Supporting Information

Beuchat et al. 10.1073/pnas.0913207107

SI Materials and Methods

Plant Materials. The *Arabidopsis* accessions have been described previously (1). Seeds were stratified 2–4 days at 4 °C before transfer into a 16 h light/8 h dark cycle with ≈120 μE intensity on 0.5× Murashige and Skoog media containing 1% sucrose. Transgenic plants were generated according to standard procedures (2). For functional assays, typically ≥18 seedlings were measured per line. Plates were scanned on a flat-bed scanner, and images were processed using ImageJ software (version 1.36b) to determine root length at 9 days after germination. The *brx* loss of function lines, brx^S, *brx-2*, and Uk-1 have been described previously (2, 3). brx^S is an introgression of the natural Uk-1 *brx* null allele into the Sav-0 background; *brx-2* is a bona fide T-DNA insertion null mutant in the Col-0 background.

Molecular Biology. Molecular biology procedures, such as cloning of constructs, RT-PCR, quantitative PCR, sequencing, or genotyping, were carried out according to standard procedures (4). All transgenic constructs for expression of *BRX* family genes and variants under control of the *35S* promoter were created in the pMDC32 binary vector (5), except the *AtBRX-AtBRXL2* hybrid construct and the $AtBRX^{Eil\text{-}0}$ and $AtBRX^{Lc\text{-}0}$ constructs, which were cloned in pTCSH1 (2). For haplotyping, *AtBRX* DNA fragments were subcloned into standard plasmid vectors, and two clones per accession were sequenced. No ambiguities between clones were observed.

Phylogenetics, Molecular Evolution, and Population Genetics. Multiple sequence alignments of genomic or protein sequences were conducted in MUSCLE 3.7 with default settings (6). Phylogenetic trees of genomic and protein sequences were estimated using Neighbor-Joining (NJ) and Bayesian inference (BI). The NJ phylogeny was generated in Matlab, with pairwise distances between sequences corrected by the Jukes-Cantor model. The BI phylogenetic trees were estimated using MrBayes 3.1.2 (7) with two simultaneous Markov Chain Monte Carlo chains run for 1,000,000 generations and sample frequency of every 100 generations with burn-in = 2,500. Appropriate substitution models were selected using Modeltest 3.7 (8). The generalized time reversible evolutionary model (GTR; Nst = 6) was applied, and the prior probability distribution on the parameters of the model was set to be mixed for amino acid sequences. The haplotype sequences of *AtBRX* and *AtBRXL1* were aligned in ClustalW 2.0.9, and their BI phylogeny was estimated according to the Hasegawa-Kishino-Yano (Nst = 2) evolutionary model. Estimators of polymorphism were calculated as two sequence groups (worldwide and Umkirch) in SITES software (9).

On the basis of the aligned coding sequences the rates of nonsynonymous substitutions per nonsynonymous site, Ka, the rates of synonymous substitutions per synonymous site, Ks, and the resulting ratio Ka/Ks were estimated with the maximum likelihood model of sequence evolution as implemented in PAML software (10). Ka/Ks ratio was also computed with the maximum likelihood model in PAML (10) within a 60 amino acid window by sliding one codon at a time across the aligned 503 amino acid BRX proteins. Sixty amino acids is the smallest window size for Ka, Ks calculation on each window with nonzero sites. Adjustment of window size to 70, 80, or 90 amino acids gave similar results.

Genetic polymorphism data were analyzed using DnaSP software (version 5.10.00; September 07, 2009) (11). Briefly, genetic polymorphism based on low-frequency nucleotide polymorphisms was calculated as Watterson's estimator θ_W, and intermediate frequency polymorphisms as $\theta\pi$. A potential skew in the frequency spectrum of polymorphisms was evaluated by computing Fu and Li's *D*, Fay and Wu's *H*, and Tajima's *D* as implemented in DnaSP software. Genetic differentiation between populations was calculated as the average over all F_{ST} values between all worldwide accessions combined and the Uk accessions.

Pairwise linkage disequilibrium (R^2) between loci genotyped in the Eil-0 × Lc-0 recombinant inbred line population described previously (4) was calculated using GGT 2.0 software (12).

1. Shindo C, Bernasconi G, Hardtke CS (2008) Intraspecific competition reveals conditional fitness effects of single gene polymorphism at the *Arabidopsis* root growth regulator BRX. *New Phytol* 180:71–80.
2. Mouchel CF, Briggs GC, Hardtke CS (2004) Natural genetic variation in *Arabidopsis* identifies BREVIS RADIX, a novel regulator of cell proliferation and elongation in the root. *Genes Dev* 18:700–714.
3. Rodrigues A, et al. (2009) The short-rooted phenotype of the brevis radix mutant partly reflects root abscisic acid hypersensitivity. *Plant Physiol* 149:1917–1928.
4. Sibout R, Plantegenet S, Hardtke CS (2008) Flowering as a condition for xylem expansion in *Arabidopsis* hypocotyl and root. *Curr Biol* 18:458–463.
5. Curtis MD, Grossniklaus U (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol* 133:462–469.
6. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797.
7. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
8. Posada D, Buckley TR (2004) Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and Bayesian approaches over likelihood ratio tests. *Syst Biol* 53:793–808.
9. Hey J, Wakeley J (1997) A coalescent estimator of the population recombination rate. *Genetics* 145:833–846.
10. Yang Z, Nielsen R (2000) Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol Biol Evol* 17:32–43.
11. Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452.
12. van Berloo R (2008) GGT 2.0: Versatile software for visualization and analysis of genetic data. *J Hered* 99:232–236.

