RESEARCH PAPER

Crop-model assisted phenomics and genome-wide association study for climate adaptation of *indica* rice. 2. Thermal stress and spikelet sterility

Michael Dingkuhn¹*, Richard Pasco², Julie Mae Pasuquin², Jean Damo², Jean-Christophe Soulié¹, Louis-Marie Raboin¹, Julie Dusserre¹, Abdoulaye Sow³, Baboucarr Manneh³, Suchit Shrestha² and Tobias Kretzschmar²

¹ Cirad, Umr AGAP (Dept BIOS) and Upr AIDA (Dept ES), F-34398, Montpellier, France
² IRRI, CESD Division, DAPO Box 7777, Metro Manila, Philippines
³ Africa Rice Center, Sahel Station, PB 96, St Louis, Senegal

* Correspondence: m.dingkuhn@irri.org, dingkuhn@cirad.fr

Received 14 February 2017; Editorial decision 20 June 2017; Accepted 27 June 2017

Editor: Greg Rebetzke, CSIRO Agriculture and Food

Abstract

Low night and high day temperatures during sensitive reproductive stages cause spikelet sterility in rice. Phenotyping of tolerance traits in the field is difficult because of temporal interactions with phenology and organ temperature differing from ambient. Physiological models can be used to separate these effects. A 203-accession *indica* rice diversity panel was phenotyped for sterility in ten environments in Senegal and Madagascar and climate data were recorded. Here we report on sterility responses while a companion study reported on phenology. The objectives were to improve the RIDEV model of rice thermal sterility, to estimate response traits by fitting model parameters, and to link the response traits to genomic regions through genome-wide association studies (GWAS). RIDEV captured 64% of variation of sterility when cold acclimation during vegetative stage was simulated, but only 38% when it was not. The RIDEV parameters gave more and stronger quantitative trait loci (QTLs) than index variables derived more directly from observation. The 15 QTLs identified at $P<1 \times 10^{-5}$ (33 at $P<1 \times 10^{-4}$) were related to sterility effects of heat, cold, cold acclimation, or unexplained causes (baseline sterility). Nine annotated genes were found on average within the 50% linkage disequilibrium (LD) region. Among them, one to five plausible candidate genes per QTL were identified based on known expression profiles (organ, stage, stress factors) and function. Meiosis-, development- and flowering-related genes were frequent, as well a stress signaling kinases and transcription factors. Putative epigenetic factors such as DNA methylases or histone-related genes were frequent in cold-acclimation QTLs, and positive-effect alleles were frequent in cold-tolerant highland rice from Madagascar. The results indicate that epigenetic control of acclimation may be important in *indica* rice genotypes adapted to cool environments.

Key words: Candidate genes, cold and heat tolerance, epigenetics, heuristics, male sterility, *Oryza sativa* L., RIDEV crop model.

Introduction

Spikelet sterility of rice, in particular male sterility, has received scientific attention for three reasons: (i) yield losses due to sensitivity to abiotic stresses such as heat (Jagadish *et al.*, 2007), cold (Dingkuhn *et al.*, 1995; Dingkuhn, 1995), and drought (Sheoran and Sain, 1996); (ii) its importance in hybrid seed production (Wang *et al.*, 2013); and (iii) genetic
sterility barriers hampering intersubspecific (Long et al., 2008) and interspecific (Garavito et al., 2010) crosses in breeding. This study focused on thermal-stress-induced sterility and its genomic associations.

Heat and cold stress cause male sterility in different ways and at different stages of reproductive development. Cold stress disrupts cellular development processes at the microspore stage, notably meiosis (De Storme and Geelen, 2014). Pollen infertility may result from deregulation or sugar starvation of meiosis itself (Sheoran and Sain, 1996) or disturbed anther, tapetum, or pollen wall development (Wang et al., 2013). Heat damage is rare at the microspore stage because the shoot apex is located near the bottom of the canopy under shaded conditions (Julia and Dingkuhn, 2013). The most heat-sensitive phase occurs at anthesis and can impair anther dehiscence, pollination, pollen germination or pollen tube growth (Jagadish et al., 2007). Given the diversity of vulnerable processes, many genes and pathways are involved in sterility, depending on the causative stress and its phenological timing (De Storme and Geelen, 2014). Tolerance to heat also depends on constitutive pollen production, affecting probability of successful pollination under stress (Jagadish et al., 2007). The female spikelet organs appear to be more resilient to thermal stresses (Wolters-Arts et al., 2013) and the male parent chiefly determines the extent of heat-induced sterility (Fu et al., 2015). Female organs, however, play a major role in the interspecific sterility barrier (Garavito et al., 2010).

Unrelated to stress-induced sterility, empty spikelets may also occur under plant-scale assimilate shortage, spikelet-level assimilate import restrictions or unfavorable topological position of spikelets within the panicle (Yang and Zhang, 2010). Such positional effects may be enhanced by physiological stresses causing poor panicle exertion from the surrounding leaf sheaths, inducing spikelet sterility when peduncles stay short (O’Toole and Namuco, 1983; Julia and Dingkuhn, 2013). Consequently, thermal spikelet sterility can occur at pre-floral stages through impeded development of male organs or indirectly through inhibited peduncle elongation, or at anthesis when pollination occurs.

We study here the incidence of unfilled spikelets, summarily termed sterility, on 203 diverse indica rices (the indica subset of the ORYTAGE diversity panel; http://ricephenonetwork. irri.org/diversity-panels/orytage-diversity-panels) grown in ten climatically contrasting environments. A heuristic, model-based approach was used to disaggregate the observed sterility into causal components dependent on the environmental conditions they were associated with during critical phenological stages. Proof of concept for this approach was previously provided by Reymond et al. (2003) for drought, Rebollolo et al. (2015) for early vigor, and Yin et al. (2005) and Dingkuhn et al. (2017) for phenology, yielding quantitative trait loci (QTLs) that would not have been discovered with untransformed phenotype data.

The RIDEV model, first published by Dingkuhn et al. (1995) and Dingkuhn and Miezan (1995) (and under this name by Wopereis et al., 2003), was used in revised form (V2) in the companion study (Dingkuhn et al., 2017) for phenology traits. Here it was further improved to simulate thermal sterility estimated from microclimate-based organ temperatures (shoot apex at microspore stage and panicle at anthesis) by using rice microclimate vs sterility relationships published elsewhere (Julia and Dingkuhn, 2012, 2013). We also introduced an acclimation parameter implementing a hardening effect of stress exposure during earlier phenological stages. Such mechanisms were reported by Shimono et al. (2010).

This study aimed at (i) heuristically phenotyping the ORYTAGE indica rice diversity panel using RIDEV V2, thereby considering fitted crop parameter values as genotypic traits; (ii) conducting genome-wide association studies (GWAS) for RIDEV V2 parameters and, by comparison, for simpler phenotypic index variables previously reported (Dingkuhn et al., 2015a); and (iii) identifying potential candidate genes.

Materials and methods

Experiments and germplasm were described by Dingkuhn et al. (2015a,b) and the companion paper (Dingkuhn et al., 2017). Genotyping and GWAS are described in the companion paper. Essential information is repeated here.

Experimental design

Experiment in Senegal

The experiment was conducted at the AfricaRice Sahel Station at Ndiaye, Senegal (16°12′N, 16°16′W, 8 m above sea level; asl) with an augmented design with six blocks, two factors (six sowing dates, 203 genotypes) and four replicated checks (IR64, Sahel 108, N22, Chomrong). Subplots were 1 m×1 m and pre-germinated seed was sown at 20 cm×20 cm. Fields were flooded throughout. Fertilizer inputs were 120–60–60 (N–P–K, kg ha⁻¹). Sowing dates were 7 February 2009, 7 March 2009, 7 April 2009, 17 July 2009, 17 September 2009 and 19 October 2009.

