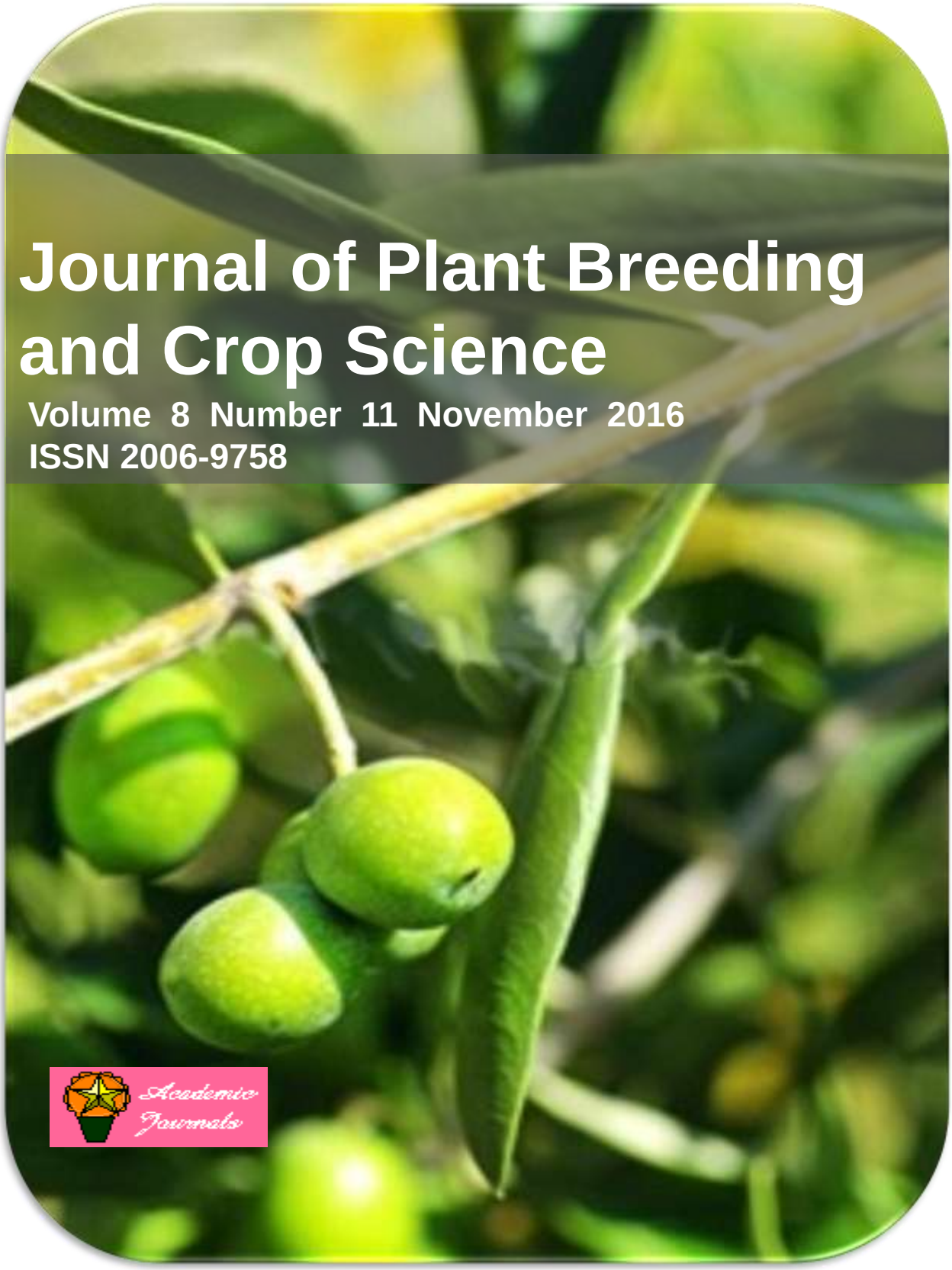


Journal of Plant Breeding and Crop Science

Volume 8 Number 11 November 2016
ISSN 2006-9758

Academic Journals

ABOUT JPBCS

The **Journal of Plant Breeding and Crop Science (JPBCS)** is published monthly (one volume per year) by Academic Journals.

The Journal of Plant Breeding and Crop Science (JPBCS) (ISSN: 2006-9758) is an open access journal that provides rapid publication (monthly) of articles in all areas of the subject such as Sustainable use of plant protection products, Agronomic and molecular evaluation of recombinant inbred lines (RILs) of lentil, Pollen behaviour and fertilization impairment in plants, Development of a fast and reliable ozone screening method in rice etc.

The Journal welcomes the submission of manuscripts that meet the general criteria of significance and scientific excellence. Papers will be published shortly after acceptance. All articles published in JPBCS are peer-reviewed.

Contact Us

Editorial Office: jpbcs@academicjournals.org

Help Desk: helpdesk@academicjournals.org

Website: http://www.academicjournals.org/journal/JPBCS

Submit manuscript online http://ms.academicjournals.me/

Journal of Plant Breeding and Crop Science

Table of Contents: Volume 8 Number 11 November, 2016

ARTICLES

academicJournals

Vol. 8(11), pp. 228-239, November 2016
DOI: 10.5897/JPBCS2016.0602
Article Number: 3E70E4E61314
ISSN 2006-9758

http://www.academicjournals.org/JPBCS

Journal of Plant Breeding and Crop Science

Full Length Research Paper

Radio-sensitivity of cowpea to ultra-violet radiation by pollen treatment

Festus Olakunle Olasupo[1,2]*, Christopher Olumuyiwa Ilori[1] and Anna Abimbola Muyiwa[2]

[1]Department of Crop Protection and Environmental Biology, University of Ibadan, Ibadan, Nigeria.
[2]Crop Improvement Division, Cocoa Research Institute of Nigeria, Idi-Ayunre, Ibadan, Nigeria.

Received 4 August, 2016; Accepted 20 September, 2016

The technique of pollen irradiation for mutation induction technique has been successfully demonstrated in many crop species. Therefore, the mutagenic effects of UV irradiated pollen on cowpea accessions were investigated. Pollen grains of eight cowpea accessions were irradiated with 30,000µWs/cm^2 UV for 60, 120, 180, 240, 300 and 360 minutes. Emasculated flowers of each accession were self-pollinated with irradiated pollen to evaluate the effects of pollen mutagenesis on seed setting in the M_1 generation. Harvested seeds from M_1 plants were advanced to M_2 to evaluate the effects of UV radiation on seed germination, plant survival and for mutant selection in the M_3 generation. Data were analyzed using descriptive statistics. Pollen irradiation with UV for a short period (60 min) increased seed setting in all the cowpea accessions at M_1 except IB-Y-1 where it reduced seed setting by 28.6%. Observed LD_{50} of UV rays among the cowpea lines ranged from 142.6 to 210.1 minutes. No significant difference was observed in seed germination for all treatments at M_2 except irradiation for 120 minutes in IT90K-284-2. A trend similar to seed germination was observed in seedling survival at M_2. The three-primary leaf and four-primary leaf mutants selected at M_3 generation reverted back to two-primary leaf seedlings at M_4. Low mutation frequencies recorded in this study shows that cowpea is considered less amenable to the application of UV irradiation as a practical breeding method.

Key words: Cowpea mutagenesis, pollen irradiation, ultra-violet induced mutant, seed germination, seed setting.

INTRODUCTION

The technique of induced mutation in plant using pollen (male gametophytes) as the starting material was initiated with maize pollen irradiation using X-ray (Stadler, 1939). This was followed by investigations on the procedures and methods for generating and studying induce mutations with particular emphasis on pollen mutagenesis in maize (Neuffer, 1957; Amano and Smith, 1965; Mottinger, 1970) and in barley (Devreux et al., 1972). The use of gamma radiation of pollen for induced mutation in plant breeding was first report in *Nicotiana* (Pandey, 1975, 1978). Subsequently, the possible application of this procedure for scientific studies and crop improvement has been explored by several authors. Pollen mutagenesis by gamma rays have been

*Corresponding author. E-mail: festusolasupo@gmail.com.

demonstrated in *Nicotiana* (Grant et al., 1980; Pandey, 1980a; 1980b), in maize (Pandey, 1983; Sanford et al., 1984a, b), in rice (Chin and Gordon, 1989a, b), in *Arabidopsis* (Yang et al., 2004; Naito et al., 2005), in *Triticum* (Bie et al., 2007), in citrus (Yahata et al., 2010), in winter squash (Kurtar and Balkaya, 2010), in Persian walnut (Grouh et al., 2011), in melon (Godbole and Murthy, 2012) and in cotton (Yue and Zou, 2012). Many other authors have attempted the use of irradiated pollen for egg transformation (Chin and Gorgon, 1989a).

Irradiated pollen can mediate the transfer of limited genetic materials, instead of complete genome from the pollen donor (male parent) to their progenies. This hypothesis was first suggested by Pandey (1975). Thus, in a maternal background, a few paternal characteristics from the pollen parent were apparently incorporated and expressed in the M_1 plants and the eggs were said to be "transformed". He defined "egg transformation" as the transfer of limited intact genes rather than the total pollen genome to the egg cell. He suggested that transformation could be the result of diploid parthenogenesis of the egg induced by irradiated pollen, followed by the incorporation of paternal chromatin during embryogenesis. The biological basis to these assumptions had been explained by several authors. First, the pollen germination and pollen tube growth are highly tolerant to irradiation (Vassileva-Dryanovska, 1966; Gillissen, 1978). In addition to this, despite extensive nuclear aberrations caused by irradiation, fusion of the fragmented paternal genetic materials with the egg nuclei was still possible (Brewbaker and Emery, 1962). Therefore, the use of irradiated pollen for limited gene transfer was considered as a self-contained "DNA injection" system in which the irradiated pollen is both a source of donor DNA fragment as well as a vector for delivering the genetic fragments to the embryo sac (Chin and Gordon, 1989a). The authors proposed this technique, as a natural and rapid means of transferring a few or single genes into plants without resorting to the use of recombinant DNA technology. Moreover, it has been suggested that cross-pollination with irradiated pollen may nevertheless be useful in practical plant breeding by causing a shift in the segregation ratio towards the maternal phenotype in the second (M_2) generation (Chin and Gordon, 1989b).

Production of haploids and doubled haploids are important aspect of plant breeding by which homozygous lines are obtained for hybrids generation. The technique of pollen irradiation by gamma rays, UV or X-rays is the most widely used method to induce in situ parthenogenetic haploid plants (Kosmrlj et al., 2013). This technology has been demonstrated and proven to be effective in many crop species elucidated by Kurtar and Balkaya (2010). In addition to these, haploid production by irradiated pollen has also been reported in winter squash (Kurtar and Balkaya, 2010); in citrus (Yahata et al., 2010); in melon (Gonzalo et al., 2011; Godbole and Murthy, 2012); in Persian walnut (Grouh et al., 2011) and in pumpkin (Kosmrlj et al., 2013).Physical mutagens (ionizing and non-ionizing radiation) may be considered as the most suitable option for pollen mutagenesis in some plants for certain reason. For example, in cowpea, chemical mutagens are unsuitable for pollen treatment simply because of the hydrophobic nature of their pollen. Consequently, the pollen grains rupture and lose viability as soon as it get in contact with aqueous medium. Among the radiation mutagens used for pollen treatment, application of UV radiation has been reported by few authors. Pollen mutagenesis by UV irradiation was reported in maize (Gavazzi and Sanguineti, 1983). In legume (faba bean), Haliem et al. (2013) investigated the mutagenic potential of UV on meiotic-pollen mother cells, pollen grains and seed yield.

UV ray is an electromagnetic radiation that does not carry enough energy per quantum to ionize atoms or molecules. It has longer wavelength (100 to 400 nm) with low penetration power into plant tissues when compared with ionizing radiations. According to Mba et al. (2012), UV radiation is classified based on their wavelengths into three forms, ultraviolet A (UVA) 315-400 nm, ultraviolet B (UVB) 280 to 315 nm and ultraviolet C (UVC) 100 to 280 nm. UVC has been implied to be the most energetic and biologically damaging among the three. The mutagenic effect of UV is due to its ability to react with DNA and other biological molecules such as bases in DNA molecules and other aromatic amino acids of proteins. UVB and UVC produce pyrimidine dimers on reacting with DNA, while UVA produces very few of these. Therefore, UVB and UVC are mostly used for mutagenic treatments. The pyrimidine dimers produced form lesions that interfere with transcription and DNA replication, lead to mutations, chromosomal rearrangements and lethality. Moreover, it has been established that exposure of crop plant cells under natural condition of growth and development to UVB resulted in excessive production of free radicals, reactive oxygen species (Agrawal et al., 2009) which can induce structural changes in DNA, such as chromosomal rearrangement, strand breaks, base deletions, pyrimidine dimers, cross-links and base modifications, mutations and their genotoxic effects (Gill and Tuteja, 2010; Haliem et al., 2013).When compared to seed mutagenesis, pollen mutagenesis is of greater advantage. First, when seed is used as the starting material for mutagenesis it always leads to the production of chimeric tissues. This problem can be overcome by starting with pollen because pollen mutagenesis involves mutagenic treatment, usually in the form of irradiation, to the pollen prior to hand pollination, while the female tissue remains free of somatic damage. The M_1 plant arising from pollination with mutated pollen is non-chimeric and will be hemizygous for any uniquely induced mutation (Yang et al., 2004). The dominant mutations in this case will be expressed in the M_1 while recessive mutations will be expressed in the M_2. Because of the absence of chimera, fewer M_2 seeds are needed per

Table 1. Quantity of cowpea seeds advanced to M_3 generation after mutagenesis

Accession	Number of seeds planted in each treatment						
	0(Control)	60 min	120 min	180 min	240 min	300 min	360 min
IB	500	800	800	580	470	400	300
IB-Y-1	500	500	460	350	0	0	0
IB-BPC	500	1000	1000	570	500	500	500
IB-CR	500	600	450	300	300	250	0
IT86D-719	500	555	435	345	300	250	200
IT86D-1010	500	550	550	300	250	0	0
IT89KD-374-57	500	750	700	500	400	250	0
IT90K-284-2	500	700	550	480	450	300	250

plant for screening than for an equivalent seed mutagenesis experiment. Further limitations of seed treatment are the possible occurrence of separated female and male germ cell primordial as well as somatic selection against mutant cells during plant development (Gavazzi and Sanguineti, 1983). Mutagenesis of male gametophyte also has another unique advantage, since mutations are directly passed onto the next generation in a hemizygous state and large numbers of pollen grains (haploid nuclei) can be mutated (Yang et al., 2004). However, information on pollen irradiation for mutation induction in cowpea is scarce and there is paucity of report on the application of UV radiation in cowpea pollen mutagenesis. Moreover, there is the need to determine optimal UV radiation dosage for pollen mutagenesis in cowpea for the purpose of widening its genetic base. Therefore, the objective of this study is to evaluate the effects of UV irradiated pollen used for self-pollination on the M_1, M_2 and M_3 generations of eight cowpea accessions.

MATERIALS AND METHODS

Mutagenesis by pollen irradiation with UV rays

Eight cowpea accessions used in this study were the cultivar, Ife Brown (IB) and its mutant derivatives (IB-Y-1, IB-CR and IB-BPC) and four elite cultivars (IT86D-719, IT86D-1010, IT89KD-374-57 and IT90K-284-2). The cowpea accessions IB, IB-Y-1, IB-CR and IB-BPC, were collected from the Genetics unit of the Department of Crop Protection and Environmental Biology (CPEB), University of Ibadan, Nigeria, while the four elite cultivars were obtained from the Genetic. Resources Centre of the International Institute of Tropical Agriculture (IITA), Ibadan. The cowpea accessions were raised at the rooftop garden of CPEB Department, University of Ibadan. Matured (opened) flowers of each of the cowpea accessions were separately harvested into labeled air-filled transparent nylon bags in the morning (07:00 to 08:00 h). The flowers were stored in the refrigerator at 10°C until further use. Pollen from these flowers was carefully collected from dehisced anthers with the aid of sterile forceps into cell-wells separately. The cell-wells were sealed with paper tape immediately to avoid pollen contaminations. UV irradiation of cowpea pollen was carried out at the Genetics laboratory of CPEB Department. The pollen from each of the cowpea accessions (in cell-wells) were exposed to 30,000µWs/cm^2 ultraviolet (UV) radiation for 60, 120, 180, 240, 300 and 360 min. Pollen used for control treatment in all the accessions were harvested into cell-wells but not irradiated. For each radiation treatment, 20 freshly emasculated pre-anthesis flower buds of the original parents were hand-pollinated (selfed) using irradiated pollen grains in the evening time (18:00 to 19:00 h). Each flower pollinated with irradiated pollen was tagged with appropriate label. Data were collected on the number of seed set in each treatment at M_1 generation. At maturity, all dry pods in each treatment were harvested into labeled envelopes and their seeds were prepared for advancement to M_2 generation.

Determination of Ultra-violet lethal dosage 50% (LD_{50})

The LD_{50} was determined by adopting the method described in Olasupo et al. (2016). Percentage seed setting at M_1 generation following hand pollination with UV irradiated pollen were calculated for each treatment. The difference in percentage seed setting between each treatment and control was calculated and expressed as percentages of control. A graph of the absorbed dose was plotted against the percentage difference (Dosage Effect Curve) for each accession to show the damage due to mutagenic treatment by UV on cowpea pollen. By inserting the 'line of best fit' and reading off the dose corresponding to 50% reduction, the LD_{50} was obtained or more precisely, calculated using the linear equation: $y = mx + c$

Evaluation of M_2 generation plants following UV mutagenesis

The M_2 plants were raised at the rooftop garden of CPEB Department, University of Ibadan. All the pods produced by M_1 plants from each treatment were harvested together in an envelope, dried and the seeds were prepared for advancement to M_2 generation. The quantity of seeds used to raise the M_2 plants was determined by seed setting of M_1 plants in each treatment (Table 1). Further screening for mutant phenotypes in the M_3 and M_4 generations were carried out at the Teaching and Research Farm of the University of Ibadan.

Data collection

In the M_1 generation, data were taken on percentage germination and number of surviving plants from each treatment. Data were also collected from 10 M_2 plants selected randomly from each treatment on the following parameters: primary leaf area, seedling height at 3 weeks, terminal leaflet area and plant height at 6 weeks. Screening

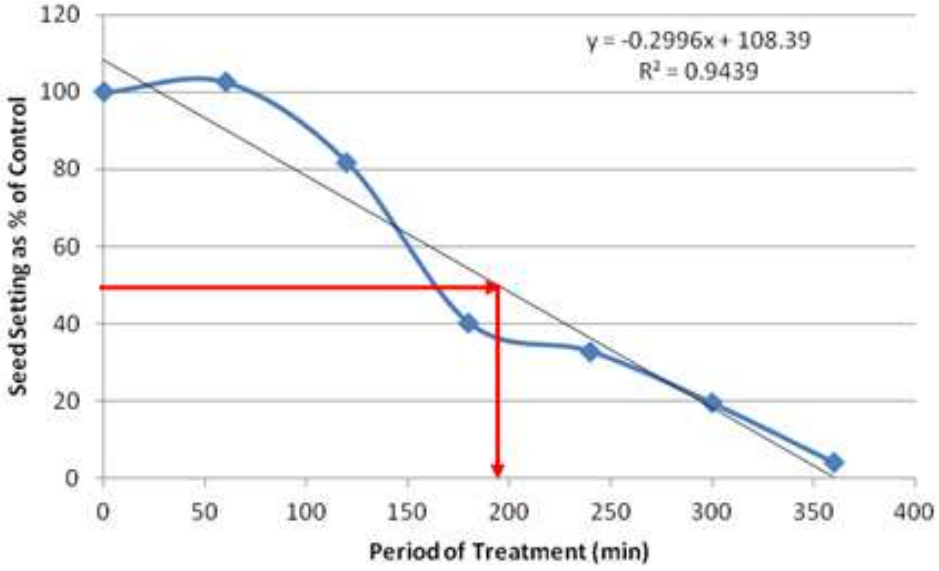


Figure 1. Percentage of reduction in seed setting from IB cowpea pollen exposed to UV irradiation.

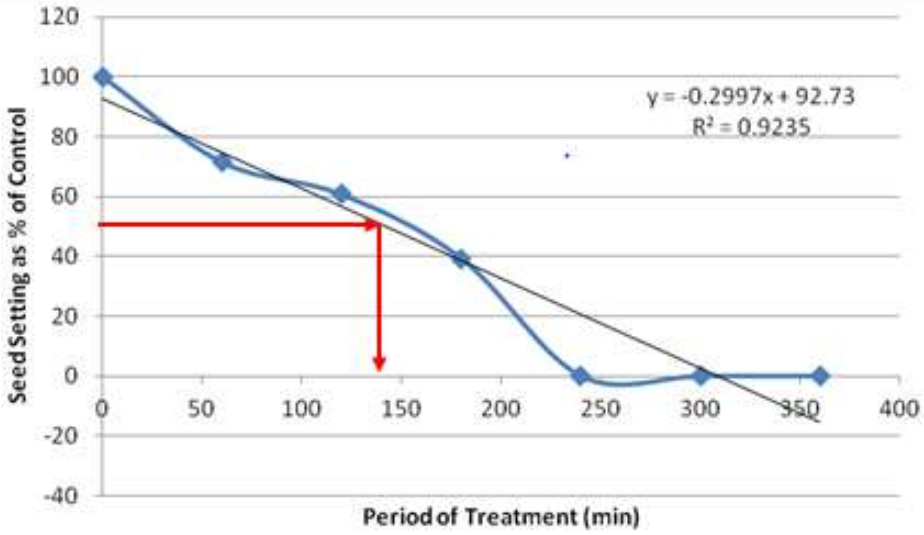


Figure 2. Percentage of reduction in seed setting from IB-Y-1 cowpea pollen exposed to UV irradiation.

for mutant was carried out by scoring the plants for any change in the phenotype when compared with the parent plants (control treatments). Descriptive statistics was used for data analyses with the aid of Microsoft Excel 2010.