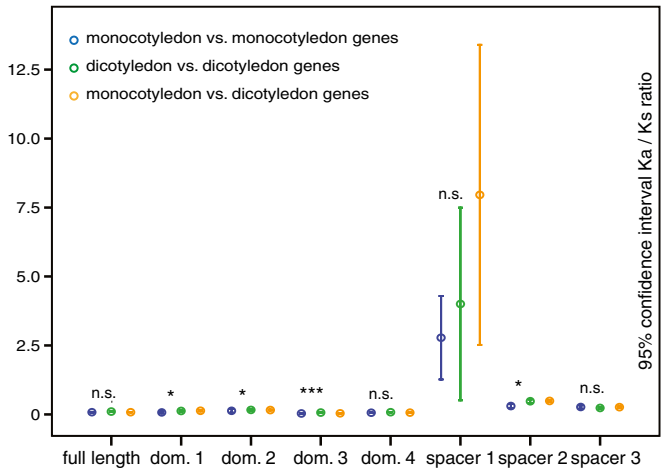

Fig. S1. Pairwise Ka/Ks calculated within mono- or dicotyledon proteins or between the two groups. Significance is indicated for intragroup comparison between Ka/Ks of monocotyledon and dicotyledons. dom., domain; n.s., not significant. *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$. Full statistics are given in Table S1.

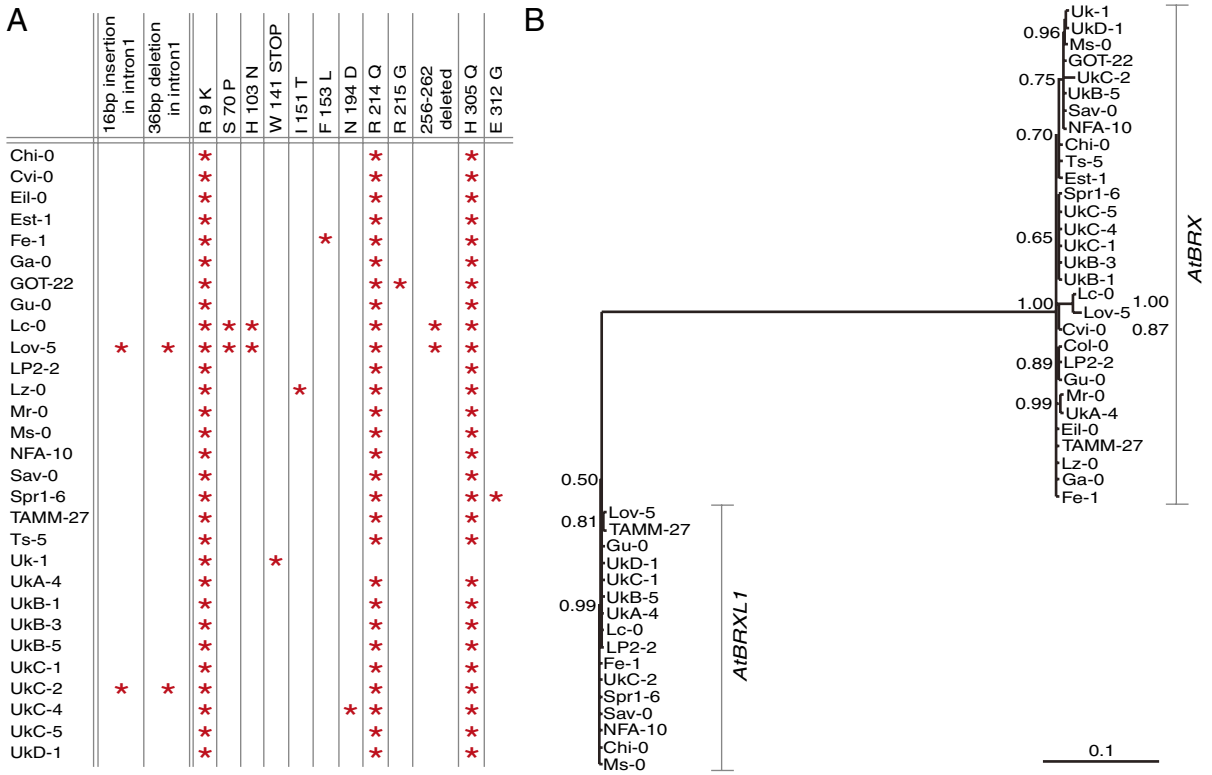

Fig. S2. Haplotype analysis of *AtBRX* and *AtBRXL1*. (*A*) Major features of *AtBRX* haplotype sequences in comparison with the Col-0 reference genome sequence. (*B*) Phylogenetic tree of the determined *AtBRX* and *AtBRXL1* haplotypes.

Table S1. Summary of statistical analysis of Ka/Ks

Table S1 (XLS)

Table S2. Tabular presentation of *AtBRX* haplotype polymorphisms in *Arabidopsis* accessions as compared with the Col-0 reference genome allele

Top labels indicate exons (E; shaded in blue) and introns (I), as well as the nucleotide position after ClustalW alignment of all sequences (maximum length: 2,268 bp). Bottom labels indicate the size of indels in bp (+, insertion; -, deletion)

Table S2 (XLS)

Table S3. List of additional accessions genotyped for the seven amino acid deletion observed in the Lc-0 and Lov-5 alleles of AtBRX

Table S3 (XLS)

Table S4. Average primary root length for Eil-0 × Lc-0 recombinant inbred lines, grouped according to *AtBRX* parental genotype

Table S4 (XLS)