Genotyping of Endosperms to Determine Seed Dormancy Genes Regulating Germination through Embryonic, Endospermic or Maternal Tissues in Rice

Xing-You Gu¹, Jinfeng Zhang², Heng Ye, Lihua Zhang, and Jiuhuan Feng

Plant Science Department, South Dakota State University, Brookings, SD 57007

²Present address: Plant Gene Engineering Lab, DBN Agr. Sci&Tech Research Institute, Beijing 100085, China
Running title: Genes for types of seed dormancy

Key Words: Seed dormancy, quantitative trait locus, endosperm, segregation distortion, rice

\(^1\)Corresponding author: Xing-You Gu; mailing address: Box 2140C, Brookings, SD 57007, USA; phone: +1(605) 688-6908; fax: +1(605) 688-4452; email: Xingyou.gu@sdstate.edu
Seed dormancy is imposed by one or more of the embryo, endosperm and maternal tissues that belong to two generations and represent two ploidy levels. Many quantitative trait loci (QTL) have been identified for seed dormancy as measured by gross effects on reduced germination rate or delayed germination in crop or model plants. This research developed an endosperm genotype-based genetic approach to determine specific tissues through which a mapped QTL regulates germination using rice as a model. This approach involves testing germination velocity for partially after-ripened seeds harvested from single plants heterozygous for a tested QTL and genotyping endosperms from individual germinated and non-germinated seeds with a co-dominant DNA marker located on the QTL peak region. Information collected about the QTL includes genotypic frequencies in germinated and/or non-germinated subpopulations; allelic frequency distributions during a germination period; endosperm or embryo genotypic differences in germination velocity; and genotypic frequencies for gametes involved in the double fertilization to form the sampled seeds. Using this approach, the seed dormancy loci \( SD_{12}, SD_{1-2}, \) and \( SD_{7-1} \) were determined to regulate germination through the embryo, endosperm and maternal tissues, respectively; \( SD_{12} \) and \( SD_{1-2} \) acted additively on germination velocity in the offspring tissues; and \( SD_{12} \) was also associated with the preferential fertilization of male gametes in rice. This new genetic approach can be used to characterize mapped genes/QTL for tissue-specific functions in endospermic seeds and for marker-assisted selection of QTL alleles before or immediately after germination in crop breeding.
INTRODUCTION

Seed dormancy developed on the mother plant (i.e., primary dormancy) is imposed by one or more of the embryo, endosperm and maternal tissues that belong to two (maternal & offspring) generations and represent two (diploid & triploid) ploidy levels. Genotypic variation of seed dormancy exists in natural populations and crop germplasm as an adaptive mechanism for seed-bearing plants to regulate the timing of germination. The natural variation in seed dormancy or germination-related traits has been associated with multiple quantitative trait loci (QTL) in barley (e.g., Ullrich et al. 1993), wheat (e.g., Anderson et al. 1993), rice (e.g., Lin et al. 1998), oats (Fennimore et al. 1999), sorghum (e.g., Lijavetzky et al. 2000), Arabidopsis (e.g., Alonso-Blanco et al. 2003), lettuce (e.g., Argyris et al. 2005), sunflower (e.g., Gandhi et al. 2005), rye (Masojć et al. 2007), oilseed rape (Schatzki et al. 2013), and peach (Blaker et al. 2013). Of the QTL reported in the listed and other research, only those in peach were detected based on marker-trait associations in the same generation (i.e., F2 seeds, Blaker et al. 2013), the remaining loci in all other species were claimed based on associations between plant genotypes and germination capabilities of seeds from the plants in mapping populations, which ignored the difference in generation or genotype between seed component tissues. Technically, it is difficult to genotype both plants and individual seeds from the plants in a mapping population. Statistical models, which combine maternal with offspring (embryo or endosperm) genotypes to map QTL associated with seed traits (e.g., Zhang et al. 2004), have not been used for research on seed dormancy. Thus, additional research is needed to determine specific tissues in which a QTL underlying gene is expressed to impose seed dormancy. This information is critical to further understand cellular and molecular
mechanisms of seed dormancy and is also useful to develop selection strategies to
manipulate germinability (e.g., the resistance to pre-harvest sprouting in cereal crops) in
breeding programs using the QTL alleles.

Experimental approaches that have been used to infer the involvement of a seed
component tissue in dormancy imposition can be grouped into somatic, molecular and
genetic categories. Somatic approaches have to resort to exercising embryos or physically
removing the maternal tissues testa and pericarp (also known as seed & fruit coats)
covering caryopses in grass species. The isolated embryos or naked caryopses were
germinated on a selected medium to infer the presence of embryo or coat-imposed
dormancy, as did in Arabidopsis, barley, oat, rice and wheat (Takahashi 1963, Morris et
various methods for gene tissue-specific expression analyses, such as RNA *in situ*
hybridization and *GUS* reporter assay, which require DNA/mRNA sequences of a known
gene to prepare probes (vectors) for hybridization (transformation). Using molecular
approaches, genes cloned from the rice *qLTG3-1* and *Sdr4* and the wheat *QPhs.ocs-3A1
*(TaMFT)* QTL were implicated to regulate dormancy or low-temperature germination in
approaches start with crosses between genotypes different in seed dormancy or QTL
allele to obtain hybrid F$_1$ or F$_2$ seeds. F$_1$ seed samples from reciprocal crosses were tested
for the difference in germinability to infer a maternal tissue effect on seed dormancy, as
reported for rice and wheat (Nair *et al.* 1965, Noll *et al.* 1982, Flintham 2000). The F$_2$
seed samples were partially after-ripened to test embryo genotypic frequencies for a QTL
marker in the germinated and/or non-germinated subpopulations; the QTL is inferred to
be involved in the maternal tissue-imposed dormancy if the genotypic/allelic frequencies keep constant between the subpopulations, or involved in the offspring tissue-imposed dormancy if the dormancy-reducing allele has a higher frequency in the germinated than in the non-germinated subpopulation (Gu et al. 2008). Using the embryo genotype-based genetic approach, the qSD7-1 and qSD12 QTL were associated with maternal and offspring tissue-imposed dormancies, respectively, in rice.

Advantages of the genetic over the somatic and molecular approaches are non-destructive testing and applicability to a large sample of seeds. Concerns about the genetic approaches are some uncertainties in inferences for tissue specificity or about interference by a linked segregation distortion locus (Gu et al. 2008). First, a maternal tissue effect on germination inferred by the difference between reciprocal F₁ s may be confounded with a cytoplasmic or endospermic effect, because hybrid seeds have female cytoplasm in all their cells and two of the three chromosomal sets in an endosperm cell from the female parent. The uncertainty about a cytoplasmic effect can be resolved by the embryo genotype-based genetic approach, as F₂ seeds self-pollinated from an F₁ plant are identical in cytoplasm. Second, the embryo genotype-based approach cannot distinguish an embryonic from an endospermic effect. An embryo is developed from a fertilized egg (2n) fused between the egg cell (n) in an embryo sac and one of the two sperms (n) in a pollen tube, while the endosperm is developed from the primary endosperm nucleus (3n) fused between two genetically identical polar nuclei (n) in the sac and the other sperm. Thus, from an F₁ plant heterozygous for dormancy locus (D/d), the F₂ seeds with the Dd genotype embryo have either DDd or Ddd genotype endosperm. Theoretically, if the dormancy gene expresses in the triploid tissue, the endospermic effect, which could be
confounded with the maternal tissue effect as inferred by the reciprocal F\textsubscript{1} approach or
cannot be detected by the embryo genotype-based approach, could be estimated by the
association between germination velocity of individual F\textsubscript{2} seeds and their endosperm
genotypes. In case that the gene expresses in the diploid offspring tissue, the endosperm
genotypes can be transformed into embryo genotypes to estimate the embryonic effect.

