

***New Phytologist* Supporting Information**

Article title: Combining UAV-RGB high-throughput field phenotyping and genome-wide association study to reveal genetic variation of rice germplasms in dynamic response to drought stress

Authors: Zhao Jiang, Haifu Tu, Baowei Bai, Chenghai Yang, Biquan Zhao, Ziyue Guo, Qian Liu, Hu Zhao, Wanneng Yang, Lizhong Xiong and Jian Zhang

Article acceptance date: 17 June 2021

The following Supporting Information is available for this article:

Fig. S1 Illustration of manual leaf-rolling rating and sample patches for each degree.

Fig. S2. Performance evaluation of DCNN-LRS.

Fig. S3 Histogram of brightness distribution of the image patches in the 2019 test dataset.

Fig. S4 Genetic variations of *OsZIP12* and *OsRCI-25* were significantly associated with LWI_D20_PM of rice at the reproductive stage under field conditions.

Notes S1 Visual leaf-rolling rating and plant water content measurements.

Notes S2 Configuration information of the imaging platform.

Notes S3 GCP management and georeferencing process in Agisoft PhotoScan.

Notes S4 Image data augmentation.

Notes S5 Website to download training and validation datasets and the DCNN-LRS model.

Notes S6 Examples of reported genes with accessioned by LWI_D20_PM.

Supplemental Figures

















Sample patches													
Rater 1 	1	1	1	2	2	2	3	3	3	4	4	4	5
Rater 2 	1	1	2	2	2	3	3	3	4	4	4	5	5
Rater 3 	1	2	2	2	3	3	3	4	4	4	5	5	5
Average LRS_m	1	1.33	1.67	2	2.33	2.67	3	3.33	3.67	4	4.33	4.67	5

Fig. S1 Illustration of manual leaf-rolling rating and sample patches for each degree. The LRS_m averaged by 3 raters includes 13 discrete values ranged from 1 to 5.

