Published August 30, 2016

RESEARCH

Farmer Participatory Early-Generation Yield Testing of Sorghum in West Africa: Possibilities to Optimize Genetic Gains for Yield in Farmers' Fields

H. Frederick W. Rattunde,* Sebastian Michel, Willmar L. Leiser, Hans-Peter Piepho, Chiaka Diallo, Kirsten vom Brocke, Bocar Diallo, Bettina I. G. Haussmann, and Eva Weltzien

ABSTRACT

The effectiveness of on-farm and/or on-station early generation yield testing was examined to maximize the genetic gains for sorghum yield under smallholder famer production conditions in West Africa. On-farm first-stage yield trials (augmented design, 150 genotypes with subsets of 50 genotypes tested per farmer) and second-stage yield trials (replicated α-lattice design, 21 test genotypes) were conducted, as well as on-station α-lattice first- and second-stage trials under contrasting phosphorous conditions. On-farm testing was effective, with yield showing significant genetic variance and acceptable heritabilities (0.56 in first- and 0.61 to 0.83 in second-stage trials). Predicted genetic gains from on-station yield trials were always less than from direct testing on-farm, although on-station trials under low-phosphorus and combined over multiple environments improved selection efficiencies. Modeling alternative designs for on-farm yield testing (augmented, farmer-as-incomplete-block, multiple lattice, and augmented p-rep) indicated that acceptable heritabilities (0.57 to 0.65) could be obtained with all designs for testing 150 progenies in 20 trials and 75 plots per farmer. Ease of implementation and risk of errors would thus be key criteria for choice of design. Integrating results from on-station and on-farm yield testing appeared beneficial as progenies selected both by on-farm and on-station first-stage trials showed higher on-farm yields in second-stage testing.

H.F.W. Rattunde, International Crops Research Institute for the Semi-Arid Tropics, BP 320, Bamako, Mali; S. Michel, Dep. for Agrobiotechnology (IFA-Tulln), Institute for Biotechnology in Plant Production, Univ. of Natural Resources and Life Sciences, Vienna (BOKU), Konrad-Lorenz-Str. 20, 3430 Tulln, Austria; W.L. Leiser, State Plant Breeding Institute, Univ. of Hohenheim, Stuttgart, Germany; H.-P. Piepho, Biostatistics Unit, Institute for Crop Science, Univ. of Hohenheim, Stuttgart, Germany; B.I.G. Haussmann, Institute of Plant Breeding, Seed Science and Population Genetics, Univ. of Hohenheim, Stuttgart, Germany; C. Diallo, Institut Polytechnique Rural de Formation et de Recherche Appliquée, BP 6, Katibougou, Mali; K. vom Brocke, CIRAD, FOFIFA, SRR Antsirabe BP 230, Antsirabe 110, Madagascar; B. Diallo, Institute d'Economie Rural, BP 262, Bamako, Mali; E. Weltzien, International Crops Research Institute for the Semi-Arid Tropics, BP 320, Bamako, Mali. Received 15 Dec. 2015. Accepted 21 Mar. 2016. Assigned to Associate Editor Jode Edwards. *Corresponding author (f.rattunde@gmail.com).

Abbreviations: BLUE, best linear unbiased estimate; BLUP, best linear unbiased prediction; DAP, diammonium phosphate; FPB, formal plant breeding programs; G × E, genotype-by-environment; GGE, genotype-genotype-by-environment; G × Y, genotype-by-year; –P, low phosphorus; +P, high phosphorus (fertilizer added to trial); p-rep, partially replicated; PGND, population Guinea Naine Diversifié; $RSE_{St:Fa}$, relative selection efficiency for indirect on-station versus direct on-farm yield testing; VD_{BLUP}, variance of differences of two best linear unbiased predictions; WCA, West and Central Africa.

Sorghum [*Sorghum bicolor* (L.) Moench] is cultivated on >42 million hectares worldwide, of which nearly 60% is in Africa and the largest share in West and Central Africa (WCA; FAOSTAT, 2014). For West African farmers, sorghum is a staple crop due to its adaptation to low soil fertility (Leiser et al., 2012), climate variability (Haussmann et al., 2012), and heat and drought tolerance (Henzell and Jordan, 2009), attributes attained through several

Published in Crop Sci. 56:2493–2505 (2016).
doi: 10.2135/cropsci2015.12.0758

thousand generations of farmer and natural selection since its domestication. In the context of increasing demographic pressures, farmers in West and Central Africa are looking for new sorghum varieties offering increased grain yield while ensuring adaptation to agroclimatic conditions and maintaining specific grain qualities for processing and consumption (Weltzien et al., 2008a; vom Brocke et al., 2010). A major challenge for effectively breeding for increased sorghum yield in WCA is the complexity of environments and the associated genotype-by-environment (G × E) interactions across small-holder farmer conditions. Even within a geographically targeted agroecosystem, diversity for factors such as soil type and depth, the timing and manner of weeding and fertilizer application, and the date of sowing result in important G × E interactions (Rattunde et al., 2013).

Furthermore, WCA sorghum farmers typically rely on low-input cropping methods, where limited plant-available phosphorous on highly weathered soils is a major production constraint (Buerkert et al., 2001; Leiser et al., 2012). Formal plant breeding programs (FPB), in contrast, are conducted in experiment stations typically managed with higher inputs and timely weeding and fertilizer applications. Breeding programs, even those targeting low-input production systems, typically prefer to carry out the initial selection stages under favorable research station conditions where heritabilities, genetic variance, and repeatabilities are high compared with the more heterogeneous and lower-yielding on-farm conditions (Dawson et al., 2008). Preliminary selection of progenies under high-yielding FPB conditions, however, may reduce the genetic variance and selection intensities in subsequent on-farm testing, resulting in lower genetic progress for performance under farmers' low-input conditions (Bänziger and Cooper, 2001). Therefore, direct selection for yield performance on-farm using on-farm, farmer-managed trials could achieve higher gains than selection under more favorable but nonrepresentative on-station conditions (Atlin et al., 2001; Ceccarelli and Grando, 2007). Also many farmers are interested to see new breeding materials and are willing to conduct trials, and thereby provide access to the large number of test environments required to sample the target population of environments in which new varieties need to perform (Haussmann et al., 2012; vom Brocke et al., 2014).

Early-generation on-farm yield testing, however, can have greater within-field, site-to-site, and year-to-year heterogeneity that may reduce its advantage over FPB for achieving genetic gains for small holder farmers (Atlin et al., 2001). Additionally, testing a large number of genotypes on-farm can be difficult (Mangione et al., 2006) and must match both farmers' time and land availability and researchers' logistical capacity. Thus, simple and robust multiple environment trial designs for on-farm progeny yield testing in the early-generation are needed to achieve gains that would exceed those of FPB programs (Atlin et al., 2001; Bänziger and Cooper, 2001).

The goal of this study was to identify the most promising options for on-farm and/or on-station testing of sorghum progenies for grain yield to maximize the genetic gains for yield under smallholder famer production conditions in Mali, West Africa. Statistical analysis and modeling methods were applied to data from on-farm and on-station sorghum breeding trials conducted by the Institut d'Economie Rural (IER), the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), and collaborating development and farmer organizations in Mali. The specific objectives were to (i) determine the usefulness (repeatabilities) of current unreplicated early-generation and replicated advanced-generation on-farm sorghum trials, (ii) assess the extent and type of G × E interaction for panicle yield under on-farm testing conditions, (iii) compare the responses to selection using on-station versus on-farm performance data for enhancing yields in farmers' fields, and (iv) assess the utility of alternative trial designs for on-farm yield testing with a large number of sorghum genotypes via simulation studies.

MATERIALS AND METHODS

Plant Materials

A random-mating Guinea-race sorghum population was generated by crossing 13 *Guinea*-race landrace varieties from Mali and Burkina Faso to a source of genetic male-sterility (ms3), and subsequent backcrossing the landraces once (10 accessions) or twice (3 accessions) before bulking and random mating for two cycles prior to initiating mass selection (Rattunde et al., 1997; Weltzien et al., 2007). Although this population was more than 3 m tall due to all donor parents having nondwarf stem internode lengths of 20 to 30 cm, the source of genetic male-sterility contributed dwarfing genes that enabled selection of short segregants in the random-mated population. A short statured population (population Guinea Naine Diversifié, PGND) was initiated in 2004 by random-mating more than 225 single panicle selections or bulks of farmers' selections of short stem-internode plants (Rattunde et al., 2009).