Third, genotypic frequencies for a dormancy gene in a germinated or non-germinated
subpopulation of partially after-ripened F\textsubscript{2} seeds can be also affected by a linked
segregation distortion locus (SDL). An SDL is often associated with preferential
dysfunction/fertilization of the male or female gamete (Lyttle 1991). Fortunately, gamete
genotypic frequencies of an SDL for both male and female sides can be estimated based
on frequencies of the four endosperm (but not the three embryo) genotypes in a random
sample of F\textsubscript{2} seeds. The estimates can help determine if a distorted segregation observed
in a germinated or non-germinated subpopulation arises from the selection for less or
more dormancy seeds, a linked SDL, or both.

This research was devoted to develop an endosperm genotype-based genetic approach
to characterize functions of previously mapped seed dormancy QTL in regulating
germination through the embryo, endosperm or maternal tissues. Genetic and seed
biology principles for the new approach are same as those described for the embryo
genotype-based approach (Gu \textit{et al.} 2008), except that genotyping information is
collected from endosperms, rather than embryos. It is relatively easy to genotype
embryos, because DNA samples for the genotyping can be prepared from embryonic
leaves or seedlings after a germination test and there are only three embryo genotypes for
a locus in an F\textsubscript{2} seed population of a diploid species. Thus, challenges to this research
would be techniques used to extract endospermic DNAs from individual seeds without a negative impact on standard germination testing and to display all four endosperm genotypes for a locus in an F\textsubscript{2} seed population. In this research, rice (*Oryza sativa* L.) was used as a model to develop a marker-genotyping system to meet the technical challenges, and the seed dormancy QTL \(SD_{1-2}\), \(SD_{7-1}\) and \(SD_{12}\), which were isolated as single Mendelian factors into the same genetic background in the previous research, were selected to demonstrate the efficacy of the new genetic approach. This article summarized techniques and analytic strategies for the endosperm genotype-based genetic approach, presented new information about the three selected loci, and discussed implications of some discoveries from this research.

### MATERIALS AND METHODS

**Parental lines and F\textsubscript{2} seed populations**

**Isogenic lines:** Four isogenic lines, including the recipient parent EM93-1, were used to develop hybrid F\textsubscript{1}s that are heterozygous for each of the \(SD_{1-2}\), \(SD_{7-1}\) and \(SD_{12}\) loci or both \(SD_{7-1}\) and \(SD_{12}\) (Table 1). EM93-1 is a line of cultivated rice (*Oryza sativa* subsp. *indica*) with a semi-dwarf plant height controlled by the *semidwarf1* (*sd\textsubscript{1}\textsuperscript{-2}* ) gene and is homozygous for the dormancy-enhancing allele at \(SD_{1-2}\) and the dormancy-reducing alleles at both \(SD_{7-1}\) and \(SD_{12}\). The isogenic lines (IL) were developed by introducing single chromosome (chr) segments from SS18-2, a line of wild-like weedy rice (*O. sativa*), into the recipient genetic background. Specifically, IL\textsubscript{sd\textsubscript{1-2}} has an introgression segment of \(~3000\) kilo base pairs (Kb) in physical length containing both \(Sd\textsubscript{1}\) and \(sd\textsubscript{1-2}\) on chr 1 (Ye *et al.* 2013); IL\textsubscript{SD\textsubscript{7-1}} has an introgression segment of \(~2\) Kb intragenic to the pleiotropic gene \(SD_{7-1}/Rc\) on chr 7 and the introgression converts the mutant allele in
EM93-1 \((sd_7, /rc)\) into a functional allele for both seed dormancy and red pericarp color
(Gu et al. 2011); and IL\(_{SD12}\) has an introgression segment of ~200 Kb containing the \(SD_{12}\) allele on chr 12 (Gu et al. 2010).

**F\(_2\) seed populations:** About 20 F\(_1\) hybrid plants from each of the four crosses were
grown in a greenhouse, verified for genotypes at \(SD_{1-2}\), \(SD_{7-1}\) and \(SD_{12}\) using markers on
the introgression segments (Table 1), and self-pollinated to produce F\(_2\) seed populations.
Seeds were harvested from individual F\(_1\) plants at 40 days (d) after flowering, air-dried at
the greenhouse for 3 d, and stored at a freezer (-20 °C) to maintain the primary dormancy.

**Seed after-ripening, germination and subpopulations**
Seed samples from the freezer were after-ripened at the room temperature (24-25 °C) for
0 to 35 d to release part of the primary dormancy prior to germination testing. The time
period of a partially after-ripening treatment varied with seed populations and was
determined in preliminary experiments to manipulate relative sizes of germinated and
non-germinates subpopulations. Two or more independent germination experiments were
conducted for each of the four F\(_2\) populations and an experiment consisted of 600-2000
fully developed seeds. About 60 seeds derived from an F\(_1\) plant were distributed in a 9-
cm Petri dish lined with a filter paper, soaked with 8-ml de-ionized water, and incubated
at 30 °C and 100% relative humidity in dark. Germination (radicle emerged >3 mm)
counting started at 48 h after imbibition and continued every 12 or 24 h for 7 or ~10 d.
Germinated seeds were transferred to new Petri dishes to collect the endosperm tissue.
All germinated seeds from an experiment were formed a germinated subpopulation, while
seeds that did not germinate and were not contaminated within 7 or 10 d were grouped as
a non-germinated subpopulation.
DNA micro-extraction and marker genotyping

For a germinated subpopulation, a newly germinated intact seed (spikelet) was cut in a cross section at 1/3 to the endosperm end. The embryo-less portion was cleaned by removing the maternal tissues and the endosperm tissue transferred into a 1.5 ml centrifuge tube and stored in a -20 °C freezer for DNA extraction. This method was also used to sample the endosperm tissue from non-germinated or dry seeds. In preliminary experiments to genotype both endosperm and embryo tissues, the sectioned seeds with an emerged radicle were transferred to 24-well cell culture plates (Corning® 15.6 mm diameter) lined with wetted filter papers and placed in an illuminated growth chamber for several days to collect embryonic leaves for DNA extraction.

Endospermic DNA was extracted using methods modified from Chunwongse et al. (1993) and Kang et al. (1998). Primarily, the endosperm tissue was incubated in 200 µl of lysis buffer (0.5% SDS & 20 µg Proteinase K) at 37 °C for 1 h and ground with a spatula. The crude sample was mixed with 400 µl of 2% CTAB solution [2% (w/v) cetyltrimethylammonium bromide, 100 mM Tris-HCl (pH 8.0), 20 mM EDTA (pH 8.0), 1.4 M NaCl, & 1% polyvinylpyrrolidone] and phase-separated with the 24:1 ratio of chloroform:isoamyl alcohol mixture. DNA in the aqueous phase was precipitated with cold isopropanol and washed with 70% ethanol. Air-dried DNA was dissolved in 50 µl TE buffer (10 mM Tris-HCl & 1 mM EDTA at pH 8.0) and treated with 1 µL RNaseA (10 mg/µl) to remove RNA. Alternatively, endosperm samples were placed in a 96-well plate and heated in 80 µl of lysis buffer (0.1 M NaOH) on a Thermocycler set at 95 °C for 15 min. The lysing solutions were mixed with 80 µl of neutralization buffer (10 mM Tris-HCl & 0.1 M HCl at pH2.0) and the supernatant was used as DNA templates for
PCR. DNA extraction from the embryonic leaves was conducted using the previously described method (Gu et al. 2008).

The simple sequence repeat (SSR) or insertion/deletion (InDel) markers on partial high-resolution maps for the $SD_{1.2}$ (Ye et al. 2013), $SD_{7.1}$ (Gu et al. 2011) or $SD_{12}$ (Gu et al. 2010) regions were selected to optimize a genotyping system. The selected markers are co-dominant, different in size by 10 or more bp between the two alleles, and capable to distinguish four endosperm genotypes of F$_2$ seeds by regular PCR and gel electrophoresis. The PCR amplification was performed using a 20-ml volume containing 40 ng of DNA template, 4 µl of 5× Green GoTaq$^\text{®}$ reaction buffer (Promaga), 200 µM of dNTP, 50 nM of each primer, and 2 units of Taq polymerase, in a BIO-RAD Thermocycler. The PCR program was initiated at 94 °C for 5 min., followed by 35 cycles of denature at 94 °C for 30 s, annealing at 55 °C for 30 s and elongation at 72 °C for 45 s, and ended with a final elongation at 72 °C for 7 min. PCR products were separated in a 6% non-denaturing polyacrylamide gel at ~300 V for ~3 h, and displayed and recorded using the AlphaEaseFC$^\text{TM}$ (Alpha Innotech) gel imaging system.

