular variation, half of the candidate genes fell directly within the corresponding map interval, and the other half were all located within 175 kb (Table 4). Given their relatively low costs for generating dense marker coverage, GBS- and similar next-generation-sequencing-based approaches clearly hold promise for the rapid mapping and identification of functional nucleotide variation (see also Spindel *et al*. 2013).

Conclusions

To the extent that it has been explored at the genetic level, parallel evolution in crop species has often pointed to shared genetic bases for shared domestication traits (Paterson 2002; Lenser & Theißen 2013). Domestication traits are also commonly associated with a relatively few QTL of large effect (Paterson 2002), and recent work suggests that human selection may specifically favour crop alleles whose phenotypic effects are robust across different genetic backgrounds (Doust *et al*. 2014). In contrast, the present study suggests a different genetic architecture for weediness traits, with multiple small-to-moderate effect QTL and relatively few loci shared across weed strains. It is possible that in the absence of domestication selection, agricultural weeds have evolved in a temporally and spatially heterogeneous selective environment that more closely resembles that of wild species than crops. If this is true, we would predict that future QTL mapping of weediness traits in other crop relatives will reveal similar patterns of genetically independent, polygenic adaptation to the agricultural environment.

Acknowledgements

This research was funded through a grant from the NSF Plant Genome Research Program (IOS-1032023) to KMO, ALC and YJ. The authors thank Tracy Bianco, Howard Black, Mia Hodges, Linfeng Li, Michael Lin, Bill Luebke, Jianbing Ma, Marcos André Nohatto, Linda Small and Coretta Torrence for their excellent technical support. Helpful comments on the manuscript were provided by members of the Olsen Laboratory group and those of Justin Fay (Washington Univ.), Allison Miller (St. Louis Univ.) and Christine Edwards (MO Botanical Garden).

References

Akasaka M, Ushiki J, Iwata H, Ishikawa R, Ishii T (2009) Genetic relationships and diversity of weedy rice (*Oryza sativa* L.) and cultivated rice varieties in Okayama Prefecture, Japan. *Breeding Science*, **59**, 401–409.

Arendt J, Reznick D (2008) Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends in Ecology & Evolution*, **23**, 26–32.

Basu C, Halfhill MD, Mueller TC, Stewart CN (2004) Weed genomics: new tools to understand weed biology. *Trends in Plant Science*, **9**, 391–398.

Beavis WD (1998) QTL analyses: power, precision, and accuracy. In: *Molecular Dissection of Complex Traits* (eds Paterson AH), pp. 145–162. CRC Press, Boca Raton, Florida.

Bres-Patry C, Lorieux M, Clément G, Bangratz M, Ghesquière A (2001) Heredity and genetic mapping of domestication-related traits in a temperate japonica weedy rice. *Theoretical and Applied Genetics*, **102**, 118–126.

Broman KW, Sen S (2009) *A Guide to QTL Mapping with R/qtl*. Springer, New York.

Broman KW, Wu H, Sen S, Churchill GA (2003) *R/qtl*: QTL mapping in experimental crosses. *Bioinformatics*, **19**, 889–890.

Burgos NR, Norman RJ, Gealy DR, Black H (2006) Competitive N uptake between rice and weedy rice. *Field Crops Research*, **99**, 96–105.

Burgos NR, Singh V, Tseng TM *et al*. (2014) The impact of herbicide-resistant rice (*Oryza sativa* L.) technology on phenotypic diversity and population structure of US weedy rice. *Plant Physiology*, **166**, 1208–1220.

Cai H-W, Morishima H (2000) Genomic regions affecting seed shattering and seed dormancy in rice. *Theoretical and Applied Genetics*, **100**, 840–846.

Caicedo AL, Williamson SH, Hernandez RD *et al*. (2007) Genome-wide patterns of nucleotide polymorphism in domesticated rice. *PLoS Genetics*, **3**, 1745–1756.