More than 1000 progenies from the PGND population were derived by farmers practicing single-plant selection for panicle characteristics such as grain density, threshability, size, and hardness in isolation plots on their farms in either 2007 or 2008. These progenies were observed in nurseries at the ICRISAT-Samanko Research Station for plant height, panicle appreciation, and disease resistance in 2009 as $S_{2.1}$ (2007 selections) or $S_{1.0}$ (2008 selections) progenies. A total of 100 short- and 50 longer-internode progenies in either $S_{2.0}$ and $S_{3.1}$ generation were retained for early-generation yield testing in 2010 under both on-farm and on-station trials.

A set of 21 progenies were selected from among the 100 short-internode early-generation progenies for evaluation in advanced generation testing. Selection was conducted only among the short-internode progenies as they were of higher priority for developing dual-purpose grain-fodder varieties

and, by limiting the height variation among test entries, the neighbor effects could be reduced. The 21 selected progenies jointly represent the 12 highest yielding genotypes identified from the on-farm trial results and the 12 highest yielding genotypes in the on-station trials, with three genotypes being common to both sets.

The 12 progenies with highest yields from the 2010 on-farm testing were identified by using the best linear unbiased estimates (BLUEs) of panicle yields based on a REML analysis in GENSTAT in each of the four short-internode sets (25 progenies each). Individual progeny yields were expressed as relative yield, dividing the test entry BLUE for panicle yield by the mean of two check varieties' yield BLUEs within each trial. The overall performance of each genotype was computed as an index over all trials of that set by summing the relative individual trial yield ratios, weighted by its repeatability estimate, using 4, 3, 2, and 2 trials with useable data for the sets A, B, C, and D, respectively. Subsequently, three genotypes were selected in each of the four sets of 25 genotypes based on rank for yield index values within each set, resulting in a 12% selection intensity.

Selection of the 12 progenies with highest yield from the 2010 on-station testing of 100 short-internode progenies was also done using grain yield data from a low (–P) and a high (+P) phosphorous-managed trial. A simple index composed of standardized ($\mu = 0$, $\sigma = 1.0$) BLUEs for grain yield in –P and +P trials, with each value weighted by the respective trial repeatability, was computed. The highest ranking 12 genotypes for their yield index values were retained for advanced testing.

All selected progenies were advanced by selfing and bulking panicles within families from plants of similar height and panicle aspects. The 21 progenies used for the advanced generation yield testing in 2011 and 2012 were thus in the S_3 or S_4 generation.

Design of On-Farm Trials

The 2010 first-stage yield trials and the 2011 and 2012 second-stage trials were conducted with farmers in southern Mali, where sorghum-based production systems predominate (Fig. 1; Table 1). Trials were conducted each year in three regions; the Dioila and Koutiala regions, located 150 and 300 km east of Bamako, both with more intensified production systems with a long history of cotton (*Gossypium hirsutum* L.) production, and the Mande zone, 80 km southwest of Bamako, where relatively little cotton is produced and the production system is less intensified (Weltzien et al., 2007). Trials were coordinated in the Dioila region by a farmer union of cooperatives, Union Locale des Producteurs de Cereals de Dioila (ULPC), in the Koutiala region by a local NGO, Association Malienne d'Eveille pour un Development Durable (AMEDD), and in the Mande region by the farmer organization Association des Organisations Paysans Professionels (AOPP).

First-Stage Unreplicated On-Farm Yield Trials, 2010

The 150 genotypes in the $S_{2.0}$ or $S_{3.1}$ generation were divided into six sets of 25 genotypes, two sets of longer stem-internode and four sets with shorter stem-internode genotypes. Each farmer was randomly allocated two of the six sets (Supplemental Fig. S1). An augmented design was used by adding to each subset the two check varieties; 'Tieble', an adapted landrace, and 'Lata', an elite bred variety. Both check varieties were added to each subblock of five test entries, with random allocations to plots within each subblock. Each set was sown in one contiguous block of 35 plots.

All farmers ridged their fields using animal traction as is the common practice in the major sorghum-growing zones of Mali, with distance between ridges varying between 50 and 80 cm. Plots consisted of a single row of 6 m length sown on the ridge, with 30 cm spacing between hills. The plots were thinned to two plants per hill. The second set of genotypes followed the first set on the same ridges, with a 1.5 m alley between sets. Farmers applied an equivalent of 100 kg ha^{-1} diammonium phosphate (DAP) and 50 kg ha^{-1} urea on their trials and hand weeded their trials according to their own schedule and practice.

A total of 34 trials were conducted in 20 villages (Fig. 1; Table 1). Panicle yield was determined by weighing harvested panicles dried at ambient temperatures for at least 2 wk in all trials, both on-farm and on-station. Threshed grain yield was not used in this study, so as to facilitate rapid data collection and minimize errors due to grain loss during threshing in the on-farm trials.

Second-Stage Replicated On-Farm Trials, 2011 and 2012

The 21 selected short-internode genotypes along with two (2011) or four (2012) check varieties were tested in α-lattice designs (incomplete block size 5) with three replicates in 2011 and 2012 (Supplemental Fig. S1). Single-row plots of 6 m were used, with each replication forming a single band of 25 rows. The three replicates were sown following one another on the same ridges, with 1.5 m alley between replicates. A total of 38 trials were conducted in 13 villages (Fig. 1; Table 1).

Design of On-Station Trials

On-station yield trials conducted at the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) station at Samanko, Mali, evaluated the same genotypes tested on-farm in the that year (Table 1). In 2012, two additional sites were used for testing; the Institut d'Economie Rurale (IER) Sotuba and Kolombada experiment stations (Fig. 1). All on-station trials used an α-lattice design with four replicates. Plots consisted of two 3 m rows with 0.75 m distance between rows and 0.30 m between hills. Hills were thinned to two plants per hill.

The yield trials at ICRISAT-Samanko were conducted under both +P and –P conditions in each year (2010–2013). The +P trials were conducted in experimental fields with yearly applications of DAP (100 kg ha^{-1}) as basal fertilizer. The –P trials were conducted in a field continuously cropped since 2006 with no applications of phosphorous-containing inorganic fertilizers, but did receive an amount of nitrogen equivalent to that applied to the +P fields. Thus, the –P trials received 37 kg ha^{-1} urea within the first 2 wk after sowing, incorporated in the ridge adjacent to the sown row, and both –P and +P trials received urea (50kg ha^{-1}) as topdressing split into two applications at approximately three and 6 wk after sowing. The –P and +P trials were planted on the same day in adjacent fields.