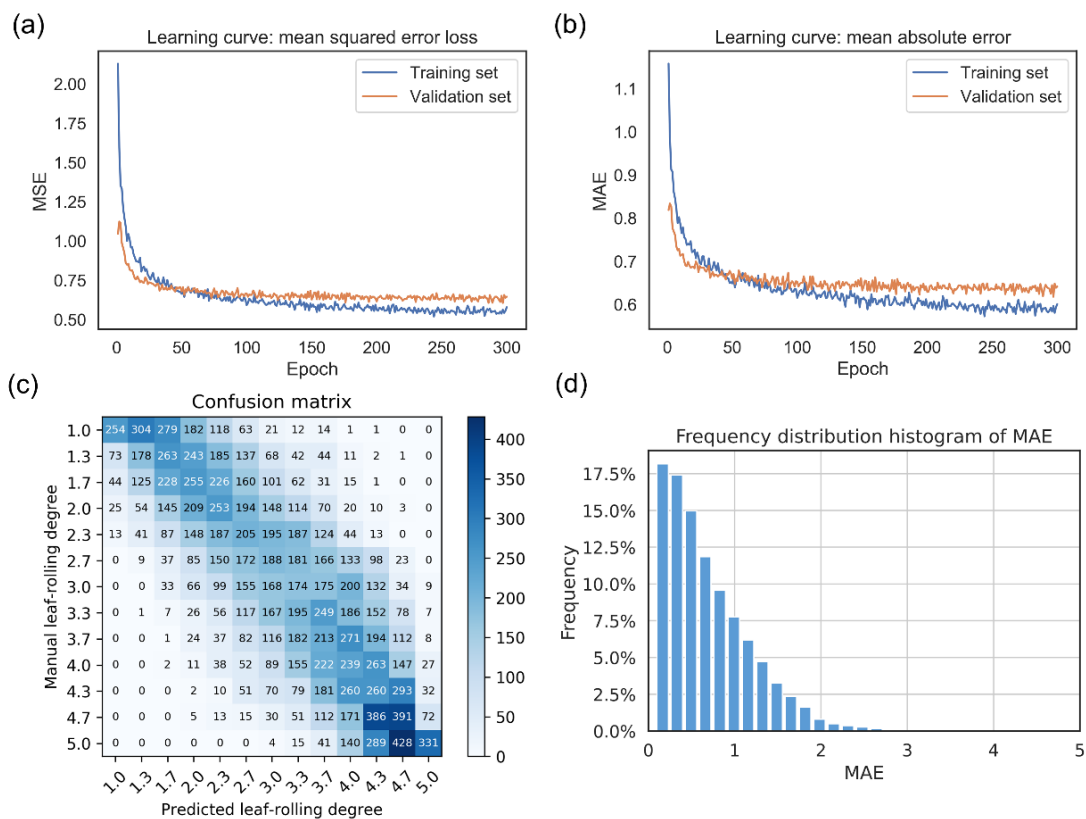


Fig. S2. Performance evaluation of DCNN-LRS. (a) Learning curve of mean squared error (MSE) loss. (b) Learning curve of mean absolute error (MAE). (c) Confusion matrix of discretized LRS_uav and LRS_m. (d) Frequency distribution histogram of MAE. The MSE and MAE decreased in the early training epochs, and then gradually stabilized in the later training epochs. And the training set performed slightly better than the validation set. Compared to the discrete LRS_m trait, the LRS_uav trait with continuous measurement was more objective and to quantify the leaf-rolling severity and more suitable for GWAS analysis. The LRS_uav was discretized to the severity scale corresponding to the manual score to verify whether the model

can replace the manual measurement. The confusion matrix of the discretized LRS_uav and LRS_m indicates an obvious linear relationship between the predicted leaf-rolling degree and manual leaf-rolling degree. Most samples were distributed in a buffer of the main diagonal because of the fuzzy boundary between the successive leaf-rolling degrees.

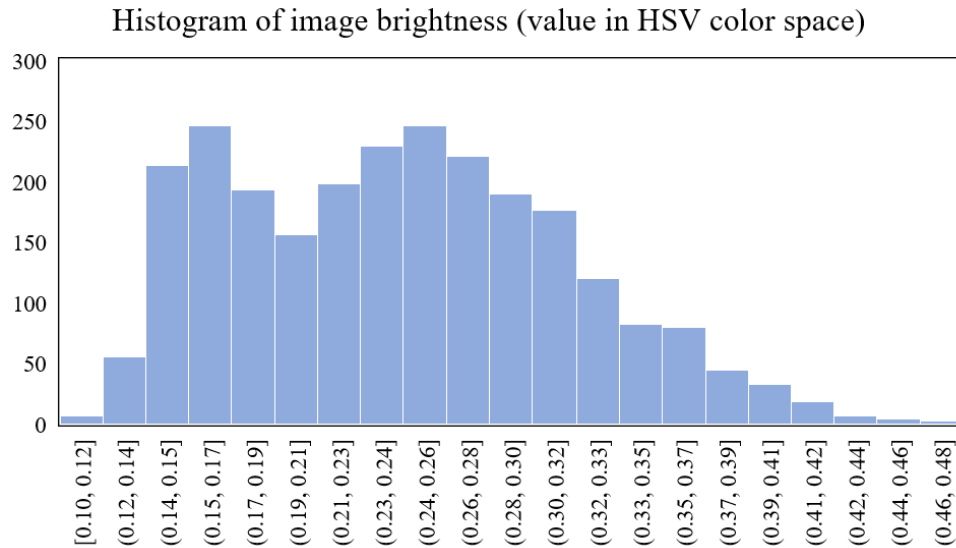


Fig. S3 Histogram of brightness distribution of the image patches in the 2019 test dataset. Averaged values in the HSV color space were calculated to indicate the brightness of image patches in the 2019 test dataset. The brightness varied a lot from 0.1 to 0.45 among the image patches collected in multiple flights, but the accuracy of the test dataset was not influenced, indicating the robustness of DCNN-LRS in different illumination conditions.