**Data analysis and genetic inferences**

**Genotypic and allelic frequencies:** Seeds (F$_2$) self-pollinated from a hybrid F$_1$ plant are identical for the maternal tissue genotype (Dd), but vary in embryo (DD, Dd & dd) and endosperm (DDD, DDD, Ddd & ddd) genotypes for a locus with two alleles that enhance (D) or reduce (d) seed dormancy (Table 2). The endosperm genotypic frequencies $F_{DDD}$, $F_{DDd}$, $F_{Ddd}$ and $F_{ddd}$ in a subpopulation or a random sample were calculated as described in Table 2, and used to estimate the overall allelic frequencies $F_D$ and $F_d$ in the sample and the genotypic frequencies for male ($F_{D^+} & F_{d^+}$) and female ($F_{D^-} & F_{d^-}$) gametes.
involved in the double fertilization to develop the F_2 seeds. The observed genotypic and allelic frequencies may or may not follow Hardy-Weinberg Equilibrium, depending on responses of the dormancy gene to the selection for germinated or non-germinated subpopulations and/or the presence or absence of a segregation distortion locus (SDL) in the gene-containing region (Gu et al. 2008). Thus, Chi-squared testing was used to determine the fitness of endosperm genotypic frequencies to the 0.25:0.25:0.25:0.25 expectation and binomial testing used to determine the fitness of allelic and gamete genotypic frequencies to the 0.5 expectation. Genetic inferences from the statistical tests were: 1) the gene was involved in offspring tissue-imposed dormancy and responded to the selection when genotypic frequencies deviate from the expectations with \( F_d \) higher in the germinated than in the non-germinated subpopulation; 2) the gene was involved in maternal tissue-imposed dormancy and did not responded to the selection when genotypic frequencies fit the expectations with \( F_d \) constant across subpopulations; and 3) there is a SDL in the dormancy gene-containing region when genotypic frequencies deviate from the expectations in the joined population of germinated and non-germinated seeds or in a random sample.

Binomial testing was also used to determine the equity or difference in genotypic frequency between male and female gametes (e.g., \( F_d^\prime = F_d^\prime \)). Standard error used for a test was calculated as \( \sqrt{2F_d \times F_d / N} \), where \( N \) is the number of genotyped seeds in a subpopulation or random sample. Significant differences in the tests were used to infer underlying mechanisms, such as a dosage effect of the dormancy gene on germination velocity in endosperms (\( F_d^\prime > F_d^\prime \)), or a differentiation of the SDL in preferential fertilization between male and female gametes (\( F_d^\prime > F_d^\prime \)).
Germination distributions: A sample of partially after-ripened seeds is characterized by germination heterogeneity, *i.e.*, some germinate earlier than others and some dormant seeds never germinate in an experiment. The heterogeneity occurs in a sample of seeds from a pure line (due to non-genetic factors) and a segregating (*e.g.*, F₂) population (due to both genetic and non-genetic factors). To quantify the heterogeneity in an experiment, the daily counted germination data from all samples were used to develop a germination distribution:
\[ y_j = \Sigma N_{gj} / N \times 100\% \]  
where, \( y_j \) is the cumulative germination rate at day \( j \) (\( j = 2 \) to 7 or 10), \( \Sigma N_{gj} \) is the summation of seeds germinated from day 2 to day \( j \), and \( N \) is the total number of seeds tested in the experiment. The germination distribution was compared with allelic frequency distributions in the germinated subpopulation to infer if the dormancy gene is involved in the regulation of germination through the offspring or maternal tissues.

For experiments genotyped for both germinated and non-germinated seeds, the germination distribution was calculated for individual endosperm genotypes:
\[ y_{ij} = \Sigma N_{gij} / (\Sigma N_{gij} + N_{ng} \times F_{ngi}) \times 100\% \]  
where, \( y_{ij} \) is the cumulative germination percentage for the \( i \)th endosperm genotype (\( i = 0, 1, 2 \) & 3, the copy number of \( D \) allele in the genotype) on day \( j \); \( \Sigma N_{gij} \) is the summation of the \( i \)th genotypic seeds germinated from days 2 to \( j \); \( N_{ng} \) is the total number of non-germinated seeds (including those with missing genotyping data) in the experiment; and \( F_{ngi} \) is the \( i \)th genotypic frequency in the non-germinated subpopulation. The difference in germination distribution pattern among the genotypes was used to infer if the gene is
involved in the genetic control of seed dormancy imposed by the embryo, endosperm or maternal tissues.

**Genetic effect estimation:** For genes involved in the regulation of germination through the embryo or endosperm tissue, their component effects in germinated subpopulations were estimated using linear regression models for embryo genotypes:

\[ y_{ik} = \mu + a x_i + d z_i + \varepsilon_{ik} \]  (3)

or for endosperm genotypes (Mo 1988):

\[ y_{ik} = \mu + a x_i + d_1 z_{i1} + d_2 z_{i2} + \varepsilon_{ik} \]  (4)

where: \( y_{ik} \) is the incubation time required to complete germination for seed \( k (k=1 \ to \ Ng) \), the number of genotyped germinated seeds) of the \( i \)th endosperm genotype; \( \mu \) is the model mean; \( x_i \) is the dummy variable for the additive component and is coded as -1, 0, and 1 for the embryo genotypes \( dd, Dd \) and \( DD \), respectively, or coded as -1.5, -0.5, 0.5 and 1.5 for the endosperm genotypes \( ddd, Ddd, DDd \), and \( DDD \), respectively; \( z_i \) is the dummy variable for the dominancy component and is coded as -0.5 for both \( dd \) and \( DD \) or 0.5 for \( Dd \); \( z_{i1} \) is the dummy variable for the 1st dominance component of the two \( D \) alleles over the \( d \) allele in the \( DDd \) genotype and is coded as 1 for \( DDD \) or 0 for the remaining three endosperm genotypes; \( z_{i2} \) is the dummy variable for the 2nd dominance component of the \( D \) allele over the two \( d \) alleles in the \( Ddd \) genotype and is coded as 1 for \( Ddd \) or 0 for the remaining three endosperm genotypes; \( a, d, d_1 \) and \( d_2 \) are partial regression coefficients for corresponding variables and the estimates of gene additive or dominance effects; and \( \varepsilon_{ik} \) is the residual effect of the model. Regression analysis was implemented using the REG procedure of SAS® 9.3 (SAS, 2011) with a stepwise selection set at the significance level of probability <5%.
For the germinated subpopulation segregating for both $SD_{7.1}$ and $SD_{12}$, two-way analysis of variance was used to detect their main and interactional effects on germination velocity. The variance analysis was conducted based on embryo genotypes using the two-factor factorial model:

\[ y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk} \]  

where, $y_{ijk}$ is the incubation time required to complete germination for the $k$th seed having the $i$th genotype at $SD_{7.1}$ and the $j$th genotype at $SD_{12}$, $\mu$ is the model mean; $\alpha_i$ and $\beta_j$ are main effects of $SD_{7.1}$ and $SD_{12}$, respectively; $(\alpha\beta)_{ij}$ is the interactional effect between the two loci; and $\epsilon_{ijk}$ is the error term of the model.

**RESULTS**

Reliability of the marker genotyping system to distinguish endospermic genotypes

The markers RID12, RM315/3602 and SD12m13, which were selected to tag $SD_{7.1}$, $SD_{1-2}$ and $SD_{12}$, respectively, could separate both embryo and endosperm genotypes in F$_2$ seed populations. As shown on gel images for RID12 (Fig. 1A), the two homozygous ($DDD$ & $ddd$) endosperms displayed one of the two alleles (bands), which are same as the homozygous ($DD$ & $dd$) embryos; whereas the two heterozygotes ($Ddd$ & $Ddd$) showed both bands (co-dominance) with one brighter than the other (dosage effect), which are different from the heterozygous ($Dd$) embryos that have the two bands equal in signal intensity. All the selected markers were consistent in gel image patterns (i.e., co-dominance & dosage effect) in F$_2$ seed populations (Fig. 1B). Of the 11 germination experiments, 9 had >90% germinated or non-germinated seeds genotyped with this marker-genotyping system, and the success rate was similar to that (93%) for a random sample of seeds segregating for $SD_{12}$ (Table 3). These results demonstrated that the
quality of endospermic DNAs from germinated seeds is good for genotyping with the
PCR-based markers and the genotypes are readily converted into embryo genotypes.