RESULTS

Effect of UV irradiated pollen on the M_1 generation

Following cowpea pollen treatment with 30,000 µWs/cm^2 UV radiation for 60, 120, 180, 240, 300 and 360 min, mutagenic effect of the treatments as observed in the seed setting at M_1 generation, are presented in Figures 1 to 8. As revealed by the radiation dosage effect curves, treatment of fresh pollen grains with 30,000 µWs/cm^2 UV rays for up 60 min before pollination increased seed setting in all the cowpea accessions used in this study except IB-Y-1 where it was reduced by 28.6% (Table 2). Percentage increase in seed setting was highest (9.2%) in IB-BPC at the UV treatment for 60 min, followed by IT86D-1010 (8.3%). Further pollen irradiation with 30,000 µWs/cm^2 UV above 60 min before pollination reduced

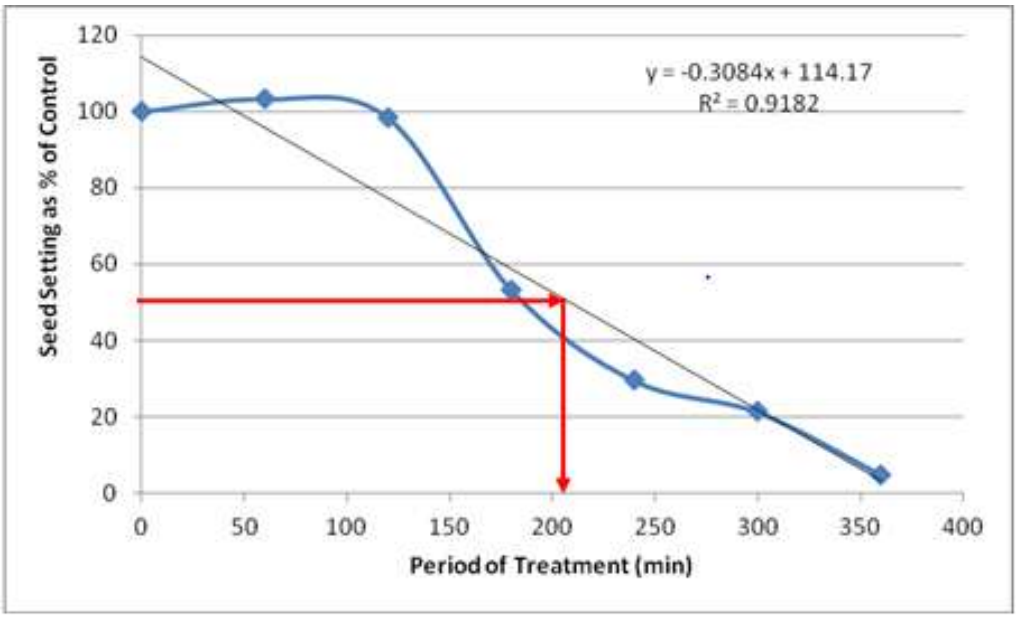


Figure 3. Percentage of reduction in seed setting from IB-CR cowpea pollen exposed to UV irradiation.

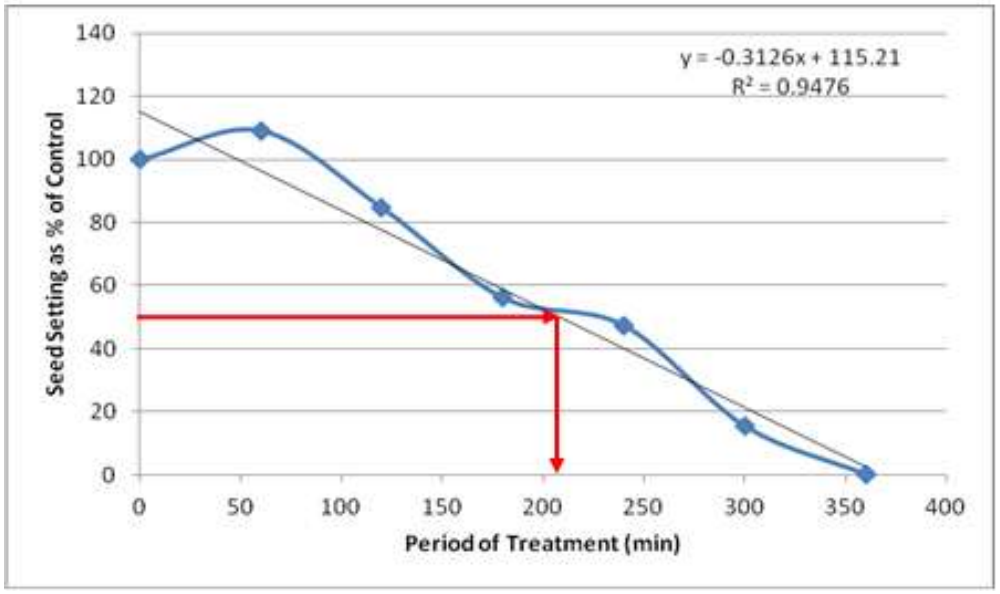


Figure 4. Percentage of reduction in seed setting from IB-BPC cowpea pollen exposed to UV irradiation.

seed setting of all cowpea accession used in this study with a highest effect (R^2 x100 = 94.8%) recorded in IB-BPC (Figure. 4). There was variation in the observed UV radiation LD_{50} of pollen among the eight cowpea lines evaluated. The lowest LD_{50} (142.6 min) and the highest LD_{50} (210.1 min) were observed in IB-Y-1 and IT90K-284-2 respectively.

Effects of UV irradiated pollen on cowpea in the M_2 generation

Following hand pollination of cowpea with UV irradiated pollen, the results of seed germination and survival of M_2 plants are presented in Figures 9 and 10, respectively. Seed germination percentage was based on the number

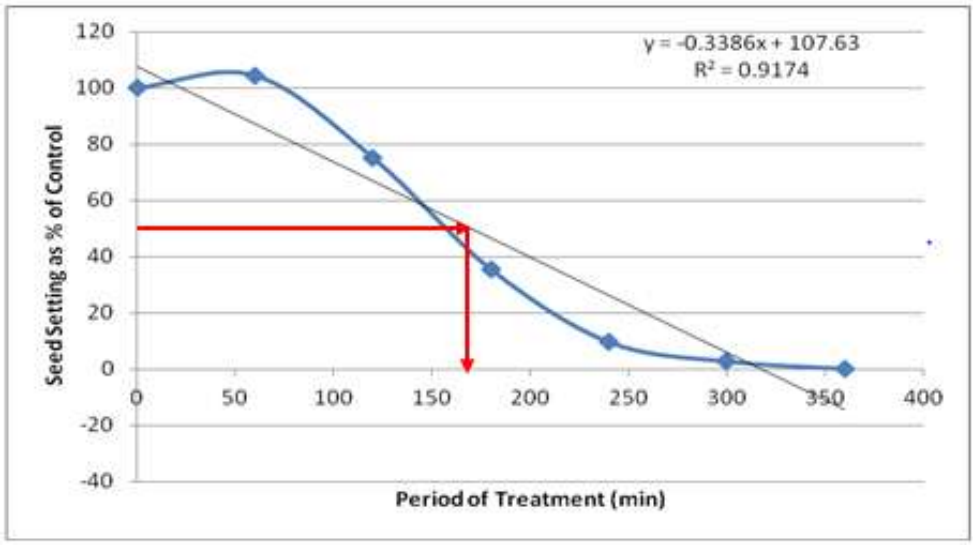


Figure 5. Percentage of reduction in seed setting from IT86D-719 cowpea pollen exposed to UV irradiation.

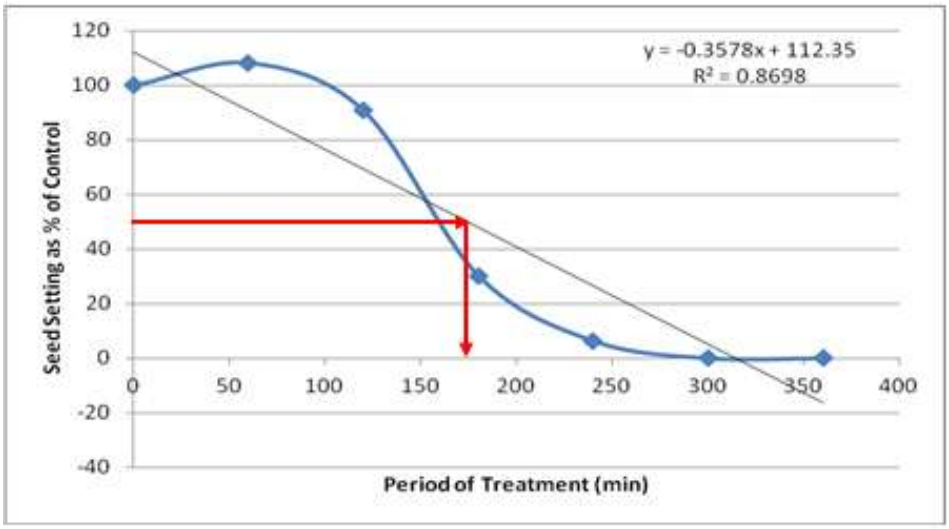


Figure 6. Percentage of reduction in seed setting from IT86D-1010 cowpea pollen exposed to UV irradiation.

of seeds which germinated out of the total number of seeds sown (in percentage) for each treatment. Percentage seedling survival was based on the proportion of the seeds planted which became established plants at each treatment level. No significant difference was observed in the percentage seed germination in all treatments across the eight cowpea accessions except radiation treatment 120 min in IT90K-284-2. Similarly, pollination with UV irradiated pollen did not reveal significant difference in the percentage plant survival cowpea plants in all the treatments except in IB-BPC, IT86D-719 and IT86D-1010. Lower percentage seedling survivals were observed at radiation level 240 min in IT86D-1010 and treatment 300 min in IB-BPC and IT86D-719 when compared with control treatments.

Spectra and frequencies of mutations in the M_3 generation

Based on the observed phenotypic changes, only three mutant spectra were selected across the treatments in

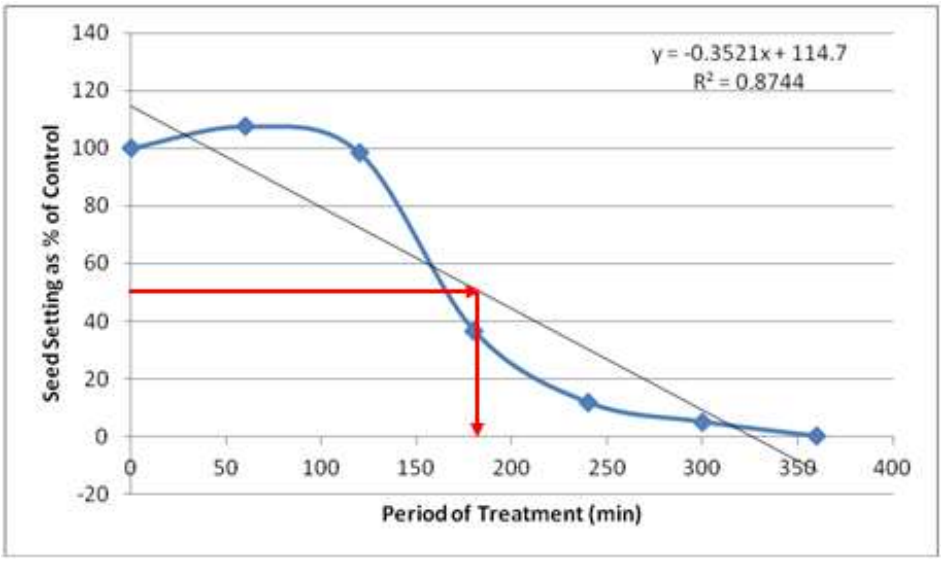


Figure 7. Percentage of reduction in seed setting from IT89KD-374-57 cowpea pollen exposed to UV irradiation.

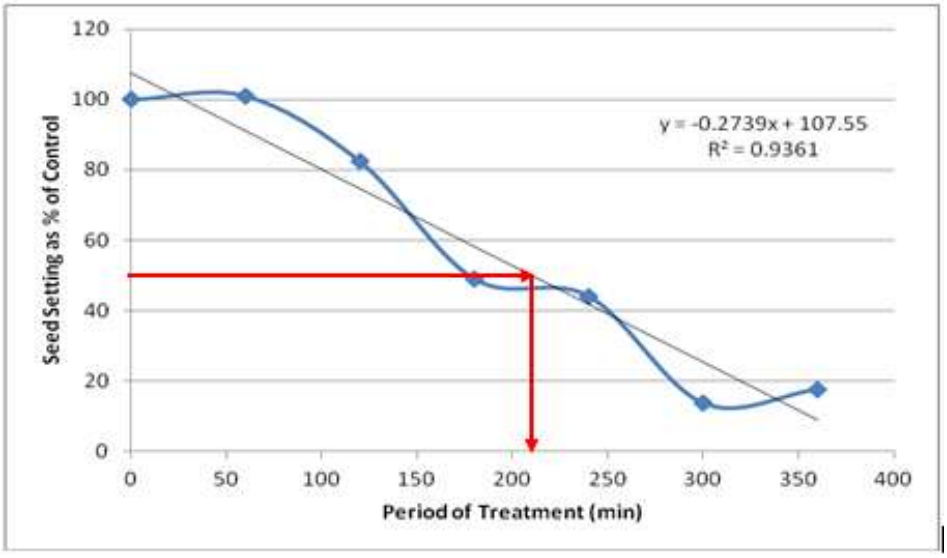


Figure 8. Percentage of reduction in seed setting from IT90K-284-2 cowpea pollen exposed to UV irradiation.

the M_3 plant generation (Table 3). Yellow albino mutant seedlings which died after a week of germination were observed in IB-Y-1 at 60, 120 and 180min treatment levels. A four-primary leaf mutant (Figure 12) was observed in IT90K-284-2 at 360 min, while three-primary leaf mutants (Figure 11) were observed in IT86D-719, IT86D-1010 and IT90K284-2. The three-primary leaf and four-primary leaf mutants produced the normal trifoliate secondary leaves. However, these mutants were not stable, but reverted back to normal (two-primary leaf plants). The three-primary leaf mutant reverted to normal plant when advanced from M_3 to M_4 generation, while the four-primary leaf mutant reverted back to three-primary leaf in the M_4 and finally to normal (two-primary leaf) plant in the M_5 generation. The mutation frequencies in the M_3 generation were very low with no mutant phenotype observed in most treatments. Highest mutation frequency (0.02%) was observed in IT90K-284-2 at 60 min period of UV radiation treatment. This result suggests that UV irradiation of pollen has low efficiency for mutation

Table 2. Effect of UV irradiated pollen on cowpea seed setting in the M_1 generation

Cowpea accession	Percentage increase in seed setting at different periods of UV irradiation						LD_{50} (min)
	60 min	120 min	180 min	240 min	300 min	360 min	
IB	2.7	-18.1	-59.7	-67.1	-80.5	-96.0	194.9
IB-Y-1	-28.6	-39.3	-60.7	-100.0	-100.0	-100.0	142.6
IB-CR	3.3	-1.6	-46.7	-70.5	-78.7	-95.1	208.1
IB-BPC	9.2	-15.3	-43.9	-52.7	-84.7	-100.0	208.6
IT86D-719	4.3	-25.0	-64.7	-90.5	-97.4	-100.0	170.2
IT86D-1010	8.3	-9.2	-69.8	-93.7	-100.0	-100.0	174.3
IT89KD-374-57	7.5	-1.7	-63.3	-88.3	-95.0	-100.0	183.8
IT90K-284-2	1.0	-17.6	-51.0	-55.9	-86.3	-82.4	210.1

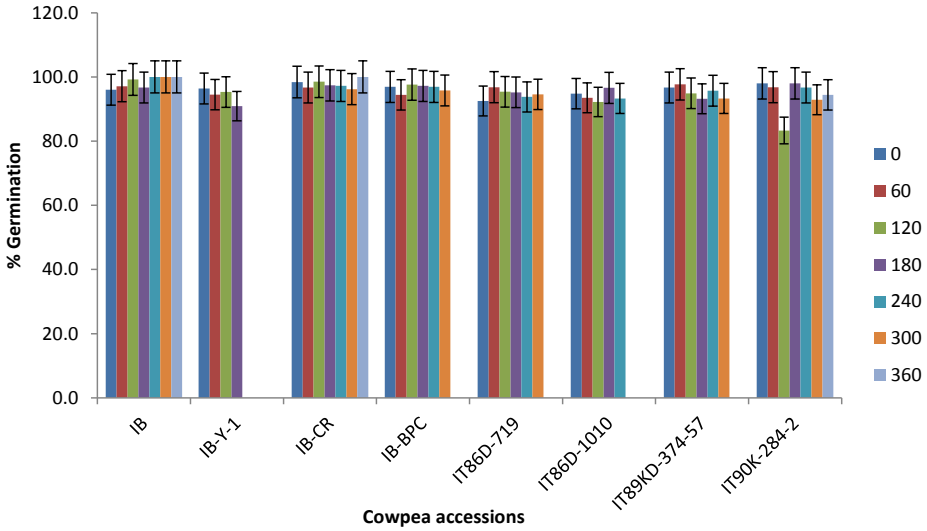


Figure 9. Effects of UV Rays on the Germination of Cowpea Seeds at M_2 Generation

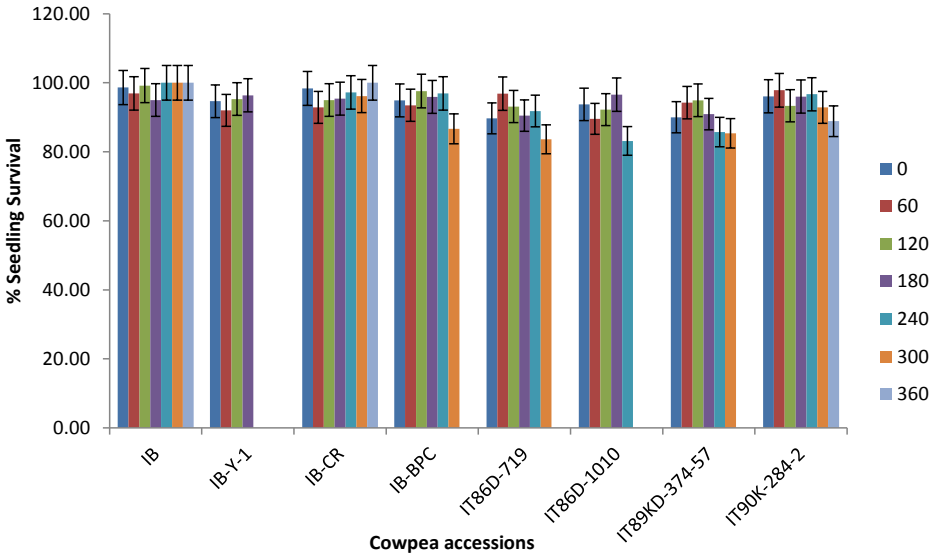


Figure 10. Effects of UV Rays on the Seedling Survival of Cowpea at M_2 Generation.

Table 3a. Mutation spectra and frequencies in the M_3 generation following UV irradiation of cowpea pollen at different treatment periods

Cowpea accession	Period of UV treatment (min)	Mutation Spectra				Mutation Frequency (%)
		Total Number of Plants	Yellow Seedling	Four Primary Leaves	Three Primary Leaves	
IB	0	500	0	0	0	0
	60	800	0	0	0	0
	120	800	0	0	0	0
	180	580	0	0	0	0
	240	470	0	0	0	0
	300	400	0	0	0	0
	360	300	0	0	0	0
IB-Y-1	0	500	0	0	0	0
	60	500	2	0	0	0.004
	120	460	2	0	0	0.0043
	180	350	2	0	0	0.0057
	240	NS	0	0	0	0
	300	NS	0	0	0	0
	360	NS	0	0	0	0
IB-CR	0	500	0	0	0	0
	60	1000	0	0	0	0
	120	1000	0	0	0	0
	180	570	0	0	0	0
	240	500	0	0	0	0
	300	500	0	0	0	0
	360	500	0	0	0	0
IB-BPC	0	500	0	0	0	0
	60	600	0	0	0	0
	120	450	0	0	0	0
	180	300	0	0	0	0
	240	300	0	0	0	0
	300	250	0	0	0	0
	360	NS	0	0	0	0

NS = No survived plant .

induction in cowpea.

DISCUSSION

Treatment of fresh pollen grains with 30,000 µWs/cm^2 UV rays for up 60 min before pollination appeared to enhance seed setting in cowpea, while exposure of pollen to the radiation for longer period (>60 min) has inhibitory effect on seed setting. Similar observation was made by Haliem et al. (2013) that UVB irradiation of pollen for 60 min had a positive effect on seeds yielded following radiation treatment from their molecular analysis. The result of radiation dosage effect curves shows that seed setting in cowpea following pollination with UV irradiated pollen appeared to be dose dependent. Plant cells and nucleic acids are damaged when exposed to high level of UV radiation which may cause the pollen grains to lose viability. Britt (1995), Ravanat et al. (2001) and Lagoda (2012) reported that UV radiation has deleterious effects on cellular DNA which may be either mutagenic or toxic and the induced damage can lead to cell death due to photochemical damage. The variability observed in LD_{50} could be as a result of genotypic variation among the cowpea accessions.