Population-level spikelet sterility was measured on plants at the center of subplots. All panicles of eight hills were collected, counted, and spikelets stripped off, bulked, separated into filled and unfilled spikelets, oven-dried and weighed. Mean spikelet weight was determined from 200 filled and 200 unfilled spikelets. Flowering dates were recorded.

Experiment in Madagascar

Experimental sites were Ambihotromby (Hautes Terres, Région d’Antsirabe, 19°52′S, 46°59′E, 1494 m asl) and Ivory (19°32′S, 46°24′E, 1494 m asl). The experiment, identical at both sites, was a randomized complete block design with three replications. Elemental plots (variety) were 0.6 m×1.4 m (3 × 7 hills planted at 20 cm×20 cm, with the central row used for observations). Check varieties were as in Senegal but cv. Chomrong was absent.

Pre-germinated seed was sown on 3 December 2009 and 6 December 2010 at Ivory and on 11 December 2009 and 15 December 2010 at Ambihotromby in a seedling nursery and transplanted as single plants 15–20 d after sowing. Plots were kept flooded at 5–10 cm. Fertilizer was applied as green manure (5 t ha⁻¹), 81 kg ha⁻¹ N as urea, 27 kg ha⁻¹ P as triple superphosphate and 49 kg ha⁻¹ K as KCl.

Spikelet sterility was measured as described for Senegal but on the five central plants on the middle row of the plots.

Germlasm

The population was taken from the ORYTAGE species-wide (O. sativa L.) diversity panel of Cirad (http://ricephenonetwork. irri.org/diversity-panels/orytage-diversity-panels) which is also a resource for the GRiSP Global Rice Phenotyping Network (http:// ricephenonetwork.irri.org). A subsample of 203 indica accessions was grown (see Supplementary Table S1 at JXB online).
Genotyping was conducted at Diversity Arrays Technology Pty Ltd. (Australia) using genotyping by sequencing (GBS) as described by Courtois et al. (2013) and Dingkuhn et al. (2017). Data cleaning involved removing sequences having more than one hit on the pseudomolecules, which were discarded, as well as markers with call rates ≤80% or minor allele frequency ≤2.5%. After data cleaning and before imputation of the GBS resource, the rate of missing data was 6.5%. The final data set had 16 232 markers (8 214 DArT and 8 018 single nucleotide polymorphism (SNP) markers), giving an average density of one marker per 24 kb. There were nine gaps of more than 500 kb devoid of markers on chromosomes 2, 4, 6, 7, 8, and 11, of which two were of approximately 1 Mb.

GWAS was conducted for each trait using a mixed model with control of the structure and kinship under Tassel v5.0 (Bradbury et al., 2007). Kinship was computed using Tassel v5. The genetic structure of the panel is presented in Supplementary Fig. S1. The threshold to declare an association highly significant was set at $P < 1 \times 10^{-5}$. Associations having $1 \times 10^{-5} < P < 1 \times 10^{-4}$ are presented in Supplementary Table S2. The 50% decay genomic distance of linkage (LD) was calculated according to Sved (1971). Supporting SNPs were defined as SNPs with $P < 1 \times 10^{-3}$ located within ±LD. Quantile–quantile plots for the GWAS are presented in Supplementary Fig. S2.

**RIDEV V2 model**

The model is described in detail in the companion paper (Dingkuhn et al., 2017) and only essential information is provided, namely on spikelet sterility simulations.

**General**

The RIDEV V2 model is an improved version of RIDEV described and validated by Dingkuhn et al. (1995), Dingkuhn and Miezan (1995), and later described as an agronomic decision tool by Wopereis et al. (2003). It simulates the duration of phenological phases, the timing of panicle initiation (PI) and flowering (FL) events, and the incidence of cold- or heat-induced sterility.

The model can be used in simulation mode (phenology/sterility prediction) or heuristic (reverse) mode to estimate genotypic parameters. Table 1 presents the variables and parameters used in this study.

**Spikelet sterility**

Similar to the original RIDEV (Dingkuhn, 1995), RIDEV V2 simulates three sources of sterility: (i) an unexplained baseline level set by parameter $K_{SterBase}$, (ii) chilling set by $K_{SterCold}$ and $K_{SterSlope}$, and (iii) heat set by $K_{SterHeat}$. Baseline sterility is a constant spikelet fraction. Cold sterility is driven by the mean minimum water temperature ($T_{min}$) during the microspore stage (about the booting stage), estimated to occur between 0.3 and 0.8 of the RPR stage (equal to 120–320 °C d before flowering in this study, or about 9–23 d in a warm environment). Heat sterility is driven by the panicle temperature ($T_{pan}$) during the time of day of anthesis and averaged for the period −60 to +60 °C d around flowering, a period of ca. 9 d. The timing of these phases is determined by the phenology modules described in the companion paper. The phenological windows were chosen wide to account for uncertainty.

Two examples of observed sterility response to low $T_{min}$ are given in Fig. 1B. Figure 1C describes the broken-stick linear models implemented to simulate the three different components of sterility, and the equation that provides total sterility. For data supporting the phenological and factor hypotheses for thermal sterility, refer to Dingkuhn et al. (1995). The C++ source code of the present version of RIDEV is given in Supplementary Protocol S1.

**Development**

RIDEV simulates crop development from sowing until physiological maturity, whereby crop establishment is included in the basic vegetative phase (BVP). The BVP is followed by a photoperiod-sensitive phase (PSP) ending at PI, a reproductive phase (RPR) ending at FL and a maturation phase (MATU). Each of these phases except PSP has a genotypic thermal-time budget. Duration of PSP depends on photoperiod and thermal time using the Impatience model (Dingkuhn et al., 2008). Thermal time calculation uses genotypic cardinal temperatures $T_{base}$ and $T_{opt}$ (Dingkuhn and Miezan, 1995) and a standard diurnal temperature pattern. Depending on

---

Table 1. **RIDEV V2 parameters and input/output variables used in this study**

<table>
<thead>
<tr>
<th>Parameter/variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scenario parameters</strong></td>
<td></td>
</tr>
<tr>
<td>Latitude</td>
<td>xx,yy, decimal (degrees)</td>
</tr>
<tr>
<td>SowingDate</td>
<td>dd/mm/yyyy</td>
</tr>
<tr>
<td>Transplanting</td>
<td>0 or 1, binary</td>
</tr>
<tr>
<td>DDTTransplantingShock</td>
<td>0 ... 200 (default=100), development lag in thermal time units (°C d)</td>
</tr>
<tr>
<td>POP</td>
<td>Plant population density, 50 000 ... 1 000 000 (plants ha−1)</td>
</tr>
<tr>
<td><strong>Input variables (variables in brackets only needed for heat sterility under flooded cultivation)</strong></td>
<td></td>
</tr>
<tr>
<td>Tmin</td>
<td>Daily minimum air temperature at 2 m (°C)</td>
</tr>
<tr>
<td>Tmax</td>
<td>Daily maximum air temperature at 2 m (°C)</td>
</tr>
<tr>
<td><strong>Crop parameters for phenology (can be optimized by R-Genoud)</strong></td>
<td></td>
</tr>
<tr>
<td>$K_{SterBase}$</td>
<td>Baseline sterility that cannot be explained by thermal stresses, 0 ... 1 (typically 0.1 to 0.2) (fraction)</td>
</tr>
<tr>
<td>$K_{SterCold}$</td>
<td>Critical $T_{min}$ at microspore stage for onset of cold stress, 20 ... 0 (°C)</td>
</tr>
<tr>
<td>$K_{SterSlope}$</td>
<td>Sensitivity of chilling response: $T_{w(min)}$ for sterility to rise from zero to 100%</td>
</tr>
<tr>
<td>$K_{SterHeat}$</td>
<td>Critical $T_{min}$ for onset of heat stress, at TOA during the flowering period, 25 ... 40 (°C)</td>
</tr>
<tr>
<td>AcclimSterC</td>
<td>Acclimatic effect of low $T_{min}$ during BVP (&lt;18 °C), reducing $K_{SterCold}$ (effective at microspore stage), 0 ... 15</td>
</tr>
<tr>
<td><strong>Output variables</strong></td>
<td></td>
</tr>
<tr>
<td>SterBase</td>
<td>Baseline (non-stress related) sterility (%)</td>
</tr>
<tr>
<td>SterCold</td>
<td>Cold sterility (%)</td>
</tr>
<tr>
<td>SterHeat</td>
<td>Heat sterility (%)</td>
</tr>
<tr>
<td>SterTot</td>
<td>Total sterility (non-additive) (%)</td>
</tr>
</tbody>
</table>
location in the canopy of the shoot apex, temperatures of air or floodwater are used.