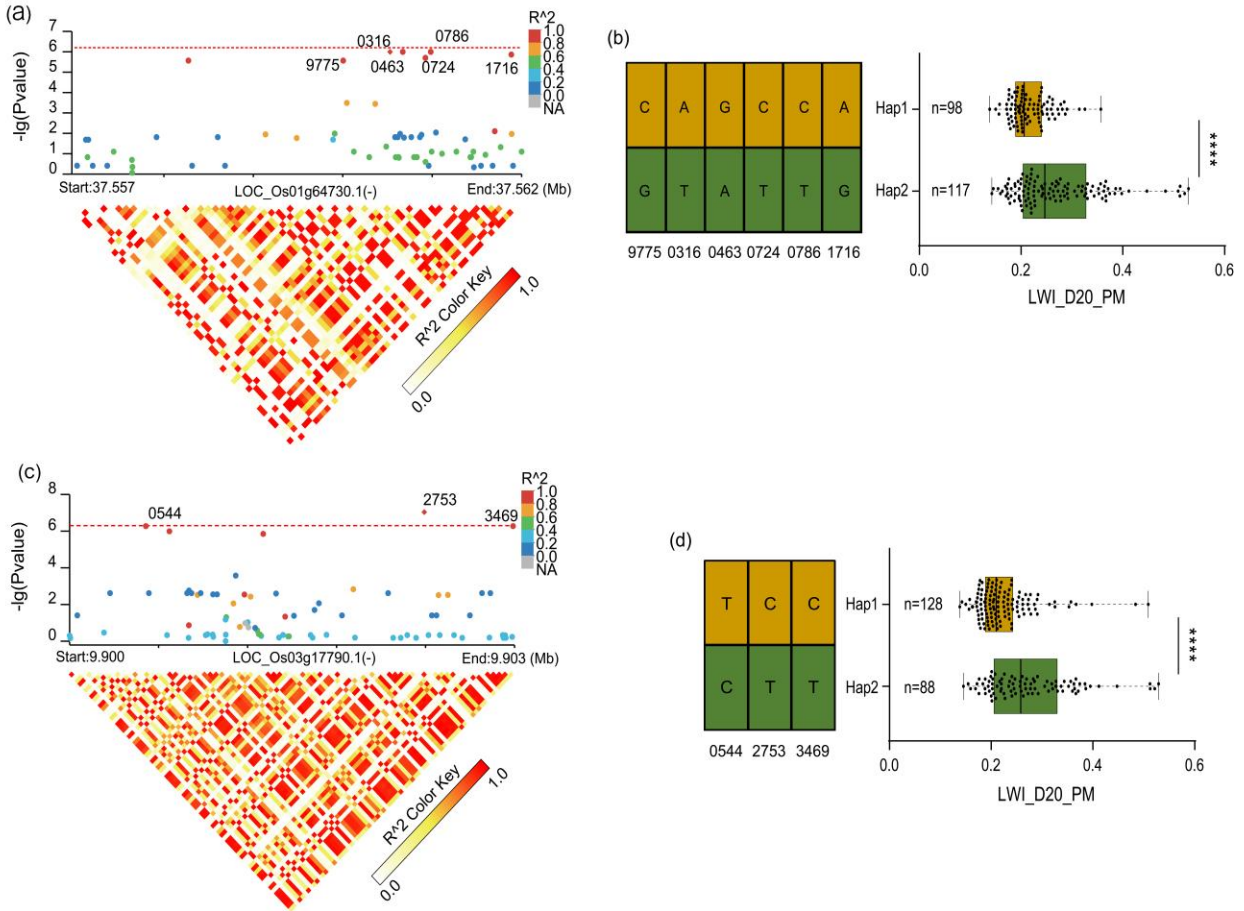


Fig. S4 Genetic variations of *OsbZIP12* and *OsRCI-25* were significantly associated with LWI_D20_PM of rice at the reproductive stage under field conditions. *OsbZIP12*(a), *OsRCI-25*(c) Manhattan plot and pairwise LD analysis. The most significant variation in the gene is shown with a red diamond, and other variations are colored to display their LD with the most significant variation. Red dashed line indicates suggestive threshold. Haplotype groups of *OsbZIP12*(b), *OsRCI-25*(d) in the association population based on the significant variation. LWI_D20_PM of rice distribution of each haplotype group is showed by box-plot. In the boxplots (b, d), the black line within the box is the median, the top and bottom of the box represent the 0.75 and 0.25 quartiles, respectively. The whiskers extend to data no more than 1.5 times the interquartile range, and data are indicated by dots. “n” denotes the number of accessions in each haplotype group. Statistical significance was determined by Welch’s *t*-test with “**” representing $P < 0.0001$.**

Supplemental Notes

Notes S1 Visual leaf-rolling rating and plant water content measurements

The manual leaf-rolling score (LRS_m) was visually rated for 240 plots (120 accessions) using a “1” to “5” scale with “1” being the first evidence of rolling (less than 20% leaves are rolling) and “5” being a closed cylinder of leaves (near 100% leaves are rolling) (O’Toole & Cruz, 1980), and the intermediate scores “2”, “3”, “4” being more than 40%, 60%, 80% leaves are rolling, respectively. The rating scores were taken from the average values of three raters at two time windows, early morning (6:00 -- 7:30 AM) and late afternoon (5:40 -- 7:10 PM), and UAV images were simultaneously collected at the two time windows. The averaged LRS_m value of the three raters