However, the results obtained from percentage seed germination and seedling survival at M_2 generation showed that pollination with UV irradiated pollen at various treatment levels did not produce significant effects in cowpea. This suggests there appeared to be reversions of most UV induced mutagenic changes in the

Table 3b. Mutation spectra and frequencies in the M_3 generation following UV irradiation of cowpea pollen at different treatment periods

Cowpea accession	Period of UV treatment (min)	Mutation Spectra				Mutation frequency (%)
		Total number of plants	Yellow seedling	Four primary leaves	Three primary leaves	
IT86D-719	0	500	0	0	0	0
	60	550	0	0	0	0
	120	435	0	0	0	0
	180	345	0	0	1	0.0029
	240	300	0	0	0	0
	300	250	0	0	0	0
	360	200	0	0	0	0
IT86D-1010	0	500	0	0	0	0
	60	550	0	0	1	0.0018
	120	550	0	0	0	0
	180	300	0	0	0	0
	240	250	0	0	0	0
	300	NS	0	0	0	0
	360	NS	0	0	0	0
IT89KD-374-57	0	500	0	0	0	0
	60	750	0	0	0	0
	120	700	0	0	0	0
	180	500	0	0	0	0
	240	400	0	0	0	0
	300	250	0	0	0	0
	360	NS	0	0	0	0
IT90K-284-2	0	500	0	0	0	0
	60	700	0	0	2	0.0029
	120	550	0	0	5	0.009
	180	480	0	0	6	0.0125
	240	450	0	0	2	0.0044
	300	300	0	0	1	0.0033
	360	250	0	1	4	0.02

NS = No survived plant.

plant genetic materials from M_1 to M_2 generation. The null effect of UV radiation on seed germination and seedling survival in the M_2 generation could also be attributed to the repairs of some induced damages to the DNA by certain biochemical mechanisms present in plants. This is consistent with observed reversion of the three-primary leaf and four-primary leaf seedling mutants when advanced from M_3 to M_4 and from M_4 to M_5 generation respectively. Since plants are unique in the obligatory nature of their exposure to UV, Britt (1995) hypothesized therefore that they may have evolved particularly efficient mechanisms for the elimination of UV-induced DNA damages and mutations. In addition to this, it has been reported that the production of reactive oxygen species in plant tissues as a response to ionizing and nonionizing radiations is controlled by the very efficient enzymatic and non-enzymatic antioxidant defense systems which serve to keep down the levels of free radicals, permitting them to perform useful biological functions without too much damage and act as a cooperative network employing a series of redox reactions (Gill and Tuteja, 2010; Karuppanapandian et al., 2011). Leguminous plants, especially faba bean had been reported to have high antioxidant activity due to the fact that they contained phenolic and flavonoid compounds (Aly and El-Beltagi, 2010; Chaieb et al., 2011; Ismael et al., 2012; Haliem et al., 2013). This may be responsible in part, for the low mutagenic effects of UV irradiation observed in this study.

Low frequencies of mutations recorded by phenotyping

Figure 11. Three-primary leaf seedling mutant of IT86D-719.

Figure 12. Four-primary leaf seedling mutant of IT90K-284-2.

in this study revealed that cowpea plant is considered less amenable to the application of UV irradiated pollen as a practical breeding method. This may not be unexpected since UV radiation has low penetration power and less mutagenic potential in plants when compared to the ionizing radiations. The weak mutagenic effect of UV, even at higher doses has been suggested to be the result of the occurrence of a dark repair system in plant cells (Britt, 1995; Gavazzi and Sanguineti, 1983). Although the results obtained in this study corroborates the findings of Chin and Gordon (1989b) who concluded that the method of pollination with irradiated pollen has not been promising and may not be a useful technique of mutation breeding in rice. However, recent study of the mutagenic potential of UVB irradiation on meiotic-pollen mother cells, pollen grains and seed yield of faba beans revealed wide range of mutagenic action on the frequency and type of chromosomal anomalies, fertility of pollen grains and seed yield productivity based on irradiation dosages, while SDS-PAGE and RPAD-PCR analyses of seeds yielded from irradiated seedlings from UV-B dose for 60 minutes recorded bio-positive effects (Haliem et al., 2013). Therefore, cytogenetic and molecular analysis of the M_2 and M_3 cowpea populations would reveal detail information on the mutagenic potency of UV irradiation of cowpea pollen.

Conclusion

Pollen mutagenesis by UV radiations has not been able to produce any stable cowpea mutant in this study. The method of pollination with UV irradiated pollen has not been promising and this suggests that it may not be an efficient technique for cowpea mutation breeding. However, there is the need to further investigate the mutagenic effect of pollination with irradiated pollen in cowpea using other physical mutagens.

Conflict of Interests

The authors have not declared any conflict of interests.

REFERENCES

Agrawal SB, Singh S, Agrawal M (2009). Ultraviolet-B induced changes in gene expression and antioxidants in plants. Adv. Bot. Res. 52:47-86.

Aly AA, El-Beltagi HES (2010). Influence of ionizing irradiation on the antioxidant enzymes of Vicia faba L. *Grasas Y Aceites.* 61(3):288-294.

Amano E, Smith HH (1965). Mutations induced by ethylmethanesulfonate in maize. Mutat. Res. 2:344-351.

Bie T, Cao Y, Chen P (2007). Mass production of intergeneric chromosomal translocations through pollen irradiation of *Triticum durum-Haynaldia villosa* Amphiloid. J. Integr. Plant Biol. 49(11):1619-1626

Brewbaker JL, Emery GC (1962). Pollen Radiobotany. Rad. Bot. 1:101-154.

Britt AB (1995). Repair of DNA damage induced by ultraviolet radiation. Plant Physiol.108:891-896.

Chaieb N, González JL, López-Mesas M, Bouslama M, Valiente M (2011). Polyphenols content and antioxidant capacity of thirteen faba bean (Vicia faba L.) genotypes cultivated in Tunisia. Food Res. Int. 44(4):970-977.

Chin S, Gordon GH (1989a). Pollination with irradiated pollen in rice – *Oryza sativa* L.I. First (M_1) generation. Heredity 63:163-170.

Chin S, Gordon GH (1989b). Pollination with irradiated pollen in rice

– *Oryza sativa* L.II. Second (M_2) generation. Heredity 63:171-179.
Devreux M, Donini B, Scarascia-Mugnozza GT (1972). Genetic effects of gametophyte irradiation in barley III. Frequecies and types of mutations induced. Rad. Bot. 12:87-98.
Gavazzi G, Sanguineti MC (1983). Pollen mutagenesis in maize. In Mulcahy DL, Ottaviano E (Eds.) Pollen: Biology and implication for plant breeding, (pp 352-357). Elsevi Biomedical, New York.
Gill SS, Tuteja N (2010). Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. Plant Physiol. Biochem. 48(12):909-930.
Gillissen LJW (1978). Post-X-irradiation effects on pollen germination in-vivo. Env. Expt. Bot. 18:81-86.
Godbole M, Murthy HN (2012). Parthenogenetic haploid plants using gamma irradiated pollen in snapmelon (Cucumis melo var. momordica). Plant Cell Tissue Organ Cult.109:167-170.
Gonzalo MJ, Claveria E, Monforte AJ, Dolcet-Sanjuan R (2011). Parthenogenic haploids in melon: Generation and molecular characterization of a doubled haploid line population. J. Am. Soc. Hortic. Sci. 136:145-154.
Grant JE, Pandey KK, Williams EG (1980). Pollen nuclei after ionizing irradiation for egg transformation in *Nicotiana*. NZ. J. Bot. 18:339-341.
Grouh MSH, Vahdati K, Lotfi M, Hassani D, Biranvand NP (2011). Production of haploids in Persian walnut through parthenogenesis induced by gamma-irradiated pollen. J. Am. Soc. Hortic. Sci. 136:198-204.
Haliem EA, Abdullah H, AL-Huqail AA (2013). Oxidative Damage and Mutagenic Potency of Fast Neutron and UV-B Radiation in Pollen Mother Cells and Seed Yield of Vicia faba L. BioMed Res. Int. 2013: Article ID 824656, 12p.
Ismael DS, Timoracka M, Vollmannova A (2012). Influence of variety, locality and soil contamination on total polyphenol content and antioxidant activity of faba bean grains. J. Microbiol. Biotechnol. Food Sci. 2012(1):931-941.
Karuppanapandian T, Moon J, Kim C, Manoharan K, Kim W (2011). Reactive oxygen species in plants: their generation, signal transduction, and scavenging mechanisms. Aust. J. Crop Sci. 5(6):709-725.
Kosmrlj K, Murovec J, Bohanec B (2013). Haploid induction in hull-less seed pumpkin through parthenogenesis induced by X-ray-irradiated pollen. J. Am. Soc. Hortic. Sci. 138(4):310-316.
Kurtar ES, Balkaya A (2010). Production of in vitro haploid plants from in situ induced haploid embryos in winter squash (*Cucurbita maxima* Duchesne ex Lam.) via irradiated pollen. Plant Cell Tissue Organ Cult. 102:267-277.
Lagoda PJI (2012). Effects of radiation on living cells and plants. In Shu QY, Forster BP, Nakagawa H (Eds.), Plant Mutation Breeding and Biotechnology, Joint FAO/IAEA Division of nuclear Techniques in Food and Agriculture International Atomic Energy Agency, Vienna, Austria. pp. 123-134.
Mba C, Afza R, Shu QY (2012). Mutagenic Radiations: X-Rays, Ionizing Particles and Ultraviolet. In Shu QY, Forster BP, Nakagawa H (Eds.), Plant Mutation Breeding and Biotechnology, Joint FAO/IAEA Division of nuclear Techniques in Food and Agriculture International Atomic Energy Agency, Vienna, Austria. pp. 83-90.
Mottinger JP (1970). The effects of X-rays on the bronze and shrunken loci in maize. Genetica 64:259-271.
Naito K, Kusaba M, Shikazono N, Takano T, Tanaka A, Tanisaka T, Nishimura M (2005). Transmissible and nontransmissible mutations induced by irradiating *Arabidopsis thaliana* pollen with gamma rays and carbon ions. Genetica 169:881-889.
Neuffer MG (1957). Additional evidence on the effect of x-ray and ultraviolet radiation on mutation in maize. Genetica 42(3):273-282.
Olasupo FO, Ilori CO, Forster BP, Bado S (2016). Mutagenic effects of gamma radiation on eight accessions of cowpea (*Vigna unguiculata* [L] Walp.). Am. J. Plant Sci. 7:339-351.
Pandey KK (1975). Sexual transfer of specific genes without gametic fusion. Nature 256:310-313.
Pandey KK (1978). Gametic gene transfer in *Nicotiana* by means of irradiated pollen. Genetica 49:53-69.
Pandey KK (1980a). Further evidence for egg transformation in *Nicotiana*. Heredity 45:15-29.
Pandey KK (1980b). Parthenogenetic diploidy and egg transformation induced by irradiated pollen in *Nicotiana*. NZ J. Bot. 18:203-207.
Pandey KK (1983). Evidence for gene transfer by the use of sub-lethal irradiated pollen in *Zea mays* and theory of occurrence by chromosome repair through somatic recombination and gene conversion. Mol. Gen. Genet. 191:358-365.
Ravanat J, Douki T, Cadet, J (2001). Direct and indirect effects of UV radiation on DNA and its components. J. Photochem. Photobiol. B: Biol. 63:88-102.
Sanford JC, Chyi YS, Reisch BI (1984a). An attempt to induce “egg transformation” in *Lycopersicon esculentum* Mill. Using irradiated pollen. Theor. Appl. Genet. 67:553-558.
Sanford JC, Chyi YS. Reisch BI (1984b). Attempted “egg transformation” in *Zea* mays L. using irradiated pollen. Theor. Appl. Genet. 68:269-275.
Stadler LJ (1939). Genetic studies with ultra-violet radiation. In: Punnrtt, R.C. ed. Proceedings of the seventh international genetics congress. Cambridge, UK: Cambridge University Press. pp. 269-276.
Vassileva-Dryanovska OA (1966). Fertilization with irradiated pollen in *Tradescantia*. Rad. Bot. 6:469-479.
Yahata M, Yasuda K, Nagasawa K, Harusaki S, Komatsu H, Kunitake H (2010). Production of haploid plant of Banpeiyu’ pummel [*Citrus maxima* (Burm.) Merr.] by pollination with soft x-rayirradiated pollen. J. Jap. Soc. Hortic. Sci. 79:239-245.
Yang C, Mulligan BJ, Wilson ZA (2004). Molecular genetic analysis of pollen irradiation mutagenesis in *Arabidopsis*. New Physiol. 164:279-288.
Yue J, Zou J (2012). Study of radiation effects on upland cotton (*Gossypium hirsutum* L.) pollen grain irradiated by 60Co- ray. J. Agric. Sci. 4(7):85-94.

academicJournals

Vol. 8(11), pp. 240-249, November 2016
DOI: 10.5897/JPBCS2016.0568
Article Number: 5F1CA7B61317
ISSN 2006-9758

http://www.academicjournals.org/JPBCS

Journal of Plant Breeding and Crop Science

Full Length Research Paper

Analysis of genotype x environment interaction and seed yield stability of sesame in Northern Ethiopia

Fiseha Baraki[1]*, Yemane Tsehaye[2] and Fetien Abay[2]

[1]Crop Research Case Team, Humera Agricultural Research Center, Tigray, Ethiopia.
[2]Department of Crop and Horticultural Science, Mekelle University, Ethiopia.

Received 7 February, 2016; Accepted 22 April, 2016

The experiment was conducted in Northern Ethiopia from 2011-2013 under rain fed conditions in a total of seven environments vis. E1, E2, E3, E4, E5, E6 and E7. The objective of the study was to evaluate the adaptability and stability of sesame genotypes across environments. 13 sesame genotypes were evaluated and the experiment was laid out in a Randomized Complete Block Design with three replications. The average grain yield of the genotypes was 742.9 Kg/ha with the outstanding genotypes being G4 (926.8 kg/ha), G1 (895.1 kg/ha) and G12 (832.7 kg/ha) respectively, and low the yielding genotype was G9 (614.3 kg/ha). The combined ANOVA for grain yield showed significant effects of the genotypes, environments and genotype x environment interaction. According to the additive main effect and multiplicative interaction bi-plot (AMMI bi-plot) and Genotype x Environment interaction bi-plot (GGE bi-plot) G12 was the most stable, and G7, G8 and G9 were the unstable genotypes. Furthermore, the Genotype main effects and GGE bi-plot showed E5 as the most discriminating and representative environment. The GGE bi-plot also identified two different growing environments, the first environment containing E4 and E6 (in the Dansha area) with the wining genotype G1; and the second environment encompassing E1, E2, E3, E5 and E7 (in the Humera, Dansha and Sheraro areas) with winning genotype of G4.

Key words: AMMI bi-plot, environment, GEI, GGE bi-plot.

INTRODUCTION

Sesame (*Sesamum indicum* L.) is an ancient oil seed crop known and used by man.It is not clearly known where Sesame originated and different scholars declare different regions as the centers of origin of this crop. However, Ethiopia is recognized as a center of diversity for this crop owing to the highly diversified sesame types present in the country.

According to FAOSTAT (2012), Ethiopia is the sixth largest sesame producer in the world and third in Africa following Myanmar, India, China, Tanzania and Uganda.

*Corresponding author. E-mail: fish051bar@gmail.com.

Table 1. Agro-climatic and soil characteristics of the experimental sites.

Location	Latitude (°N)	Longitude (°E)	Altitude (m)	Annual rain fall (mm)	Min - Max Temp (°C)	Soil texture		
						Clay (%)	Silt (%)	Sand (%)
Humera	14°15'	36°37'	609	563.2	18.8-37.6	35.66	25.66	38.66
Sheraro	14°24'	37°45'	1028	676.7	18.8-34.9	21	27.28	51.71
Dansha	13°36'	36°41'	696	888.4	28.7(mean)	_	_	_

The productivity of sesame in Northern Ethiopia (specifically in the study area) is very low (525 kg/ha) compared with the national average yield of about 757 kg/ha in 2012/13production year (CSA, 2013) as well as with the world average yields, especially countries like Mozambique which produce up to 1500 kg/ha (Buss, 2007). On-top of this problem the quality of sesame seeds is deteriorating from time to time which may negatively affect the important traits of Humera sesame such as its seed color aroma and seed size uniformity. Sesame genotypes grown in Ethiopia, including the released varieties, are highly variable when grown across locations. Hence, it is important to test different newly introduced genotypes or released varieties across locations.

According to Ceccarelli (2012) the response of genotypes across environments may be with no interaction, quantitative interaction or qualitative interaction. GEI (Genotype x environment interaction) occurs when different genotypes respond differently to different environments and it is familiar in agricultural research (Allard and Bradshaw, 1964). Sesame is a short day plant and sensitive to photoperiod, temperature, moisture stress, water logging and different management practices and its yield and yield attributes are not stable and vary widely over different environments. Hence, this experiment was undertaken to identify stable and high yielding genotype(s) and to recommend best Sesame genotype(s) for the different sesame growing areas so as to boost sesame production and productivity in these (study areas) sesame belts.

In the presence of significant GEI, there are a number of univariate and multivariate stability measures used to identify stable and high yielding genotypes. Additive main effect and multiplicative interaction (AMMI) is important to analyze multi-environment trials (METs) data and it interprets the effect of the genotype (G) and environments (E) as additive effects and the GEI as a multiplicative component (which are sources of variation) and submits it to principal component analysis (Zobel et al., 1988). Another multivariate stability measure called Genotype main effects and Genotype x Environment interaction (GGE) effects is also important to identify mega-environments, the “which-won-where” pattern, and to evaluate genotypes and test environments (Yan et al., 2007).

MATERIALS AND METHODS

Procedure

The experiments were conducted in North western and Western Tigray, Northern Ethiopia, under rain fed condition, from 2011-2013 in Humera and Dansha areas, and in 2013 cropping season in Sheraro (a total of seven environments); E1, E2, E3 are 2011, 2012, 2013 growing seasons respectively in Humera; E4, E5, E6 are 2011, 2012, 2013 growing seasons respectively in Dansha; and E7 is 2013 growing season in Sheraro. Some characteristics of the study areas are given in Table 1. Thirteen sesame genotypes (Acc#031 (G1), Oro (9-1) (G2), NN-0079-1(G3), Acc-034 (G4), Abi-Doctor (G5), Serkamo (G6), Acc-051-020sel-14 (G7), Tate (G8), Acc-051-02sel-13 (G9), Adi (G10), Hirhir (G11), Setit-1 (G12), Humera-1(G13)), brought from WARC (Werer Agricultural Research Center), were sown in RCBD with three replications. Each genotype was randomly assigned and sown in a plot area of 2.8 m by 5m with 1m between plots and 1.5 m between blocks keeping inter and intra row spacing of 40 cm and 10 cm, respectively. Each experimental plot received all management practices equally and properly as per the recommendations for the crop.

Statistical analysis

Homogeneity of residual variances was tested prior to a combined analysis over locations in each year as well as over locations and years (for the combined data) using Bartlet's test (Steel and Torrie, 1980). Accordingly, the data collected were homogenous and all data showed normal distribution.

A combined analysis of variance was performed from the mean data of all environments to detect the presence of GEI and to partition the variation due to genotype, environment and genotype x environment interaction. Moreover, mean comparison using Duncan's Multiple Range Test (DMRT) was performed to explain the significant differences among means of the genotypes. GenStat 16th edition (GenStat, 2009) statistical software was used to analyze the combined mean of the different traits of the genotypes. The model employed in the analysis was;

$$Y_{ijk} = \mu + G_i + E_j + B_k + GE_{ij} + \varepsilon_{ijk}$$

where: Y_{ijk} is the observed mean of the i^{th} genotype (G_i) in the j^{th} environment (E_j), in the k^{th} block (B_k); μ is the overall mean; $G_i$is effect of the i^{th} genotype; $E_j$is effect of the j^{th} environment; B_k is block effect of the i^{th} genotype in the j^{th} environment; GE_{ij} is the interaction effects of the i^{th} genotype and the j^{th} environment; and ε_{ijk} is the error term.

A bi-plot showing the genotype and environmental means against

Table 2. Mean Squares for different agronomical traits recorded on sesame genotypes across locations.

Source of variation	d.f	YLD (kg/ha)	DF	DM	LCBZ (cm)	NB	NC	PH (cm)
Rep	2	732	7.1	4.9	82.2	0.1	31.3	74.8
Genotype	12	208413**	16.5**	27.9**	120.1**	1.4**	81.1**	429.7**
Env	6	329874**	60**	221.8**	2792.2**	9.4**	1268.2**	16284.7**
Gen*Env	72	24149**	4.1**	8.3**	64.0**	0.4**	60.9**	189.1**
Residual	180	2707	1.3	2.4	25	0.1	16.1	50.7

*,** statistically significant at (p<0.05 and p<0.01) respectively, ns= non –significant; d.f= degree of freedom, YLD=Grain Yield, DF=Days to 75% flowering, DM= Days to 75% maturity, LCBZ= Length of capsule bearing zone, NB=Number of branches, NC=Number of capsules and PH=Plant height.

Interaction Principal component analysis one (IPCA1) (AMMI1 bi-plot), and Interaction Principal component analysis one (IPCA1) against Interaction Principal component analysis two (IPCA2) (AMMI2 bi-plot) was also performed using AMMI model using GenStat software.