Water and apex temperature
The shoot apex of irrigated rice is submerged until booting (stem elongation) stage. Daily minimum and maximum water temperature \( T_{w} \) are simulated from leaf area index as described in the companion paper (Dingkuhn et al., 2017). Apex temperature \( T_{\text{apex}} \) is equal to \( T_{w} \) while the apex is under the water line. Internode elongation lifts the apex nearly to the canopy top and the apex assumes air temperature.

Panicle temperature
Data from an extensive study by Julia and Dingkuhn (2013), describing the rice Panicle temperature \( T_{\text{pan}} \) relationship with microclimate in multiple field environments in Senegal and the Philippines by infrared imagery, were revisited to generate an empirical model predicting \( T_{\text{pan}} \). Multiple, stepwise, linear regression of 2941 infrared-based temperature difference measurements \( (T_{\text{p}}-T_{\text{a}(2m)}) - (T_{\text{pan}}) \) observations vs simultaneous crop and weather observations at 2 m resulted in the following model:

\[
T_{\text{D}} = 0.782 + 0.422 T_{\text{a}} - 0.0443 RH - 0.00287 R_{\text{g}} - 8.05 Z - 6.59 LTR
\]

where \( T_{\text{D}} \) is the temperature difference between air (at 2 m) and panicle temperature; \( T_{\text{a}(2m)} \) is air temperature at 2 m; \( RH \) is air relative humidity; \( R_{\text{g}} \) (W m\(^{-2}\)) is solar radiation; \( Z \) (m m\(^{-1}\)) is the ratio of ground height of the center of the panicle to plant height; and LTR is light transmission ratio of the canopy as calculated from LAI using the Lambert–Beer law and \( K_{\text{a}}=0.5 \). This model had an adjusted \( R^{2} \) of 0.82, standardized coefficients of 0.63 (\( T_{\text{a}} \)), -0.30 (\( RH \)), -0.14 (\( R_{\text{g}} \)), -0.19 (\( Z \)), and -0.13 (LTR). The F-value was 2613. Figure 2 shows for check cultivars the calculated vs observed TD for Senegal (November and February sowings).

Time of day of anthesis
Prediction of time of day of anthesis (TOA) is needed to compute \( T_{\text{pan}} \) at anthesis, which happens during 2 h usually in the morning, but depends on the weather. RIDEV uses finding of Julia and Dingkuhn (2012) that mean \( T_{\text{min}}(\text{air}) \) during the basic vegetative phase \( (1995) \) reported 18 °C to be the critical temperature below which cold-induced sterility occurs in many rice cultivars. We hypothesized that this was caused by cold acclimation at high altitude in Madagascar whereas in Senegal, periods with cold nights were short and seasonal (Dingkuhn et al., 2015b). An optional function was implemented that reduces \( K_{\text{SterCold}} \) (temperature below which cold-induced sterility occurs) when the crop experienced cold at the vegetative stage, using the genotypic coefficient \( AcclimSterC \). The decrease of \( K_{\text{CritSter}} \) following acclimation is the product of \( AcclimSterC \) and the number of degrees Celsius by which mean \( T_{\text{min}(\text{air})} \) is below 18 °C during the basic vegetative phase (BVP). Dingkuhn et al. (1995) reported 18 °C to be the critical \( T_{\text{min}} \) below which cold-induced sterility occurs in many rice cultivars.

The equation calculating cold sterility is as follows, with the additional component of cold acclimation shown in bold:
Cold sterility (fraction spikelets) = \min(1, \max(0, K_{SterCold} - (AcclimSterC \times \max(0, 18 - T_{min,mean}(BVP)) - T_{max,mean}(booting stage)) / \text{KSterSlope}))

(AcclimSterC thus reduces the cold sensitivity (KCritSter) of sterility during the microspore stage in cases where the crop experienced cold conditions during BVP.)

RIDEV variables and parameters

Table 1 presents an overview of RIDEV V2 variables and parameters used in the sterility study. A complete list is presented in Supplementary Table S3. As weather input, daily air $T_{min}$ and $T_{max}$ ($^\circ$C), solar radiation $R_s$ (MJ), and minimum and maximum relative humidity HR (%) were used. The following crop parameters were used to simulate Tapex (cold stress at microspore stage) and $T_{pan}$ (heat stress at anthesis) which determine sterility:

- $Z$: the relative position $Z$ of the panicle relative to plant tops (ground height), set to default=0.8
- $POP$: plant population density (plants ha$^{-1}$), whose value depended on the experiment
- $PLA_{init}$: initial leaf area per plant, set to default=0.0001 (m$^2$)
- $LRGR_{max}$: leaf area relative growth rate, set to 1.3 (unitless)
- $LAI_{max}$: maximal leaf area index, set to 5 (unitless)

(The last four parameters were needed to implement a standard LAI time course that served to estimate LTR, affecting microclimate.)

Genotypic parameters for crop phenology were estimated by optimization procedures as described in the companion paper (Dingkuhn et al., 2017: Run3 which simulated phenotype best). The genotypic parameters for sterility, the object of the present study, were estimated the same way with phenology parameters fixed to values estimated previously.

Parameters and ranges for estimation of genotypic values

The genotypic parameters estimated by optimization, thereafter considered as genotypic traits, were as follows:

- $K_{SterBase}$: unexplained sterility (fraction), estimation boundaries 0...1
- $K_{SterCold}$: critical $T_{min,apex}$ for cold sterility ($^\circ$C), estimation boundaries 20...5
- $K_{SterHeat}$: critical $T_{pan}$ at TOA for heat sterility ($^\circ$C), estimation boundaries 20...40
- $K_{SterSlope}$: temperature interval below $K_{SterCold}$ causing 100% sterility ($^\circ$C). No slope parameter was used for the heat response of sterility and the model used a fixed coefficient of 5 (sterility attains 100% if $T_{pan}$ at TOA attains the value of $K_{SterHeat} + 5^\circ$C). Estimation boundaries 2...15
- $AcclimSterC$: acclimation response of $K_{SterCold}$ to chilling during BVP, estimation boundaries 0...15

Parameters were either fixed to default values or optimized by R-Genoud (Dingkuhn et al., 2017) within physiologically relevant boundaries, with all parameters optimized in one procedure. The
accuracy of predicted vs observed sterility across environments was evaluated by normalized-least square error method (cost function). R-Genoud varied parameter values to minimize the cost function. Table 2 presents the four combinations of optimized parameters used in this study (estimation runs), namely Run1 and Run2 without considering acclimation and Run3 and Run4 with it. In Run2 and Run4, KSterSlope was fixed to the default value to reduce parameter number. (A smaller number sometimes improved GWAS associations but reduced accuracy of phenotype prediction.)