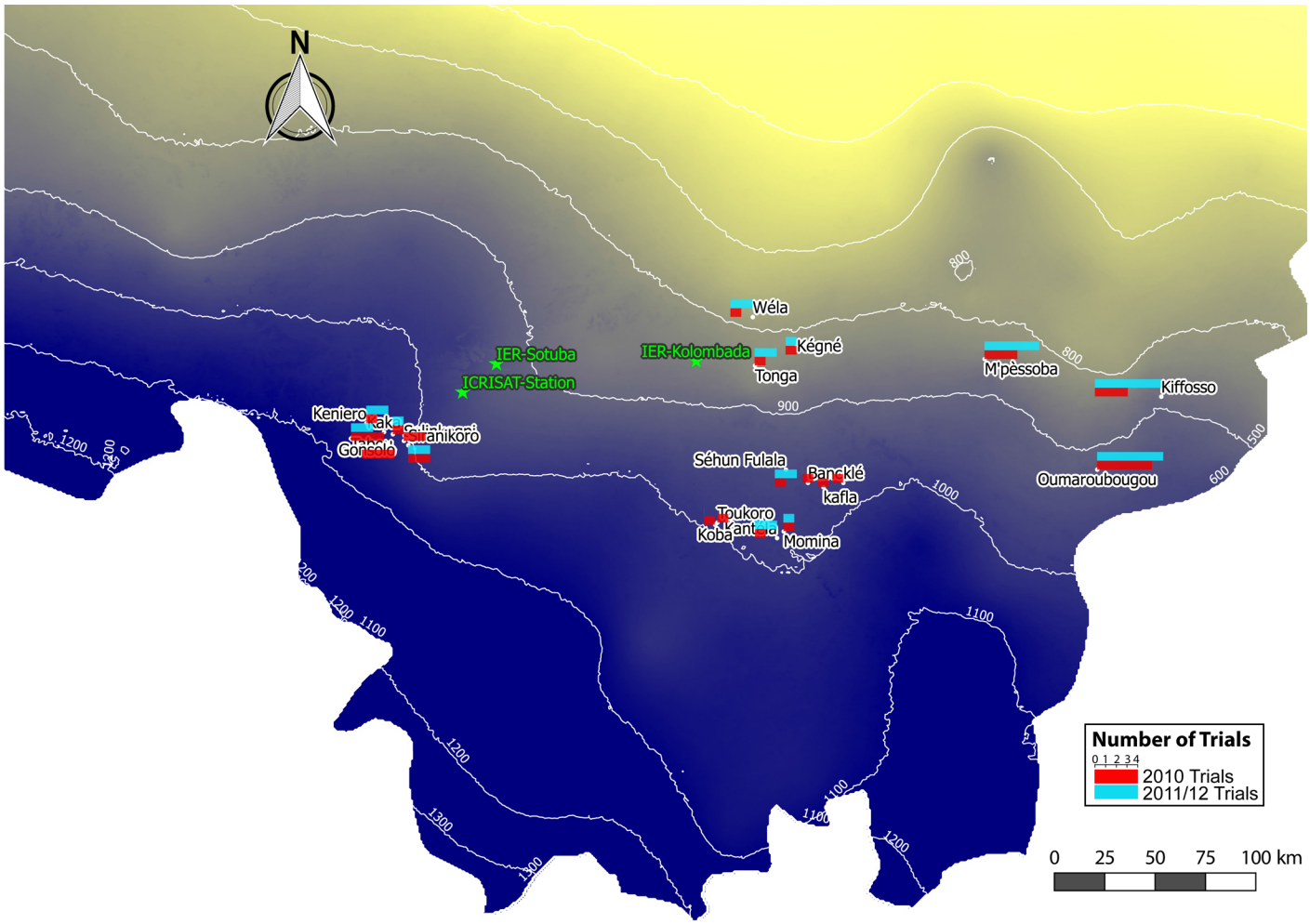

Fig. 1. Locations and number (indicated by bar length) of first-stage (red) and second-stage (blue) sorghum yield trials conducted in south central Mali. Isohytes show the annual precipitation (in mm) averaged for the period 1950 to 2000 (Credit: www.worldclim.org).

Table 1. The number of progenies tested in total and in individual trials and the number of on-farm and on-station trials conducted and analyzed for first- and second-stage testing.

Testing stage	No. of progenies tested			Year	No. of on-farm trials		No. of on-station trials	
	Total	per on-farm trial	per on-station trial		Conducted	Analyzed	Conducted	Analyzed
First	150	50	50 long-internode 100 short-internode	2010	34	20	2 long-internode 2 short-internode	2 long-internode 2 short-internode
Second	21	21	21	2011	19	16	2	1
				2012	19	17	4	3
				2013	0	0	2	2

First-Stage On-Station Yield Trials, 2010

The 100 short- and 50 long-stem internode genotypes were tested separately in "short" and "tall" trials, respectively, to reduce neighbor effects. Ten check varieties were included in each trial, with eight varieties occurring once per replication and the two varieties used in the on-farm trials occurring three (tall-trial) or four (short-trial) times. An α-lattice design with four replicates and incomplete blocks of four plots was used for all trials. The first two replicates of the –P short and –P tall trials occurred in the portion of the –P field that was limed (1.5 t ha^{-1}), whereas the third and fourth replicates occurred in the part of the field that received gypsum (0.2 t ha^{-1}).

Second-Stage On-Station Yield Trials, 2011, 2012, and 2013

The 21 genotypes and four check varieties were evaluated in α-lattice designs (incomplete block size 5) with four replicates in 2011, 2012, and 2013. Trials were conducted each year at ICRISAT-Samanko under both +P and –P conditions and at the IER-Sotuba and IER-Kolombada experiment stations in 2012 with the same fertilization as the +P trials at ICRISAT-Samanko.

Estimation of Quantitative-Genetic Parameters

A linear mixed model was set up for an analysis across trials, following the methodology presented by Piepho et al. (2003):

$$y_{ijkl} = \mu + g_i + t_j + (gt)_{ij} + r_{jk} + b_{jkl} + \varepsilon_{ijkl} \quad [1]$$

with y_{ijkl} designating the panicle yield of the *i*th genotype in the *j*th trial, *k*th replicate, and *l*th block, and the model consisting of the grand mean (μ), the effect of the *i*th genotype (g_i), the *j*th trial (t_j), the effect of the *k*th replicate within the *j*th trial (r_{jk}), the effect of the *l*th block within the *k*th replicate in the *j*th trial (b_{jkl}), the interaction of *i*th genotype with the *j*th trial ($gt)_{ij}$, and the residual effect (ε_{ijkl}). Model [1] was modified for the first-stage on-farm trials by replacing the replicate effect r_j by a set effect s_j to account for the random allocation of two out of six subsets to each farmer. Variance components and best linear unbiased predictions (BLUPs) of genotypic effects were separately derived for each year as well as each series of on-station and on-farm trials. The on-station trials with short and intermediate internode genotypes were analyzed together within the common phosphorous treatment, with +P or –P trials analyzed separately. All design and treatment effects, as well as their interaction, were assumed to be random. Computations for estimating repeatabilities in single environments were based on BLUEs of genotypic effects using Formula 19 from Piepho and Möhring (2007), whereas heritabilities over multiple environments were based on BLUPs of genotypic effects (Cullis et al., 2006; Piepho and Möhring, 2007):

$$h^2 = 1 - \frac{VD_{BLUP}}{2\sigma_G^2} \quad [2]$$

where σ_G^2 is the genotypic variance and VD_{BLUP} the mean variance of a difference of the BLUPs.

For the combined analysis across the 2 yr of the advanced generation trials on-farm or on-station, the corresponding model was:

$$y_{ijkl} = \mu + g_i + e_j + (ge)_{ij} + r_{jk} + b_{jkl} + \varepsilon_{ijkl} \quad [3]$$

In this random effect model, y_{ijkl} is the observation of the *i*th genotype, in the *j*th environment, in the *l*th block of the *k*th replicate of the experimental design. The intercept is given by μ and the interaction of the *i*th genotype with the *j*th environment is denoted as $(ge)_{ij}$, where each environment was a combination of both year and trial within year effects. The experimental design is accounted for by the block b_{jkl}, replicate r_{jk}, and residual ε_{ijkl} effects.

To illustrate patterns of G × E interaction, a genotype-genotype-by-environment (GGE) biplot analysis was conducted in Genstat 17 (VSN International, 2014). The means of the two check varieties Grinkan and Ngolofing were dropped from the analysis since they were not present in all datasets.

A stability analysis was conducted for yield performance of 21 selected genotypes over the second-stage on-farm trials of 2011–2012 using the stability variances (Shukla, 1972) and linear regression of each genotype on the average yield of all genotypes in the studied environments (Finlay and Wilkinson, 1963), which are strongly related to the static and dynamic concept of stability, respectively.

Comparison of Response to Selection On-Station vs. On-Farm

The genetic correlations between on-farm and on-station performance were estimated from

$$y_{ijklm} = \mu + g_i + s_j + (st)_{jk} + (gs)_{ij} + (gst)_{ijk} + r_{jkl} + b_{jklm} + \varepsilon_{ijklm} \quad [4]$$

where s_j is the site effect with the two levels *station* and *farm*. By imposing the unstructured variance-covariance structure on $g_i + (gs)_{ij}$, with the genetic variances on the diagonal and the covariance between both sites on the off-diagonal, it was possible to calculate the genetic correlation between the on-station and on-farm performance as

$$r_{Station;Farm} = \frac{\sigma_{Station;Farm}}{\sqrt{\sigma^2_{Station} \times \sigma^2_{Farm}}} \quad [5]$$

The relative selection efficiency of indirect selection on-station versus direct selection on-farm ($RSE_{St:Fa}$) was estimated with the same selection intensities, using the formula

$$RSE_{St:Fa} = \frac{h_{Station}}{h_{Farm}} \times r_{Station;Farm} \quad [6]$$

where $RSE_{St:Fa}$ is the relative selection efficiency based on the square root of the heritabilities on-station ($h_{Station}$) and on-farm (h_{Farm}) and the genetic correlation between on-farm and on-station performance ($r_{Station;Farm}$).

Comparison of Trial Designs for On-Farm Testing

Different trial designs were considered that could conceivably respond to the following four requirements for effective early-generation on-farm progeny yield-testing. All trial designs should (i) represent the target population of environments, (ii) evaluate a large number of progenies, (iii) limit the number of plots per individual farmer's trial, and (iv) result in trial repeatability sufficient to effectively discriminate among the progenies under test. Each design was modeled assuming 150 genotypes under test, 20 participating farmers, and each farmer provided an area in his or her field accommodating 75 single-row plots.