A GGE bi-plot was also executed using GGE bi-plot in the Meta analysis of GenStat 16th edition. This methodology uses a bi-plot to show the factors (G and GE) that are important in genotype evaluation and that are also the sources of variation in GEI analysis of MET (Multi-environment trial) data (Yan, 2001). Moreover, mean comparison using Duncan's Multiple Range Test (DMRT) was performed to explain the significant differences among means of genotypes and their traits.

RESULT AND DISCUSSION

Combined Anova and estimation of variance components

The results obtained from the combined analysis of variance of all the evaluated traits and genotypes is illustrated in Table 2. The genotype, environment and genotype x environment interaction (GEI) variance were decomposed to provide a general overview in relation to the evaluated traits and overall performance of the genotypes (Tables 2 and 3). Accordingly, the genotypes, the environments and the genotype x environment interaction components showed highly significant variation (p<0.001) for all agronomic traits. This statistical difference confirms that the difference of the traits was due to both the main and interaction effects. Zerihun et al. (2011) also found similar results of the genotype, environment and genotype x environment interaction effects in barley land races. On top of the genetic variability, the ANOVA also revealed that the environments (both locations and growing seasons) on which the experiments were conducted were different from one another in treating the genotypes. Moreover, it also indicates that the response of the genotypes were unstable and fluctuated in their trait expression with change in the environments. This phenomenon clearly confirms the existence of GEI in this study. For most of the traits the contribution of environment for the overall variance was high (ranging from 29.5% for grain yield to 77.7% for plant height) followed by genotype × environment interaction and genotype respectively.

Similar results were reported by Hagos (2009); Ahmed and Ahmed (2012). With respect to grain yield, the greatest source of variation was mainly the inherent genetic component meaning genotypic effect (37.3 %) (Table 3) which is similar to the results reported by Zenebe and Hussien (2009) and John et al. (2001)..

Agronomic performance of Sesame genotypes

The average grain yield of the tested sesame genotypes over the seven environments was 742.9 kg/ha. G4 had the highest average grain yield (926.8 kg/ha) followed by G1 (895.1 kg/ha) while G9 was the lowest yielding genotype (614.3 kg/ha) (Table 4). G4 had early flowering (39.7 days) and early maturing (84.7 days). On the contrary, G8, G1, and G2 were late flowering genotypes (Table 4). Similarly G2 was the latest maturing genotype (89.6 days) followed by G8 and G11 which took on average of about 87.8 days each to reach maturity. The shortest (43.01cm) and longest (51.27 cm) average length of capsule bearing zone was recorded from G8 and G1 respectively. G1 also had the highest number of branches and number of capsules whereas G10 and G2 had the lowest (table 3). G1 (119.5 cm) and G6 (119.9 cm) were the genotypes with longest stature.

Variance estimate of grain yield of the genotypes

The combined ANOVA for grain yield revealed that there were highly significant variation (p<0.01) among the genotypes, environments (year, location, year x location) and genotype by environment interaction (Genotype x Year, Genotype x Location and Genotype x Year x Location) (Table 5). These significant variations of the genotypes, environments and the GEI indicated that the response of the genotypes were unstable and fluctuated in their grain yield with change in environment and these

Table 3. Combined Sum of Squares for agronomic traits of Sesame genotypes evaluated during 2011-2013.

Source of variation	d.f	YLD	DM	DF	LCBZ	NB	NC	PH
Replication	2	1464(0.0)	14.2(1.3)	10.0(0.4)	164.4(0.6)	0.2(0.2)	62.6(0.4)	149.7(0.1)
Genotype	12	2500959(37.3)	197.6(17.9)	334.7(12.4)	1441.0(5.2)	16.8(14.4)	973.2(6.1)	5157.0(4.1)
Env	6	1979243(29.5)	359.8(32.5)	1331.0(49.1)	16753.4(61.0)	56.3(48.1)	7609.0(47.7)	97708.7(77.7)
Gen*Env	72	1738701(25.9)	295.6(26.7)	596.1(22.0)	4609.3(16.8)	32.0(27.4)	4383.6(27.5)	13617.1(10.8)
Residual	180	487308(7.3)	238.5(21.6)	438.0(16.2)	4486.6(16.3)	11.6(9.9)	2907.5(18.2)	9125.1(7.3)
Total	272	6707676	1105.7	2709.9	27454.7	116.9	15935.9	125757.5

¥ No out of parenthesis and inside parenthesis are SS and % SS of traits respectively; Gen=genotype; Env=environment.

Table 4. Combined mean yield and related traits of sesame genotypes over all environments.

Genotype	YLD	DF	DM	LCBZ	NB	NC	PH
Acc#031	895.1^{b}	42.2^{a}	86.7^{bcd}	51.3^{a}	2.3^{a}	28.88^{a}	119.5^{a}
Oro(9-1)	638.1^{hi}	42.1^{a}	89.6^{a}	44.1^{cd}	1.7^{ef}	21.42^{e}	106^{d}
NN-0079-1	740.4^{e}	41.0^{bc}	86.5^{cd}	49.7^{ab}	1.8^{cde}	24.5^{bcd}	118.4^{a}
Acc-034	926.8^{a}	39.7^{f}	84.7^{e}	48.8^{ab}	2.2^{a}	27.2^{ab}	108.3^{cd}
Abi-Doctor	662.6^{gh}	40^{ef}	85.7^{d}	51.1^{a}	1.6^{fg}	23.3^{cde}	112.4^{bc}
Serkamo	711.5^{ef}	41.4^{b}	87.0^{bc}	47.7^{ab}	1.7^{def}	23.5^{cde}	119.9^{a}
Acc-051-020sel-14	687.5^{fg}	41.2^{bc}	87.4^{bc}	47.11^{bc}	1.8^{bcd}	25.1^{bc}	108.8^{cd}
Tate	655.2^{gh}	42.8^{a}	87.8^{b}	43.01^{d}	2.1^{a}	24.6^{bcd}	110.4^{cd}
Acc-051-02sel-13	614.3^{i}	40.2^{def}	87.2^{bc}	49.04^{ab}	1.6^{fg}	24.2^{cd}	112.5^{bc}
Adi	697.6^{f}	40.5^{c-f}	87.2^{bc}	48.5^{ab}	1.5^{g}	22.1^{de}	108.8^{cd}
Hirhir	791.5^{d}	41.0^{bcd}	87.8^{b}	47.9^{ab}	2.0^{bc}	24.6^{bcd}	111.8^{bc}
Setit-1	832.7^{c}	40.7^{b-e}	86.4^{cd}	49.5^{ab}	1.9^{bc}	24.4^{cd}	112.1^{bc}
Humera-1	805.1^{cd}	40.9^{bcd}	86.9^{bc}	47.3^{b}	2.0^{b}	25.8^{bc}	116.4^{ab}
Mean	742.9	41.07	86.98	48.1	1.9	24.6	112.7
LSD	83.83	1.8	2.5	8.04	0.409	6.47	11.5
CV(%)	7	2.8	1.8	10.4	13.7	16.4	6.3

Means followed by the same letter are not statistically different from each other (DMRT, at 5%).

phenomenon clearly confirmed the presence of GEI in this study. Figure 1 depicts clearly the fluctuation of the genotypes across the environments. The grain yield of the thirteen genotypes was highly fluctuating over the seven environments showing highest grain yield cross-over interaction from environment to environment.

Among the environments the highest seed yield (1131 kg/ha) was observed from genotype G1in environment seven (E7) and the lowest seed yield (395.9 kg/ha) was recorded from genotype G9 in environment two (E2) (Table 6).

AMMI model

The AMMI model is fully informative for both the main effect as well as for the multiplicative effects, for clearly understanding GEI (Zobel et al., 1988). In addition to the usual ANOVA the ANOVA from the AMMI model for grain yield also detected significant variation ($p<0.001$) for both the main and interaction effects indicating the existence of a wide range of variation between the genotypes, years (seasons), locations and their interactions.

AMMI1 bi-plot analysis

The AMMI bi-plot analysis provides a graphical representation to summarize information on main effect and interaction effects of both genotypes and environments at the same time.The AMMI1 bi-plot containing the genotype and environment means against interaction principal component analysis one (IPCA1) scores is illustrated in Figure 2. As indicated in Figure 1 the displacement along the abscissa reflected differences in main effects, whereas displacement along the ordinate

Table 5. Combined ANOVA for grain yield (kg/ ha) of sesame genotypes.

Source of variation	d.f	SS	MS
Rep	2	1464	732
Genotype	12	2500959	208413**
Year	2	1180209	590104**
Location	2	722088	361044**
Genotype x Year	24	388067	16169**
Genotype x Location	24	702189	29258**
Year x Location	2	76946	38473**
Genotype x Year x Location	24	648446	27019**
Residual	180	487308	2707
Total	272	6707676	24661

d.f= Degree of freedom, SS=Sum of squares, MS= Mean squares.

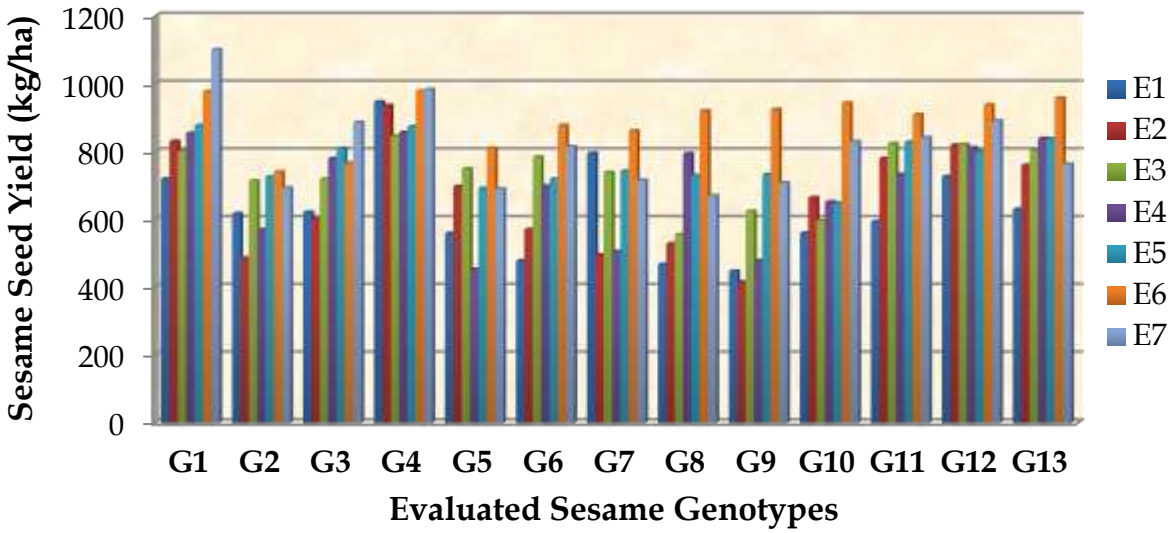


Figure 1. Mean seed yield of 13 sesame genotypes across seven environments, where: E1, E2, E3...refers to the environments and G1, G2, G3...refers to the genotypes.

exhibited differences in interaction effects. Genotypes and environments with IPCA1 greater than zero are classified as high yielding genotypes and favorable environments whereas those with IPCA1 lower than zero are classified as low yielding genotypes and unfavorable environments (Yan and Thinker, 2006). Accordingly genotypes such as, G1, G4, G11, G12 and G13 were the genotypes with above average mean grain yield as they laid-down on the right side of the vertical line (grand mean of the genotypes and environments). Conversely, genotypes G2, G5, G6, G7, G8, G9 and G10 had below grand mean because they laid down to the left side of the vertical line. Exceptionally, G3 laid down very close to the vertical line, indicating the mean yield of G3 was highly similar over all environments and parallel to the grand mean of all genotypes. G4 followed by G1 had higher mean yield in the favorable environments, whereas G9 and G2 had lower mean yield in the unfavorable environments. Regardless of their contribution for the interaction, G8 and G5 fall on the same vertical line (ideal) showing their similarity in their mean yield. G1 and G10 which laid down on the same horizontal line had similar contribution in the interaction component despite of their yield performance.

Regarding the environments, E5, E6 and E7 had above the grand mean grain yield and were considered as favorable environments. On the other hand, E1, E2 and E4 had below average grain yield and were considered as unfavorable environments. E3 laid down very close to the grand mean line indicating that genotypic yield in E3

Table 6. Grain yield recorded from 13 sesame genotypes in each of seven environment and overall genotypic mean.

Genotype	Gen	Test Environments							
Name	Code	E1	E2	E3	E4	E5	E6	E7	Mean
Acc#031	G1	736.8 ^{c}	840^{ab}	801.8^{bc}	852.2 ^{a}	**911.5 ^{a}**	**992.5 ^{a}**	**1131 ^{a}**	895.1^{b}
Oro(9-1)	G2	612.6 ^{d}	470.9^{fg}	711.8 ^{e}	582.6^{ef}	701.2 ^{e}	725.7 ^{g}	661.7 ^{f}	638.1^{hi}
NN-0079-1	G3	606.1 ^{d}	610.6 ^{de}	721.7 ^{e}	787.9^{abc}	832.3^{bc}	750^{fg}	873.9 ^{c}	740.4^{e}
Acc-034	G4	964 ^{a}	930.6 ^{a}	846.4 ^{a}	868.4^{a}	890.3^{ab}	987^{ab}	1001 ^{b}	926.8^{a}
Abi-Doctor	G5	562.2 ^{d}	698 ^{cd}	752.7 ^{d}	453.1 ^{g}	684.5 ^{e}	805.2^{ef}	682.4^{ef}	662.6^{gh}
Serkamo	G6	470.4 ^{e}	577.9 ^{e}	793.3 ^{c}	709.7 ^{cd}	717.5 ^{e}	887.1^{cd}	825 ^{cd}	711.5^{ef}
Acc-051-020sel-14	G7	815.3 ^{b}	465.3^{fg}	738.7 ^{de}	501.5^{fg}	712.3 ^{e}	850.3^{de}	728.8^{ef}	687.5^{fg}
Tate	G8	454.7 ^{e}	510^{ef}	552.6 ^{g}	775.9^{abc}	699.2 ^{e}	923.3^{abc}	670.4^{ef}	655.2^{gh}
Acc-051-02sel-13	G9	438.5 ^{e}	395.9 ^{g}	618.9 ^{f}	475 ^{g}	751.8 ^{de}	924.5^{abc}	695.4^{ef}	614.3^{i}
Adi	G10	558.5 ^{d}	686.8 ^{cd}	596.2 ^{f}	646.6 ^{de}	604.6 ^{f}	946.7^{abc}	843.6 ^{c}	697.6^{f}
Hirhir	G11	587.2 ^{d}	794.6^{bc}	827.1^{ab}	742.6^{bc}	828.1^{bc}	913.2^{bcd}	848.1 ^{c}	791.5^{d}
Setit-1	G12	746.5^{bc}	811.2 ^{b}	820.5^{abc}	806.9^{ab}	812.3 ^{cd}	932.3^{abc}	898.8 ^{c}	832.7^{c}
Humera-1	G13	630.1 ^{d}	770.4^{bc}	803.3^{bc}	839.9 ^{a}	876.3^{abc}	960.9^{abc}	755 ^{de}	805.1^{cd}
Mean	-	629.5	658.6	737.3	695.5	770.9	892.2	816.6	742.9
LSD ±)	-	69.7	100.5	26.4	86.7	69.5	65.2	78.1	83.83
CV(%)	-	6.6	9.1	2.1	7.4	5.3	4.3	5.7	7

*Means followed by the same letter are not statistically different from each other (DMRT, at 5%); *Bolded yield is highest seed yield of genotypes in their respective environments; *Underlined yield is lowest seed yield of genotypes in their respective environments.

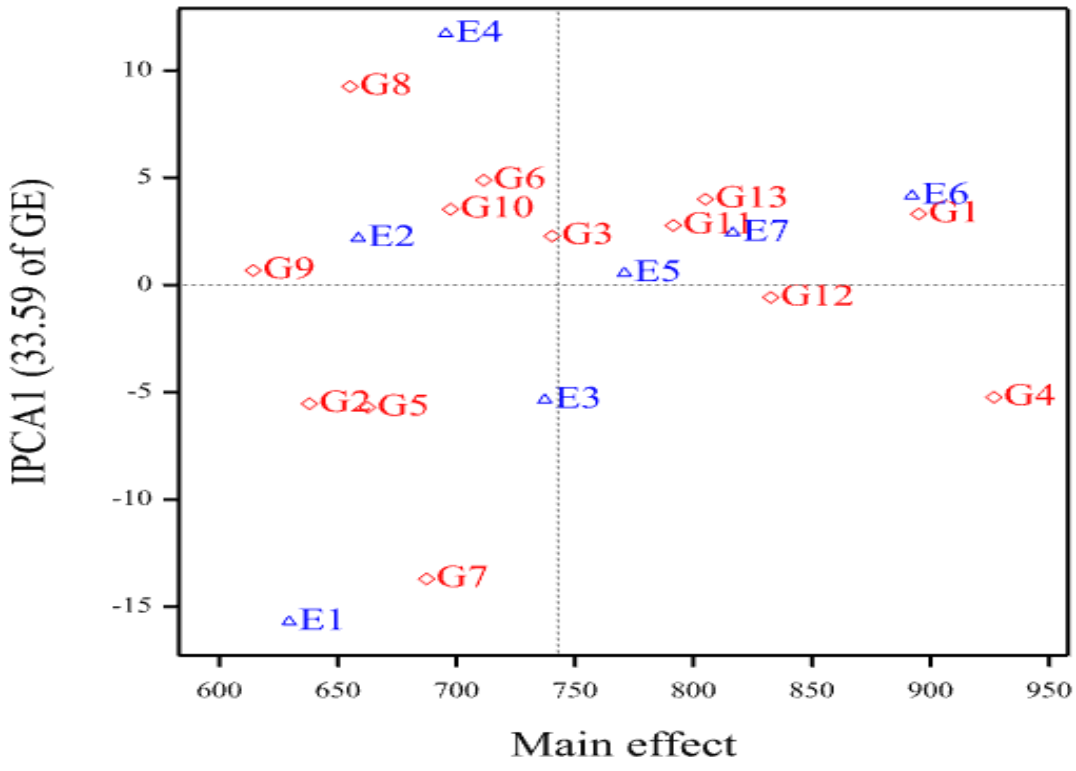


Figure 2. AMMI1 bi-plot showing Genotype and Environmental means against IPCA1. Where the environments are represented by (E) and the genotypes by (G) within detail description of the environments and the genotypes in the material and method part.

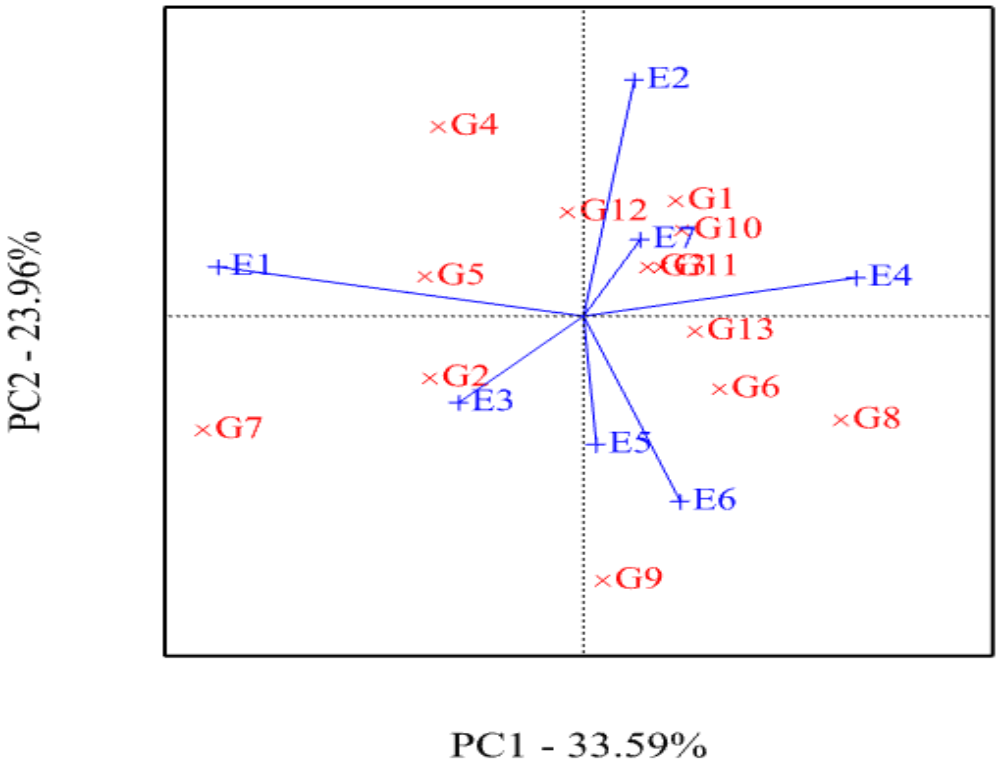


Figure 3. AMMI2 bi-plot showing PC1 versus PC2 indicating the stability of the Genotypes, where the environments are represented by (E) and the genotypes by (G) with in detail description in the material and method part.

represents the overall genotypic mean across all environments.