**SD$_{7.1}$ had neither endospermic nor embryonic effect on germination**

Three germination experiments were conducted for 0-(Ex. #1), 3-(Ex. #2) or 10-(Ex. #3)
day after-ripened (DAR) seeds derived from the hybrid F$_{1,SD_{7.1}}$ (Table 1). Germination
rate was 6% for Ex. #1 and 22% for Ex. #2, and germinated seeds were genotyped with
the marker RID12. In the germinated subpopulations, four endosperm genotypes fit the
expectation (Table 3) and frequencies for the dormancy-enhancing ($F_{SD_{7.1}}$) and -reducing
($F_{sd_{7.1}}$) alleles distributed around 0.5 during the germination period (Fig. 2A). Thus, the
selection for early germinated seeds did not alter the genetic equilibrium at SD$_{7.1}$, and
endosperm or embryo genotypic variation did not contribute to the phenotypic variation
in germination velocity.

Ex. #3 yielded 77% germinated seeds and both germinated and non-germinated seeds
were genotyped. Endosperm genotypic frequencies in the two subpopulations fit the
expectation (Table 3), which confirmed the observations in Ex. #1 and 2. In addition, the
four genotypes of seeds were same in germination distribution pattern (Fig. 3A),
indicating that they were identical in the degree of seed dormancy. Since SD$_{7.1}$ had
neither endospermic nor embryonic effect on germination in the three experiments, the
dormancy gene should express in the maternal tissue(s) during seed development.

**SD$_{1.2}$ was associated with an endospermic effect on germination**

Three germination experiments were conducted for 1-(Ex. #1 & 3) and 10-(Ex. #2) DAR
seeds derived from the hybrid F$_{1,SD_{1.2}}$ (Table 1). Germination rate was 25% for Ex. #1
and 55% for Ex. #2 at the 7th day after imbibition, and germinated seeds were genotyped
with the markers RM315 or RM3602. In the germinated subpopulation of Ex. #1, endosperm genotypic frequencies dramatically deviated from the expectation, with $F_{sd1-2sd1-2}$ (0.38) > $F_{SD1-2SD1-2sd1-2}$ (0.30) > $F_{SD1-2SD1-2SD1-2}$ (0.20) > $F_{SD1-2SD1-2SD1-2}$ (0.13) (Table 3), and the frequency of the dormancy-reducing allele $F_{sd1-2}$ varied from 0.81 to 0.64 during the germination period from days 2 to 7 (Fig. 2B). The genotypic segregation pattern and higher $F_{sd1-2}$ estimates in the germinated subpopulation indicate that $SD_{1-2}$ expressed in the offspring tissue. However, endosperm genotypic frequencies fit the 0.25:0.25:0.25:0.25 expectation (Table 3) and $F_{sd1-2}$ distributed around 0.5 during the germination period (Fig. 2B) in the germinated subpopulation of Ex. #2. Results from Ex. #2 suggested that the inhibitory effect of $SD_{1-2}$ on germination was released during the after-ripening period of 10 d and seed dormancy left in the sample could be due to the other factors in the genetic background.

Ex. #3 (24%) was a repeat of Ex. #1 and both germinated and non-germinated seeds were genotyped for endosperms. Segregation distortion was observed in the germinated ($F_{sd1-2}$=0.65) and non-germinated ($F_{sd1-2}$=0.45) subpopulations, but not in their joined population ($F_{sd1-2}$=0.50, Table 3). In the germinated subpopulation, patterns for both genotypic ($F_{sd1-2SD1-2sd1-2}$ > $F_{SD1-2SD1-2sd1-2}$ > $F_{SD1-2SD1-2SD1-2}$ > $F_{SD1-2SD1-2SD1-2}$) and allelic (Fig. 2B) frequency distributions were similar to those observed in Ex. #1. In contrast, the non-germinated subpopulation displayed an opposite genotypic frequency distribution pattern ($F_{sd1-2SD1-2sd1-2}$ < $F_{SD1-2SD1-2sd1-2}$ < $F_{SD1-2SD1-2SD1-2}$ < $F_{SD1-2SD1-2SD1-2}$), with the allelic frequency in favor of the dormancy-enhancing allele ($F_{SD1-2}$=0.55). In addition, the four endosperm genotypes differed from each other in germination velocity, with $sd1-2sd1-2SD1-2$ > $SD_{1-2sd1-2}$. 

The distinct genotypic difference indicates that $SD_{1.2}$ regulates germinability through the endosperm tissue.

It is noted that gamete genotypic frequencies for $sd_{1.2}$, which were estimated based on germinated subpopulations of Ex. #1 and 3, were greater ($P<0.025$) in the female ($F_{sd1.2} =$ 0.68-0.69) than in the male ($F_{sd1.2} =$ 0.57-0.60) (Table 3). This result indicates that endosperm genotypes of the early germinated seeds consisted of more $sd_{1.2}$ allele donated from the female than from the male gametes. The difference between $F_{sd1.2}''$ and $F_{sd1.2}'$ has nothing to do with allelic differentiation in gamete fertility or preferential fertilization, because gamete genotypic frequencies in the joined population of Ex. #3 fit the Mendelian expectation (Table 3). In fact, the difference between $F_{sd1.2}''$ and $F_{sd1.2}'$ is determined by frequencies for the two heterozygous endosperm genotypes (refer to formula in Table 2), which are different in the copy number (or dose) of the dormancy-reducing or -enhancing allele. In both of the germinated subpopulations, the frequency was greater for $SD_{1.2}sd_{1.2}SD_{1.2}sd_{1.2} (~0.3)$, which has two germination-promoting alleles from the female gametes, than for $SD_{1.2}SD_{1.2}sd_{1.2} (~0.2)$, which has one germination-promoting allele from the male gametes (Table 3). Therefore, the observed difference between $F_{sd1.2}''$ and $F_{sd1.2}'$ can be accounted for by a dosage effect of the $sd_{1.2}$ ($SD_{1.2}$) allele in endosperms on germination promotion (inhibition).

It was estimated based on model (4) for endosperm genotypes that the $SD_{1.2}$ locus consisted of only an additive effect on the time to germination (0.31 to 0.47 d) in the germinated subpopulations of Ex. #1 and 3 (Table 4).

$SD_{1.2}$ was associated with an embryonic effect on germination and a genetic differentiation in gamete preferential fertilization.
Three germination experiments were conducted for 7-(Ex. #1), 10-(Ex. #3), or 14-(Ex. #2) DAR seeds derived from the hybrid $F_{1\_SD12}$ (Table 1). Germination rate was 20% for Ex. #1 and 36% for Ex. #2 at the 7th day after imbibition, and the germinated seeds were genotyped with the marker SD12m13. In germinated subpopulations of the two experiments, endosperm genotypic frequencies dramatically deviated from the expectation (Table 3) and the frequency of the dormancy-reducing allele $F_{sd12}$ varied from 0.68 at day 3 to 0.60 at day 7 (Fig. 2C). The $F_{sd12}$ estimates reduced gradually with the increase in incubation time in these two experiments, indicating that the genotypic variation partly contributed to the germination heterogeneity and the dormancy gene expressed in the offspring tissue.