Description of Alternative Designs

Design 1: Augmented design. The 150 progenies were subdivided into six subsets, each containing 25 genotypes. For each set, a different randomization was created using an α-lattice design with 25 genotypes per replicate and a block size containing five test entries. Pairs of sets of 25 test entries were allocated to farmers in a diallel manner over 15 farmers, assuring that each combination occurred at least once. The remaining five farmers were assigned sets such that each set occurred at least once. Two check cultivars were assigned to each block, with 10 blocks per farmer trial, resulting in 70 plots per farmer. Five additional check plots were randomly allocated to the 10 blocks per farmer to reach the limit of 75 plots. Finally, all blocks were randomized. This design is similar to the design used in the unreplicated on-farm trials in 2010, with the modification of adding five plots to obtain a total of 75 plots per farmer trial.

Design 2: Multiple lattice design. The 150 progenies were randomly assigned to one of five subsets, each subset containing

30 genotypes. Each farmer trial evaluates one subset in an α-lattice design with two replicates. Five incomplete blocks occurred within each replicate, with random allocation of six test genotypes and one check variety to each block, and five additional check plots were randomly allocated to different blocks over the two replicates to complete the total of 75 plots available per farmer. Each subset is tested by a total of four farmers. The combined analysis over all five subsets is feasible due to the common check genotypes (Piepho et al., 2006).

Design 3: Farmers-as-incomplete-blocks design. This design, described by Atlin et al. (2002), randomly allocated half of the 150 progenies to one farmer, and the remaining 75 genotypes to a second farmer, with each pair of farmers' trials comprising the full set. This randomization provided a resolvable design, where every genotype occurs once per pair of farmers and 10 times in the entire design over the 10 pairs of farmers. This design has only one replication per trial and no inclusion of check varieties.

Design 4: Augmented partially replicated design. This design combines both an augmented design (Federer, 1961) and a partially replicated (p-rep) design (Cullis et al., 2006) by replicating a subset of genotypes in an α-lattice design at each farmers' location, and subsequently assigning the remaining unreplicated genotypes to the incomplete blocks (Williams et al., 2011). We used a 15% replication level with 10 genotypes replicated twice, and 55 genotypes unreplicated in each farmer trial, totaling 75 plots per trial.

Methods Used for Determining Utility of Each Design and Comparing Designs

The four different experimental designs were created in CycDesign (VSN International, 2014), leading to different fixed and random effect design matrices, which were extracted and used in a mixed model analysis. In this mixed model analysis, the trial main effect was modeled as fixed for the subsequent comparison of different experimental designs, whereas all other effects were considered random. Variance components were derived from analysis of panicle yields in the 2010 early-generation on-farm trials (Table 2) and constrained to their initial parameter estimates in the comparison of the alternative trial designs. In cases where an effect had to be dropped to set up an appropriate model for the design under consideration, the variance of that effect was added to the residual variance.

Table 2. Genotypic variance (σ^2_G), genotype × trial interaction variance (σ^2_{GT}), replicate variance (σ^2_{REP}), block variance (σ^2_{BLOCK}), and residual variance (σ^2_e) values and percentage of total variance used for simulating the efficiency of alternative trial designs for early-generation on-farm progeny yield evaluation.

Parameter	Value	Percentage of total variance
σ^2_G	703.4 ± 156.2	13.2
σ^2_{GT}	1091.2 ± 193.9	20.4
σ^2_{REP}	409.7 ± 227.7	7.7
σ^2_{BLOCK}	974.5 ± 166.2	18.2
σ^2_e	2167.0 ± 166.2	40.5

To assess the usefulness of each alternative design, we estimated the heritability and the mean variance of a difference. The broad-sense heritability was estimated based on Model [2], which maximizes the expected response to selection in early-generation trials when genotypic effects are random (Cullis et al., 2006). Specific estimates of mean variance of a difference of the BLUPs were determined for each design according to the fixed and random effect matrices of each design.

Analogues to previous analyses, Model [1] was used to estimate the heritability for the augmented and multiple lattice designs. Following Williams et al. (2011), the replicate effect was dropped from Model [1] for analyzing the augmented p-rep design.

A simple linear mixed model was used for the farmer-as-incomplete-block design:

$$y_{ij} = \mu + g_i + f_j + \varepsilon_{ij}, \quad [7]$$

where y_{ij} is the yield of the *i*th genotype in the *j*th farmer's field. Both the grand mean μ and the effect of the *j*th farmer f_j were fixed, whereas the effect of the *i*th genotype g_i was random. It has to be taken into consideration that, in the analysis of this kind of trial design, the variance of the residual effect ε_{ij} comprises all the genotype × farmer interaction, replicate and block variances, as well as the unexplained stochastic variation. Both a resolvable and unresolvable incomplete block design were evaluated, where in the former a combination of two on-farm trials contains each genotype exactly once.

All analyses were conducted with the software language R v.2.10.0 (R Development Core Team, 2011) for statistical computing and graphics. Estimation of variance components and other mixed model applications were implemented with v.3 of the ASReml software package for R (Butler et al., 2009).

RESULTS

Twenty of the 34 first-stage on-farm trials sown in 2010 could be analyzed (8 Koutiala, 7 Dioila, 5 Mande), and the remaining 14 were unusable due to losses caused by uncontrolled animal grazing, flooding, or bird feeding on sown seed resulting in low plant stands (Table 1). The replicated second-stage on-farm trials, in contrast, had only four failures out of a total of 38 trials over 2011 and 2012.

Yield results were obtained from the on-station +P trials in all years, whereas yield results from –P testing were obtained only in 2010 and 2013. Both the 2011 and 2012 –P trials suffered such severe damage from sorghum midge, *Contarinia sorghicola* (Coquillett), that they were dropped from the study.

Repeatability and Mean Yield Levels of On-Farm Trials

Panicle yield levels of on-farm trials ranged from approximately 100 g m^{-2} to over 350 g m^{-2} in the 2010 first-stage (Fig. 2A) as well as the second-stage trials in 2011 and 2012 (Fig. 2B). The trials were somewhat skewed toward lower productivity, with most trials having mean panicle yields <200 g m^{-2} in both the first- and second-stage trials (Fig. 2). The mean on-farm panicle yield was highest in

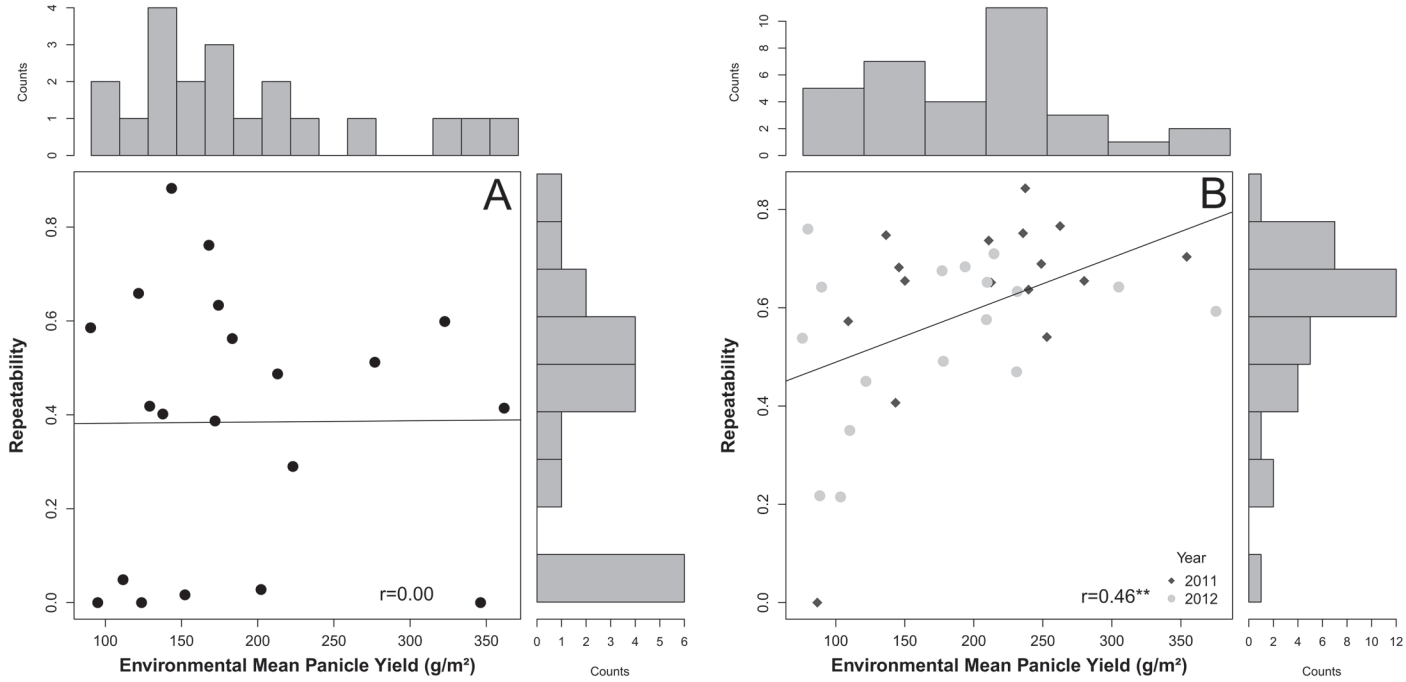

Fig. 2. Relationship between individual trial repeatability and trial mean productivity level of (A) first-stage augmented design trials in 2010, and (B) second-stage α-lattice design trials in 2011 and 2012.