AMMI 2 bi-plot: The AMMI 2 bi-plot, containing IPCA1 in the X-axis and IPCA2 in the Y-axis, is plotted in Figure 3. The first interaction principal component (IPC1 or PC1) contained 33.59% and the second interaction principal component (IPC2 or PC2) explained about 23.96% and the two interaction principal components cumulatively explained about 57.55% of the sum of squares of the genotype by environment interaction of the genotypes (Figure 3). Purchase (1997) stated that the closer the genotypes to the origin are the more the stable and the furthest genotypes from the origin are the more the unstable ones. In addition the closer the genotypes to the given vector of any environment is the more adaptive to that specific environment and the farthest the genotypes to the given vector of any environment is the less adaptive to that specific environment. Accordingly, genotypes G7, G8, G9 and G4 are far apart from the bi-plot origin indicating these genotypes as the more responsive and contributed largely to the interaction component and considered as specifically adapted genotypes. On the other hand, G11, G12, G13and G3 were the genotypes with least contribution to the interaction component as they are located near the bi-plot origin, indicating their wider adaptability (Figure 3). Regarding the adaptability of the genotypes in the environments; genotypes G1, G3, G10 and G11 were adaptive to E7; and genotypes G2, G5, G12 and G13 were adaptive to environments E3, E1, E2 and E4, respectively.

GGE Bi-plot

The GGE bi-plot used in this study constitutes a summed up of 74.99% total variance of the first two principal components. As indicated by Yan and Thinker (2006), the similarity between two environments as well as genotypes is determined by both the length of their vectors and the cosine of the angle between them (Figure 4). E1 is at about 90^{o} with E4 and E6 indicating that it had no correlation with these environments and could produce less similar information about the tested genotypes (Figure 4). But the other environments had vectors that were linked with less than $90^{o,}$ indicating, these environments were positively correlated with each other. E2 had longest vector and small IPCA2 and that was relatively the most representative and discriminating environment and considered as the ideal environment for

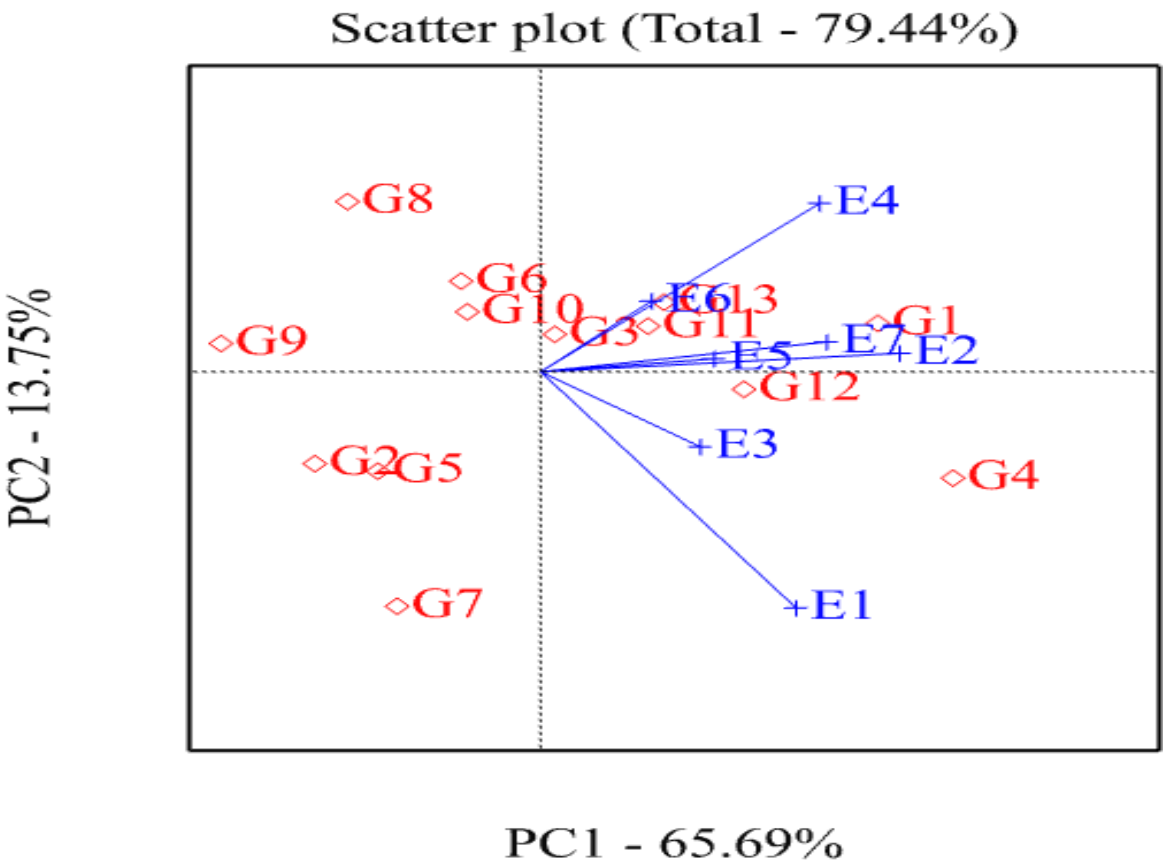


Figure 4. The environment-vector view of the GGE bi-plot to show similarities among test environments.

generally adapted genotypes. Hence, Genotypes with above average yield in this environment had above average yield in all environments. E1 and E4 were the most discriminating but least representative environments which were with little information of the genotypes and favorable for specifically adapted genotypes. Exclusively, E6 was neither discriminating nor representative environment. To clearly display graphically, the 'which-won-where' pattern of a polygon view of GGE bi-plot is exhibited in Figure 5. The polygon was formed by connecting the vertex genotypes that were furthest away from the bi-plot origin such that all other genotypes were included in the polygon. From the polygon view of bi-plot analysis (Figure 5) the bi-plot showed there were two different sesame growing environments. The one environment includes the high yielding environments (E4 and E6), which were in the Dansha area with the winning genotype G1; the second environment contained the low to medium yielding environments (E1, E2, E3, E5 andE7), which were under Humeraand Sheraro areas with a vertex genotype G4. The other vertex genotypes (G7, G8 and G9) without any environment in their sectors were not the highest yielding genotypes at any environment rather they were the poorest genotypes of all or some environments.

CONCLUSION AND RECOMMENDATION

The combined ANOVA showed significant differences among the sesame genotypes in this study for meangrain yield across environments. The results also showed that the environments were highly variable with respect to climatic and/or edaphic factors. This GEI in turn indicated that, the performance or ranking of the genotypes was variable across environments and it was difficult to identify superior genotype for all environments or locations. The GGE bi-plot identified two sesame growing environments; the area of Dansha (E4 and E6) with G1 as a winning genotype, and the other environment encompassing Sheraro (E7) and Humera and Dansha (E1,E2, E3andE5) with G4 as a wining genotype.

The AMMI bi-plot and GGE bi-plot of grain yield data identified G12 as the most stable and widely adapted genotype for grain yield while, G4 and G1 were specifically adapted in the favorable environments

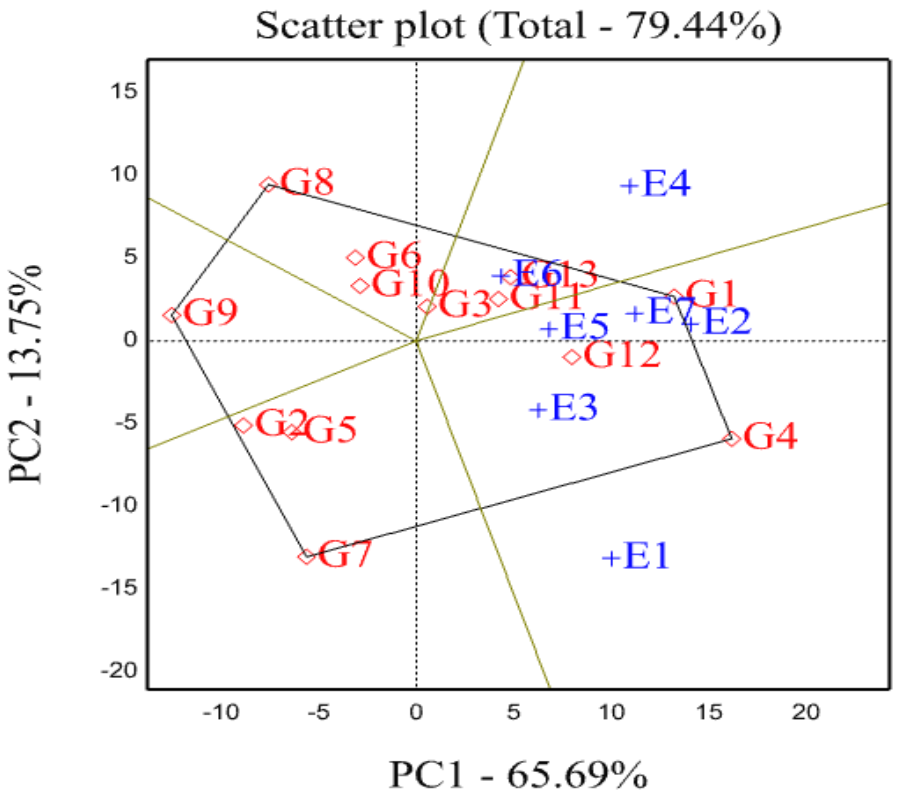


Figure 5. The which-won-where view of the GGE bi-plot.

Conflict of Interests

The authors have not declared any conflict of interests.

ACKNOWLEDGEMENTS

The authors cordially acknowledge the research members of crop department in Humera Agricultural Research Center and for Public and Private Partnership Organization project (PPPO) for financial support.

REFERENCES

Ahmed MBS, Ahmed FA (2012). Genotype X season interaction and characters association of some Sesame (*Sesamumindicum* L.) genotypes under rain-fed conditions of Sudan. Afr. J. Plant Sci. 6(1):39-42.

Allard RW, Bradshaw AD (1964). Implication of genotype by environmental interaction in applied plant breeding. Crop Sci. 4(5):503-506.

Buss J (2007). Sesame production in Nampula: Baselinesurvey report, Mozambique. pp. 2-20.

Ceccarelli S (2012). Plant breeding with farmers – a technical manual.ICARDA, Syria.

CSA (2013).Agricultural Sample Survey. Report on Area and Production, Volume III, Addis Ababa, Ethiopia.

FAOSTAT (2012).(Food and Agriculture Organization of the United Nations). Available at: http://faostat.fao.org/

GenStat (2009).GenStat for Windows (14th Edition) Introduction. VSN International, Hemel Hempstead.

Hagos T (2009). Genotype by Environment Interaction and yield stability of Sesame (*Sesamum indicum* L.) genotypes under North Western and Western lowland Tigray, M.Sc. thesis, Mekelle University, Ethiopia.

John A, Subbaraman N, Jebbaraj S (2001). Genotype by Environment Interaction in Sesame (*Sesamum indicum* L.): Sesame and safflower newsletter no. 16, Institute of Sustainable Agriculture, FAO, Rome.

Steel R, Torrie J (1980).Principles and Procedures of Statistics a Biometrical Approach. 2nd ed. McGraw-Hill, Inc. pp. 471-472.

Yan W (2001). GGE bi-plot windows application for graphical analysis of multi-environment trial data and other types of two-way data. Agron. J. 93:1111-1118.

Yan W, Kang MS, Ma B, Woods S, Cornelius PL (2007). GGE bi-plot vs. AMMI analysis of genotype-by-environment data. Crop Sci. 47:643-653.

Yan W, Thinker NA (2006). bi-plot analysis of multi-environment trial data: Principles and applications. Can. J. Plant Sci. 86:23-645.

Zenebe M, Hussien M (2009). Study on Genotype X Environment Interaction of Oil Content in Sesame (*Sesamum indicum* L.). Middle-East J. Sci. Res. 4:100-104.

Zerihun J, Amsalu A, Fekadu F (2011). Assessment of yield stability and disease responses in Ethiopian Barley (*Hordeium vulgare* L.) Landraces and Crosses. Int. J. Agric. Res. 6:754-768.

Zobel RW, Wright MJ, Gauch HG (1988). Statistical analysis of a yield trial. Agron. J. 80:388-393.

academicJournals

Vol. 8(11), pp. 250-256, November 2016
DOI: 10.5897/JPBCS2016.0586
Article Number: D33659B61320
ISSN 2006-9758

http://www.academicjournals.org/JPBCS

Journal of Plant Breeding and Crop Science

Full Length Research Paper

Combining ability and gene action in sesame (*Sesamum indicum* L) elite genotypes by diallel mating design

Walter Okello Anyanga*, Patrick Rubaihayo, Paul Gibson and Patrick Okori

Makerere University, Kampala, P. O. Box 7062. Uganda.

Received 10 May, 2016; Accepted 29 August, 2016

An 8 x 8 diallel cross mating design with the parents and F1s was used to estimate gene action and combining ability for yield and yield characters in sesame (Sesamum indicum L.). The experiment was conducted at the National Semi-Arid Resources Research Institute (NaSARRI), Serere, Uganda during the second season of 2013. The traits involved: days to 50% flowering, days to maturity, height to first branch, height to first capsules, plant height, length of the capsule fruiting zone and yield per plant. The genetic variance of combining ability was separated into general (GCA) and specific (SCA) combining ability variance components. Baker's ratio, coefficients of genetic determination (CGD) were determined for narrow and broad sense heritabilities. High GCA effects were recorded for days to flowering, height of first capsule and number of branches, branches per plant and capsule length. Desirable SCA effects were recorded for Sesim2//5181 x Renner 1-3-1-1 for reduced height of first branch (-9.48*), Sesim1 x Sesim2 for number of branches per plant (0.901*), capsules on branches (20.75) and yield per plant (2.42*). High Baker's ratio was recorded for most of the traits except for yield per plant (0.233). CGDbs was high for most of the traits except for days to maturity (0.064), plant height (0.346), capsules on main stem (0.358) and capsule width (0.286). These results suggested that both additive and non-additive gene actions played a greater role in these traits. Sesim 1 had the highest number of desirable traits scored for GCA effects and could be considered as a parent for crossing to produce desirable progeny. Sesim 1 x Sesim 2, Sesim 1 x Ajimo A1-5 and Sesim2//5181 x SPS1438-1-6-4 recorded positive significant SCA effects for yield per plant and therefore they could be recommended for hybrid seed production.**

Key words: Baker's ratio, coefficient of genetic determination, gene effects, yield, yield components, sesame.

INTRODUCTION

Sesame (*Sesamum indicum* L.), commonly known as gingelly, til, benniseed and simsim is a member of the order Tubiflorae and family Pedaliaceae (Ahmed and Adam, 2014). It is normally called 'Simsim' in Eastern Africa and it is a traditional and important oilseed crop in Uganda (Auckland, 1970). In 2011, Uganda was the 7th in world sesame production with 173,000 tons of seed. The highest in the world was Myanmar, India and China with 861,573; 769,000 and 605, tons, respectively (FAOSTAT, 2015). In Uganda, sesame ranks second in importance to

*Corresponding author. E-mail: waltanyanga@hotmail.com.

groundnuts and it is predominantly grown in the North and North Eastern parts of the country, although there is a considerable amount grown in Eastern Uganda (Anyanga and Obong, 2001). Sesame is gaining considerable importance in Uganda on account of its economic value and especially its export potential, as a non-traditional export crop from Uganda. The national yield is reported to be about 600 kg/ha (FAOSTAT, 2009). Brigham (1985) reported average yields of 1,564 kg/ha in USA. This disparity in yields is caused by various constraints including lack of improved varieties. It is called the "Queen of oil seeds" because of its excellent quantities of the seed, oil and meal. Sesame is highly nutritive (oil 50% and protein 25%). It is an important annual oilseed crop in the tropics and warm sub-tropics where it is mainly grown in small plots as source of edible oil and one of the ingredients in food products. The seed is also rich in protein, vitamins including minerals and lignans such as sesamolin and sesamin (Moazzami and Kamal-Eldin, 2006). Sesame oil has medicinal and pharmaceutical value and is being used in many health care products (Coulman et al., 2005). The seed contains 50 to 60% oil and 25% protein with antioxidants lignans such as sesamolin, sesamol, sesamin which impart to it a high degree of resistance against oxidative rancidity and gives it a long shelf life (Ashri, 1989). It has been used as an active ingredient in antiseptics, bactericides, viricides, disinfectants, moth repellants and anti-tubercular agents (Bedigian et al., 1985). It is a source of calcium, tryptophan, methionine and many minerals (Johnson et al., 1979).

Although, it is a self-pollinated crop, there is some extent of cross-pollination which has a good scope for exploitation of heterosis. Further, an understanding of the combining ability and gene action is a prerequisite for any successful breeding programme. For breaking the yield barrier and evolving varieties with high yield potential, it is desirable to combine the genes from genetically diverse parents. There are several techniques for evaluating the varieties or cultivars or lines in terms of their combining ability and genetic makeup. Of these, Diallel, partial Diallel and line x tester techniques are in common use.

The concept of combining ability analysis gives precise estimates of the nature and magnitude of gene actions involved in the inheritance of quantitative characters, which facilitate the identification of parents with good general combining ability effects and crosses with good specific combining ability effects.

Successful breeding programme depends on the variability available among the different genotypes and in-depth understanding of the underlying gene action and genetic architecture of traits related to yield. Selection of parents based on their performance *per se* alone may not always be a sound procedure, since phenotypically superior genotypes may yield inferior hybrids and/or poor recombinants in the subsequent segregating generations (Banerjee and Kole, 2009). It is very important to identify parents with high GCA value for the trait to be improved (Banerjee and Kole, 2009). Griffing (1956) provides an efficient estimation of combining ability and the nature of gene action involved. General combining ability is largely due to additive genetic effects and additive x additive epistasis, while specific combining ability is largely a function of non-additive dominance and other types of epistasis.

Knowledge on the nature of the combining ability effects and their resulting variances has a paramount significance in deciding on the selection procedure for exploiting either heterosis or obtaining new recombinants of desirable types in sesame (Solanki and Gupta, 2003). Combining ability is helpful to identify the desirable parents for producing better recombinants (Muhammad and Sedeck, 2015). No breeding research work has been done in Uganda regarding combining ability and gene action on sesame. This can also be observed by limited publications on sesame breeding work in Africa.

The objectives of the present study were to estimate general and specific combining ability, and assess the nature and magnitude of gene action controlling the inheritance of seed yield and some of its components in Uganda's sesame germplasm.

MATERIALS AND METHODS

Experimental site

The trial was planted at the National Semi-Arid Resources Research Institute (NaSARRI), Serere district, Uganda (Latitude 1°31'N; Longitude 33°27'E;and altitude 1,140 masl) during the second rains season of 2013. Uganda has rainy seasons with the first rainy season planted between March to May and the second season planted between July to September. The climate is sub-humid with a mean rainfall of 1350 mm which is bi-modally distributed with higher peak in April-May and lower peak in September-October. There is a very dry season from December to March, during which the temperatures rise to a maximum of around 95°F. The institute has mainly sandy loam soils and some clay loams with tropical savannah grassland vegetation.

Experimental materials

Eight diverse sesame genotypes including local and exotic lines (Table 1) that were previously characterized morphologically in the breeding programme at NaSARRI, Serere and showed contrasting traits were crossed in an 8 x 8 diallel mating design without the reciprocals resulting in 28 F_1 combinations. These hybrids along with their parents were grown in a completely randomized block design with three replications. The plots were single rows in 2 m length with spacing of 30 x 10 cm. In the experimental design, Sesim 2 variety was used as border row planted at the beginning and the end of each replication to control the border effects.

Data collection

The data recorded were: number of days to 50% flowering, days to physiological maturity, number of primary branches per plant, height

Table 1. Selected genotypes for diallel mating design.

S/N	Genotype/line	Country of origin	Main characteristics
1	Sesim 1	Uganda	Many branches from the middle, late maturing
2	Sesim 2	Uganda	Many branches from the middle, late maturing
3	Ajimo A1-5	Uganda	Many branches from the middle, late maturing
4	Sesim2//5181	Breeding line developed from a cross between Ugandan (Sesim 2) and Thailand (5181) lines	Hairy, few branches, branches from up
5	Ajimo A1-6//7029	Breeding line developed from a cross between Ugandan (Ajimo A1-6) and Thailand (7029) lines	Medium early, medium height
6	SPS1438-1-6-4	China	Early maturing, tall, branches from down, long capsule zone formation, long capsules,
7	4036-1-10-2	China	Medium maturity, normally two to four branches per plant, long thick capsules
8	Renner 1-3-1-1	USA	Early maturing, short capsules, branches from the middle of the stem

to first branch on the main stem, height to first capsule on the main stem, length of the capsule formation zone, number of capsules on the branches, number of capsules on the main stem, total number of capsules per plant, plant height and seed yield per plant. Five plants were selected from the centre of the row for data collection except for days to 50% flowering and days to maturity which was on the plot basis.

Statistical analysis

The characters under study were subjected to analysis of variance for a completely randomized block design on an entry mean basis. The combining ability analysis was carried out following Griffings (1956) method 4 which includes only the direct F_1 crosses without parents and reciprocals. The estimates of GCA effect for the parents and the SCA effects for the crosses were calculated according to Singh and Chaudhary (1985).

The model followed was:

$$Y_{ij} = \mu + g_i + g_j + s_{ij} + \frac{1}{bc}\sum_{k}\sum_{l} \varepsilon_{ijkl} \quad i,j=1,\ldots p,$$

Where, Y_{ij} is the observation of cross (x_{ij}), μ is the population mean, g_i and g_j are the general combining ability effect for the i^{th} and j^{th} parents, S_{ij} is the specific combining ability effect of the cross between the i^{th} and j^{th} parents such that $S_{ij} = S_{ji}$ and ε_{ijkl} is the experimental error due to environmental effect associated with the $ijkl^{th}$. The variance component was calculated using the formula: $\sigma^2_g=(MS_{gca}-MS_{error})/(p-2)$; $\sigma^2_s=(MS_{sca}-MS_{errror})/1$

Where, MS_{gca}= variance due to GCA; MS_{sca}= variance due to SCA; MS_{error}= error variance.