Germination rate was 26% for Ex. #3 at the 7th day after imbibition and both germinated and non-germinated (50%) seeds were genotyped. Segregation distortion for endosperm genotypes was observed in the germinated but not in the non-germinated subpopulation (Table 3) and the $F_{sd12}$ distribution pattern during the germination period was similar to those in the first two experiments (Fig. 2C). In addition, the four endosperm genotypes of seeds displayed three germination distribution patterns, with the two heterozygotes ($SD_{12sd12sd12} & SD_{12SD_{12sd12}}$) similar in germination velocity (Fig. 3C). The association between germination distribution patterns and embryo genotypes is an indication that $SD_{12}$ is involved in embryo dormancy. It was estimated based on model (3) for embryo genotypes that the $SD_{12}$ locus consisted of only an additive effect on the time to germination (0.22 to 0.53 d) in the germinated subpopulations of Ex. #1, 2 and 3 (Table 4).
It is noted that gamete genotypic frequencies for the dormancy-reducing allele $sd_{12}$ in
the joined population of Ex. #3 deviated from 0.5 for the male ($F_{sd_{12}} = 0.588$) but not for
the female ($F_{sd_{12}'} = 0.50$) gametes (Table 3). The deviation suggests that there could be a
segregation distortion locus (SDL) in the $SD_{12}$-containing region affecting the
fertilization of male gametes. To prove the hypothesis, a random sample of 451 F$_2$
was genotyped (Ex. #4). Four endosperm genotypes in this sample also deviated from the
expected equal frequency, with the allelic frequency in favor of $sd_{12}$ in the male
($F_{sd_{12}} = 0.57$), but not in the female ($F_{sd_{12}'} = 0.52$) gametes (Table 3). The genetic
disequilibrium observed in the sizable random sample clearly indicates that the $SD_{12}$
locus is associated with a genetic differentiation in fertilization capability and male
gametes with the dormancy-enhancing allele $SD_{12}$ tended to be less competitive in
fertilization. The associated effect on gamete preferential fertilization contributed only
part to the genetic disequilibrium in the germinated subpopulations, because the $F_{sd_{12}'}$
estimates (0.61-0.67) in the above-stated three experiments were numerically greater than
that (0.57) in the random sample.

**Selection for early or late germination broke the genetic equilibrium for $SD_{12}$, but
not for $SD_{7-1}$, in seed populations segregating for both loci**

Two germination experiments were conducted for 14-(Ex. #1) or 35-(Ex. #2) DAR seeds
derived from the dihybrid $F_{1,SD_{7-1}SD_{12}}$ (Table 1). Germination rate for the 14-DAR seed
samples was 29%, which was lower than those for the 10- and 14-DAR seed samples
segregating only for $SD_{7-1}$ (77%) and $SD_{12}$ (36%), respectively. The reduced germination
rate indicates that pyramiding of the dormancy-enhancing alleles at $SD_{7-1}$ and $SD_{12}$
lengthened the dormancy duration. Germinated seeds in Ex. #1 were genotyped for
endosperms with the markers RID12 and SD12m13. In the germinated subpopulation, the joint frequency distribution for the two loci was dramatically biased in favor of the four genotypes homozygous (sd\textsubscript{12}sd\textsubscript{12}sd\textsubscript{12}) for the dormancy-reducing allele at SD\textsubscript{12}, the two (SD\textsubscript{12}SD\textsubscript{12}sd\textsubscript{12} & SD\textsubscript{12}sd\textsubscript{12}sd\textsubscript{12}) groups of genotypes heterozygous for the SD\textsubscript{12} locus were similar in frequency (Fig. 4A), and allelic frequency distributions displayed two patterns, i.e., $F_{sd12}$ varied from 0.74 to 0.68 while $F_{sd7-1}$ was constant around 0.5 during the germination period (Fig. 4B), which were similar to those observed in the germinated subpopulations segregating only for the SD\textsubscript{7-1} or SD\textsubscript{12} locus (Fig. 2A & C). These results indicate that SD\textsubscript{12} played a major role in regulating germinability through the embryo, not endosperm tissue in the digenic system.

It appeared that SD\textsubscript{7-1} also interacted with SD\textsubscript{12} in embryos to influence germination. For example, the variation in the time to germination among the three SD\textsubscript{12} genotypes was smaller in the sd\textsubscript{7-1}sd\textsubscript{7-1} than in the SD\textsubscript{7-1}sd\textsubscript{7-1} or SD\textsubscript{7-1}SD\textsubscript{7-1} backgrounds in Ex. #1 (Fig. 4D). However, the two-way analysis of variance based on model (5) revealed that only the main effect of SD\textsubscript{12} was significant (see Table S1). The main effect was contributed by the additive component of the SD\textsubscript{12} locus, based on model (3) for embryo genotypes (Table 4).

Germination rate was 81% for Ex. #2 and non-germinated seeds were identified for endosperm genotypes at SD\textsubscript{7-1} and SD\textsubscript{12}. A relatively high selection strength (as indicated by the germination rate) for non-germinated seeds in this experiment is expected to suppress less dormant genotypes in the subpopulation. As expected, the joint frequency distribution for SD\textsubscript{7-1} and SD\textsubscript{12} was biased in favor of the four genotypes that are homozygous (SD\textsubscript{12}SD\textsubscript{12}SD\textsubscript{12}) for the dormancy-enhancing allele at SD\textsubscript{12}, in particular the
digenic genotype \((sd_{7.1}sd_{7.1}sd_{7.1}SD_{12}SD_{12}SD_{12})\) (Fig. 4C). Segregation distortion was detected for \(SD_{12}\) but not for \(SD_{7.1}\), with the \(F_{sd12}\) estimate (0.35) lower than the Mendelian expectation (0.5) (Table 3). The results from the non-germinated subpopulation support that the \(SD_{12}\)-controlled embryo dormancy played a major role in inhibiting germination, and also suggest that \(SD_{7.1}\) may influence embryo dormancy as well in the digenic system, as implied by the observation in the germinated subpopulation of Ex. #1 (Fig. 4D).

In addition, a numerical difference in genotypic frequency for the \(SD_{12}\) locus between male and female gametes, as estimated based on the non-germinated subpopulation, was observed \((F_{sd12}' < F_{sd12}'' \text{ or } F_{SD12}' > F_{SD12}'')\), which was opposite to the observation in the germinated subpopulation of Ex. #1 \((F_{sd12}' > F_{sd12}'')\) (Table 3). The observations in the digenic system also confirmed that the associated effect of \(SD_{12}\) on male gamete preferential fertilization contributed only part to the genetic disequilibrium in germinated subpopulations segregating for the major seed dormancy locus.

**DISCUSSION**

**Applications of the endosperm genotype-based genetic approach**

An endosperm genotype-based genetic approach was developed to determine whether previously mapped seed dormancy genes/QTL regulate germination through the embryo, endosperm or maternal tissues. This approach involves techniques: 1) partially after-ripening a segregating population of seeds to separate relatively less dormant from more dormant genotypes by germination testing, 2) sampling endospermic DNAs from individual germinated and non-germinated seeds, and 3) identifying all endosperm genotypes using regular PCR-based co-dominant markers. After DNA sampling, the
newly germinated seeds can be selected to develop progeny lines. Thus, the techniques, similar to those reported for pre-germination screening of genotypes (Chunwongse et al. 1993, Kang et al. 1998), can be also used for early selection of QTL alleles in breeding programs. Information collected using this approach includes: 1) endosperm and embryo genotypic frequencies for tested loci in germinated and non-germinated subpopulations (Table 3); 2) association between incubation times required for individual seeds to complete germination and allelic frequencies in the germinated subpopulation (the association was more informative before the 7th day, as shown in Figs. 2B & C); 3) genotypic differences for the degree of dormancy in a segregating population of seeds (Fig. 4); 4) the dosage effect for a dormancy gene with an additive effect on germination in endosperms; and 5) genotypic frequencies for both male and female gametes that were involved in the double fertilization to form the seed sample. Segregation distortion loci were frequently reported for mapping populations (Xu et al. 1997, Faris et al. 1998); the endosperm genotype-based genetic approach can be also used to characterize SDL and artificial mutants for functions in gamete development and fertilization.