2011, the year with the lowest average total and lowest August rainfall amounts (Table 3). The on-station mean panicle yields were lower in the –P trials (191 g m^{-2}) compared with the +P trials (286 g m^{-2}) for the 2010 first-stage trials, with the 2011 and 2012 +P trials (284 g m^{-2}) similar to +P in 2010.

The repeatability estimates of individual on-farm yield trials ranged from 0 to more than 0.8 in both the unreplicated augmented first-stage trials (Fig. 2A) and replicated, α-lattice, second-stage trials (Fig. 2B). The first-stage trials had a mean repeatability of 0.38, with 65% of trials having repeatabilities superior to 0.3. The mean repeatability of second-stage trials was 0.59, with 91% of trials having repeatabilities superior to 0.3.

Whereas there was no relationship between repeatability and productivity level among the unreplicated first-stage trials (Fig. 2A), a weak positive relationship existed among the replicated second-stage trials which, after dropping trials with repeatability estimates less than 0.3, became nonsignificant.

Table 3. On-farm trial minimum, maximum, and mean best linear unbiased predictions of genotypic panicle yield (GY) and mean rainfall total for the year (Rain_Year) and for the month of August (Rain_Aug) over the three zones of on-farm testing by year.

Parameter	First-stage trial 2010	Second-stage trial 2011	Second-stage trial 2012
GY_Min, g m^{-2}	123	116	149
GY_Mean, g m^{-2}	185	206	176
GY_Max, g m^{-2}	236	263	205
Rain_Year, mm	1100	786	1127
Rain_Aug, mm	296	204	375

Heritability and G × E Interaction across a Series of On-Farm Trials

Combined analysis of panicle yields over all 20 first-stage on-farm yield trials showed significant genotypic variance for panicle yield (Table 4). The genotype × trial interaction variance component was also significant and of a similar magnitude to the genotypic variance component, resulting in a heritability estimate of intermediate magnitude.

Table 4. Estimated quantitative-genetic parameters and relative selection efficiency ($RSE_{St:Fa}$) of indirect selection for panicle yield on-station under phosphorus fertilized (+P), nonfertilized (–P), or combined over both +P and –P conditions versus direct on-farm selection in first-stage yield testing in 2010.†

Site	σ^2_G	σ^2_{GT}	h^2	r_P	r_G	$RSE_{St:Fa}$
Station (+P)	4045 ± 1187	1078 ± 9560	0.66	0.36	0.65	0.71
Station (–P)	1698 ± 6070	376 ± 491	0.53	0.38	0.84	0.82
Station (combined)	2987 ± 443	658 ± 186	0.78	0.46	0.71	0.84
Farm	511 ± 109	617 ± 116	0.56			

† Genotypic variance (σ^2_G), genotype × trial interaction variance (σ^2_{GT}), heritability (h^2), rank correlation of on-station and on-farm predictions of genotypic effects (r_P), genotypic correlation between on-station and on-farm trials (r_G) relative selection efficiency of indirect selection on-station versus direct selection on-farm ($RSE_{St:Fa}$).

Combined analyses within year over the second-stage on-farm trials also showed significant genetic variance (Table 5). Genotype × trial interactions were significant in both years, with the genotypic variance exceeding genotype × trial interaction variance (1:0.5) in 2011 but not in 2012 (1:1.36). Panicle yields in 2012 were lower (Table 3), with genotypic variance being reduced more than the genotype × trial interaction as compared with the 2011 estimates.

Combined analysis of the second-stage on-farm trials over 2011 and 2012 showed presence of significant genotypic variance, but also significant and important genotype-by-year (G × Y) as well as genotype-by-trial within year interactions, with variance component ratios of 1:1:1.3 for $\sigma^2_G : \sigma^2_{GY} : \sigma^2_{GYT}$ (data not shown). The presence of G × Y interaction can be also seen in the GGE biplot, where 2011 and 2012 test environments are separated into fairly distinct megaenvironments (Fig. 3).

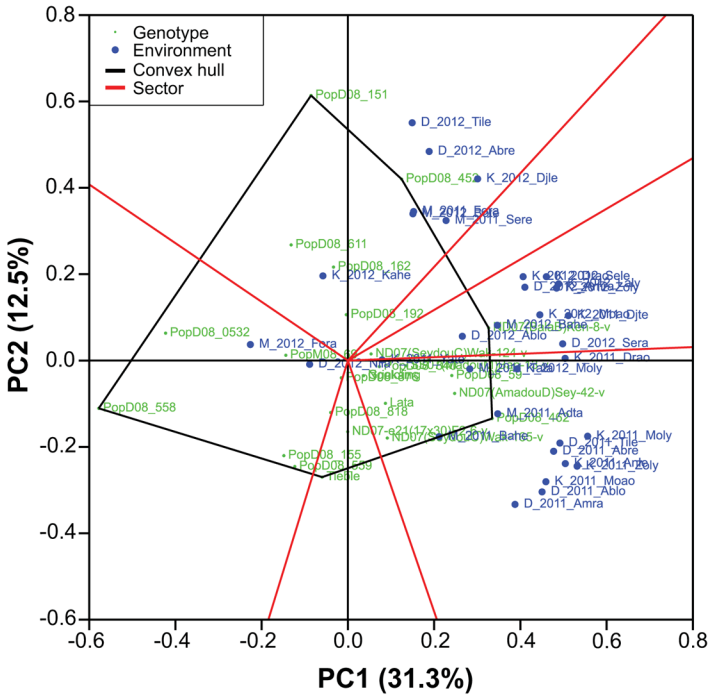

Fig. 3. Genotype-genotype-by-environment biplot of panicle yield assessed in 33 on-farm trial environments over 23 genotypes (green) in 2011 and 2012. Environment labels are coded to show the zone (D = Diolia, K = Koutiala, M = Mande), the year (2011, 2012), and the farmer ID.

Comparison of Response to Selection On-Station vs. On-Farm

The relative selection efficiency ($RSE_{St:Fa}$) for indirect (on-station) versus direct (on-farm) first-stage yield testing was less than 1.0 for both +P and –P on-station conditions, although it was somewhat higher under –P conditions due to the higher genetic correlation with on-farm values as compared with the +P on-station results (Table 4). The $RSE_{St:Fa}$ from using combined +P and –P on-station performance for selection was marginally superior to that of –P alone, with a higher heritability and a genetic correlation that was intermediate between those of –P and +P conditions.

The $RSE_{St:Fa}$ estimates for the second-stage testing were all less than 1.0 for individual trials (Table 5), with high heritabilities of on-farm yields being an important determinant. Using the combined analysis over the three on-station trials in 2012 resulted in a $RSE_{St:Fa}$ estimate that was higher than for the individual trials. The 2013 on-station yield results under +P conditions gave an $RSE_{St:Fa}$ estimate similar to those from the +P trials of 2012, whereas the –P results gave a higher $RSE_{St:Fa}$ estimate than any of the preceding +P trials.

Table 5. Estimated quantitative-genetic parameters and relative selection efficiency ($RSE_{St:Fa}$) of indirect on-station selection versus direct on-farm selection in second-stage trials conducted in 2011 and 2012, using combined analyses over years, and with supplementary phosphorus (+P) and nonfertilized (–P) on-station trials in 2013 correlated with the combined 2011–2012 on-farm performances.†

Year	Site	σ^2_G	σ^2_{GE}	h^2	r_P	r_G	$RSE_{St:Fa}$
2011	Samanko-station (+P)‡	3567 ± 1898		0.54	0.19	0.61	0.47
	On-farm	1109 ± 3640	548 ± 116	0.93			
2012	Samanko-station (+P)	5416 ± 1829		0.86	0.59	0.71	0.73
	Kolombada-station (+P)	1035 ± 1350		0.25	0.57	0.98	0.29
	Sotuba-station (+P)	3647 ± 1549		0.69	0.64	0.90	0.76
	Station combined§	1687 ± 911	2284 ± 799	0.56	0.83	0.99	0.81
	On-farm	249 ± 87	339 ± 66	0.83			
2011–2012	Station combined§	1962 ± 861	1937 ± 638	0.67	0.73	0.91	0.78
	Samanko-station (+P)	3962 ± 1545	817 ± 757	0.77	0.58	0.61	0.57
	On-farm combined¶	416 ± 132	562 ± 68	0.91			
2013	Samanko-station (+P)	2500 ± 875		0.83	0.33	0.50	0.68
	Samanko-station (–P)	240 ± 127		0.56	0.36	0.99	0.92
	Station combined§	519 ± 273	429 ± 241	0.57	0.38	0.68	0.64

† Genotypic variance (σ^2_G), genotype × trial (single years) or genotype × environment (across years) interaction variance (σ^2_{GE}), heritability (h^2), rank correlation of on-station and on-farm predictions of genotypic effects (r_P), genotypic correlation between on-station and on-farm trials (r_G).