Cao Q, Lu B-R, Xia H *et al*. (2006) Genetic diversity and origin of weedy rice (*Oryza sativa* f. spontanea) populations found in North-eastern China revealed by simple sequence repeat (SSR) markers. *Annals of Botany*, **98**, 1241–1252.

Chung J-W, Park Y-J (2010) Population structure analysis reveals the maintenance of isolated sub-populations of weedy rice. *Weed Research*, **50**, 606–620.

Conte GL, Arnegard ME, Peichel CL, Schluter D (2012) The probability of genetic parallelism and convergence in natural populations. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 5039–5047.

Darwin C (1859) *On the Origin of Species by Means of Natural Selection*. Murray, London.

Doust AN, Lukens L, Olsen KM *et al.* (2014) Beyond the single gene: how epistasis and gene-by-environment effects influence crop domestication. *Proceedings of the National Academy of Sciences, USA*, **111**, 6178–6183.

Ellstrand NC, Heredia SM, Leak-Garcia JA *et al.* (2010) Crops gone wild: evolution of weeds and invasives from domesticated ancestors. *Evolutionary Applications*, **3**, 494–504.

Elshire RJ, Glaubitz JC, Sun Q *et al.* (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, **6**, e19379.

Fatokun CA, Menancio-Hautea DI, Danesh D, Young ND (1992) Evidence for orthologous seed weight genes in cowpea and mung bean based on RFLP mapping. *Genetics*, **132**, 841–846.

Fujino K, Wu J, Sekiguchi H *et al.* (2010) Multiple introgression events surrounding the *Hd1* flowering-time gene in cultivated rice, *Oryza sativa* L. *Molecular Genetics and Genomics*, **284**, 137–146.

Fujino K, Yamanouchi U, Yano M (2013) Roles of the *Hd5* gene controlling heading date for adaptation to the northern limits of rice cultivation. *Theoretical and Applied Genetics*, **126**, 611–618.

Fuller DQ, Denham T, Arroyo-Kalin M *et al.* (2014) Convergent evolution and parallelism in plant domestication revealed by an expanding archaeological record. *Proceedings of the National Academy of Sciences, USA*, **111**, 6147–6152.

Gao H, Jin M, Zheng X-M *et al.* (2014) Days to heading 7, a major quantitative locus determining photoperiod sensitivity and regional adaptation in rice. *Proceedings of the National Academy of Sciences, USA*, **111**, 16337–16342.

Garris AJ, Tai TH, Coburn J, Kresovich S, McCouch S (2005) Genetic structure and diversity in *Oryza sativa* L. *Genetics*, **169**, 1631–1638.

Grimm A, Fogliatto S, Nick P, Ferrero A, Vidotto F (2013) Microsatellite markers reveal multiple origins for Italian weedy rice. *Ecology and Evolution*, **3**, 4786–4798.

GRIN: Germplasm Resources Information Network (2015) Germplasm Resources Information Network. Available from http://www.ars-grin.gov/npgs/. USDA-ARS. Accessed September 15, 2011.

Gross BL, Reagon M, Hsu S-C *et al.* (2010a) Seeing red: the origin of grain pigmentation in US weedy rice. *Molecular Ecology*, **19**, 3380–3393.

Gross BL, Steffen FT, Olsen KM (2010b) The molecular basis of white pericarps in African domesticated rice: novel mutations at the *Rc* gene. *Journal of Evolutionary Biology*, **23**, 2747–2753.

Gu X-Y, Kianian SF, Hareland GA, Hoffer BL, Foley ME (2005) Genetic analysis of adaptive syndromes interrelated with seed dormancy in weedy rice (*Oryza sativa*). *Theoretical and Applied Genetics*, **110**, 1108–1118.

Gu X-Y, Foley ME, Horvath DP *et al.* (2011) Association between seed dormancy and pericarp color is controlled by a pleiotropic gene that regulates abscisic acid and flavonoid synthesis in weedy red rice. *Genetics*, **189**, 1515–1524.