The genetic approach demonstrated in rice can be extended to the other cereal crops to characterize tissue-specific functions of mapped genes/QTL controlling seed dormancy and some other seed traits. Before DNA markers became prevalent for QTL mapping, Robertson (1952) combined morphological markers with an A-B (normal-supernumerary chromosome) translocation system to determine the tissue-specific function of the maize viviparous-5 mutant. The mutant distinguishes itself from the wild type by white endosperm and albino embryonic leaves and was concluded to regulate germination through the embryo, not the endosperm tissue. The A-B system was used in the reported
research to help separate the four endosperm genotypes, because the morphological markers are dominant and have only two phenotypes for a locus. Nowadays, many co-dominant DNA markers can be selected to genotype endosperms in many species. We also tried a quantitative real-time PCR method to genotype endosperms with a dominant DNA marker in the SD$_{1-2}$ region (data not shown) and found out that genotyping with regular PCR for a co-dominant marker was more reliable and cost-effective.

Compared with the endosperm genotype-based genetic approach, the embryo genotype-based genetic approach (Gu et al. 2008) provides less information, but it is relatively easy to conduct and can be used to distinguish an embryo from a maternal tissue effect on germination for a dormancy gene. If a seed dormancy gene functions in both maternal and offspring tissues, only the inhibitory effect on germination expressed in the offspring tissue(s) can be detected by the endosperm or embryo genotype-based genetic approach, because F$_2$ seeds from single F$_1$ plants do not segregate for the maternal tissue genotype (Table 2). For this case, a joint maternal-offspring model (e.g., Zhang et al. 2004) could be used to estimate effects of a seed dormancy gene on germination through different component tissues.

**Genetic controls for types of seed dormancy**

Natural variation for such a complex, adaptive trait as seed dormancy involves multiple genetic controls in different seed component tissues. This research provided evidence that SD$_{12}$, SD$_{1-2}$ and SD$_{7-1}$ are involved in the genetic controls of seed dormancy through the embryo, endosperm and maternal tissues, respectively. SD$_{12}$ has the largest effect on seed dormancy in rice and its allele from SS18-2 delays germination (Gu et al. 2004). The major effect of SD$_{12}$ on germination was contributed mainly by the additive component in
embryos (Table 4) and contributed to the germination heterogeneity of seeds on the
heterozygous \((SD_{12}sd_{12})\) plants. Therefore, marker-assisted selection for \(SD_{12}\) alleles can
be conducted before (or after) germination using the endosperm (or embryo) genotype-
based genetic approach, and phenotypic selection for an early or late germinated
subpopulation from a segregating population of partially after-ripened seeds can be used
to enrich the dormancy-reducing or -enhancing allele. \(SD_{12}\) was also associated with the
preferential fertilization of male gametes on the heterozygous plants. The associated
effect is defined as the preferential fertilization, not gamete fertility, because plants
homozygous or heterozygous for the \(SD_{12}\) dormancy-enhancing allele had a normal seed
setting rate (>95%) in this and previous research (Gu et al. 2004 & 2008). Both the major
effect on embryo dormancy and the association with gamete preferential fertilization
must have restrained the distribution of the dormancy-enhancing allele during rice
domestication. This also explains why the \(SD_{12}\) locus was not detected in the other
mapping populations reported so far. Further research will elucidate if the association
arises from a pleiotropic effect of \(SD_{12}\) or a gene linked to \(SD_{12}\).
\(SD_{1-2}\) could be the first naturally occurring gene that has been unambiguously
associated with endosperm-imposed dormancy. Two analytic strategies were used to
recognize \(SD_{1-2}\)’s endospermic effect on germination: 1) a distinct difference in
germination rate among the four endosperm genotypes in a segregating population of
partially after-ripened seeds (Fig. 3B), and 2) a significant difference in genotypic
frequency between female and male gametes estimated based on the germinated
subpopulation \((F_{sd1-2} \sim F_{sd1-2'}\) in Table 3). It appeared that the inhibitory effect of \(SD_{1-2}\) on
germination lasts only for several days after seed maturation or harvest. The previous
research delimited \(SD_{1.2}\) to a short genomic region encompassing the gene \(Sd1\) (or \(GA20ox-2\)) (Table 1, Ye et al. 2013). \(GA20ox-2\) encodes GA20-OXIDASE2 catalyzing the second-to-last step of the gibberellin (GA) biosynthesis and the loss-of-functional mutation (\(sd1\)) reduces the GA level in vegetative tissues and plant height (Ashikari et al. 2002; Monna et al. 2002; Spielmeyer et al. 2002). The dormancy-enhancing (\(SD_{1.2}\)) and -reducing (\(sd_{1.2}\)) alleles at the \(SD_{1.2}\) QTL couple with the GA functional and loss-of-functional alleles, respectively, and \(SD_{1.2}\)’s effect on inhibiting germination could be compensated by GA application (Ye et al. 2013). The endosperm aleuronic cell layer is known to plays a central role in GA signaling to synthesize hydrolytic enzymes for cell wall weakening and food reserve mobilization during and post germination (Steber 2007). It is likely that \(GA20ox-2\) is the QTL underlying gene. Research is being conducted to prove the hypothesis and to identify mechanisms for dormancy development and germination by \(SD_{1.2}\) expressed in endosperms.

\(SD_7.1\) was confirmed to control maternal tissue-imposed dormancy using the perfect isogenic system (IL\(SD_7.1\) is heterozygous only for a 2-kb \(SD_7.1\) intragenic segment, Table 1). The \(SD_7.1\) functional allele was isolated from the “red” weedy rice line SS18-2 and encodes a transcription factor that promotes the abscisic acid (ABA) biosynthesis in early development seeds to induce primary dormancy and also activates the flavonoid biosynthesis pathway in the lower epidermal cell layer of the pericarp tissue to produce red pigments (Gu et al. 2011). Since an endospermic or embryonic effect on germination was not detected in the isogenic background (Figs. 2A & 3A), the \(SD_7.1\)-regulated dormancy-inducing events must also occur in the maternal tissue. \(SD_7.1\) and \(SD_{12}\) control seed dormancy through different tissues, but their effects on delaying germination is
cumulative, because it took a longer after-ripening period for seeds from the digenic
system (F_{1\_SD7-1SD12} plants) to reach similar germination rates than seeds from the
monogenic systems (F_{1\_SD7-1} & F_{1\_SD12} plants). Questions to be addressed include how
genes functioning in different tissues work in the same genetic system to lengthen the
dormancy duration, and if a gene controlling coat-imposed dormancy is also involved in
embryo dormancy, and *vice versa*.

**ACKNOWLEDGEMENTS**

Funding for this research was supported by grants from Natural Science Foundation (IOS
1021382 & 0641376) and United States Department of Agriculture-National Research
Initiative (2008-35301-19058) and -National Institute of Food and Agriculture (2013-
03572), and by South Dakota Agricultural Extension Station and Drought Center.

**LITERATURE CITED**

Analysis of natural allelic variation at seed dormancy loci of *Arabidopsis
thaliana*. Genetics 164: 711-729.

regions associated with resistance to pre-harvest sprouting in wheat. Crop Sci. 33:
453-459.

loci associated with seed and seedling traits in Lactuca. Theor. Appl. Genet. 111:
1365-1376.


Robertson, D. S., 1952 The genotype of the endosperm and embryo as it influences vivipary in maize. Genetics 38: 580-583.