Gene action

Since the parents used in the crosses were considered fixed, coefficient of genetic determination were used to estimate total genetic variability, broad sense and narrow sense coefficient of genetic determination (CGD) by the formula below:

Baker's ratio = $(2\sigma^2_{gca})/(2\sigma^2_{gca} + \sigma^2_{sca})$
$CGD_{BS} = (2\sigma^2_{gca} + \sigma^2_{sca})/(2\sigma^2_{gca} + \sigma^2_{sca} + \sigma^2_e)$
$CGD_{NS} = (2\sigma^2_{gca})/(2\sigma^2_{gca} + \sigma^2_{sca} + \sigma^2_e)$

Both Genstat 14th Edition (Payne et al., 2010) and Excel computer programmes were used to analyze the data.

RESULTS AND DISCUSSION

General combining ability

The general combining ability (GCA) effects of the parents are presented in Table 2. Sesim 1 had highly significant effect (P<0.001) for days to flowering, height to first branch, branches per plant and capsule width. It was also significant (P<0.05) for height to first capsule, plant height, capsules on branches, length of capsule zone and capsule length. This indicated that Sesim1 as a parent contributed much to the progenies and therefore is a desirable parent for crossing in order to improve those traits in a crossing programme. Sesim 2 was desirable for number of branches per plant such that it can only be used to increase the number of branches in the progeny. Ajimo A1-5 had desirable GCA effects on number of branches, plant height and capsules on the branches. It is second to Sesim 1 parent as it recorded significance to three desirable traits in its general combining ability. Sesim2//5181 and 4036-1-10-2 did not show any desirable trait as there was no significant level of difference in its contribution. They are therefore not desirable parents to be used in the crossing programme to improve yield and yield components but could be used for improving some other traits such as pest resistance. Ajimo A1-6//7029 had highly significant negative effect (P<0.001) on days to flowering and highly positive significant (P<0.001) effect on capsule length and width. Significant negative effect on days to flowering by Ajimo A1-6//7029 implies that it can be used as a parent to reduce days to flowering and maturity in progenies to adapt to effect of climate change.

Table 2. General combining ability effects on various yield components in sesame.

Parents	Days to Flowering	Plant height (cm)	Height of first branch (cm)	Height 1st capsules (cm)	Branches / plant	Capsules on main stem	Capsules on branches	Length of capsule zone (cm)	Capsule length	Capsule width	Yield/ plant	Score
Sesim 1	1.44***	5.09*	-6.49***	-3.65*	0.81***	0.69ns	7.57*	6.26*	0.10*	0.07***	0.39ns	8
Sesim 2	2.33***	3.59ns	5.02*	9.45***	0.42*	-1.36ns	3.77ns	-3.88ns	-0.08*	0.02ns	0.21ns	3
Ajimo A1-5	1.66***	5.26*	0.57ns	4.85**	0.58**	0.12ns	7.46*	1.36ns	0.00ns	-0.02ns	0.30ns	4
Sesim2//5181	1.72***	-0.25ns	8.52***	8.61***	-0.34ns	-3.21**	-6.51*	-10.71***	-0.03ns	0.03ns	0.26ns	2
Ajimo A1-6//7029	-1.52***	-3.44ns	0.99ns	-0.56ns	-0.40*	-3.07*	-0.04ns	-2.94ns	0.18***	0.07***	0.47ns	3
SPS1438-1-6-4	-4.48***	-3.53ns	-7.28***	-11.83***	-0.86***	-2.69*	-6.34*	9.01**	0.16***	0.04*	0.45ns	6
4036-1-10-2	0.16ns	-1.31ns	0.86ns	-1.88ns	-0.25ns	1.53ns	-4.96ns	0.73ns	0.00ns	-0.01ns	0.29ns	0
Renner 1-3-1-1	-1.29**	-5.80*ns	-2.96ns	-5.61**	0.11ns	1.68ns	0.35ns	-0.47ns	-0.29***	-0.02ns	-0.29ns	1
SE_{GCA}	0.37	2.49	1.97	1.78	0.19	1.20	2.91	2.68	0.05	0.02	0.49	

Positive effect on capsule length and width would mean some increase in the number of seeds per capsule which indirectly contributes to yield. SPS1438-1-6-4 had highly negative significant (P<0.001) effect on days to flowering thus reducing flowering time and good to escape drought and other abiotic stresses. It had high significant (P<0.001) negative effect on height of first branch and first capsule and positive significant (P<0.5) effect on the length of the capsule fruiting zone. This is desirable since the number of capsules would be increased through low capsule setting and increased capsule length and thus increased yield. Renner 1-3-1-1 had high negative significant (P<0.01) effect on days to 50% flowering, height of first capsule, plant height and capsule length.

Sesim 1 was the best parent with the highest number of significant GCA effects followed by SPS1438-1-6-4. Thirugnana et al. (2006) reported significant results of general combining ability for all the characters in sesame for the experiment they evaluated under both normal and flood environments except 1000 seed weight under normal conditions. Significant results of GCA suggest the role of additive genetic effects in the inheritance of these characters. Ravindran and Raghinam (1996) reported predominance of additive gene effects on traits like branches per plant, days to 50% flowering, 1000 seed weight and height to first capsule. Early flowering could be a desirable selection criterion if the reproductive period was long enough to increase productivity or if the shorter time to flowering resulted in a concomitant decrease in time to maturity without decreasing the yield to a significant level or if it helps escape the terminal drought. Since none of the parents showed desirable combining ability for majority of the component characters, multiple crosses involving more than two parents would be appropriate techniques to be employed in the development of hybrids and or selection of superior recombinants in the segregating generation (Banerjee and Kole, 2009).

Specific combining ability

The results of specific combining ability (SCA) are presented in Table 3. Ajimo A1-5 x Renner 1-3-1-1 had positive significant (P<0.05) effect on the height of first branch (10.83) which is not desirable because few capsules would then be set on the main stem thus resulting in less yield per plant and more so if few or no branches are developed. Meanwhile, Sesim 2//5181 x Renner 1-3-1-1 had negative significant (P<0.05) SCA effect (-9.48) which is desirable as the capsules start developing from lower part of the main stem and therefore more capsules are formed on the plant. Sesim 1 x Sesim 2, Sesim1 x Ajimo A1-5 and Sesim 2//5181 x SPS 1438-1-6-4 showed positive significant (P<0.05) effect on yield per plant. Sesim 1 x Sesim 2 had positive significant (P<0.05) effect on days to flowering, number of branches per plant and number of capsules on branches. None of the hybrids exhibited superior SCA effects for all the characters indicating that there was no any cross that was superior for all the traits. Ravindran and Raghinam (1996) recorded seeds per capsule and capsule length to exhibit a preponderance of non-additive gene effects. Non-additive gene effects are non-fixable and have a consequence in slowing down selection progress but could be good for hybrid

Table 3. Estimates of specific combining ability effects on yield and yield components in sesame.

S/N	Cross	Days to flowering	Maturity	Height 1st branch (cm)	Height 1st capsules (cm)	Plant height (cm)	Branches/ plant	Capsules on main stem	Capsules on branches	Capsules zone (cm)	Capsule width	Capsule length	Yield/ plant (gm)
1	4036-1-10-2 x Renner 1-3-1-1	-0.19	0.71	1.54	5.94	5.33	-0.06	-0.08	0.02	-1.58	0.02	-0.08	0.41
2	Ajimo A1-5 x 4 036-1-10-2	-0.13	-.0.03	-2.65	0.48	-0.02	0.134	-0.85	3.57	1.25	-0.03	-0.00	-0.5
3	Ajimo A1-5 x Ajimo A1-6//7029	-0.12	0.21	-3.79	-3.52	1.11	-0.38	1.41	7.99	4.26	0.02	-0.01	-1.28
4	Ajimo A1-5 x Renner 1-3-1-1	-0.02	0.01	10.83*	7.2	2.46	-0.22	-1.34	-2.06	-3.88	-0.01	0.06	0.8
5	Ajimo A1-5 x Sesim2//5181	-0.02	-0.63	4.33	1.32	-6.09	-0.78	-4.45	-14.88*	-5.3	-0.04	-0.00	-2.62*
6	Ajimo A1-5 x SPS1438-1-6-4	1.18	1.2	-3.66	1.08	7.89	0.41	2.31	9.63	6.64	0.05	0.01	2.01
7	Ajimo A1-6//7029 x 4036-1-10-2	0.38	0.24	-2.08	-6.46	-3.63	0.11	0.67	-0.59	3.89	-0.02	0.02	0.76
8	Ajimo A1-6//7029 x Renner 1-3-1-1	2.17**	-0.38	-1.26	-0.06	-0.14	0.09	0.19	-4.56	1.09	0.03	0.06	0.23
9	Ajimo A1-6//7029 x SPS1438-1-6-4	-0.98	-0.20	5.71	3.83	-6.41	0.39	-5.5*	-5.2	-11.06	-0.05	-0.1	-1.43
10	Sesim 2 x 4036-1-10-2	0.37	-0.77	3.22	3.2	-4.66	-0.04	-3.04	-8.06	-7.51	0.00	0.11	-0.66
11	Sesim 2 x Ajimo A1-5	0.37	-0.80	2.51	-2.2	-4.92	0.13	-1.63	-10.15	-6.13	-0.03	-0.02	-0.70
12	Sesim 2 x Ajimo A1-6//7029	-1.46	0.14	1.41	6.21	9.07	-0.22	3.23	2.36	1.83	0.02	0.04	1.73
13	Sesim 2 x Renner 1-3-1-1	-0.69	0.28	-4.63	-6.4	-3.17	-0.06	0.48	3.96	2.36	-0.01	-0.22*	-1.39
14	Sesim 2 x Sesim2//5181	0.30	0.60	8.21	6.71	5.58	0.05	1.03	2.82	-2.39	-0.04	-0.09	-0.67
15	Sesim 2 xSPS1438-1-6-4	-0.50	0.13	-2.97	-5.18	-8.44	-0.76	-1.2	-11.67	1.21	0.05	-0.01	-0.71
16	Sesim 1 x 4036-1-10-2	0.74	0.57	7.07	2.64	-4.56	-0.43	-2.08	-8.86	-5.65	-0.02	-0.13	-2.02
17	Sesim 1 x Ajimo A1-5	-1.26	0.05	-7.57	-4.36	-0.42	0.70	4.55	5.9	3.16	0.04	-0.03	2.28*
18	Sesim 1 x Sesim2	1.59*	0.41	-7.76	-2.35	6.54	0.90*	1.13	20.75**	10.63	0.02	0.18	2.42*
19	Sesim 1 x Sesim2//5181	-0.13	-0.43	5.39	0.15	-4.22	-0.34	-4.68	-6.31	-10.54	0.04	-0.14	-1.26
20	Sesim 1 x SPS1438-1-6-4	-0.94	-0.61	2.87	3.92	2.66	-0.82*	1.08	-11.47	2.41	-0.08	0.11	-1.40
21	Sesim2//5181 x 4036-1-10-2	-0.86	0.39	-5.29	-2.28	6.48	0.05	3.14	7.54	10.66	0.02	0.06	1.53
22	Sesim2//5181 x Renner 1-3-1-1	-1.07	-0.56	-9.48*	-6.22	-4.73	0.36	2.33	0.57	3.86	0.01	0.19	0.97
23	Sesim2//5181 x SPS1438-1-6-4	1.78*	0.63	-3.16	0.32	2.99	0.66	2.64	10.26	3.71	0.03	-0.03	2.07*
24	SPS1438-1-6-4 x 4036-1-10-2	-0.33	-1.10	-1.81	-3.52	1.06	0.24	2.24	6.38	-1.06	0.04	0.10	0.49
25	SPS1438-1-6-4 xRenner 1-3-1-1	-0.216	-0.06	3.01	-0.46	0.25	-0.12	-1.57	2.07	-1.86	-0.04	0.00	-1.02
	SE_{SCA}	0.78	0.61	4.13	3.72	5.21	0.40	2.51	6.09	5.60	0.04	0.10	1.21

production. Yield per plant exhibited almost equal importance of both additive and non-additive gene effects. Thirugnana et al. (2006) recorded higher SCA variance than GCA variance for number of branches, number of capsules, number of seeds per capsule, 1000 seed weight and seed yield under normal conditions as compared to analysis under floods. They also reported that the magnitude of GCA variance was higher than that of SCA variance for days to 50% flowering and plant height under normal conditions, indicating the predominance of additive and additive x additive type of gene action. The SCA is considered to be the best criterion for selection of superior hybrids (Ahmed and Adam, 2014). From the results of this study, it could be concluded that both additive and non-additive gene action were important for improving seed yield in sesame.

Table 4. General combining ability and specific combining ability, ratios of additive (σ^2gca) to non-additive (σ^2sca) gene effects on yield and its components.

Variables	Days to flowering	Maturity	Height 1st branch (cm)	Plant height (cm)	Branches/ plant	Capsules on main stem	Capsules on branches	Capsule zone (cm)	Capsule length	Capsule width	Yield/plant (gm)
GCA	31.93***	0.65ns	144.91	105.7*	1.71***	22.42*	178.01**	191.17***	0.12***	0.00*	3.01ns
SCA	1.25	0.44	39.58	34.89ns	0.28ns	9.68ns	99.96*	45.71ns	0.01ns	0.01ns	2.97*
Baker's ratio	0.96	1.0	0.70	1.0	0.89	0.82	0.44	0.96	0.91	1.0	0.23
CGD_{bs}	0.92	0.06	0.69	0.356	0.69	0.36	0.62	0.51	0.78	0.29	0.57
CGD_{ns}	0.88	0.06	0.48	0.356	0.61	0.30	0.27	0.49	0.71	0.29	0.13

Significant amount of dominance (non-additive) variance for any character is generally a prerequisite for exploitation of heterosis.

Effects of GCA and SCA were analyzed (Table 4). GCA effects were found to be highly significant (P<0.001) for days to flowering, branches per plant, length of capsule zone and significant (P<0.05) for plant height, capsules on main stem and capsule width, indicating preponderance of additive gene action. Praveenkumar et al. (2012) revealed that additive gene action was pre-dominant for plant height, number of secondary branches per plant, number of capsules per plant, 1000 seed weight and number of seeds per capsule. SCA showed significant effect on only capsules on branches and yield per plant, thus indicating major action of non-additive gene action for these traits. Murty (1975) reported that general combining ability variance was larger than specific combining ability variance for all the characters except of oil, indicating the predominance of additive gene action. Saravanan and Nadarajan (2003) reported that the GCA variance was greater than the SCA variance for eight traits including days to 50% flowering, plant height, number of primary branches per plant and phyllody incidence, indicating preponderance of additive gene action for those traits, while the SCA variance was greater than GCA variance for number of capsules per plant, number of seeds per capsule, single plant yield, indicating predominance of dominant gene action for these traits. Mothilal and Manoharan (2005) reported that non-additive gene action was involved in the expression of characters viz. number of capsules on branches, 1000 seed weight and seed yield per plant. They observed over dominance for number of branches, number of seeds per capsule and seed yield per plant. Thirugnana et al. (2006) reported the magnitude of GCA variance was higher than that of SCA variance for days to 50% flowering and plant height under normal conditions. Praveenkumar et al. (2012) also recorded that the proportion of GCA variances were higher as compared to SCA. Aladji et al. (2014) showed that the values of GCA/SCA ratios had SCA variance higher than GCA variance component except for number of seeds per capsule and days to maturity. They recorded that the SCA variance was more than GCA variance, indicating the role of non-additive gene action for the inheritance of date to flowering, duration of maturation, plant height, number of branches, number of capsules per plant and capsule length.

As recorded in Table 4, additive gene action controlled days to flowering, plant height, branches per plant, capsules on main stem, capsules on branches, length of capsule zone, capsule length and capsule width. Non-additive gene action controlled capsules on branches and yield per plant only. For selected traits controlled by non-additive gene action, it would therefore be desirable to maintain a certain degree of heterozygosity to exploit the additive gene effects and recurrent selection involving crossing desirable segregannts alternated with selection in order to increase the magnitude of additive genetic variance and at the same time to maintain heterozygosity.

Pedigree method is applied for selection of desirable traits that are controlled by additive gene action since they are fixable. Selection of progenies for these traits is done at early stages of segregation to produce pure lines. Meanwhile, traits controlled by non-additive (dominance and epistasis) gene action are good for hybrid production where dominance is important. The selection of these traits are normally delayed to later stages of segregation after some selfing in order to reduce the number of genes that heterozygous and genes would then be fixed to homozygosity.

Baker's ratio was high for most of the traits (Table 4) except yield per plant (0.233) indicating that the parents contributed much to the progenies. Coefficient of genetic determination in

broad sense (CGDbs) was higher for most traits except days to maturity, plant height, capsules on main stem and capsule width (0.29). Coefficient of genetic determination in narrow sense (CGDns) was high for most traits except days to maturity (0.06), capsules on main stem, capsules on branches, capsule width and yield per plant, implying that these traits have high heritability and the parents contributed highly to those traits. Where CGDns is low, it indicates non-additive (dominance or epistasis) behavior playing an important role. This also suggests that the progeny could not be predicted on the GCA values of the parents.

Conclusion

These results suggested that both additive and non-additive gene actions played a greater role in these traits. Sesim 1 had the highest number of desirable traits scored for GCA effects and could be considered as a parent for crossing to produce desirable progeny. Sesim 1 x Sesim 2, Sesim 1 x Ajimo A1-5 and Sesim2//5181 x SPS1438-1-6-4 recorded positive significant SCA effects for yield per plant and therefore they could be recommended for hybrid seed production.

Conflict of Interests

The authors have not declared any conflict of interests.

ACKNOWLEDGEMENTS

Special thanks to RUFORUM for funding this project. Awio Brunno is also acknowledged for supporting in the statistical analysis.

REFERENCES

Aladji-Abatchoua MMI, Noubissie- Tchiangam JB, Njintang-Yanou N (2014). Genetic analysis of seed yield component in sesame (*Sesamum indicum* L.) at Mora (Cameroon). Sch. Acad. J. Biosci. 2(5):318-325.

Ahmed SBM, Adam SI (2014). Combining ability for yield and yield components in six parents and their 15 F1 hybrids of sesame (*Sesamum indicum* L.) in half diallel mating design. J. Plant Breed. Crop Sci. 6(12):179-184.

Anyanga WO, Obong YY (2001). Sesame (*Sesamum indicum* L.) In: J.K. Mukiibi (ed.) Agriculture in Uganda. Crops 11:97-102.

Ashri A (1989). Sesame breeding objectives and approaches. In: Oil Crops: Sunflower, Linseed and Sesame, Ed. Omran, A., IDRC-MR205E, IDRC, Ottawa.

Auckland AK (1970). Sesame Improvement in East Africa. In: Crop Improvement in East Africa, Chapter Ed. C.L.A. Leakey. 6:157-177.

Baker RJ (1978). Issues in diallel analysis. Crop Sci. 18(4):533-536.

Banerjee PP, Kole PC (2009). Combining ability analysis for seed yield and some of its component characters in sesame (*Sesamum indicum* L.). Int. J. Plant Breed. Genet. 3(1):11-21.

Bedigian D, Seigher DS, Harlan JR (1985). Sesamin, sesamolin and the origin of sesame. Biochem. Syst. Ecol. 13(2):133-139.

Brigham RD (1985). Status of sesame research and production in Texas and the USA. Sesame and safflower: status and potentials. FAO Plant Prod. Prot. 66:73-74.

Coulman KD, Liu Z, Hum WQ, Michaelides J, Thompson LU (2005). Whole sesame is as rich a source of mammalian lignin precursors as whole flaxseed. Nutr. Cancer 52:156-165.

FAOSTAT (2009). Food and Agricultural Organization of the United Nations.

FAOSTAT (2015). Food and Agricultural Organization of the United Nations.

Griffings B (1956). Concept of general and specific combining ability in relation to diallel crossing system. Austr. J. Biol. Sci. 9:463-493.

Johnson LA, Suleiman TM, Lusas EW (1979). Sesame protein, a review and prospectus. J. Am. Oil Chem. Soc. 56(3):463-468.

Moazzami AA, Kamal-Eldin A (2006). Sesame seed is arich source of dietary lignans. J. Am. Oil Chem. Soc. 8:719-723.

Mothilal C, Manoharan V (2005). Diallel analysis for the estimation of genetic parameters in Sesame (*Sesamum indicum* L.). Agric. Sci. Digest 25(2):133-135.

Muhammad SH, Sedeck FS (2015). Combining ability and heterosis estimates in sesame. World Appl. Sci. J. 33(5):690-698.

Murty DS (1975). Hetrosis, combining ability and reciprocal effects for agronomic and chemical characters in sesame. Theor. Appl. Genet. 45:294-299.

Payne RW, Harding SH, Murray DA, Soutar DM, Baird DB, Glaser AR, Thompson R, Webster R (2010). The Guide to Genstat Release 14, Part 2: Statistics. Hemel Hempstead: VSN International, U.K.

Praveenkumar KM, Nadaj HL, Patil RK and Deshpande SK (2012). Combining ability and gene action studies in inter-mutant hybrids of sesame (*Sesamum indicum* L.). Karnataka J. Agric. Sci. 25(1):1-4.

Ravindran RG, Raghinam AD (1996). Combining ability analysis on sesame. Sesame and Safflower Newslett. 11:70-75.