ranged from 1 to 5 (Fig. S1). Based on the dynamic LRS_m traits, the leaf-rolling days (LRD) was manually recorded by counting the days required for the leaves to become irreversibly rolling as observed in the morning. More specifically, the manually recorded LRD trait referred to the number of days from the drainage day to the day when the LRS_m value stability reached 3.67 in the morning. The LRD could reflect the variation of rice drought resistance very well based on pretesting of 30 drought-resistant accessions, 30 drought-sensitive accessions, and 30 moderate drought-resistant accessions, which were selected for the evaluation of UAV-based traits.

The fresh weight (FW) of each plot (50 plots were used) was weighed immediately after harvest of above-ground plants at the mild drought stress stage in 2019. Then the dry weight (DW) was weighed after drying, and the plant water content (PWC) was calculated according to the difference between FW and DW. The manually measured FW_m, DW_m, and PWC_m traits (Table S3) were used as ground truth to build non-destructive models with UAV-based traits.

Notes S2 Configuration information of the imaging platform.

The consumer-grade RGB camera mounted on the UAV has a 35.8 mm × 23.9 mm CMOS sensor with a resolution of 42 million pixels (7952 × 5304). The camera aperture was set to f/5.6, and the shutter was set to 1/1000 to avoid movement blur. The ISO was adjusted manually before every flight to ensure proper exposure. The images were stored in a high-speed SD card as compact JPEG files. The flight parameters were set in the DJI GS Pro software. The flying speed was set to 1.5 m/s and the camera trigger interval was set to 2s to provide 80% front overlap and 70% side overlap between images. The same route configuration was repeated in every flight to capture images with the same resolution.