Traits analysed by GWAS

Two types of traits were analysed: (i) index variables directly calculated from measurements and (ii) estimated RIDEV parameters. The index variables were described in detail by Dingkuhn et al. (2015a) and are summarized as follows:

Ster_20C (fraction spikelets)
Estimates sterility incidence at Tw(min)=20 °C, based on the observed correlations of sterility vs Tw(min) at micropore stage across all environments.

Ster_Slope (°C−1)
Correlations were used to estimate the rate of increase of sterility for T_{w(min)} below 18 °C. Frequency distributions are presented in Supplementary Fig. S4. The other traits used for GWAS were the RIDEV parameters listed in Table 2, KSterBase, KSterCold, KSterHeat, KSterSlope, and AcclimSterC.

Searches for annotated genes associated with the QTLs detected

Annotated genes within ±100 kbp of a peak SNP were extracted from Phytozone (Joint Genome Institute JGI), https://phytozone.jgi.doe.gov) and MSU database (Michigan SU, http://rice.plantbiology.msu.edu/). Functional and expression-related information was extracted from GeneVisible (https://genevisible.com/), PiaNet (http://aranet.mpimp-golm.mpg.de/), Gramene (http://archive.granene.org/), UniProt (http://www.uniprot.org/), RiceXPro (http://ricexpro.dna.affrc.go.jp/), RiceChip (http://www.ricechip.org/), and gabi (https://www.gabipd.org/). Genes within ±39 kbp (estimated LD; Dingkuhn et al., 2017) were considered as likely candidates if supported by function.

Statistics

Linear regression analyses and frequency distributions were conducted with SigmaPlot 12.3 (Systat Sofwares Inc.).

Results

Sensitivity analysis for crop parameters

The effect of crop parameter variation on sterility simulations was evaluated for the four most contrasting environments (see Supplementary Fig. S5). Each environment caused different patterns of parameter effects. Consequently, there was sufficient environmental diversity in this study to disaggregate component traits of sterility (parameters optimized). The only similarity in environment response was for KSterCold (critical T) vs KSterSlope (1/ slope) for the cold response, and we thus conducted estimation runs for either both parameters or only KSterCold (Table 2).

Estimated RIDEV parameters

Estimated RIDEV parameter value distributions are presented in Fig. 3. The histograms describe the entire panel as previously characterized for phenology (Dingkuhn et al., 2015b) and sterility phenomics (Dingkuhn et al., 2015a) with 203 accessions. KSterBase was not included because its effect is environment independent.

KSterBase (unexplained baseline sterility) had a normal distribution centered at 0.10–0.15 but showed an upward tail constituted by only a few accessions, with a maximal value of 0.5 (Fig. 3A). The vast majority of accessions had baseline sterility below 0.2 (20% of spikelets). For KSterSlope (T_{min} interval within which sterility increases from zero to 100%) (Fig. 3B), a value of <2.5 °C was observed for about half of the accessions, indicating a steep response of sterility to temperature. The other accessions showed a broad distribution of higher KSterSlope values up to 10.5 °C. The critical temperature KSterCold (Fig. 3C) showed normal distribution centered at about 18 °C, with only few accessions outside the 17–20 °C range. For the KSterHeat critical temperature (Fig. 3D), the vast majority of accessions ranged between 29 and 33 °C but some had higher values. Since extreme heat occurred rarely, KCritHeat estimations above 35 °C are uncertain and mean that heat-induced sterility was absent.

Cold-acclimation parameter AcclimSterC was estimated with either the KSterSlope co-optimized (Run3, Fig. 3E) or set to the value 5 (Run4, Fig. 3F). Most accessions (70% for Run3) had estimated values ≤2, 20% were between 1 and 2, and only 10% had greater values between 2 and 12. Fixing the value of KSterSlope had little effect on this distribution (Run4), but both estimations of AcclimSterC were retained because the GWAS results were different. The parameter distributions indicate that most accessions showed little acclimation, but a small subset showed strong acclimation.

Table 2. RIDEV parameter estimation runs

Fields marked × indicate inclusion in estimation run.

<table>
<thead>
<tr>
<th>Parameter estimation run</th>
<th>Environments</th>
<th>Crop parameters estimated</th>
<th>Figures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Senegal</td>
<td>Madagascar</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>4</td>
<td>×</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>4</td>
<td>×</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>4</td>
<td>×</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>4</td>
<td>×</td>
</tr>
</tbody>
</table>
Variation in spikelet sterility explained by RIDEV

The RIDEV model explained about two-thirds of observed variation in sterility across all accessions and environments (Run3 and Run4). Regression analyses of simulated \( Y \) vs observed sterility \( X \), sterility expressed as a fraction, gave the following. Senegal subset: \( Y = 0.045 + 0.717X, R^2 = 0.70\text{****}, n=1098 \); Madagascar subset: \( Y = 0.037 + 0.622X, R^2 = 0.61\text{****}, n=865 \). All environments: \( Y = 0.043 + 0.652X, R^2 = 0.64\text{****}, n=1963 \). For check variety Sahel 108 alone, the correlation was as follows: All environments: \( Y = 0.019 + 0.738X, R^2 = 0.77\text{***}, n=48 \).

The one-third unexplained sterility (or one-fourth for the check) was probably caused by the high intrinsic biological variation of sterility incidence, which responds to many factors, and error propagation from phenology to sterility simulations, causing errors in the timing of sensitive phenological phases.

Several alternative hypotheses to acclimation were tested but did not further improve predictions. When setting \( \text{AcclimSterC=0} \) (disabling acclimation), RIDEV explained only 38% of observed sterility: \( Y = 0.083 + 0.456X, R^2 = 0.38\text{***}, n=1963 \).

Consequently, the phenomenon captured with \( \text{AcclimSter} \) impacted strongly on spikelet sterility.

Genomic associations for baseline sterility

Associations for unexplained baseline sterility are presented in Fig. 4 (details in Table 3). Two major QTLs \( P < 1.0 \times 10^{-5} \) were observed for index variable \( \text{Ster}_20C, Sb2(1)i \) (chr2) and \( Sb11i \) (chr11). A minor QTL having \( P = 4.0 \times 10^{-5} \) \( (Sb6(1)i) \) supported by nine SNPs was located on chr6. \( \text{KSterBase} \), conveying similar information as \( \text{Ster}_20C \), yielded two major QTLs on chr6 \( (Sb6(2)m) \) and chr8 \( (Sb8m) \). None colocalized with QTLs for \( \text{Ster}_20C \). However, there were similarities in the Manhattan plots (Fig. 4A vs 4B). Minor QTLs on chr4, 6 and 8 colocalized in both diagrams, and major RIDEV QTL \( Sb8m \) (Fig. 4B) was also discernible in Fig. 3A. Associations changed slightly when the model parameter \( \text{KSterSlope} \) was fixed to a constant value to avoid over-parameterization (Fig. 4C). While most peaks diminished, QTL \( Sb6(2)m \) strongly increased to attain \( P = 4.0 \times 10^{-8} \).

**Fig. 3.** Frequency distributions of RIDEV parameter values for spikelet sterility for the ORYTAGE indica panel.
Genomic associations for sterility response to temperature

The correlation-derived index variable Ster_Slope gave three QTLs on chr2, 3 and 7 (Fig. 5A). Major QTL Ss2i was supported by 22 SNPs and had high allele frequency (21%) whereas Ss3i and Ss7i were isolated SNPs with low allelic frequency (see Supplementary Table S2). Minor QTL Ss1i on chr1 was supported by 13 SNPs ($P=2.5 \times 10^{-5}$).