## Table 1 Genotypic and phenotypic information about isogenic lines and hybrid F₁s used to develop F₂ seed populations

<table>
<thead>
<tr>
<th>Parental line or F₁</th>
<th>Genotype</th>
<th>Segment/marker</th>
<th>Pericarp color</th>
<th>Plant height</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. IL₅₁₂-2 (EM93-1)</td>
<td>SD₁₂SD₅₁₂sd₇₁sd₇₁sd₁₂sd₁₂</td>
<td>Recipient</td>
<td>White (rcrc)</td>
<td>Semidwarf (sd₁sd₁, ~80 cm)</td>
</tr>
<tr>
<td>B. IL₅₁₂</td>
<td>sd₁₂sd₁₂sd₇₁sd₇₁sd₁₂sd₁₂</td>
<td>~3000 Kb</td>
<td>White (rcrc)</td>
<td>Tall (Sd₁Sd₁, ~100 cm)</td>
</tr>
<tr>
<td>C. IL₅₇-1</td>
<td>SD₁₂SD₁₂SD₇₁SD₇₁sd₁₂sd₁₂</td>
<td>~2 Kb</td>
<td>Red (RcRc)</td>
<td>Semidwarf (sd₁sd₁)</td>
</tr>
<tr>
<td>D. IL₅₁₂</td>
<td>SD₁₂SD₁₂sd₇₁sd₇₁SD₁₂SD₁₂</td>
<td>~200 Kb</td>
<td>White (rcrc)</td>
<td>Semidwarf (sd₁sd₁)</td>
</tr>
<tr>
<td>F₁₅₁₂ (A×B)</td>
<td>SD₁₂sd₁₂sd₇₁sd₇₁sd₁₂sd₁₂</td>
<td>RM315/3602</td>
<td>White (rcrc)</td>
<td>Tall (Sd₁sd₁, ~90 cm)</td>
</tr>
<tr>
<td>F₁₅₇-1 (A×C)</td>
<td>SD₁₂SD₁₂SD₇₁SD₇₁sd₁₂sd₁₂</td>
<td>RID12</td>
<td>Red (RcRc)</td>
<td>Semidwarf (sd₁sd₁)</td>
</tr>
<tr>
<td>F₁₅₁₂ (A×D)</td>
<td>SD₁₂SD₁₂sd₇₁sd₇₁SD₁₂sd₁₂</td>
<td>SD₁₂m13</td>
<td>White (rcrc)</td>
<td>Semidwarf (sd₁sd₁)</td>
</tr>
<tr>
<td>F₁₅₇-1SD₁₂ (C×D)</td>
<td>SD₁₂SD₁₂SD₅₁₂SD₇₁SD₁₂sd₁₂</td>
<td>RID12 &amp; SD₁₂m13</td>
<td>Red (RcRc)</td>
<td>Semidwarf (sd₁sd₁)</td>
</tr>
</tbody>
</table>

**a** EM93-1 was the recipient of single introgression segments from a line of “red” weedy rice in isogenic lines (IL) for the SD₁₂, SD₇₁ or SD₁₂ locus. Upper or lower case letters indicate dormancy-enhancing (SD) or -reducing (sd) alleles at the three loci.

**b** Physical lengths of the introgression segments in kilo base pairs (Kb) or DNA markers selected to tag the loci.

**c** Rc, red pericarp color gene, which belongs to the same locus as SD₇₁, with the functional allele responsible for both red pigment and enhanced seed dormancy (Gu et al. 2011).

**d** sd₁, semidwarf1 gene located on the SD₁₂-containing region, with the EM93-1-derived allele responsible for both reduced plant height and enhanced seed dormancy (Ye et al. 2013).
Table 2 List of genotypes for seed component tissues and endosperm genotypic frequencies in an F₂ seed population segregating for a dormancy locus (\(D/d\))

<table>
<thead>
<tr>
<th>Female gamete</th>
<th>Male gamete</th>
<th>Genotypic frequency&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>d (0.5)</td>
<td>D (0.5)</td>
</tr>
<tr>
<td>d (0.5)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>ddd (dd) [Dd]&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Ddd (Dd) [Dd]</td>
</tr>
<tr>
<td></td>
<td>(F_{dd} = N_{dd}/N ) (0.25)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>(F_{Ddd} = N_{Ddd}/N ) (0.25)</td>
</tr>
<tr>
<td>D (0.5)</td>
<td>DDd (Dd) [Dd]</td>
<td>DDD (DD) [Dd]</td>
</tr>
<tr>
<td></td>
<td>(F_{DDD} = N_{DDD}/N ) (0.25)</td>
<td>(F_{DDD} = N_{DDD}/N ) (0.25)</td>
</tr>
</tbody>
</table>

Genotypic frequency<sup>b</sup>

\[
F_d' = F_{ddd} + F_{Ddd} \\
F_D' = F_{Ddd} + F_{DDD}
\]

Overall allelic frequency

\[
F_d = F_{ddd} + \frac{1}{2}F_{Ddd} + \frac{1}{2}F_{DDd} \\
F_D = F_{DDD} + \frac{1}{2}F_{DDd} + \frac{1}{2}F_{Ddd}
\]

<sup>a</sup> Genotypes for the endosperm (triploid), embryo (parentheses) and maternal (brackets) tissues of a seed at a dormancy locus with the two functionally differentiated alleles \(D\) and \(d\). \(F_{ddd, Ddd, DDD}\) and \(F_{DDD}\) are genotypic frequencies for the \(ddd, Ddd, DDD\) and \(DDD\) endosperms, respectively, estimated based on the number of seeds for individual genotypes \((N_{ddd}, N_{Ddd}, N_{DDD} \& N_{DDD})\) and the population size \((N)\).

<sup>b</sup> Genotypic frequencies for male \((F_d' \& F_D')\) and female \((F_d'' \& F_D'')\) gametes involved in the double fertilization to form the seed population.

<sup>c</sup> The value in the parentheses is the Mendelian expectation for the endosperm (0.25) or gamete (0.5) genotypic frequency, or the overall allelic frequency (0.5).
### Table 3 Summary of genotypic and allelic frequencies for the $SD_{71}, SD_{12}$ or $SD_{12}$ locus in F$_2$ seed subpopulations or joined populations

<table>
<thead>
<tr>
<th>Experiment$^a$</th>
<th>Subpopulation (genotyped seeds)$^b$</th>
<th>Endosperm genotypic frequency$^c$</th>
<th>$\chi^2$ value (probability)$^c$</th>
<th>Allelic/gametic frequency$^d$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$F_{dd}$</td>
<td>$F_{Dd}$</td>
<td>$F_{DD}$</td>
</tr>
<tr>
<td><strong>$SD_{71}$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex. $^1$ (720, 0 DAR, 6%)</td>
<td>G (43, 100%)</td>
<td>0.256</td>
<td>0.256</td>
<td>0.256</td>
</tr>
<tr>
<td>Ex. $^2$ (810, 3 DAR, 22%)</td>
<td>G (176, 98%)</td>
<td>0.256</td>
<td>0.216</td>
<td>0.307</td>
</tr>
<tr>
<td>Ex. $^3$ (631, 10 DAR, 77%)</td>
<td>G (481, 99%)</td>
<td>0.262</td>
<td>0.247</td>
<td>0.249</td>
</tr>
<tr>
<td>NG (138, 93%)</td>
<td>0.283</td>
<td>0.232</td>
<td>0.246</td>
<td>0.239</td>
</tr>
<tr>
<td>G+NG (619)</td>
<td>0.267</td>
<td>0.244</td>
<td>0.249</td>
<td>0.241</td>
</tr>
<tr>
<td><strong>$SD_{12}$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex. $^1$ (1091, 1 DAR, 25%)</td>
<td>G (250, 92%)</td>
<td>0.376</td>
<td>0.296</td>
<td>0.200</td>
</tr>
<tr>
<td>Ex. $^2$ (1309, 10 DAR, 55%)</td>
<td>G (674, 93%)</td>
<td>0.261</td>
<td>0.236</td>
<td>0.242</td>
</tr>
<tr>
<td>Ex. $^3$ (1096, 1 DAR, 24%)</td>
<td>G (280, 100%)</td>
<td>0.396</td>
<td>0.300</td>
<td>0.207</td>
</tr>
<tr>
<td>NG (801, 99%)</td>
<td>0.205</td>
<td>0.236</td>
<td>0.257</td>
<td>0.302</td>
</tr>
<tr>
<td>G+NG (1081)</td>
<td>0.254</td>
<td>0.253</td>
<td>0.244</td>
<td>0.249</td>
</tr>
<tr>
<td><strong>$SD_{12}$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex. $^1$ (1980, 7 DAR, 20%)</td>
<td>G (381, 98%)</td>
<td>0.367</td>
<td>0.234</td>
<td>0.257</td>
</tr>
<tr>
<td>Ex. $^2$ (639, 10 DAR, 34%)</td>
<td>G (215, 99%)</td>
<td>0.381</td>
<td>0.195</td>
<td>0.228</td>
</tr>
<tr>
<td>Ex. $^3$ (633, 14 DAR, 26%)</td>
<td>G (161, 99%)</td>
<td>0.372</td>
<td>0.186</td>
<td>0.298</td>
</tr>
<tr>
<td>NG (234, 50%)</td>
<td>0.269</td>
<td>0.209</td>
<td>0.291</td>
<td>0.231</td>
</tr>
<tr>
<td>G+NG (395)</td>
<td>0.544$^{ns}$</td>
<td>0.588$^{ns}$</td>
<td>0.499$^{ns}$</td>
<td>0.499$^{ns}$</td>
</tr>
<tr>
<td>Ex. $^4$ (a random sample of 484 seeds)</td>
<td>0.287</td>
<td>0.231</td>
<td>0.285</td>
<td>0.196</td>
</tr>
</tbody>
</table>