‡ Data available from only one on-station trial in 2011.

§ Combined analysis across all on-station trials in that year.

¶ Combined analysis of the 2011–2012 on-farm data.

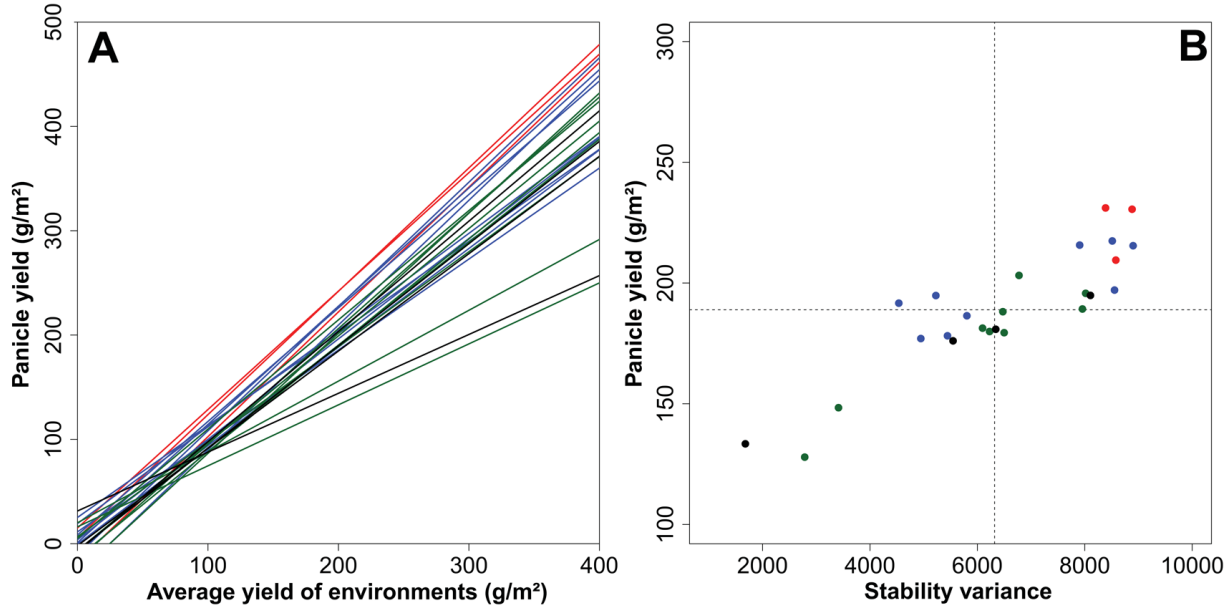

Fig. 4. Yield performance of progenies selected in first-stage only on-farm (green), only on-station (blue), or both on-farm and on-station (red), and checks (black) in 33 replicated on-farm trials 2011–2012 (A) regressed on mean yields and (B) plotted by stability variance.

Examination of the second-stage on-farm yield performances of the 21 test progenies classified by their first-stage selection history revealed that the three progenies included in the top 12 ranked progenies from both on-farm and on-station yield testing had shown consistently superior mean yield performance across zones and years relative to the mean yields of the remaining progenies selected only on the basis of on-station index values or on-farm index values (Fig. 4; Supplemental Table S1). Furthermore, the mean performance of the three progenies selected both on-station as well as on-farm were distinctly superior to the performance of the check varieties. The overall mean yield of the nine progenies selected only on the basis of first-stage on-station testing showed consistent superiority ($P < 0.05$) to the mean of the nine progenies selected only on the basis of on-farm index values.

Comparison of Alternative Trial Designs for Early-Generation On-Farm Testing

Modeling the four alternative trial designs for usefulness for early-generation on-farm yield testing indicated that the heritability estimates were fairly close among all four designs (Table 6). The farmer-as-incomplete-block and the augmented p-rep design had slightly higher heritabilities relative to the augmented design, whereas the heritability of the multiple lattice design was slightly lower.

The alternative designs differed considerably however for variance of differences of two best linear unbiased predictions (VD_{BLUP}) (Table 6). The augmented p-rep design had a VD_{BLUP} estimate that was markedly lower than all other designs.

Table 6. The mean variance of a difference of two best linear unbiased predictions (VD_{BLUP}) and heritability (h^2) estimates of alternative experimental designs for early-generation on-farm trials modeled with parameter values in Table 2 and comparable testing resources.

Design	VD_{BLUP}	h^2	Relative h^2†
			%
Augmented design	596	0.59	–
Multiple lattices	624	0.57	96
Farmer-as-incomplete-block‡	570	0.61	103
Augmented p-rep§ design	505	0.65	111

† Relative to augmented design.
‡ Resolvable incomplete block design.
§ Partially replicated.

DISCUSSION

Farmers' involvement in on-farm sorghum variety selection and single-plant selection in segregating populations has been shown to be effective for identifying "good variety fits" to specific contexts in Burkina Faso (vom Brocke et al., 2010) and Nicaragua (Trouche et al., 2011). The possibility of achieving genetic gains for a complex and environmentally sensitive trait like yield with on-farm testing of early-generation progenies, however, has been questionable due to the obscuring effects of uncontrolled within-field, site-to-site, and year-to-year heterogeneity (Atlin et al., 2001).

Feasibility of Early-Generation On-Farm Sorghum Yield Testing in Mali

Repeatabilities and Yield Levels for Individual On-Farm Trials

The majority of our on-farm yield trials, both first-stage and second-stage trials, had individual trial repeatabilities for panicle yield that were acceptable for discrimination among the progenies under farmer-managed conditions. The replicated second-stage trials had both higher repeatabilities on average and a lower rate of failed experiments than the first-stage trials, as would be expected. Acceptable on-farm repeatabilities for sorghum grain yield have been reported (Weltzien et al., 2007, 2008b; vom Brocke et al., 2010; Rattunde et al., 2013) but in trials with fewer and genetically fixed genotypes in much larger plots. The large range of repeatabilities, from 0 up to 0.8, in this and other studies (vom Brocke et al., 2010; Rattunde et al., 2013; vom Brocke et al., 2014) reflects observed differences for within-trial heterogeneities for water stagnation or drought, presence of trees or the parasitic weed *Striga hermonthica* (Delile) Benth., and uneven weeding or manure application, not to mention soil and slope gradients. Thus, many on-farm yield trials need to be conducted to obtain a sufficient pool of trials with useful levels of repeatability for effectively sampling the diversity of environments.

The range of on-farm trial mean grain yields, from approximately 70 g m^{-2} to >200 g m^{-2} (assuming a 70% threshing index), samples the range of productivity encountered by Malian sorghum farmers. These trials enabled discrimination among test genotypes in poor as well as better productivity conditions since there was little correlation between productivity and repeatability levels (Fig. 2). Our provision of fertilizer to farmers may have contributed somewhat to obtaining acceptable repeatability levels, as was also observed by vom Brocke et al. (2014), yet low yield levels due to soil type, manner of fertilizer application, and other factors, were well represented in the set of trials. The inclusion of low-yielding on-farm trials in our combined analysis was important for representing the large portion of sorghum farmers in West Africa whose yields are below 1 t/ha (vom Brocke et al., 2010).

Heritability and G × E Interactions for Panicle Yield over Multienvironment Trials

A major challenge of on-farm selection is to accurately predict and rank genotypes for yield performance over a diverse population of test environments, particularly in the presence of large G × E interaction (Atlin et al., 2001). Our first-stage trials, with their large range of productivity levels (Fig. 2), sowing dates (18 June to 5 August), soil types, rainfall, and agronomic practices (timing and manner of weeding and fertilizer application) represented the diversity of sorghum production environments in the target zone. Despite the presence of significant G × E interaction for panicle yield, significant genotypic variance and a heritability estimate (h^2 = 0.56) sufficiently large to permit discrimination among genotypes for yield performance were obtained (Table 4). The heritability estimate, based on a total of 20 farmers' trials with six to eight per genotype set, relied on the assumption that the residual variance for checks and test genotypes in the augmented design (Möhring et al., 2014) were the same.