Haley CS, Knott SA (1992) A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity*, **69**, 315–324.

Hammer K (1984) Das Domestikationssyndrom. *Die Kulturpflanze*, **32**, 11–34.

Harlan JR (1992) *Crops and Man*. American Society of Agronomy, Inc. Crop Science Society of America, Inc, Madison.

Hovav R, Chaudhary B, Udall JA, Flagel L, Wendel JF (2008) Parallel domestication, convergent evolution and duplicated gene recruitment in allopolyploid cotton. *Genetics*, **179**, 1725–1733.

IRRI: International Rice Research Institute and International Board for Plant Genetic Resources (1980) *Descriptors for Rice, Oryza sativa L.* IBPGR-IRRI Rice Advisory Committee. International Rice Research Institute, Manila, Philippines.

Kane NC, Rieseberg LH (2008) Genetics and evolution of weedy *Helianthus annuus* populations: adaptation of an agricultural weed. *Molecular Ecology*, **17**, 384–394.

Kawahara Y, de la Bastide M, Hamilton JP *et al.* (2013) Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice*, **6**, 4.

Konishi S, Izawa T, Lin SY *et al.* (2006) An SNP caused loss of seed shattering during rice domestication. *Science*, **312**, 1392–1396.

Larson G, Piperno DR, Allaby RG *et al.* (2014) Current perspectives and the future of domestication studies. *Proceedings of the National Academy of Sciences, USA*, **111**, 6139–6146.

Lenser T, Theißen G (2013) Molecular mechanisms involved in convergent crop domestication. *Trends in Plant Science*, **18**, 704–714.

Li C, Zhou A, Sang T (2006) Rice domestication by reducing shattering. *Science*, **311**, 1936–1939.

Lin Z, Li X, Shannon LM *et al.* (2012) Parallel domestication of the *Shattering1* genes in cereals. *Nature Genetics*, **44**, 720–724.

Londo JP, Schaal BA (2007) Origins and population genetics of weedy red rice in the USA. *Molecular Ecology*, **16**, 4523–4535.

Losos JB (2011) Convergence, adaptation, and constraint. *Evolution*, **65**, 1827–1840.

Luo J, Liu H, Zhou T *et al.* (2013) *An-1* encodes a basic helix-loop-helix protein that regulates awn development, grain size, and grain number in rice. *Plant Cell*, **25**, 3360–3376.

Manceau M, Domingues VS, Linnen CR, Rosenblum EB, Hoekstra HE (2010) Convergence in pigmentation at multiple levels: mutations, genes and function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 2439–2450.

Meyer RS, Purugganan MD (2013) Evolution of crop species: genetics of domestication and diversification. *Nature Reviews Genetics*, **14**, 840–852.

Mispan MS, Zhang L, Feng J, Gu X-Y (2013) Quantitative trait locus and haplotype analyses of wild and crop-mimic traits in U.S. weedy rice. *G3 (Bethesda, Md.)*, **3**, 1049–1059.

Moncada P, Martínez CP, Borrero J *et al.* (2001) Quantitative trait loci for yield and yield components in an *Oryza sativa* × *Oryza rufipogon* BC 2 F 2 population evaluated in an upland environment. *Theoretical and Applied Genetics*, **102**, 41–52.

Moose SP, Dudley JW, Rocheford TR (2004) Maize selection passes the century mark: a unique resource for 21st century genomics. *Trends in Plant Science*, **9**, 358–364.

Morrell PL, Buckler ES, Ross-Ibarra J (2011) Crop genomics: advances and applications. *Nature Reviews Genetics*, **13**, 85–96.

Olsen KM, Wendel JF (2013a) Crop plants as models for understanding plant adaptation and diversification. *Frontiers in Plant Science*, **4**, 290.

Olsen KM, Wendel JF (2013b) A bountiful harvest: genomic insights into crop domestication phenotypes. *Annual Review of Plant Biology*, **64**, 47–70.