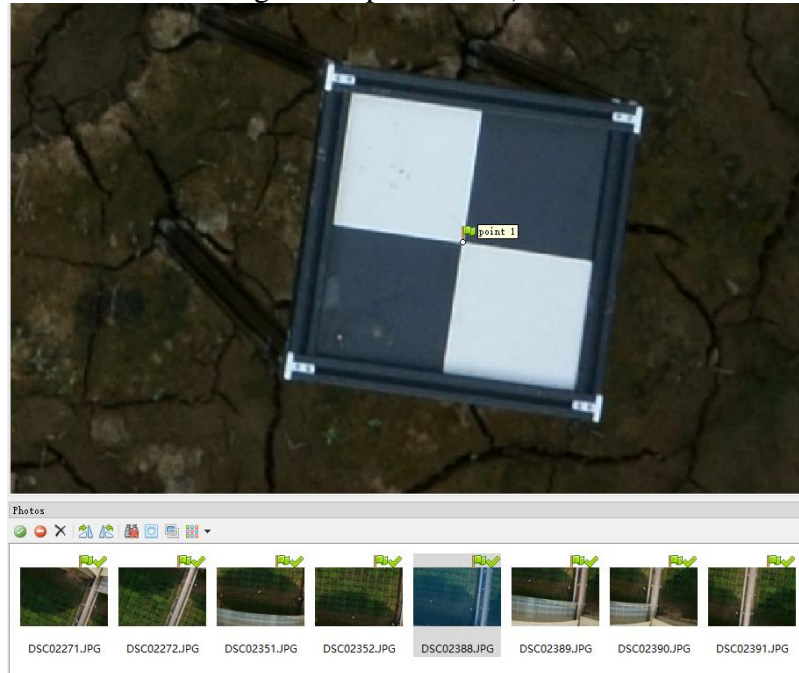
Notes S3 GCP management and georeferencing process in Agisoft PhotoScan.

The GCP coordinates were imported to the Agisoft PhotoScan software for the GCP management. The general steps of GCP management in the software are shown as follows.

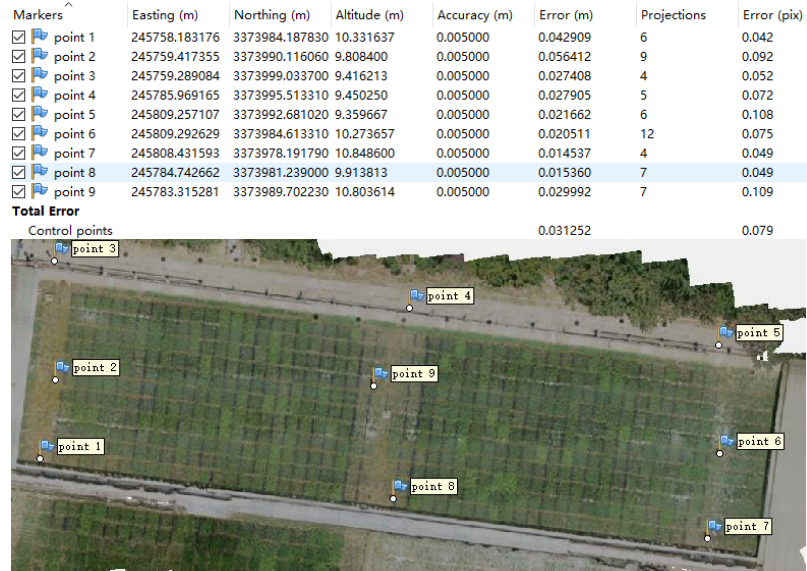
- (1) Import the GCP coordinates measured by a GNSS RTK receiver.

Markers	Easting (m)	Northing (m)	Altitude (m)
<input checked="" type="checkbox"/> point 1	245758.183176	3373984.187830	10.331637
<input checked="" type="checkbox"/> point 2	245759.417355	3373990.116060	9.808400
<input checked="" type="checkbox"/> point 3	245759.289084	3373999.033700	9.416213
<input checked="" type="checkbox"/> point 4	245785.969165	3373995.513310	9.450250
<input checked="" type="checkbox"/> point 5	245809.257107	3373992.681020	9.359667
<input checked="" type="checkbox"/> point 6	245809.292629	3373984.613310	10.273657
<input checked="" type="checkbox"/> point 7	245808.431593	3373978.191790	10.848600
<input checked="" type="checkbox"/> point 8	245784.742662	3373981.239000	9.913813
<input checked="" type="checkbox"/> point 9	245783.315281	3373989.702230	10.803614

- (2) Find out the images in which the GCPs were captured and mark the precise position of the GCPs manually. Make sure each GCP were found in at least 2 images (we actually found each GCP in at least 4 images for optimization).