The second-stage replicated α-lattice on-farm trials exhibited even higher heritability estimates for yield (Table 5) and had improved accuracy of BLUPs (data not shown) within years relative to the first-stage trials. The second-stage trials also had acceptable heritabilities within test regions and year (0.53 to 0.89), except Mande 2012, where the number of trials was small (data not shown).

Further, a mixed model analysis across a series of trials can assist in targeting a broader region comprising a population of diverse environments. This approach identifies superior genotypes with a higher precision than with a weighted combination of results from individual trials, and would presumably lead to a higher response to selection in the framework of early-generation on-farm testing (Smith et al., 2005). Regarding genotypes as a random sample out of a larger population in mixed modeling would further facilitate a selection decision based on BLUPs (Piepho et al., 2008). Their properties incorporate the correct ranking of genotypes, the possibility to exploit correlations between environments (Piepho et al., 2008) or different traits (Bauer and Léon, 2008), and the modeling of genetic relationships based on available marker data (Bauer et al., 2006). Further model optimizations for on-farm trial data analysis can provide options for enhancing the utility of these results. Modeling variance-covariance structures such as a factor-analytic or heterogeneous compound symmetry structure can improve the prediction accuracy (Piepho, 1998). Most statistical packages readily allow such mixed model analysis, even with complex data structures or unbalanced trial designs (Smith et al., 2005).

The G × E interactions over our 2011 and 2012 second-stage yield trials did not reveal zone or village as important determinants of interaction (Fig. 3), with considerable G × E interaction at the individual farmers' field level as was found in Rattunde et al. (2013). Important G × Y interactions, however, were indicated by the separation between the 2011 and 2012 trials (Fig. 3) and significant G × Y variance (data not shown). Rainfall was considerably lower in 2011 (Table 3), being reduced 20 to 26% for Koutiala and Dioila and 42% for the Mande zone relative to 2012. Rainfall amounts in the month of August were even more dramatically reduced, with 38 to 54% reductions across the three zones in 2011 relative to 2012. The 2011 reductions of excess August rainfall, and reduced waterlogging, may have contributed to that year's increased yields, with genetic variation for tolerance to

waterlogging possibly having contributed to the observed G × E interactions.

The breeding programs' ongoing advanced generation on-farm yield evaluations test the same entries over 2 yr (Weltzien et al., 2008b) based on farmers' desire to see new varieties over two subsequent years. This practice appears to be useful and justified based on the G × Y interactions observed in this study.

Relative Yield Response to Selection On-Farm vs. On-Station

The $RSE_{St:Fa}$ were <1.0 from both first- and second-stage progeny trials (Tables 4 and 5), indicating that on-station testing is expected to be less efficient than directly selecting for yield performance under farmers' conditions with the same intensity of selection. Low plant-available P is one of the major constraints to sorghum production in West Africa (Buerkert et al., 2001), with most farmers' soils below the critical level of 10 ppm available P content (Bray-1; Manu et al., 1991; Doumbia et al., 2003). Also, West African sorghum breeding materials have been found to differ significantly for adaptation to low-available P (Leiser et al., 2012). Our –P on-station trials may have enabled better assessment of genetic variance for adaptation to –P, and thereby contributed to the higher genetic correlations to on-farm performance and $RSE_{St:Fa}$ values observed for the –P versus the +P on-station testing (Tables 4 and 5).

Greater use of low-fertility testing conditions could therefore reduce the handicap that on-station conditions poorly reflect on-farm conditions (Ceccarelli, 1996; Atlin et al., 2001; Ceccarelli and Grando, 2007; Weber et al., 2012) and thereby contribute to increasing genetic gains for performance under farmer's conditions. The –P on-station trials, nevertheless, are at increased risk of attack by sorghum midge due to the delayed maturity caused by lower fertility, as evidenced by the 2011 and 2012 –P trial failures.

Nevertheless, on-station multienvironment trials, even under +P conditions, appear to offer benefit for successfully selecting for on-farm performance, as was suggested by the higher genetic correlation and $RSE_{St:Fa}$ estimates of combined on-station analyses relative to single on-station trials (Table 5). Further, the potential benefit of combining on-farm with on-station test results was suggested by the superior performance of progenies selected on the basis of both on-farm as well as on-station testing (Fig. 4; Supplemental Table S1). Also, there might be the possibility of testing larger numbers of progenies in on-station trials with corresponding higher selection intensity, unless more farmers become interested in conducting these types of trials, particularly with increased options for collecting and exchanging results with other farmers and researchers using new digital tools.

Alternative Trial Designs for On-Farm Testing

Although the augmented p-rep design displayed the highest heritability estimate for early-generation on-farm yield testing in the modeling exercise (Table 6), the other designs had heritabilities that were quite close. The slightly higher heritability of the augmented p-rep design is expected to be due to the high number of concurrence and the reduced residual variance, where some of the heterogeneity within trials could be accounted for.

The fact that all designs had acceptable heritability levels when modeled with the assumptions of 20 farmers, 75 plots per farmer, and 150 progenies, encourages West African sorghum breeding programs to pursue on-farm yield testing of larger numbers of progenies. As the heritabilities were similar, the choice of design should be made primarily based on practical aspects for implementation. For example, the farmer-as-incomplete-block design could be easier to implement, possibly with fewer errors in preparation relative to a p-rep design. The farmer-as-incomplete-block was found to be useful for on-farm testing large numbers of rice progenies (Atlin et al., 2002).

Both the resolvable and unresolvable incomplete block design gave the same result, yet a resolvable design might have an advantage when any further restrictions of the randomization are envisaged, for example, allocating two incomplete blocks to the same village, which would enable farmers to more easily see all genotypes under test. A concrete example was given by the analysis of such a trial network with sorghum hybrids in south-central Mali, in which heritabilities as high as $h^2 = 0.77$ could be achieved (H. Some, unpublished data, 2014).

Based on the results obtained in this study, sorghum and other crop breeders seeking to achieve yield gains in the context of diverse and primarily low-input production conditions and limited numbers of experiment stations within a target ecology can be encouraged to use early-generation on-farm yield testing, manage experiment station environments so as to increase genetic correlation with target population of environments, and integrate both on-farm and on-station yield performance information to optimize genetic gains for small-holder farmers.

Supplemental Material Available

Supplemental material is available with the online version of this article.

Acknowledgments

The contributions of Malian farmers and field staff of the Association des Organisations Paysans Professionels (AOPP), Union Local de Producteurs de Cereals (ULPC), Association d'Eveil Malienne pour la Developpement Durable (AMEDD), Cooperatives des semenciers du Mande (COOPROSEM), Institut d'Economie Rural and International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) for the conduct of the yield trials is greatly appreciated. The support of the McKnight

Foundation and the Bill and Melinda Gates Foundation for the farmer field testing is sincerely appreciated. The work was undertaken as a part of the CGIAR Research Program on Dryland Cereals.

References

Atlin, G.N., M. Cooper, and A. Bjornstad. 2001. A comparison of formal and participatory breeding approaches using selection theory. Euphytica 122:463–475. doi:10.1023/A:1017557307800

Atlin, G.N., T. Paris, and B. Courtois. 2002. Sources of variation in varietal selection trials with rainfed rice: Implications for the design of mother-baby trial networks. In: Quantitative analysis of data from participatory methods in plant breeding. CIMMYT, Mexico City, Mexico. p. 36–43.

Bänziger, M., and M. Cooper. 2001. Breeding for low input conditions and consequences for particpatory plant breeding: Examples from tropical maize and wheat. Euphytica 122:503–519. doi:10.1023/A:1017510928038

Bauer, A.M., and J. Léon. 2008. Multiple-trait breeding values for parental selection in self-pollinating crops. Theor. Appl. Genet. 116:235–242. doi:10.1007/s00122-007-0662-6

Bauer, A.M., T.C. Reetz, and J. Léon. 2006. Estimation of breeding values of inbred lines using best linear unbiased prediction (BLUP) and genetic similarities. Crop Sci. 46:2685–2691. doi:10.2135/cropsci2006.01.0019

Buerkert, A., A. Bationo, and H.-P. Piepho. 2001. Efficient phosphorus application strategies for increased crop production in sub-Saharan West Africa. Field Crops Res. 72:1–15 doi:10.1016/S0378-4290(01)00166-6

Butler, D.G., B.R. Cullis, A.R. Gilmour, and B.J. Gogel. 2009. Mixed models for S language environments. ASReml-R reference manual. Department of Primary Industries and Fisheries, Brisbane, Toowoomba, Australia.