Olsen KM, Caicedo AL, Jia Y (2007) Evolutionary genomics of weedy rice in the USA. *Journal of Integrative Plant Biology*, **49**, 811–816.

Paterson AH (2002) What has QTL mapping taught us about plant domestication? *New Phytologist*, **154**, 591–608.

Paterson AH, Lin YR, Li Z *et al.* (1995) Convergent domestication of cereal crops by independent mutations at corresponding genetic Loci. *Science*, **269**, 1714–1718.

Price AH (2006) Believe it or not, QTLs are accurate!. *Trends in Plant Science*, **11**, 213–216.

Reagon M, Thurber C, Gross B (2010) Genomic patterns of nucleotide diversity in divergent populations of U.S. weedy rice. *BMC Evolutionary Biology*, **10**, 180.

Reagon M, Thurber CS, Olsen KM, Jia Y, Caicedo AL (2011) The long and the short of it: SD1 polymorphism and the evolution of growth trait divergence in U.S. weedy rice. *Molecular Ecology*, **20**, 3743–3756.

Schluter D, Clifford EA, Nemethy M, McKinnon JS (2004) Parallel evolution and inheritance of quantitative traits. *American Naturalist*, **163**, 809–822.

Sen Ś, Churchill GA (2001) A statistical framework for quantitative trait mapping. *Genetics*, **159**, 371–387.

Shivrain VK, Burgos NR, Scott RC *et al.* (2010) Diversity of weedy red rice (*Oryza sativa* L.) in Arkansas, U.S.A. in relation to weed management. *Crop Protection*, **29**, 721–730.

Song B-K, Chuah T-S, Tam SM, Olsen KM (2014) Malaysian weedy rice shows its true stripes: wild *Oryza* and elite rice cultivars shape agricultural weed evolution in Southeast Asia. *Molecular Ecology*, **23**, 5003–5017.

Spindel J, Wright M, Chen C *et al.* (2013) Bridging the genotyping gap: using genotyping by sequencing (GBS) to add high-density SNP markers and new value to traditional bi-parental mapping and breeding populations. *Theoretical and Applied Genetics*, **126**, 2699–2716.

Stern DL (2013) The genetic causes of convergent evolution. *Nature Reviews Genetics*, **14**, 751–764.

Subudhi PK, Singh PK, DeLeon T *et al.* (2014) Mapping of seed shattering loci provides insights into origin of weedy rice and rice domestication. *Journal of Heredity*, **105**, 276–287.

Sweeney MT, Thomson MJ, Pfeil BE, McCouch S (2006) Caught red-handed: *Rc* encodes a basic helix-loop-helix protein conditioning red pericarp in rice. *Plant Cell*, **18**, 283–294.

Sweeney MT, Thomson MJ, Cho YG *et al.* (2007) Global dissemination of a single mutation conferring white pericarp in rice. *PLoS Genetics*, **3**, e133.

Takahashi Y, Shimamoto K (2011) *Heading date 1* (*Hd1*), an ortholog of Arabidopsis *CONSTANS*, is a possible target of human selection during domestication to diversify flowering times of cultivated rice. *Genes & Genetic Systems*, **86**, 175–182.

Takahashi Y, Teshima KM, Yokoi S, Innan H, Shimamoto K (2009) Variations in *Hd1* proteins, *Hd3a* promoters, and *Ehd1* expression levels contribute to diversity of flowering time in cultivated rice. *Proceedings of the National Academy of Sciences, USA*, **106**, 4555–4560.

Takuno S, Terauchi R, Innan H (2012) The power of QTL mapping with RILs. *PLoS ONE*, **7**, e46545.

Thomson MJ, Tai TH, McClung AM *et al.* (2003) Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between *Oryza rufipogon* and the *Oryza sativa* cultivar Jefferson. *Theoretical and Applied Genetics*, **107**, 479–493.

Thurber CS, Reagon M, Gross BL *et al.* (2010) Molecular evolution of shattering loci in U.S. weedy rice. *Molecular Ecology*, **19**, 3271–3284.