- (3) Optimize the photo alignment and minimize the GCP error in the image tie points. Build the georeferenced digital surface model (DSM), and digital ortho photo (DOM) based on the georeferenced image tie points and dense cloud.



Notes S4 Image data augmentation.

Firstly, random cropping was implemented to the patches by placing a cropping window to down-sample the original patches to a fixed resolution of 160×160 pixels. The sample distribution was balanced by setting the number of patches as 1000 for the 13 discrete LRS_m degrees during this random cropping procedure. Secondly, real-time data augmentation techniques such as random rotation (within 45 degrees) and random flipping (horizontal and vertical) were applied to the 13000 patches during the training and validation process to increase the training data diversity. Real-time data augmentation was implemented with the data generator in Keras, and technically every batch and epoch for the training and validation process consisted of newly augmented patches.

Notes S5 Websites for the training and validation datasets and the DCNN-LRS model.

The training and validation datasets collected in 2018 are available online with request: [https://www.researchgate.net/publication/350835627 UAV image patches for deep convolutional neural network leaf-rolling scorer DCNN-LRS training and validation](https://www.researchgate.net/publication/350835627_UAV_image_patches_for_deep_convolutional_neural_network_leaf-rolling_scorer_DCNN-LRS_training_and_validation). The DCNN-LRS can be downloaded as a python library through <https://pypi.org/project/DCNN-LRS/>.

Notes S6 Examples of reported genes with associated by LWI_D20_PM.

OsZIP12 encodes a basic zipper transcription factor which confers drought tolerance (Joo *et al.*, 2014). One SNP in the 5'-UTR and five variations in the 1 Kb promoter region were associated with LWI_D20_PM ($P_{\text{imm}} = 2.7 \times 10^{-6}$, 1.01×10^{-6} , 1.01×10^{-6} , 1.99×10^{-6} , 1.01×10^{-6} , 1.37×10^{-6} , respectively) (Fig. S2a). Two haplotype groups were found in the accession population based on these variations, and the Hap2 group (117 accessions) had a significantly

higher LWI_D20_PM than the Hap1 group (98 accessions) (Fig. S2b). OsEREBP2, which belongs to the AP2/ERF transcription factor family, may play a central role in regulating various abiotic stress responses (Serra *et al.*, 2013). One SNP in the *OsEREBP2* promoter was associated with LWI_D20_PM ($P_{\text{imm}} = 2.1 \times 10^{-6}$). *OsRCI2-5*, which encodes a protein containing a signal-peptide and two conservative hydrophobic domains, positively regulates drought resistance (Li *et al.*, 2014). Three promoter region variations were significantly associated with LWI_D20_PM ($P_{\text{imm}} = 5.32 \times 10^{-7}$, 9.16×10^{-8} , 5.32×10^{-7} , respectively) (Fig. S2c), and two haplotype groups were also identified. The Hap2 group had a significantly higher LWI_D20_PM than the Hap1 group (Fig. S2d).

Supplemental References

GUO L-B, QIAN Q, ZENG D-L, DONG G-J, TENG S, ZHU L-H. 2004. Genetic dissection for leaf correlative traits of rice (*Oryza sativa* L.). *Acta Genetica Sinica* **31**: 275–280.

Joo J, Lee YH, Song SI. 2014. Overexpression of the rice basic leucine zipper transcription factor *OsZIP12* confers drought tolerance to rice and makes seedlings hypersensitive to ABA. *Plant Biotechnology Reports* **8**: 431–441.