No major QTL was detected for parameter KSterCold (Fig. 5B; Sc10m was classified as minor due to absence of supporting SNPs) and one was found for KSterHeat (Fig. 5D: Sh4m). No significant association was observed for KSterSlope (Fig. 5C). However, when KSterSlope was fixed to constant value (Run2), KSterCold which then absorbed the variation of both slope and critical temperature gave three highly significant associations (Fig. 6), Sc1m and Sc6m

Fig. 4. GWAS associations for baseline sterility in the absence of thermal stress. (A) Index variable Ster_20C derived from regression analysis. (B) RIDEV parameter KSterBase estimated with co-fitting of RIDEV slope parameter (Estimation Run1, Table 2). (C) KSterBase estimated with parameter KSterSlope fixed to constant value (Run2).
Table 3. GWAS associations and putative candidate genes for major QTLs (for complete list including minor QTLs refer to Supplementary Table S2)

QTLs are identified by name, trait (for RIDEV estimation runs see Table 2), variable type and figure showing the Manhattan plot. Physical parameters for QTLs are chromosome number, locus and \( P \) for most significant SNP; number of supporting SNPs \( P<0.001 \) within LD, QTL effect on phenotype, base pair of main SNP, allele frequency (number accessions), and number of annotated genes within ±100 kb. Genes are shown in bold when located within LD (±39 kb). Distance of gene center from main SNP (kb) is provided in brackets. Gene function is provided as available from public databases.

<table>
<thead>
<tr>
<th>QTL</th>
<th>Trait, model run</th>
<th>Variable type</th>
<th>Chr</th>
<th>Locus of smallest ( P ) (bp)</th>
<th>SNPs ( P&lt;0.001 )</th>
<th>Effect</th>
<th>Base pair</th>
<th>Allele frequency (no. of cases)</th>
<th>No. genes</th>
<th>Fig Putative candidate genes: LOC_ annotation(MSU), (distance from SNP in kb)], (common name and function using public databases)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sb2(1)i</td>
<td>Ster_20C</td>
<td>Index</td>
<td>2</td>
<td>18620629</td>
<td>7.9E−06</td>
<td>22</td>
<td>0.100</td>
<td>G.A</td>
<td>161</td>
<td>29</td>
</tr>
<tr>
<td>Sb11i</td>
<td>Ster_20C</td>
<td>Index</td>
<td>11</td>
<td>26684031</td>
<td>1.0E−05</td>
<td>9</td>
<td>−0.141</td>
<td>T.A</td>
<td>15</td>
<td>175</td>
</tr>
<tr>
<td>Sb8m</td>
<td>KSterBase Run1</td>
<td>Model</td>
<td>8</td>
<td>27353014</td>
<td>7.9E−06</td>
<td>7</td>
<td>−0.135</td>
<td>C.A</td>
<td>183</td>
<td>6</td>
</tr>
<tr>
<td>Sb6(2)m'</td>
<td>KSterBase' Run2</td>
<td>Model</td>
<td>6</td>
<td>28957091</td>
<td>4.0E−08</td>
<td>4</td>
<td>−0.241</td>
<td>A.G</td>
<td>183</td>
<td>6</td>
</tr>
<tr>
<td>Sb8m'</td>
<td>KSterBase' Run2</td>
<td>Model</td>
<td>8</td>
<td>27353014</td>
<td>1.0E−05</td>
<td>6</td>
<td>−0.137</td>
<td>C.A</td>
<td>183</td>
<td>6</td>
</tr>
<tr>
<td>Ss2i</td>
<td>Ster_slope</td>
<td>Index</td>
<td>2</td>
<td>22320801</td>
<td>7.9E−06</td>
<td>22</td>
<td>0.042</td>
<td>A.T</td>
<td>153</td>
<td>37</td>
</tr>
<tr>
<td>Sh4m</td>
<td>KSterHeat Run1</td>
<td>Model</td>
<td>4</td>
<td>29495307</td>
<td>2.0E−06</td>
<td>16</td>
<td>−4.46</td>
<td>T.G</td>
<td>180</td>
<td>9</td>
</tr>
<tr>
<td>Trait, model run</td>
<td>Number of smallest ( P ) SNPs ( P &lt; 0.001 )</td>
<td>Effect</td>
<td>Base pair</td>
<td>Allele frequency (no. of cases)</td>
<td>No. genes</td>
<td>Putative candidate genes: LOC_ annotation(MSU), (distance from SNP in kB), (common name and function using public databases)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>------------------------------------------</td>
<td>--------</td>
<td>------------</td>
<td>-------------------------------</td>
<td>-----------</td>
<td>-----------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sc7m</strong></td>
<td><strong>K SterCold</strong> Run2</td>
<td><strong>7</strong></td>
<td>23251129</td>
<td>4.0E-07</td>
<td><strong>9</strong></td>
<td><strong>6.14</strong> G C 184 6 40 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>H1(1)m</strong></td>
<td><strong>AcclimSterC</strong> Run3</td>
<td><strong>11</strong></td>
<td>21187019</td>
<td>2.5E-08</td>
<td><strong>14</strong></td>
<td><strong>−4.20</strong> A T 174 15 22 7A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>H1(2)m</strong></td>
<td><strong>AcclimSterC</strong> Run4</td>
<td><strong>1</strong></td>
<td>11451409</td>
<td>2.5E-08</td>
<td><strong>3</strong></td>
<td><strong>−3.19</strong> A T 170 19 22 7B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>H2(1)m</strong></td>
<td><strong>AcclimSterC</strong> Run4</td>
<td><strong>2</strong></td>
<td>27031056</td>
<td>4.0E-06</td>
<td><strong>38</strong></td>
<td><strong>−3.48</strong> G A 178 11 17 7B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>H2(2)m</strong></td>
<td><strong>AcclimSterC</strong> Run4</td>
<td><strong>2</strong></td>
<td>34040198</td>
<td>5.0E-06</td>
<td><strong>5</strong></td>
<td><strong>−2.88</strong> C A 167 22 23 7B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>H6(m)</strong></td>
<td><strong>AcclimSterC</strong> Run4</td>
<td><strong>6</strong></td>
<td>10122260</td>
<td>(locus 1)</td>
<td><strong>11</strong></td>
<td><strong>−6.80</strong> A T 182 7 22 7B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10347284</td>
<td>(locus 2)</td>
<td><strong>11</strong></td>
<td><strong>−6.80</strong> A T 182 7 23 7B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Acclimation related QTLs**

**H1(1)m**

**H1(2)m**

**H2(1)m**

**H2(2)m**

**H6(m)**
### Table 3. Continued

<table>
<thead>
<tr>
<th>QTL</th>
<th>Trait, model</th>
<th>Variable type</th>
<th>Run</th>
<th>Locus of smallest effect</th>
<th>Allele frequency</th>
<th>Effect P (&lt;0.001)</th>
<th>No. SNPs P&gt;0.001</th>
</tr>
</thead>
<tbody>
<tr>
<td>H7m</td>
<td>AcclimSterC</td>
<td>Variable type</td>
<td>Run4</td>
<td>LOC_Os07g43870 [29], RNA binding nucleic acid splicing factor; LOC_Os07g44100 [26], MYB-like TF</td>
<td>78</td>
<td>1.0×10⁻⁵</td>
<td>13</td>
</tr>
<tr>
<td>H8m</td>
<td>AcclimSterC</td>
<td>Variable type</td>
<td>Run4</td>
<td>LOC_Os08g10070 [31], LOC_Os08g10150 [78], serine/threonine kinase; LOC_Os08g09940 [62], rRNA methyltransferase; LOC_Os08g09950 [58], LOC_Os08g10010 [23], FAD, fatty acid desaturase</td>
<td>78</td>
<td>2.0×10⁻⁷</td>
<td>2</td>
</tr>
<tr>
<td>H9m</td>
<td>AcclimSterC</td>
<td>Variable type</td>
<td>Run4</td>
<td>LOC_Os09g34150 [44], LOC_Os09g34214 [4], LOC_Os09g34270 [27], LOC_Os09g34320 [55], methyltransferase</td>
<td>78</td>
<td>5.0×10⁻⁷</td>
<td>17</td>
</tr>
<tr>
<td>H11m</td>
<td>AcclimSterC</td>
<td>Variable type</td>
<td>Run4</td>
<td>LOC_Os09g34160 [29], LOC_Os09g34230 [11], LOC_Os09g34250 [19], LOC_Os09g34270 [27], UTP:glycinamide ribonucleoside kinase (GAKinase); LOC_Os09g34320 [55], methyltransferase</td>
<td>78</td>
<td>1.3×10⁻⁹</td>
<td>17</td>
</tr>
</tbody>
</table>