| **$SD_{71}$ & $SD_{12}$** |                                       |         |         |         |         |       |         |         |
| Ex. $^1$ (1529, 14 DAR, 29%) | G (436, $SD_{12}$) | 0.477 | 0.181 | 0.216 | 0.126 | 127 (<0.0001) | 0.675$^{***}$ | 0.693$^{***}$ | 0.661$^{***}$ |
| G (436, $SD_{71}$) | 0.225 | 0.239 | 0.280 | 0.257 | 2.97 (0.396) | 0.484$^{ns}$ | 0.504$^{ns}$ | 0.463$^{ns}$ |
| Ex. $^2$ (1019, 35 DAR, 81%) | NG (121, $SD_{12}$) | 0.149 | 0.231 | 0.165 | 0.454 | 29 (<0.0001) | 0.347$^*$ | 0.314$^*$ | 0.380$^{**}$ |
| NG (121, $SD_{71}$) | 0.314 | 0.256 | 0.198 | 0.231 | 3.46 (0.326) | 0.541$^{ns}$ | 0.512$^{ns}$ | 0.570$^{ns}$ |

$^a$ Listed in the parentheses are the total number of seeds received the number of days of after-ripening (DAR) treatment prior to the germination test and the mean germination rate at the 7th day after imbibition.
Listed in the parentheses are the number and proportion of genotyped seeds from the germinated (G) or non-germinated (NG) subpopulation.

The letters D or d in subscripts represent dormancy-enhancing or -reducing alleles in endosperm genotypes, which were tested against the 0.25:0.25:0.25:0.25 expectations.

The overall allelic frequency (\(F_d\)) and the male (\(F_d'\)) and female (\(F_d''\)) gamete frequencies for the \(d\) allele are defined in Table 2. The estimates for \(SD_{12}\) in the joined population (G+NG in \(SD_{12}\) Ex. *3) are means weighted by the germination rate 26%. The superscripts indicate that the difference of the estimate from 0.5 was not significant at \(P=0.05\) (**), \(P<0.01\) (***) or \(P<0.001\) (***) and the underlined estimates indicate \(F_d' > F_d''\) for \(SD_{1-2}\) or \(F_d'' > F_d'\) for \(SD_{12}\) at the significant level of \(P<0.05\).
Table 4 Summary of estimated gene component effects of the \(SD_{1-2}\) or \(SD_{12}\) locus on the time period of incubation required for individual seeds to complete germination in germinated subpopulations\(^a\)

<table>
<thead>
<tr>
<th>Locus (subpopulation)</th>
<th>Additive effect (day)</th>
<th>Standard error (day)</th>
<th>t-value</th>
<th>Probability</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>(SD_{1-2} (SD_{1-2} \text{ Ex.} \text{ #1}))</td>
<td>0.47</td>
<td>0.08</td>
<td>5.76</td>
<td>&lt;0.0001</td>
<td>Endosperm (4)</td>
</tr>
<tr>
<td>(SD_{1-2} (SD_{1-2} \text{ Ex.} \text{ #2}))</td>
<td>0.31</td>
<td>0.08</td>
<td>4.13</td>
<td>&lt;0.0001</td>
<td>Endosperm (4)</td>
</tr>
<tr>
<td>(SD_{12} (SD_{12} \text{ Ex.} \text{ #1}))</td>
<td>0.22</td>
<td>0.11</td>
<td>2.08</td>
<td>0.0383</td>
<td>Embryo (3)</td>
</tr>
<tr>
<td>(SD_{12} (SD_{12} \text{ Ex.} \text{ #2}))</td>
<td>0.51</td>
<td>0.14</td>
<td>3.64</td>
<td>0.0003</td>
<td>Embryo (3)</td>
</tr>
<tr>
<td>(SD_{12} (SD_{12} \text{ Ex.} \text{ #3}))</td>
<td>0.53</td>
<td>0.22</td>
<td>2.41</td>
<td>0.0168</td>
<td>Embryo (3)</td>
</tr>
<tr>
<td>(SD_{12} (SD_{7-1} &amp; SD_{12} \text{ Ex.} \text{ #1}))</td>
<td>0.73</td>
<td>0.14</td>
<td>5.20</td>
<td>&lt;0.0001</td>
<td>Embryo (3)</td>
</tr>
</tbody>
</table>

\(^a\)Refer to Table 3 for additional information on the subpopulations. Gene component effects were estimated using the cited additive-dominance model for endosperm or embryo genotypes. Estimates for the component dominance effect are not listed as they are not significant at \(P=0.05\).
Figure 1 Electrophoresis patterns of endosperm genotypes for co-dominant markers. A. Comparison between endosperm and embryo genotypes of same seeds. B. Segregation patterns for four endosperm genotypes of F$_2$ seeds. RID12, RM315 and SD12m13 were selected to mark the seed dormancy loci $SD_{7-1}$, $SD_{1-2}$ and $SD_{12}$, respectively. Gel images show electrophoresis patterns for individual F$_2$ seeds from germinated subpopulations. The genotypes are indicated by combinations of the dormancy-enhancing ($D$) and/or -reducing ($d$) alleles, or by the copy number of the $D$ allele (0 to 3) at a locus.
Figure 2 Germination and allelic frequency distributions in germinated subpopulations segregating for the seed dormancy loci $SD_{7-1}$ (A), $SD_{1-2}$ (B) and $SD_{12}$ (C). Dotted lines indicate germination distributions for three independent experiments (open diamonds, circles & triangles), which were conducted for each locus using the indicated number (N) of $F_2$ seeds received given days of after-ripening (DAR) treatment. Solid lines indicate frequency distributions for the dormancy-reducing alleles ($F_{sd7-1}$, $F_{sd1-2}$ or $F_{sd12}$) in each of the germinated subpopulations. Note: the expected allelic frequency for the genetic equilibrium status is $p=0.5$ and frequencies for the dormancy-enhancing alleles (1-$p$) are not shown.
Figure 3  Cumulative germination distributions for four endosperm genotypes of seeds. Seeds were sampled from populations segregating for the seed dormancy locus $SD_{7.1}$ (A), $SD_{1.2}$ (B) or $SD_{12}$ (C). N was the number of seeds used for the germination experiment. Both germinated and non-germinated seeds were determined for endosperm genotypes, which are indicated by combinations of dormancy-enhancing ($D$) and/or -reducing ($d$) alleles. Embryo genotypes for $SD_{12}$ are listed in the parentheses (C).
Figure 4  Genotypic or allelic frequency distributions in germinated and non-germinated subpopulations segregating for \(SD_{7.1}\) and \(SD_{12}\).  

A & C. Joint frequency distributions for endosperm genotypes of the two loci. The triploid genotypes for each of the two loci are represented by combinations of the dormancy-enhancing (upper case) and/or -reducing (low case) alleles. Refer to Table 3 (\(SD_{7.1} \& SD_{12}\) Ex. #1 & 2) for single-locus information of the germinated (A) and non-germinated (C) subpopulations. 

B. Germination and allelic frequency distributions. \(N\) was the number of seeds used for the germination experiment. \(F_{sd_{7.1}}\) and \(F_{sd_{12}}\) are frequencies for the dormancy-reducing alleles at \(SD_{7.1}\) and \(SD_{12}\), respectively, in the germinated subpopulation. 

D. Genotypic differences in the time to germination. Data shown are means (circles), standard errors (bars), and sample sizes (numbers) for the nine embryo genotypes of seeds.