Saravanan S, Nadarajan N (2003). Combining ability studies in sesame. Sesame and Safflower Newslett. 18:20-24.

Singh RK, Chaudhary BD (1985). Biometrical Methods in Quantitative Genetic Analysis. Kalyani Publishers, New Delhi, P 300.

Solanki ZS, Gupta D (2003). Inheritance studies for seed yield in sesame. Sesame and Safflower Newslett. 18:25-28.

Thirugnana-Kuma S, Srinivasan R, Kubenthiran K, Eswaran R, Kumar CPS (2006). Studies on combining ability and heterosis for yield and yield component characters under normal and temporarily flooded conditions in sesame (*Sesamum indicum* L.). Sesame and Safflower Newslett. 21.

academicJournals

Vol. 8(11), pp. 257-272, November 2016
DOI: 10.5897/JPBCS2016.0614
Article Number: B93F9CF61326
ISSN 2006-9758

http://www.academicjournals.org/JPBCS

Journal of Plant Breeding and Crop Science

Full Length Research Paper

A method of estimating broad-sense heritability for quantitative traits in the type 2 modified augmented design

Frank M. You[1]*, Gaofeng Jia[1,2], Sylvie Cloutier[3], Helen M. Booker[2], Scott D. Duguid[1] and Khalid Y. Rashid[1]

[1]Morden Research and Development Centre, Agriculture and Agri-Food Canada, Morden, MB, R6M 1Y5, Canada.
[2]Crop Development Centre, Department of Plant Sciences, University of Saskatchewan, 51 Campus Drive, Saskatoon, SK, S7N 5A8, Canada.
[3]Ottawa Research and Development Centre, Agriculture and Agri-Food Canada, Ottawa, ON, K1A 0C6, Canada.

Received 20 September, 2016; Accepted 17 October, 2016

Heritability is a basic genetic parameter for quantitative traits that may determine their selection generation and intensity as well as predict their selection response and efficiency in plant breeding. Estimation of heritability varies based on experimental design. The type 2 modified augmented design (MAD2) as an unbalanced experimental design, has been proposed for evaluating numerous unreplicated test genotypes with several replicated control genotypes to adjust for soil heterogeneity. Here, we define an inter-environment correlation (r_E), that is, the mean Pearson's correlation coefficient of trait performance for test genotypes between all pairs of environments, to approximate broad-sense heritability (H^2). Computer simulation and empirical results demonstrated that r_E was consistent with H^2 estimates on a plot basis by ANOVA for non-missing data sets, and similar to those by the restricted maximum likelihood (REML)-based method for missing data sets. The r_E method was shown to generally outperform the ANOVA- and REML-based methods.

Key words: Broad-sense heritability, analysis of variance, inter-environment correlation, modified augmented design, restricted maximum likelihood, flax.

INTRODUCTION

Heritability is a basic genetic parameter for quantitative traits that may determine their selection generation and intensity as well as predict their selection response and efficiency in plant breeding. Heritability estimation varies depending on the experimental design (Holland et al., 2003). The modified augment design (MAD) has been proposed for square plots (Type 1) (Lin and Poushinsky, 1983) and specifically for rectangular plots (Type 2 - MAD2) (Lin and Poushinsky, 1985) for field evaluation of a large number of breeding lines and used in many crops such as flax (Soto-Cerda et al., 2014a; Soto-Cerda et al., 2014b; Kumar et al., 2015), wheat (Golparvar et al.,

*Corresponding author. E-mail: collinsagho@yahoo.com.

2015), potato (Schaalje et al., 1987), soybean (Lin and Voldeng, 1989), and barley (May et al., 1989; May and Kozub, 1995). This design can accommodate a large number of unreplicated test genotypes with typically only three control genotypes for error control (You et al., 2013).

Our earlier study (You et al., 2013) demonstrates that soil heterogeneity can be sufficiently adjusted for traits in MAD2 trials. The adjusted observations for test and control genotypes are expected to exclude the effect of soil heterogeneity; thus, the variation among replications of each control genotypes should be solely caused by random errors. A method based on analysis of variance (ANOVA) to approximately estimate broad-sense heritability (H^2) for this design has been developed (You et al., 2016b), in which genetic variance is calculated based on the total phenotypic variance estimated from test genotypes minus the error variance estimated from control genotypes. Because H^2 is always overestimated in single trials, joint analysis over multiple environments for heritability estimation was proposed in which the error variance is jointly estimated using the three replicated control genotypes (You et al., 2016b). However, a caveat to this estimation method is the requirement for the same control genotypes to be used in all environments or trials, and this limits its potential use for joint analysis of data from multi-environment trials with different control genotypes or experimental designs.

There are two units for measurement of phenotypic variances in heritability estimation: on a plot basis and on an entry (or genotype)-mean basis. Estimates of phenotypic variance on an entry-mean basis are always larger than those on a plot basis because the error and interaction variance components are divided by the corresponding numbers of observations (You et al., 2016b). From the perspective of breeding, usually our interest is the heritability estimated from a set of trials. Thus, heritability is usually estimated on an entry-mean basis, that is, the genotype × environment variance is divided by the number of environments. However, we found from empirical results that H^2 on an entry-mean basis was overestimated in multi-environment MAD2 trials (You et al., 2016a); thus, H^2 estimates on a plot basis were put forward and used in MAD2 trials of flax (You et al., 2016a) and in this study.

Herein, we define an inter-environment correlation (r_E) to approximately estimate H^2 in multi-environment MAD2 trials and which does not rely on control genotype data. H^2 estimates from two traditional methods, ANOVA and restricted maximum likelihood (REML), are compared with r_E using computer-simulated and empirical data sets with and without missing data.

MATERIALS AND METHODS

Simulation data

For multi-environment MAD2 trials, a linear model for an adjusted observation (y_{ij}) of *g* test genotypes at *e* environments can be written as:

$$y_{ij} = \mu + G_i + E_j + (GE)_{ij} + \varepsilon_{ij}, (i = 1, 2, \ldots, g; j = 1, 2, \ldots, e), \quad (1)$$

Where $y_{ij} \sim N(\mu, \sigma_P^2)$, $G_i \sim N(0, \sigma_G^2)$, $E_j \sim N(0, \sigma_E^2)$, $(GE)_{ij} \sim N(0, \sigma_{GE}^2)$, and $\varepsilon_{ij} \sim N(0, \sigma_e^2)$. σ_P^2, σ_G^2, σ_E^2, σ_{GE}^2, and σ_e^2 are phenotypic, genetic (G), environment (E), genotype-by-environment (G × E) interactions, and error variances, respectively. σ_e^2 is jointly estimated based on *e* environments with *t* replicated control genotypes in each trial (You et al., 2016b). Typically, *t* is three in MAD2 trials.

For a typical multi-environment MAD2 trial, the data set includes observations for *g* unreplicated test genotypes and three control genotypes at each of *e* environments (representing individual trials in multiple years or/and locations). For the three control genotypes, one main plot control has *m* replications and each of two subplot controls are replicated *n* times. The values of *g*, *m*, *n* and *e* depend on the design. Thus, a total of (*g* + *m* + 2**n*)**e* data points will be generated from a typical multi-environment MAD2 trial. For example, an individual MAD2 trial with a grid of 10 rows by 10 columns contained 100 whole plots with each whole plot being split into 5 subplots, resulting in a total of 500 subplots. The main plot control was placed in the center subplot of each whole plot and so *m* = 100. Two additional subplot controls were randomly assigned to two subplots of 5 randomly selected whole plots and thus *n* = 5 or 2*n* = 10. Test genotypes were accommodated to the remaining 390 subplots (*g* = 390), and the ratio of *m* to *g* and that of *m* to all subplots were 25.6 and 20%, respectively (see Figure 1 in You et al., 2013). To simplify the process, all simulations comprised the same percentage of test genotypes for the main plot control replicates (*m* = 0.25 * *g*) and the same number of subplot control replicates (*n* = 5).

To assess the effect of the number of environments (*e*), number of test genotypes (*g*), and theoretical heritability of a trait on H^2 estimation, simulation data sets for a total of 180 parameter combinations of H^2 (0.1 to 0.9 with an interval of 0.1), *e* (2, 4, 6, 8, and 10), and *g* (50, 100, 200, and 300) were generated. For each parameter combination, 500 simulations were replicated, which was sufficient to represent the sampling distribution of a parameter combination based on results of different simulation runs.

Data sets with 5, 10, 15 and 20% missing data for test genotypes were constructed from each complete data set generated for parameter combinations. Missing data were distributed in all environments at random. The R *sample* function was applied to all simulated observations from *e* environments to randomly assign NULL as missing values. Consequently, a total of 450,000 data sets were generated for analysis.

According to empirical results, plot-based H^2 estimates were more accurate than entry-mean-based estimates in MAD2 trials (You et al., 2016a). Accordingly, heritability estimates on a plot basis were used in this study. Given H^2, σ_{GE}^2, and σ_e^2 for a trait, σ_G^2 can be estimated as $\sigma_G^2 = (\sigma_{GE}^2 + \sigma_e^2)H^2/(1 - H^2)$ on a plot basis. Because traits may have different σ_{GE}^2 and σ_e^2, these two variances were randomly and independently generated. In addition, the environmental variance (σ_E^2) was also randomly generated. Thus, the effects of G, E, G × E, and random error can be simulated according to Equation 1 for a multi-environment MAD2 trial. Similar simulation procedures for data generation have been described previously (You et al., 2016b). The R pipeline program for simulation is available at http://probes.pw.usda.gov/bioinformatics_tools/MADPipeline/index.html.

Empirical data

Three flax biparental genetic populations, namely 243 F_6-derived

recombinant inbred lines (RILs) generated from a cross between CDC Bethune and Macbeth (BM), 90 F_6-derived RILs from a cross between E1747 and Viking (EV), and 78 F_1-derived doubled haploid lines from a cross between SP2047 and UGG5-5 (SU), plus a core collection (CC) of 391 accessions, were field evaluated at two locations in Canada (Morden, Manitoba and Kernen Farm near Saskatoon, Saskatchewan) from 2009 to 2012. The same MAD2 design was employed with the same population in all environments but the designs differed across populations. The experimental designs and phenotyping for the 22 traits in CC, 14 in BM, 19 in EV, and 11 in SU over six to eight environments have been previously described (Cloutier et al., 2010; Soto-Cerda et al., 2014b; Kumar et al., 2015; You et al., 2016b). The adjusted observations in each environment were used for H^2 estimation and evaluation.

Traditional estimation of H^2

The simulated and empirical adjusted observations were used to estimate H^2 on a plot basis (You et al., 2016a; You et al., 2016b). H^2 was estimated as $\hat{H}^2 = \hat{\sigma}_G^2 /(\hat{\sigma}_G^2 + \hat{\sigma}_{GE}^2 + \hat{\sigma}_e^2)$, where $\hat{\sigma}_G^2$, $\hat{\sigma}_{GE}^2$, and $\hat{\sigma}_e^2$ are the genetic, G × E, and error variance, respectively. These variance components were estimated using the method of moments based on both ANOVA (You et al., 2016b) and REML (Holland, 2006; Piepho and Möhring, 2011). The R (https://www.r-project.org/) package *lme4* (Bates et al., 2015) was used to calculate the REML-based variance components. A test for homogeneity of error variance across environments was performed before parameter estimations to satisfy the assumption of the model (Equation 1).

Inter-environment correlation as a H^2 statistic

Based on Equation 1, the inter-environment correlation of adjusted observations ($\boldsymbol{y}$) for two environments (E_1 and E_2) was defined as:

$$r_E(\boldsymbol{y}_1, \boldsymbol{y}_2) = \frac{COV(\boldsymbol{y}_1,\boldsymbol{y}_2)}{\sigma_{y_1}\sigma_{y_2}} = \frac{COV(\boldsymbol{\mu}+\boldsymbol{G}+\boldsymbol{E}_1+(\boldsymbol{GE})_1+\boldsymbol{\varepsilon}_1,\boldsymbol{\mu}+\boldsymbol{G}+\boldsymbol{E}_2+(\boldsymbol{GE})_2+\boldsymbol{\varepsilon}_2)}{\sigma_{y_1}\sigma_{y_2}} = \frac{\sigma_G^2}{(\sigma_G^2+\sigma_{GE}^2+\sigma_e^2)^{\frac{1}{2}}(\sigma_G^2+\sigma_{GE}^2+\sigma_e^2)^{\frac{1}{2}}} = \frac{\sigma_G^2}{\sigma_P^2} = H^2. \quad (2)$$

Here, $COV(\boldsymbol{y}_1,\boldsymbol{y}_2)$ is the covariance of adjusted observations between two environments. It is equal to σ_G^2 because G, E, and random error are hypothesized to be independent of each other and thus the covariance between them is zero, whereas $\sigma_{y_1}\sigma_{y_2} = \sigma_P^2$ because, for a single environment, the environmental variance $\sigma_E^2 = 0$. Therefore, $r_E(\boldsymbol{y}_1, \boldsymbol{y}_2)$ corresponds to an H^2estimate. To validate this definition, independent observations were simulated in two-environment MAD2 trials at given H^2 values (0.1 to 0.9 with 0.1 intervals) and for a number of test genotypes (50, 100, 200 and 300).

For a multi-environment MAD2 trial, r_E was defined as the mean inter-environment correlation coefficient of all possible pairs of environments to estimate the H^2 of a trait:

$$r_E = \sum_{i<j}^{e} r_E(\boldsymbol{y}_i, \boldsymbol{y}_j)/(\frac{e(e-1)}{2}) \ (i, j = 1, 2, \ldots, e), \quad (3)$$

Where e is the number of environments and $r_E = 0$ if $r_E < 0$.

Standard error of H^2 estimates

The delta method was implemented for MAD2 (You et al., 2016b) and used to estimate the standard error (SE) of $\hat{H}^2$ by ANOVA and REML for the empirical data sets. For simulation data, the standard deviation of 500 replicates for each parameter combination was calculated to represent the SE of three H^2 estimates ($s(\hat{H}^2)$ or $s(r_E)$).

Owing to the properties of r_E as a mean Pearson's correlation coefficient, the SE of r_E can also be approximated based on the SE formula of Pearson's correlation (Bowley, 1928; Kendall and Stuart, 1977) using:

$$SE(r_E) = \frac{1-r_E^2}{g^{1/2}}/(\frac{e(e-1)}{2})^{1/2}, \quad (4)$$

Where e is the number of environments and g is the number of test genotypes.

The bootstrap method (Efron, 1979; Xie and Mosjidis, 1997) was also used to estimate the standard error of three heritability estimates for either simulated or empirical data. Bootstrapping was performed by resampling test genotypes with replacement from the original population, and each bootstrap sample had the same size as the original population. The R *sample* function was used to generate the bootstrap samples. The standard deviation of r_E in 500 bootstrap samples was calculated to represent the standard error of r_E, denoted as $s_b(r_E)$ to distinguish from and compare with $SE(r_E)$ in Equation 4, $s(\hat{H}^2)$, or $s(r_E)$ estimated by simulation.

An R pipeline program was developed to automate the simulations and, consequently, the H^2 estimation for three heritability estimation methods. A separate R program for calculating r_E and its SE (bootstrap method) for empirical test genotype data was also created and is available at http://probes.pw.usda.gov/bioinformatics_tools/MADPipeline/index.html.

Statistical power

Statistical power (P) of three heritability estimation methods (r_E, ANOVA, and REML) was calculated for all parameter combinations to determine the minimum number of environments and test genotypes required for heritability estimation and for method comparison. The power of a trait heritability estimate is the probability of correctly rejecting the null hypothesis that heritability is zero when the true heritability of the trait is greater than zero. First, the Z score for the power of a heritability estimate was calculated using (Klein, 1974):

$$Z = Z_\alpha - E(\hat{H}^2)/\sigma_{\hat{H}^2}, \quad (5)$$

Where Z_α is the Z score of the significance level α used for a one-tailed test ($Z_{0.05} = 1.645$); $E(\hat{H}^2)$ is the expected value of the heritability, here the given or parametric heritability value; and $\sigma_{\hat{H}^2}$ is the SE of the expected $\hat{H}^2$ value, estimated by the standard deviation of 500 simulated samples. The one-tailed test was adopted to test whether a heritability estimate is greater than 0 because heritability should be always equal to or greater than zero. The statistical power was defined as the area under the standard normal curve from Z to plus infinity, and can be calculated using an R function:

$$P = 1 - pnorm(-Z), \quad (6)$$

Where *pnorm* is the R function for calculating the cumulative density of the normal distribution.

RESULTS

Genetic and phenotypic variance estimation in the definition of r_E

Simulation results showed a perfectly consistent

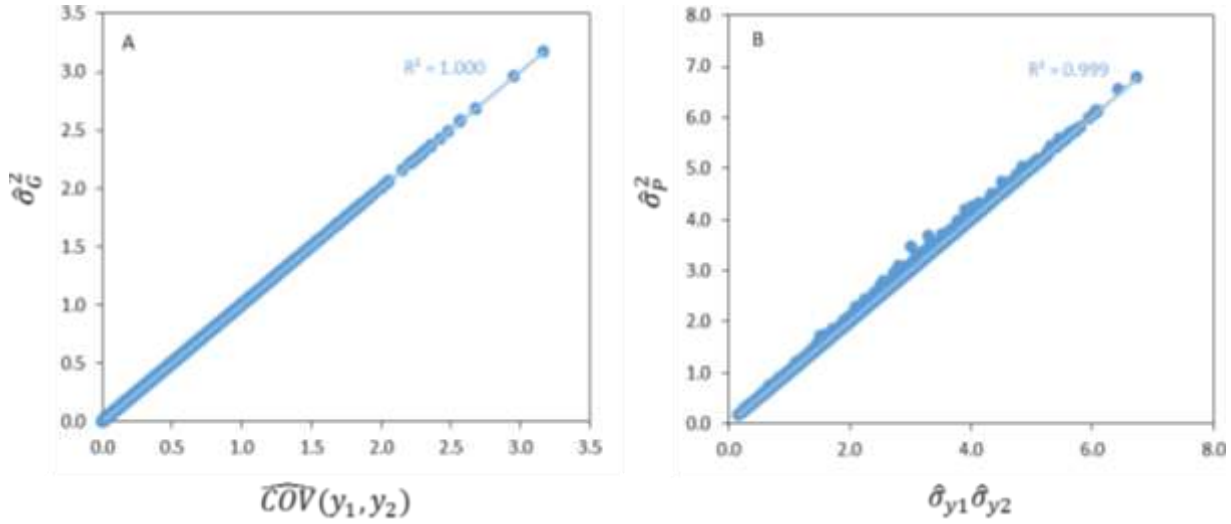


Figure 1. Covariances ($\widehat{COV}(y_1, y_2)$) of trait performance between two environments in relation to genetic variances (σ_G^2) estimated by ANOVA (A), and the product of standard deviations of trait performance from two environments ($\hat{\sigma}_{y1}\hat{\sigma}_{y2}$) in relation to phenotypic variances (σ_P^2) on a plot basis estimated by ANOVA (B).

relationship between $\widehat{COV}(\boldsymbol{y}_1, \boldsymbol{y}_2)$ and $\hat{\sigma}_G^2$ (r = 1.000) (Figure 1A) and between $\hat{\sigma}_{\boldsymbol{y1}}\hat{\sigma}_{\boldsymbol{y2}}$ and $\hat{\sigma}_P^2$ (r = 0.999) (Figure 1B) for the ANOVA-based method, confirming the definition in Equation 2.

H^2 estimates for simulated data sets without missing data

H^2 estimates by r_E, ANOVA ($\widehat{H}^2$(A)), and REML ($\widehat{H}^2$(R)) were calculated for all complete simulated data sets. r_E was perfectly correlated with the given H^2 values (r = 1.000) independent of the number of environments (Figure 2A) or test genotypes (Figure 2C). However, standard errors of r_E ($s(r_E)$) declined with increasing number of environments or test genotypes (Figure 2B and D). For a trait with H^2 equal to or greater than 0.5, the heritability can be correctly estimated at as few as two environments with 50 test genotypes at a statistical power of > 0.999 (Table 1). For traits with low heritability, a larger number of environments and test genotypes contribute positively to more reliable H^2 estimates through increasing statistical power. For an H^2 estimate of 0.2, a statistical power in excess of 0.95 can be achieved from data sets with 300 test genotypes at two environments, 100 test genotypes at four environments, or 50 test genotypes at ≥ six environments when there was no missing data (Table 1).

r_E was consistent with $\widehat{H}^2$(A) for different numbers of environments and test genotypes (Table 2) (r = 1.000). r_E also highly correlated with $\widehat{H}^2$(R) (r > 0.993) (Table 2). High similarity between $\widehat{H}^2$(A) and $\widehat{H}^2$(R) (r > 0.993) was observed (Table 2).

H^2 estimates for simulated data sets with missing data

H^2 estimates by r_E, ANOVA, and REML were calculated for all simulated data sets with missing data of 5 to 25%. The SEs of H^2 estimates by all three methods increased with the rate of missing data. Among the three estimation methods, REML yielded smaller SEs than the r_E- and ANOVA-based methods for both non-missing and missing data sets (Table 3). $\widehat{H}^2$(A) had SEs consistent with r_E for non-missing data sets (R^2 = 1.000) but larger SEs for missing data sets (Figure 3A); r_E had more consistent REs with $\widehat{H}^2$(R) (R^2 = 0.974) for missing data sets (Figure 3B).