### Genomic associations for cold acclimation

*AcclimSterC*, when co-optimized with all model parameters participating in sterility simulation (Run3), yielded a single but strong association on chr11 (*H11m*, P=2.5×10⁻⁸) (Fig. 7A). Its allelic frequency was 8% and 14 SNPs supported it (Table 3). When the KSterSlope parameter was fixed (Run4, Fig. 7B), the significance of this QTL increased to P=1.3×10⁻⁹ (labelled *H11m*), and eight new QTLs appeared on chr 1, 2, 6, 7, 8, and 9, with P between 1.0×10⁻⁷ and 1.0×10⁻¹⁰. The QTL *H2(1)m*, *H3m*, *H4m*, *H5m*, *H6m*, *H7m*, *H8m*, and *H9m* were found. For QTLs related to thermal stress, three major ones were found. For *S2i*, LOC_Os02g36924 is the *OsMADS27* flowering-related TF and LOC_Os02g36974 is a dynein protein.

### Candidate genes

In Table 3, putative candidate genes within ±100 kb and ±39 kb (±LD) of peak SNP are presented for the 18 major QTLs that had P<1×10⁻⁵, ≥3 supporting SNPs and allelic frequency ≥4% for the minor allele. Information on all 33 observed QTLs is presented in Supplementary Table S2. On average, 24 annotated genes were located within ±100kb of peak SNP, and nine within ±39 kb. For QTLs for baseline sterility, functionally plausible candidate genes located closest to peak SNP were LOC_Os02g31140 in *Sb2(1)i*, involved in cuticle development; LOC_Os11g44170 in *Sb11i*, a calmodulin binding protein (with SNP located within the gene); LOC_Os06g47880 in *Sb6(2)m* (colocalized with *Sb6(2)m*, a protein kinase controlling cell division; and LOC_Os08g43270 in *Sc7m* (colocalized with *Sc7m*), an apoptotic gene involved in ROS stress protection. Several other genes of functional interest were located within LD, such as serine/threonine kinases. In *Sc7m*, an Shd1 mRNA dehydrogenation factor (TF) (LOC_Os08g43200) and a heat shock protein involved in drought and salt tolerance (LOC_Os08g43334) were found. Genes of functional interest were also detected minor QTLs (see Supplementary Table S2).

For QTLs related to thermal stress, three major ones were found. For *S2i*, LOC_Os02g36924 is the *OsMADS27* flowering-related TF and LOC_Os02g36974 is a dynein protein.
For *Sh4m*, five plausible functional candidates were observed, including a pollen protein, an H4 histone and a Zn-finger TF involved in growth regulation. For *Sc7m*, *Loc_Os07g38730* was found near the peak SNP, a GTPase involved in reproductive development. Minor QTLs also had plausible candidate genes including RAS-family proteins, the important flowering gene *OsMADS56*, the squamosa-binding TF *OsSPL19* and a phytochrome-related gene involved in flowering control.

Eleven major QTLs were found for *AcclimSterC*. Several candidate genes potentially conveying epigenic expression control were associated with them, notably within ±20 kb a chaperone (*H11m*/*H11m′*, *Loc_Os11g36050*), a histone-like TF (*H6m′*, *Loc_Os06g17480*) and an RNA helicase (*H7m′*, *Loc_Os06g43980*). We also found serine/threonine kinases (*H2(1)m′*, *Loc_Os02g55560*; *H11m*, *Loc_Os02g44610* and *Loc_Os02g45130*), a gene involved in meiosis (*H2(2)m′*, *Loc_Os02g55570*), and an androgen-induced sister chromatid cohesion factor (*H6m′*, *Loc_Os06g17840*). Within ±50 kb were an RNA methyltransferase (*H11m*, *Loc_Os09g34320*), an RNA-splicing MYB-like TF (*H7m′*, *Loc_Os07g43950*), a RAS protein (*H7m′*, *Loc_Os07g44040*) and a fatty acid desaturase (*H8m′*, *Loc_Os08g09940*). At 50–65 kb from the peak SNP, we found further putative epigenetic genes: a histone acetyltransferase (*H1(2)m′*, *Loc_Os01g68020*), two methyltransferases at (*H1(1)m′*, *Loc_Os01g20206*; *H9m′*, *Loc_Os09g34320*), and an rRNA methyltransferase (*H8m′*, *Loc_Os08g09940*).

**Fig. 5.** Associations for spikelet sterility response to temperature. (A) Index variable Ster_Slope, derived from correlation of observed sterility vs Tmin; (B) RIDEV parameter for critical temperature KSterCold; (C) RIDEV slope parameter KSterSlope; (D) RIDEV parameter for critical temperature KSterHeat. Results are for Estimation Run1 (Table 2).

**Fig. 6.** GWAS associations for RIDEV parameter KSterCold as in Fig. 3B, but with KSterSlope fixed to constant value (Run2).
In general, for direct thermal response of sterility, many genes functionally related to flowering, flower development or flower fertility were found. For cold acclimation, genes putatively conveying epigenetic control were frequent. Both types of gene functions were less frequent in QTLs for baseline sterility.

**Discussion**

**Heuristic estimation of thermal adaptation traits**

Physiological models have previously been used to extract environment response traits from phenotype data for genetic analyses (phenology: Nakagawa et al., 2005; Yin et al., 2005; morphogenetic processes: Rebollo et al., 2015). Heuristic methods are useful when the raw phenotypic data contain underlying information that requires a physiological key to unlock. They are justified when the a priori physiological theory is sufficiently accurate.

The present study on sterility and its companion on phenology (Dingkuhn et al., 2017) fitted RIDEV to 203 accesses across diverse climatic situations. The model explained 91% of variation of time to flowering, which increased further to 94% when cold acclimation was considered. For sterility, RIDEV explained only 38% of variation but this value nearly doubled to 64% when acclimation was considered. It thus seems that cold-stress hardening during earlier developmental phases is important for sterility responses in rice. The lower predictability of sterility as compared with phenology resided partly in error propagation (sterility prediction is sensitive to timing of stress-sensitive phases), but also in third causes of unfilled spikelets such as panicle architecture and assimilate supply (Sheoran and Sain, 1996).

RIDEV parameters gave mostly different but stronger QTLs as compared with the index variables that were more directly derived from observation. This raises the question of whether the parameters reflected precisely the physiological traits they were meant to express. Parameters were not functionally independent and thus absorbed some of each other’s estimation errors (e.g. slope vs threshold parameters for thermal response). This problem is intrinsic to complex models and particularly to extracting different types of information (heat or cold sterility) from the variation of a single measured variable (unfilled spikelets), climate variation being the discriminating key. Good phenotype predictions thereby are no proof of biological accuracy.

Significant genomic associations with model parameters indicate that the model reflects a biological reality because measurements of phenotype and DNA polymorphisms are independent. The biological function of causative genes, once known, can further inform on the biological meaning of model parameters.