Ceccarelli, S. 1996. Adaptation to low/high input cultivation. Euphytica 92:203–214. doi:10.1007/BF00022846

Ceccarelli, S., and S. Grando. 2007. Decentralized-participatory plant breeding: An example of demand driven research. Euphytica 155:349–360. doi:10.1007/s10681-006-9336-8

Cullis, B.R., A. Smith, and N. Coombes. 2006. On the design of early generation variety trials with correlated data. J. Agric. Biol. Enivronmental Stat. 11:381–393. doi:10.1198/108571106X154443

Dawson, J.C., K.M. Murphy, and S.S. Jones. 2008. Decentralized selection and participatory approaches in plant breeding for low-input systems. Euphytica 160:143–154 doi:10.1007/s10681-007-9533-0

Doumbia, M.D., A. Sidibé, A. Bagayoko, M.A. Diarra, A. Bationo, R.A. Kablan, R.S. Yost, L.R. Hossner, and F.M. Hons. 2003. Recommandations specifiques d'engrais: Calibration et validation du module Phosphore du modele NuMaSS. Afr. Crop Sci. J. 11(1):17–26. doi:10.4314/acsj.v11i1.27564

FAOSTAT. 2014. FAO, Rome. http://faostat3.fao.org/home/E (accessed 10 Oct. 2014; verified 3 May 2016).

Federer, W.T. 1961. Augmented designs with one-way elimination of heterogeneity. Biometrics 17:447–473. doi:10.2307/2527837

Finlay, K.W., and G.N. Wilkinson. 1963. The analysis of adaptation in a plant-breeding programme. Crop Pasture Sci. 14:742–754. doi:10.1071/AR9630742

Haussmann, B.I.G., H. Fred Rattunde, E. Weltzien-Rattunde, P.S.C. Traoré, K. vom Brocke, and H.K. Parzies. 2012. Breeding strategies for adaptation of pearl millet and sorghum to climate variability and change in West Africa. J. Agron. Crop Sci. 198:327–339.

Henzell, R.G., and D.R. Jordan. 2009. Grain sorghum breeding. In: M.J. Carena, editor, Handbook of plant breeding. Springer, New York. p. 183–197.

Leiser, W.L., H.F.W. Rattunde, H.-P. Piepho, E. Weltzien, A. Diallo, A.E. Melchinger, H.K. Parzies, and B.I.G. Haussmann. 2012. Selection strategy for sorghum targeting phosphorus-limited environments in West Africa: Analysis of multi-environment experiments. Crop Sci. 52:2517–2527. doi:10.2135/cropsci2012.02.0139

Mangione, D., S. Senni, M. Puccioni, S. Grando, and S. Ceccarelli. 2006. The cost of participatory barley breeding. Euphytica 150:289–306. doi:10.1007/s10681-006-0226-x

Manu, A., A. Bationo, and S.C. Geiger. 1991. Fertility status of selected millet producing soils of Wset Africa with emphasis on phosphorus. Soil Sci. 152:315–320 doi:10.1097/00010694-199111000-00001

Möhring, J., E.R. Williams, and H.P. Piepho. 2014. Efficiency of augmented p-rep designs in multi-environmental trials. Theor. Appl. Genet. 127:1049–1060. doi:10.1007/s00122-014-2278-y

Piepho, H.P. 1998. Empirical best linear unbiased prediction in cultivar trials using factor-analytic variance-covariance structres. Theor. Appl. Genet. 97:195–201. doi:10.1007/s001220050885

Piepho, H.P., A. Buchse, and B. Truberg. 2006. On the use of multiple lattice designs and in plant breeding trials. Plant Breed. 125:523–528. doi:10.1111/j.1439-0523.2006.01267.x

Piepho, H.P., A. Buechse, and K. Emrich. 2003. A hitchhiker's guide to the mixed model analysis of randomized experiments. J. Agron. Crop Sci. 189:310–322. doi:10.1046/j.1439-037X.2003.00049.x

Piepho, H.-P., and J. Möhring. 2007. Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177:1881–1888 doi:10.1534/genetics.107.074229.

Piepho, H.P., J. Möhring, A.E. Melchinger, and A. Buchse. 2008. BLUP for phenotypic selection in plant breeding and variety testing. Euphytica 161:209–228.

Rattunde, H.F.W., K. vom Brocke, E. Weltzien, and B.I.G. Haussmann. 2009. Developing open-pollinated varieties using recurrent selection methods In: S. Ceccarelli et al., editors, Plant breeding and farmer participation. FAO, Rome. p. 259–273.

Rattunde, H.F.W., E. Weltzien, P.J. Bramel-Cox, K. Kofoid, C.T. Hash, W. Schipprack, J.W. Stenhouse, and T. Presterl. 1997. Population improvement of pearl millet and sorghum: Current research, impact and issues for implementation. In: Proc. of the Int. Conf. on Genetic Improvement of Sorghum and Pearl Millet. Lubbock, TX. 22–27 Sept. 1996. ICRISAT, Patancheru, India. p. 188–212.

Rattunde, H.F.W., E. Weltzien, B. Diallo, A.G. Diallo, M. Sidibe, A.O. Touré, A. Rathore, R.R. Das, W.L. Leiser, and A. Touré. 2013. Yield of photoperiod-sensitive sorghum hybrids based on guinea-race germplasm under farmers' field conditions in Mali.

Crop Sci. 53:2454–2461. doi:10.2135/cropsci2013.03.0182
R Development Core Team. 2011. R: A language and environment for statistical computing. R Foundation, Vienna, Austria. http://www.r-project.org/ (verified 3 May 2016).
Shukla, G.K. 1972. Some statistical aspects of partitioning genotype environmental components of variability. Heredity 29:237–245. doi:10.1038/hdy.1972.87
Smith, A.B., B.R. Cullis, and R. Thompson. 2005. The analysis of crop cultivar breeding and evaluation trials: An overview of current mixed model approaches. J. Agric. Sci. 143:449–462. doi:10.1017/S0021859605005587
Trouche, G., S. Aguirre Acuna, B. Castro Briones, N. Gutierrez Palacios, and J. Lancon. 2011. Comparing decentralized participatory breeding with on-station conventional sorghum breeding in Nicaragua: I. Agronomic performance. Field Crops Res. 121:19–28. doi:10.1016/j.fcr.2010.11.016
vom Brocke, K., G. Trouche, E. Weltzien, C.P. Barro-Kondombo, E. Gozé, and J. Chantereau. 2010. Participatory variety development for sorghum in Burkina Faso: Farmers' selection and farmers' criteria. Field Crops Res. 119:183–194 doi:10.1016/j.fcr.2010.07.005
vom Brocke, K., G. Trouche, E. Weltzien, C.P. Kondombo-Barro, A. Sidibé, R. Zougmoré, and E. Gozé. 2014. Helping farmers adapt to climate and cropping system change through increased access to sorghum genetic resources adapted to prevalent sorghum cropping systems in Burkina Faso. Exp. Agric. 50:284–305 doi:10.1017/S0014479713000616
VSN International. 2014. Data analytics for bioscience. VSN International, Hemel Hempstead, UK. http://www.vsni.co.uk (accessed 20 Sept. 2014; verified 3 May 2016).
Weber, V.S., A.E. Melchinger, C. Magorokosho, D. Makumbi, M. Bänziger, and G.N. Atlin. 2012. Efficiency of managed-stress screening of elite maize hybrids under drought and low nitrogen for yield under rainfed conditions in Southern Africa. Crop Sci. 52:1011–1020. doi:10.2135/cropsci2011.09.0486
Weltzien, E., K. vom Brocke, A. Touré, F. Rattunde, and J. Chantereau. 2008a. Revue et tendances pour la recherche en sélection paticipative en Afrique de l'Ouest. Cah. Agric. 17:165–171.
Weltzien, E., M. Kanouté, A. Toure, F. Rattunde, B. Diallo, I. Sissoko, A. Sangaré, and S. Siart. 2008b. Sélection participative des variétés de sorgho à l'aide d'essais multilocaux dans deux zones cibles. Cah. Agric. 17:134–139.
Weltzien, E., A. Christinck, A. Touré, F. Rattunde, M. Diarra, A. Sangare, and M. Coulibaly. 2007. Enhancing farmers' access to sorghum varieties through scaling-up participatory plant breeding in Mali, West Africa. In: C. Almekinders and J. Hardon, editors, Bringing farmers back into breeding, Experiences with participatory plant breding and challenges for institutionalisation. Agromisa Foundation, Wageningen, the Netherlands. p. 58–69.
Williams, E., H.P. Piepho, and D. Whitaker. 2011. Augmented p-rep designs. Biom. J. 53:19–27. doi:10.1002/bimj.201000102