Thurber CS, Hepler PK, Caicedo AL (2011) Timing is everything: early degradation of abscission layer is associated with increased seed shattering in U.S. weedy rice. *BMC Plant Biology*, **11**, 14.

Thurber CS, Jia MH, Jia Y, Caicedo AL (2013) Similar traits, different genes? Examining convergent evolution in related weedy rice populations. *Molecular Ecology*, **22**, 685–698.

Thurber CS, Reagon M, Olsen KM, Jia Y, Caicedo AL (2014) The evolution of flowering strategies in US weedy rice. *American Journal of Botany*, **101**, 1737–1747.

Vigueira CC, Li W, Olsen KM (2013a) The role of *Bh4* in parallel evolution of hull colour in domesticated and weedy rice. *Journal of Evolutionary Biology*, **26**, 1738–1749.

Vigueira CC, Olsen KM, Caicedo AL (2013b) The red queen in the corn: agricultural weeds as models of rapid adaptive evolution. *Heredity*, **110**, 303–311.

Wan J, Ikehashi H (1997) Identification of two types of differentiation in cultivated rice (*Oryza sativa* L.) detected by polymorphism of isozymes and hybrid sterility. *Euphytica*, **94**, 151–161.

Wei X, Xu J, Guo H *et al.* (2010) *DTH8* suppresses flowering in rice, influencing plant height and yield potential simultaneously. *Plant Physiology*, **153**, 1747–1758.

Wood TE, Burke JM, Rieseberg LH (2005) Parallel genotypic adaptation: when evolution repeats itself. *Genetica*, **123**, 157–170.

Xiao J, Li J, Grandillo S *et al.* (1998) Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon*. *Genetics*, **150**, 899–909.

Yan WH, Wang P, Chen HX *et al.* (2011) A major QTL, Ghd8, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in rice. *Molecular Plant*, **4**, 319–330.

Zhang L-B, Zhu Q, Wu Z-Q *et al.* (2009) Selection on grain shattering genes and rates of rice domestication. *New Phytologist*, **184**, 708–720.

Zhu B-F, Si L, Wang Z *et al.* (2011) Genetic control of a transition from black to straw-white seed hull in rice domestication. *Plant Physiology*, **155**, 1301–1311.

Zhu Y, Ellstrand NC, Lu B-R (2012) Sequence polymorphisms in wild, weedy, and cultivated rice suggest seed-shattering locus *sh4* played a minor role in Asian rice domestication. *Ecology and Evolution*, **2**, 2106–2113.

K.M.O., A.L.C., Y.J. and D.R.G. designed the research; X.Q., Y.L. and C.C.V. performed the research; X.Q., C.C.V., N.Y. and K.M.O. analysed the data; and K.M.O. and X.Q. wrote the manuscript.

Data accessibility

Raw phenotype data: doi:10.5061/dryad.566h9.

Raw SNP genotyping data and candidate gene sequence alignments: doi:10.5061/dryad.566h9.
Illumina HiSeq Raw reads: NCBI SRA accession SRX576894.
DNA sequences for candidate genes: NCBI GenBank accessions KR815343–KR815353.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Confidence intervals for each QTL.

Table S2 Map interval information for weedy rice QTL shown in Table 3.

Table S3 Comparison of QTL identified from 2012 (F_6) and 2013 (F_7) phenotyping for panicle length in the S population and degree of shattering in the B population.

Fig. S1 Frequency distributions of the seven traits in the RIL F_6/F_7 populations developed from S and B populations in 2012 and 2013.

Fig. S2 LOD scores of the seven phenotypic traits. Red dashed lines represent the LOD threshold for significance.

Fig. S3 Comparison of QTL identified in 2012 and 2013 phenotyping: (a) panicle length in the S population, (b) degree of shattering in the B population.

Fig. S4 Allele effect plots for (a) seedling emergence date (S population); (b) heading date (S population); (c) heading date (B population); and (d) shattering (B population).