Lafitte HR, Price AH, Courtois B. 2004. Yield response to water deficit in an upland rice mapping population: associations among traits and genetic markers. *Theoretical and Applied Genetics* **109**: 1237–1246.

Lanceras JC, Pantuwan G, Jongdee B, Toojinda T. 2004. Quantitative trait loci associated with drought tolerance at reproductive stage in rice. *Plant Physiology* **135**: 384–399.

Li L, Li N, Song SF, Li YX, Xia XJ, Fu XQ, Chen GH, Deng HF. 2014a. Cloning and characterization of the drought-resistance *OsRCI2-5* gene in rice (*Oryza sativa* L.). *Genetics and Molecular Research* **13**: 4022–4035.

Mishra KK, Vikram P, Yadaw RB, Swamy BM, Dixit S, Cruz MTS, Maturan P, Marker S, Kumar A. 2013. *qDTY12.1*: a locus with a consistent effect on grain yield under drought in rice. *BMC Genetics* **14**: 12.

Moncada P, Martínez CP, Borrero J, Chatel M, Gauch Jr H, Guimaraes E, Tohme J, McCouch SR. 2001. Quantitative trait loci for yield and yield components in an *Oryza sativa* × *Oryza rufipogon* BC₂F₂ population evaluated in an upland environment: *Theoretical and Applied Genetics* **102**: 41–52.

O'Toole JC, Cruz RT. 1980. Response of leaf water potential, stomatal resistance, and leaf rolling to water stress. *Plant Physiology* **65**: 428–432.

Price AH, Townend J, Jones MP, Audebert A. 2002. Mapping QTLs associated with drought avoidance in upland rice grown in the Philippines and West Africa. *Plant Molecular Biology* **48**: 683–695.

Robin S, Pathan MS, Courtois B, Lafitte R, Carandang S, Lanceras S, Amante M, Nguyen HT, Li Z. 2003. Mapping osmotic adjustment in an advanced back-cross inbred population of rice. *Theoretical and Applied Genetics* **107**: 1288–1296.

Serra TS, Figueiredo DD, Cordeiro AM, Almeida DM, Lourenço T, Abreu IA, Sebastián A, Fernandes L, Contreras-Moreira B, Oliveira MM, et al. 2013. *OsRMC*, a negative regulator of salt stress response in rice, is regulated by two AP2/ERF transcription factors. *Plant Molecular Biology* **82**: 439–455.

You J, Li Q, Yue B, Xue W-Y, Luo L-J, Xiong L-Z. 2006. Identification of quantitative trait loci for aba sensitivity at seed germination and seedling stages in rice. *Acta Genetica Sinica* **33**: 532–541.

Yue B, Xiong L, Xue W, Xing Y, Luo L, Xu C. 2005. Genetic analysis for drought resistance of rice at reproductive stage in field with different types of soil. *Theoretical and Applied Genetics* **111**: 1127–1136.

Yue B, Xue W, Xiong L, Yu X, Luo L, Cui K, Jin D, Xing Y, Zhang Q. 2006. Genetic basis of drought resistance at reproductive stage in rice: separation of drought tolerance from drought avoidance. *Genetics* **172**: 1213–1228.

Zhang X, Zhou S, Fu Y, Su Z, Wang X, Sun C. 2006. Identification of a drought tolerant introgression line derived from Dongxiang common wild rice (*O. rufipogon* Griff.). *Plant Molecular Biology* **62**: 247–259.

ZHOU S-X, TIAN F, ZHU Z-F, FU Y-C, WANG X-K, SUN C-Q. 2006. Identification of quantitative trait loci controlling drought tolerance at seedling stage in Chinese Dongxiang common wild rice (*Oryza rufijlogon* Griff.). *Acta Genetica Sinica* **33**: 551–558.