**Control of unexplained baseline sterility**

Rice spikelet sterility is mostly male-derived and depends on pollen production, viability and load on the stigma (Saragih et al., 2013). The developmental and genetic control of male sterility is well documented because of its importance for hybrid seed production (Wang et al., 2013). Among many genes involved in the network, OsMADS3 and OsMADS58 are located upstream, controlling OsUDT1 (a helix–loop–helix protein) through OsMSP1 (a serine/threonine kinase). They control downstream pathways for pollen exine and tapetum development. Genes associated with meiosis were reported by Deveshwar et al. (2011), including OsMSP1 and many cell-cycle controlling RAD genes.

These specific genes were not among our candidate genes. However, a protein kinase controlling the cell cycle, Loc_Os06g47880, was observed in major QTL Sb6(2)m (20 kb from SNP; P = 4.0 × 10−8). It was located in close neighborhood to three serine/threonine kinases. We frequently encountered such clusters of functionally plausible candidate genes. Serine/threonine kinases, known for control functions in reproductive development and stress responses (Becraft, 2002), were also found near the center of three minor QTLs, in one case a cluster of four such genes (QTL Sb2(2)m).

The main SNP of minor QTL Sb1m′ was located within Loc_Os01g06590 (syn. OsRHC4), an anaphase controlling ringfinger gene. It is expressed in the panicle at pistil/stamen primordia differentiation stage IV, a pre-meiotic stage (Ma et al., 2009). Within 24 kb of this gene was a helix–loop–helix TF (Loc_Os01g06640) participating in anther development (Phytozome). At 39 kb from the SNP, an arginyl-tRNA synthase (Loc_Os01g06510) was located that is up-regulated in florets during meiosis under heat and cold stress, associated with male sterility (Pan et al., 2014). This cluster of three genes in QTL Sb1m′ linked to anther/microspore development and spikelet sterility, two of them during meiosis, is probably no coincidence. Mugford et al. (2013) reported ooperon-like, functional and physical gene clusters that are co-expressed and probably common in plants.

Apart from anther and pollen development, protection from oxidative stress is important for pollen viability because spikelets are exposed to desiccation, direct radiation and fluctuating temperatures. Major QTL Sb8m had the anti-apoptotic, ROS-protective, BCL2-type gene Loc_Os08g43270 at its center. This type of gene, when transformed in rice, was reported to induce tolerance to oxidative stress and salinity (Hoang et al., 2015). Li et al. (2006) reported that apoptosis participates in tapetum development and male sterility.

Many candidate genes were involved in signaling. Two genes encoded ADB ribosylation factors (ARF GTPases: Loc_Os06g02390 in major QTL Sb6(2)m; Loc_Os02g42134 in minor QTL Sb2(2)m). Zhuang et al. (2005) reported that such genes exert developmental control through polar auxin transport in Arabidopsis. Du et al. (2011) confirmed developmental effects of ARF GTPases in rice via actin cytoskeletal organization and vesicle transport. Intracellular signaling frequently involves calcium via calmodulin. Wei et al. (2014) reported that a calcium-dependent protein kinase (OsCPK9) regulates spikelet fertility in rice. One such calmodulin-binding protein (Loc_Os11g44170) was the sole candidate gene for major QTL Sb11i (peak SNP located within the gene).
The array of candidate genes for baseline sterility was diverse but this may be due to its multiple possible causes. In general, the abundance of plausible genes at the center of the QTLs supported their physiological function.

Baseline and thermal-sterility QTLs co-localize with MADS-box and other reproduction genes

MADS-box and other reproduction genes were rare in QTLs for baseline sterility. OsELF3 (syn. HD17, Loc_Os06g05060), exerting circadian control of flowering under long days (Yang et al., 2013), was found in minor QTL Sb6(1)i. This gene, expressed in leaves during perception of the photoperiod stimulus, is also strongly expressed in anthers and pistils (http://rice.plantbiology.msu.edu/), and may thus be involved in spikelet fertility. Furthermore, Loc_Os01g06590 in Sb1m’ (SNP located within gene) is an anaphase promoting Zn-finger TF. Its up-regulation under cold stresses (http://genevisible.com/) and expression in male gametes (http://genevisible.com/) makes this gene a causative candidate for spikelet fertility/sterility. Sb1m’ was classified as a minor QTL because of low allelic frequency but had 19 supporting SNPs and $P=1.0 \times 10^{-5}$.

For thermal-stress-induced sterility, candidate genes OsMADS56 (QTL Sc1m, classified as minor due to low allelic frequency; 12 supporting SNPs; $P=7.9 \times 10^{-6}$), a ferredoxin oxidoreductase (phytochromobilin) involved in phytochrome biosynthesis and, putatively, photoperiodism (Phytozome). It shows protein–protein interaction (RiceNetDb) with the ELF3/HD17 found in QTL Sb6(1)i and is strongly expressed in the stigma (http://genevisible.com/). We furthermore found OsSPL19 in minor QTL Sh11m (heat-sterility), a squamosa-binding TF putatively involved in floral development (Xie et al., 2006). Its Arabidopsis homologue, AtSPL13 (Zhu et al., 2013), controls leaf appearance rate (Preston and Hileman, 2013) involving micro-RNA156 (Wang et al. 2008).

For cold acclimation QTLs we identified two meiosis-related candidate genes. Loc_Os01g68000, located close to

![Fig. 7. Genomic associations for RIDEV AcclinSterC parameter. (A) Parameter estimation with KSterSlope co-fitted (Run3, Table 2); (B) Estimation with KSterSlope fixed (Run4).](image-url)
OsMADS96 in major QTL H1(2)m', is also called PLA2 (PLASTOCHRON2) or LHD2 (LEAFYHEAD2) and linked to floral development (http://pathway.plantcollaborative.org). It encodes an ME12-like protein, essential for meiosis in some species but unconfirmed for rice (Anderson et al., 2004). The gene is expressed in pollen and down-regulated under drought, heat and cold (http://genevisible.com/). The other meiotic candidate gene was Loc_Os02g55570 (syn. OsSGO1; RICENCODE) in major QTL H2(2)m' associated with chromosome segregation (Phytozome) and centromer cohesion (OGRO).

These gene-QTL associations suggest linkages between genetic control of flowering and sterility.

**RAS super-family genes and other GTPases are frequent in sterility QTLs**

The RAS super-family of small GTPases consists of five groups (Ras, Rho, ARI/Sar, Ran, and Rab). They are intracellular signaling nodes activated by extracellular stimuli, controlling signaling expression for growth and differentiation (for review see Rojas et al., 2012). Three RAS GTPases (Peng et al., 2014; Xu and Cai, 2014) were associated with QTLs for thermal response of sterility. A RAS suppressor was found for major QTL H11m (cold acclimation), and ARF and RAB genes for baseline sterility QTLs. The Loc_Os01g53600 RAS-protein in minor QTL Ssli is up-regulated under drought (http://genevisible.com/). An intriguing GTPase (Loc_Os07g38730) was found in major QTL Sc7m. It has a tubulin domain, is involved in flower development and leaf senescence, and is expressed during early inflorescence development (Expression Atlas, http://www.ebi.ac.uk/).

The role of these genes in spikelet sterility merits investigation. For example, Xu and Cai (2014) reported the involvement of RAN1 in rice cold tolerance. Nielsen et al. (2008) reviewed the regulatory role of RAB and ARG GTPases in intracellular vesicle transport.

**Serine/threonine kinases are omnipresent in sterility QTLs**

Many serine/threonine kinases link external signals to cellular development or defense responses, and are thus called stress proteins (Dièdhiou et al., 2008; Sinha et al., 2011). For our 32 QTLs, 20 such kinases were genes found within 50 kb of the main SNP, and 36 within 100 kb, frequently in clusters. Most known abiotic-stress-related, receptor-like kinases respond to drought and few to cold (Gao and Xue, 2012). The regulatory function of most kinases is unknown due to their large number, however.

**Candidate genes controlling physiological adaptation to abiotic stresses**

**Heat**

The strongest association for heat was major QTL Sh4m, with five interesting genes clustered within 20 kb of the main SNP. Peak SNP was located within Loc_Os04g49430, a Zn-finger TF regulating growth through OsPIN3 (an auxin transporter involved in drought response; Zhang et al., 2012). Adjacent were an H4 histone, an RNA-binding protein, a MYB-type TF and a pollen lipoprotein, all potentially relevant. Loc_Os04g49430 combines an RCC1 domain involved in chromosome condensation and, thus, probably transcription control (http://www.ricechip.org/). It has a FYVE zinc finger domain known to be involved in lipid signaling (Wywial and Singh, 2010). Which parts of this gene cluster are involved in heat sterility merits investigation.

About 40 kb from OsSPL19 on minor QTL Sh11m we found Loc_Os11g30410, a SUMO-type ubiquitination gene related to male sterility (ALOC_Os1/RAD31 component; Phytozome). Reversible protein ubiquitation is important for heat stress adaptation (Miller and Vierstra, 2011).

**Cold stress**

Proline accumulation is a stress defense involved in cold tolerance, osmoregulation and maintenance of cytoplasmic integrity (Hayat et al., 2012). A proline synthase co-expressed protein (expressed under drought; PlaNet) was found in minor QTL Sc1m and a prolly-tRNA synthase in QTL Ss7i (minor due to few supporting SNPs) (http://ricexpro.dna.affrc.go.jp/). More candidate genes were related to oxidative stress, most frequently in baseline-sterility QTLs (Sb6(1)i and Sb6(1)m, minor; Sb6(2)m, major). Cold-sterility QTL Sc1m’ (classified as minor due to low allelic frequency) had a cluster of six glutathione transferases (GST) at its center, with one (Loc_Os01g72130) strongly expressed in anthers and ovary (http://ricexpro.dna.affrc.go.jp/). Pereira da Cruz et al. (2013) reported that rice GST genes are up-regulated under cold stress and help in controlling oxidative stress. Loc_Os01g72130 is thus a promising candidate for cold tolerance.

A candidate gene for baseline sterility (major QTL Sb2(1) i) but potentially acting through desiccation avoidance was Loc_Os02g31140, a curly-leaf factor controlling cuticle development (OGRO) strongly expressed during early florescence development (http://ricexpro.dna.affrc.go.jp/). Anther cuticle development is important for spikelet fertility (Li and Zhang, 2004). Specifically expressed during early anther development (http://ricexpro.dna.affrc.go.jp/) is also Loc_Os01g06640, a helix–loop–helix (HLH) TF detected in major QTL Sb6(2)m’ (P=4.0 × 10^-8). The bHLH gene family is associated with control of cell/organ size and development (Heang and Sassa, 2012) and stress responses (Lima et al., 2015). Fu et al. (2014) reported a bHLH gene acting as central switch in anther development.

An interesting gene at the center of QTL Ss7i (classified as minor due to few supporting SNPs) was O207g07170, a plastocyanin-domain gene expressed in pollen and flag leaves (http://genevisible.com/), and up-regulated under cold stress. Its two orthologs in Citrus clementina were identified as plantacyanins (Phytozome). Plantacyanins control directional pollen tube growth in Arabidopsis (Dong et al., 2005).
triggered by previously experienced conditions are likely caused by epigenetic control of gene expression. They are restricted to the organism’s life cycle or carried into subsequent generations. Major mechanisms are histone methylation or acetylation and DNA methylation (Chinnusamy and Zhu, 2009), or on a post-transcriptional basis RNA modifications (Balint et al., 2015). Hu et al. (2015) recently reported rice DNA methylymes related to environment-induced male sterility.

The cold-acclimation QTLs were rich in putative epigenetic genes, e.g. a histone-like TF (OsHAP3D in QTL H6m′), a histone acetyltransferase (H1(2) m′), an adenine RNA methyltransferase (H2(1)m′), an RNA methyltransferase (H8m′), three unspecified methyltransferases, an RNA helicase (H7m′) and chaperones. The H3K27 histone demethylase Loc_Os01g67970 (syn. JM1705) in QTL H1(2) m′ was described by Li et al. (2013) and Shi et al. (2015) as being involved in epigenetic control of flowering. It is predominantly expressed in rice sperm cells and pollen (http://genevisible.com/).

Although the acclimation QTLs stood out for the abundance of putative epigenetics genes, a variety of other candidate genes were also identified. Loc_Os08g10010 in QTL H8m′ encodes a fatty acid desaturase (FAD) highly expressed in the stigma (http://genevisible.com/). Cold-tolerant rices increase membrane fluidity through fatty acid desaturation when chilled (Pereira da Cruz et al., 2009). Nair et al. (2009) demonstrated that expression of OsFAD8 in rice is necessary to express cold tolerance.

Epigenetic control of responses to thermal stresses in plants is poorly understood. Its experimental study requires stress treatments at various developmental stages. Our results suggest that epigenetic control of cold adaptation is important in rice, particularly where chilling occurs early enough for acclimation. This is the case in the Madagascar highlands, and indeed we observed acclimation alleles predominantly in traditional rices from there. Radanielina et al. (2013) and Dingkuhn et al. (2015a) reported this group to be cold tolerant and genetically distinct.

**Functional intersection between phenology and spikelet sterility**

The companion paper on phenology and this study on spikelet sterility will enable the construction of physiological and genetic control networks but this is beyond this paper’s scope. It appears that phenology not only affects sterility through the timing of stress exposure (Dingkuhn, 1995). The control of both flowering and spikelet fertility depends on networks controlling development processes, involving some gene families best known to control flowering (e.g. OsMADS). This is unlikely to be an artifact of incomplete separation between traits by RIDEV because sterility and flowering date were independent measurements and the model accounted for phenology when relating sterility to temperature.

**Perspectives**

It is hoped that this research will enhance breeding for adaptation to adverse and variable climates. The present and the companion study defined physiological component traits through crop-model assisted dissection of phenomics data. They produced QTLs and candidate genes on 203 accessions. The next steps are to narrow down the candidate genes by haplotype analysis using a new 700 000-SNP array (McCouch et al., 2016) and re-sequenced genomes, and to construct integrated gene networks. For causative SNPs, once discovered, a promising shortcut to crop improvement will be CRISPR/Cas9-based genome editing (Bortesi and Fischer, 2015) to express them in elite backgrounds, both for validation and developing climate-change resilient varieties.

**Supplementary data**

Supplementary data are available at *JXB* online.

Protocol S1. C++ code for RIDEV.

Fig. S1. Genetic structure of panel.

Fig. S2. Quantile–quantile (QQ) plots for Manhattan plots.

Fig. S3. Diurnal template for temperature dynamics used for hourly calculations.

Fig. S4. Frequency distributions of index variables.

Fig. S5. Sensitivity analysis of RIDEV simulations to variation of crop parameters.

Table S1. List of germplasm.

Table S2. Complete lists of GWAS associations.

Table S3. RIDEV V2 parameters and input/output variables.

**Acknowledgements**

The authors would like to thank Drs Brigitte Courtois and Nourollah Ahmadi for valuable assistance in data analysis, as well as the excellent field staff of Africa Rice Center in Senegal and the Cirad-FOFIFA team in Madagascar.

**References**


Saragih AA, Pteh AB, Ismail MR, Mondal MMA. 2013. Pollen quality traits of cultivated (Oryza sativa L. ssp. indica) and weedy (Oryza sativa var. nivara) rice to water stress at reproductive stage. Australian Journal of Crop Science 7, 1106–1112.