The statistical power of H^2 estimates was markedly affected by missing data rates (Tables 3 and 4). Owing to the small SEs, H^2 estimates by REML had higher power than those by r_E, which was also higher than those by ANOVA. The power of r_E and $\widehat{H}^2$(R) was relatively less affected by missing data. For an H^2 estimate of 0.2 in data sets with a missing data rate of ≤5%, a statistical power over 0.95 can be achieved from data sets with 300 test genotypes at two environments, 100 test genotypes at ≤ four environments, or 50 test genotypes at ≤ eight environments for all three estimation methods (Table 4).

Estimates of r_E and $\widehat{H}^2$(R) were less affected by missing data, remaining largely constant with increasing missing data rate at different given H^2 values (Table 3). r_E was consistent with $\widehat{H}^2$ (R) for different numbers of environments and test genotypes (Table 5) (r = 0.995-0.999) but less correlated with $\widehat{H}^2$(A) (r = 0.955-0.996) (Table 5). Bias of $\widehat{H}^2$ (A) and $\widehat{H}^2$ (R) from their true heritability values was observed. At a given H^2 ≤ 0.3,

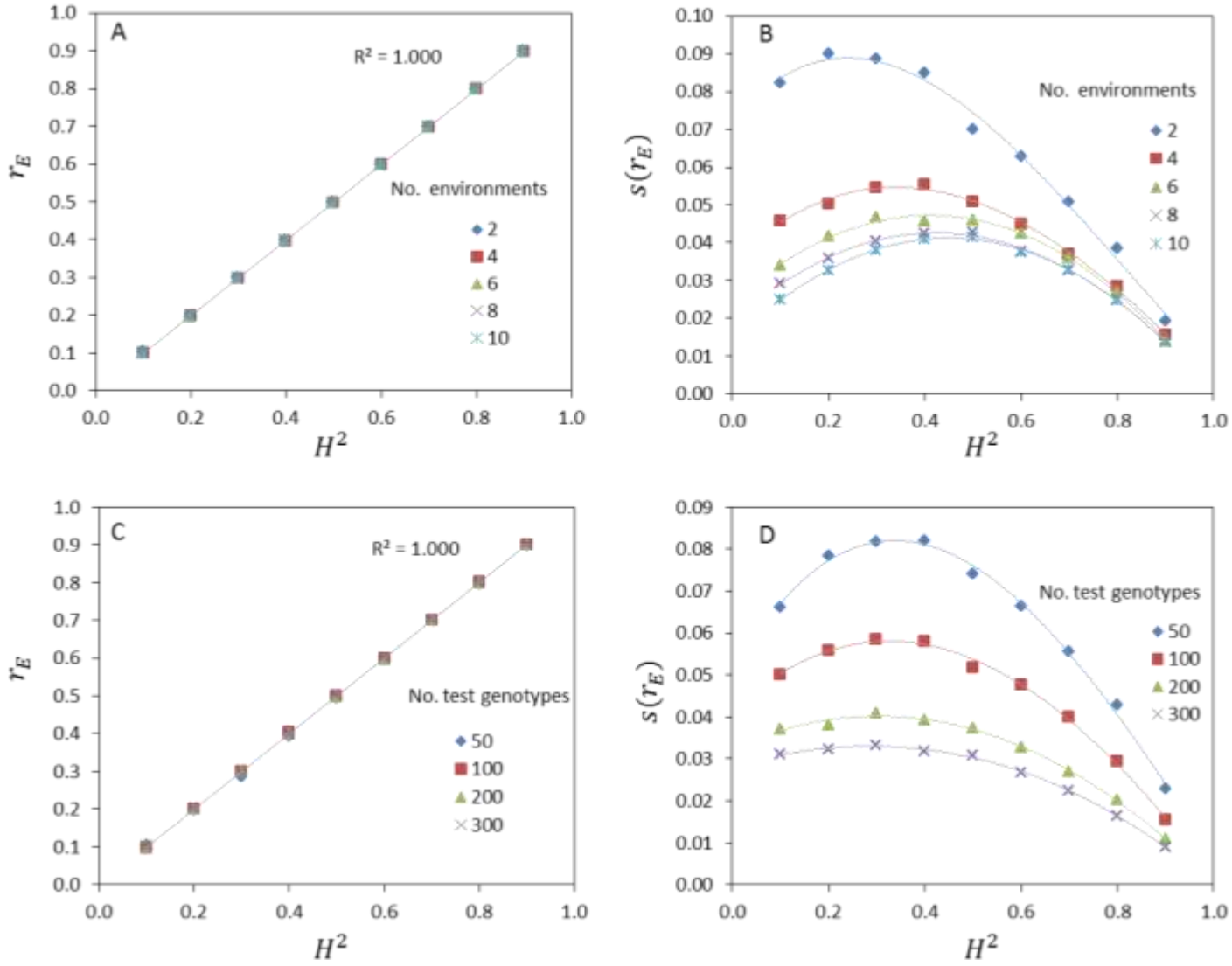


Figure 2. Inter-environment correlations (r_E) and their standard deviations ($s(r_E)$) in terms of heritability (H^2) for different numbers of environments (A and B) and test genotypes (C and D), respectively.

$\widehat{H}^2$(A) was overestimated (Table 3). However, at a given H^2 ≥ 0.4 for $\widehat{H}^2$(A) and ≥ 0.2 for $\widehat{H}^2$(R), $\widehat{H}^2$ values were slightly underestimated (Table 3). A negligible deviation of r_E estimates from their true values was also observed when given H^2 ≥ 0.3 (Table 3).

H^2 estimates for empirical data

As a case study, r_E, $\widehat{H}^2$(A), and $\widehat{H}^2$(R) and their SEs were estimated for 22 traits in the CC, 14 in the BM population, 19 in the EV population, and 11 in the SU population (Table S1). These traits varied greatly in estimated heritability (0.00-0.94), were phenotyped in a different number of environments ranging from two to eight, and evaluated in different populations, that is, one natural and three biparental populations of varying size (Table 6). Here, heritability of some traits was estimated to be 0 due to large experimental errors in data and the ANOVA based model used which resulted in negative genetic variances. The negative genetic variance was treated as 0 in calculation although the real genetic variance should be greater than 0. r_E was highly correlated to both $\widehat{H}^2$(A) (r = 0.948-0.998) and $\widehat{H}^2$ (RP) (r = 0.974-0.998) in individual populations (Table 6). Similar relationships among r_E, $\widehat{H}^2$(A), and $\widehat{H}^2$(R) were observed in different populations despite different numbers of test genotypes (Table 6). Figure 4 depicts the similar overall relationship of r_E with $\widehat{H}^2$ (A) and $\widehat{H}^2$ (R) for all 66 data points. In addition, a strong correlation between $\widehat{H}^2$(A) and $\widehat{H}^2$(R) (r = 0.995-1.000), similar to the results in the simulation data sets, was also observed in the different populations

Table 1. Statistical power (*P*) of broad-sense heritability estimates (H^2) by inter-environment correlation (r_E), ANOVA and REML for α = 0.05 (one-tailed test) as a function of number of environments and test genotypes.

No. environments	No. test genotypes	Method	*P* for H^2				
			0.1	0.2	0.3	0.4	≥ 0.5
2	50	r_E	0.765	0.533	0.769	0.940	≃1
		ANOVA	0.715	0.533	0.786	0.935	≃1
		REML	0.699	0.544	0.795	0.951	≃1
	100	r_E	0.692	0.686	0.940	0.998	≃1
		ANOVA	0.670	0.685	0.941	0.998	≃1
		REML	0.652	0.711	0.947	0.999	≃1
	200	r_E	0.555	0.890	0.999	≃1	≃1
		ANOVA	0.549	0.889	0.999	≃1	≃1
		REML	0.540	0.899	0.999	≃1	≃1
	300	r_E	0.551	0.977	1.000	≃1	≃1
		ANOVA	0.558	0.977	1.000	≃1	≃1
		REML	0.543	0.974	1.000	≃1	≃1
4	50	r_E	0.549	0.872	0.986	≃1	≃1
		ANOVA	0.567	0.872	0.983	0.999	≃1
		REML	0.535	0.875	0.989	≃1	≃1
	100	r_E	0.685	0.986	≃1	≃1	≃1
		ANOVA	0.698	0.986	≃1	≃1	≃1
		REML	0.692	0.990	≃1	≃1	≃1
	200	r_E	0.912	≃1	≃1	≃1	≃1
		ANOVA	0.892	≃1	≃1	≃1	≃1
		REML	0.896	≃1	≃1	≃1	≃1
	300	r_E	0.979	≃1	≃1	≃1	≃1
		ANOVA	0.981	≃1	≃1	≃1	≃1
		REML	0.981	≃1	≃1	≃1	≃1
6	50	r_E	0.649	0.962	0.996	≃1	≃1
		ANOVA	0.673	0.964	0.995	≃1	≃1
		REML	0.673	0.969	0.997	≃1	≃1
	100	r_E	0.908	0.999	≃1	≃1	≃1
		ANOVA	0.905	0.999	≃1	≃1	≃1
		REML	0.909	0.999	≃1	≃1	≃1
	200	r_E	0.991	≃1	≃1	≃1	≃1
		ANOVA	0.990	≃1	≃1	≃1	≃1
		REML	0.992	≃1	≃1	≃1	≃1
	300	r_E	0.999	≃1	≃1	≃1	≃1
		ANOVA	0.999	≃1	≃1	≃1	≃1
		REML	0.999	≃1	≃1	≃1	≃1
8	50	r_E	0.757	0.990	≃1	≃1	≃1
		ANOVA	0.779	0.990	≃1	≃1	≃1
		REML	0.805	0.994	≃1	≃1	≃1
	100	r_E	0.958	≃1	≃1	≃1	≃1
		ANOVA	0.953	≃1	≃1	≃1	≃1
		REML	0.959	≃1	≃1	≃1	≃1
	200	r_E	≃1	≃1	≃1	≃1	≃1
		ANOVA	0.999	≃1	≃1	≃1	≃1
		REML	0.999	≃1	≃1	≃1	≃1
	300	r_E	≃1	≃1	≃1	≃1	≃1
		ANOVA	≃1	≃1	≃1	≃1	≃1

Table 1. Contd.

		REML	≃1	≃1	≃1	≃1	≃1
10	50	r_E	0.876	0.996	≃1	≃1	≃1
		ANOVA	0.869	0.997	≃1	≃1	≃1
		REML	0.867	0.996	≃1	≃1	≃1
	100	r_E	0.992	≃1	≃1	≃1	≃1
		ANOVA	0.991	≃1	≃1	≃1	≃1
		REML	0.990	≃1	≃1	≃1	≃1
	200	r_E	≃1	≃1	≃1	≃1	≃1
		ANOVA	≃1	≃1	≃1	≃1	≃1
		REML	≃1	≃1	≃1	≃1	≃1
	300	r_E	≃1	≃1	≃1	≃1	≃1
		ANOVA	≃1	≃1	≃1	≃1	≃1
		REML	≃1	≃1	≃1	≃1	≃1

Data sets had no missing data.

Table 2. Correlation of broad-sense heritability ($\widehat{H}^2$) estimated by three methods inter-environment correlation (r_E), ANOVA (A), and REML (R) for simulated data sets without missing data.

No. of environments	No. of test genotypes	$r(r_E, \widehat{H}^2$(A))	$r(r_E, \widehat{H}^2$(R))	$r(\widehat{H}^2$(A), $\widehat{H}^2$(R))
2	50 - 300	1.000	0.992	0.992
4	50 - 300	1.000	0.997	0.997
6	50 - 300	1.000	0.998	0.998
8	50 - 300	1.000	0.998	0.998
10	50 - 300	1.000	0.998	0.998
2 - 10	50	1.000	0.991	0.991
2 - 10	100	1.000	0.997	0.997
2 - 10	200	1.000	0.999	0.999
2 - 10	300	1.000	0.999	0.999
2 - 10	50- 300	1.000	0.997	0.997

$\widehat{H}^2$(A): $\widehat{H}^2$ on a plot basis estimated by ANOVA; $\widehat{H}^2$(R): $\widehat{H}^2$ on a plot basis estimated by REML. Simulated data includes data points generated from combinations of environments (2, 4, 6, and 8), test genotypes (50, 100, 200 and 300), and heritability (0.1-0.9 in 0.1 increments) with 500 replicates for each combination.

Table 3. Effects of missing data on estimation of broad-sense heritability (H^2) and statistical power (*P*).

H^2	MDR (%)	$\widehat{H}^2 \pm s$			*P*		
		r_E	ANOVA	REML	r_E	ANOVA	REML
0.1	0	0.101 ± 0.047	0.101 ± 0.047	0.100 ± 0.045	0.835	0.836	0.836
	5	0.101 ± 0.049	0.121 ± 0.063	0.100 ± 0.047	0.821	0.792	0.821
	10	0.101 ± 0.051	0.136 ± 0.081	0.100 ± 0.048	0.830	0.823	0.831
	15	0.102 ± 0.053	0.150 ± 0.095	0.101 ± 0.050	0.812	0.749	0.815
	20	0.102 ± 0.057	0.163 ± 0.108	0.101 ± 0.053	0.806	0.710	0.811
0.2	0	0.200 ± 0.054	0.200 ± 0.053	0.199 ± 0.052	0.944	0.945	0.948
	5	0.200 ± 0.056	0.212 ± 0.064	0.198 ± 0.054	0.931	0.913	0.935
	10	0.200 ± 0.058	0.221 ± 0.077	0.198 ± 0.055	0.936	0.935	0.940
	15	0.201 ± 0.061	0.229 ± 0.088	0.199 ± 0.057	0.926	0.893	0.931
	20	0.200 ± 0.064	0.235 ± 0.099	0.198 ± 0.060	0.911	0.866	0.918
0.3	0	0.298 ± 0.057	0.298 ± 0.057	0.296 ± 0.056	0.984	0.984	0.985
	5	0.298 ± 0.059	0.303 ± 0.063	0.296 ± 0.057	0.977	0.972	0.980

Table 3. Contd.

	10	0.298 ± 0.062	0.306 ± 0.072	0.296 ± 0.059	0.982	0.981	0.984
	15	0.298 ± 0.064	0.308 ± 0.081	0.296 ± 0.061	0.973	0.963	0.978
	20	0.298 ± 0.067	0.310 ± 0.089	0.296 ± 0.064	0.968	0.951	0.973
0.4	0	0.398 ± 0.055	0.397 ± 0.054	0.395 ± 0.053	0.998	0.997	0.998
	5	0.398 ± 0.057	0.397 ± 0.058	0.395 ± 0.055	0.995	0.996	0.996
	10	0.398 ± 0.060	0.395 ± 0.064	0.394 ± 0.057	0.996	0.997	0.998
	15	0.398 ± 0.062	0.393 ± 0.071	0.395 ± 0.058	0.993	0.993	0.996
	20	0.397 ± 0.066	0.389 ± 0.078	0.394 ± 0.061	0.989	0.989	0.994
0.5	0	0.499 ± 0.052	0.498 ± 0.051	0.494 ± 0.051	≃1	≃1	≃1
	5	0.499 ± 0.053	0.493 ± 0.054	0.494 ± 0.052	≃1	≃1	≃1
	10	0.499 ± 0.056	0.488 ± 0.058	0.494 ± 0.053	≃1	≃1	≃1
	15	0.499 ± 0.059	0.482 ± 0.063	0.494 ± 0.055	0.999	≃1	≃1
	20	0.498 ± 0.061	0.475 ± 0.069	0.493 ± 0.057	0.999	0.999	≃1
0.6	0	0.598 ± 0.048	0.597 ± 0.048	0.591 ± 0.047	≃1	≃1	≃1
	5	0.598 ± 0.050	0.590 ± 0.050	0.591 ± 0.048	≃1	≃1	≃1
	10	0.597 ± 0.051	0.582 ± 0.053	0.591 ± 0.050	≃1	≃1	≃1
	15	0.597 ± 0.054	0.574 ± 0.056	0.591 ± 0.051	≃1	≃1	≃1
	20	0.597 ± 0.056	0.566 ± 0.060	0.591 ± 0.052	≃1	≃1	≃1
0.7	0	0.698 ± 0.039	0.697 ± 0.039	0.688 ± 0.040	≃1	≃1	≃1
	5	0.698 ± 0.040	0.689 ± 0.040	0.688 ± 0.041	≃1	≃1	≃1
	10	0.698 ± 0.041	0.681 ± 0.043	0.688 ± 0.042	≃1	≃1	≃1
	15	0.698 ± 0.043	0.672 ± 0.045	0.688 ± 0.042	≃1	≃1	≃1
	20	0.698 ± 0.045	0.663 ± 0.048	0.689 ± 0.043	≃1	≃1	≃1
0.8	0	0.799 ± 0.028	0.797 ± 0.028	0.784 ± 0.034	≃1	≃1	≃1
	5	0.799 ± 0.029	0.791 ± 0.030	0.784 ± 0.034	≃1	≃1	≃1
	10	0.798 ± 0.030	0.783 ± 0.031	0.785 ± 0.034	≃1	≃1	≃1
	15	0.798 ± 0.032	0.775 ± 0.033	0.785 ± 0.034	≃1	≃1	≃1
	20	0.798 ± 0.033	0.767 ± 0.036	0.785 ± 0.035	≃1	≃1	≃1
0.9	0	0.899 ± 0.016	0.898 ± 0.016	0.877 ± 0.030	≃1	≃1	≃1
	5	0.899 ± 0.016	0.894 ± 0.017	0.878 ± 0.030	≃1	≃1	≃1
	10	0.899 ± 0.017	0.889 ± 0.018	0.878 ± 0.030	≃1	≃1	≃1
	15	0.899 ± 0.018	0.884 ± 0.019	0.878 ± 0.030	≃1	≃1	≃1
	20	0.899 ± 0.018	0.879 ± 0.020	0.879 ± 0.030	≃1	≃1	≃1

MDR, Missing data rate on an entire multi-environment trial basis; H^2 was calculated based on 10,000 data points generated from combinations of various numbers of environments (2, 4, 6, 8 and 10) and test genotypes (50, 100, 200 and 300) with 500 replicates for each combination; Statistical power was the average of power estimates from 20 parametric sets of different numbers of environments and test genotypes.

(Table 6). Because the empirical data had missing values in some environments (Table S1), a slightly weaker relationship among r_E, $\hat{H}^2$(A) and $\hat{H}^2$(R), and a stronger relation between r_E and $\hat{H}^2$(R) than that between r_E and $\hat{H}^2$(A) (Table 6) were observed, which resembled the results obtained in missing data sets.

Standard error of r_E

To perform a significance test for r_E estimates, the SEs of r_E determined by Eq. 4 ($SE(r_E)$) and by the bootstrap method ($s_b(r_E)$) were compared to the corresponding simulated SEs ($s(r_E)$). Although a strong correlation of $s(r_E)$ to $SE(r_E)$ was observed (R^2 = 0.815) (Figure 5A), a high correlation (R^2 = 0.995) was obtained for $s(r_E)$ with $s_b(r_E)$ (Figure 5B) indicating the bootstrap method yielded SE estimates for r_E that were highly consistent with those obtained by simulation. The SE estimates obtained by the bootstrap method were systematically smaller (by 0.0015 on average) than those obtained by simulation. A linear regression equation, $s(r_E) = 1.0802^* s_b(r_E) - 0.0017$ (Figure 5B), may be used to adjust the bootstrap estimates.

DISCUSSION

H^2 is the ratio of genetic variance to total phenotypic variance, representing the extent with which genotypes

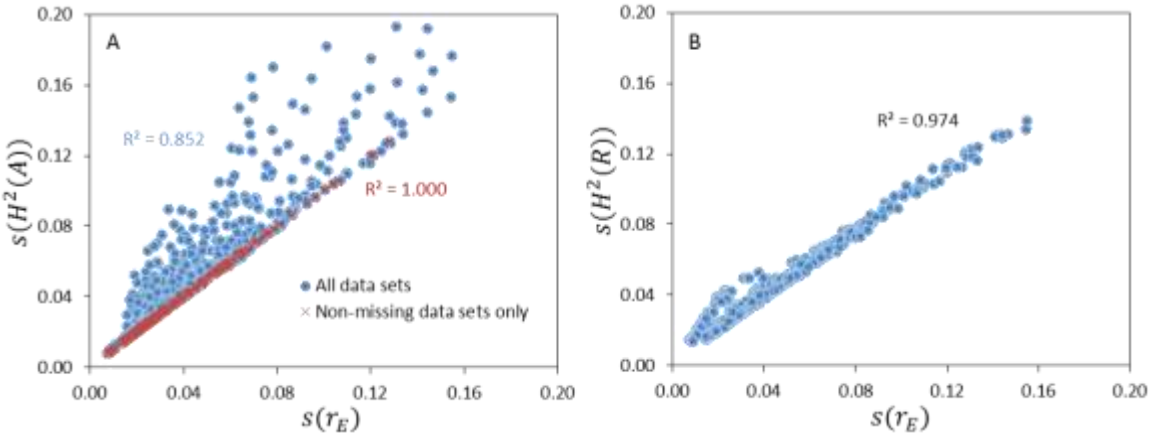


Figure 3. Relationship of standard errors (SEs) of inter-environment correlation (r_E) with SEs of H^2estimates by ANOVA (A) and REML (R). (A) SEs of r_E estimated from simulations ($s(r_E)$) vs. SEs of H^2estimates by ANOVA ($s(\hat{H}^2(A)$) using all data sets (missing and non-missing data) (blue) and non-missing data only (red). (B) $s(r_E)$) vs. $s(\hat{H}^2(R))$ using all data sets.

Table 4. Statistical power (*P*) of broad-sense heritability estimates ($\hat{H}^2$) at a given heritability of 0.2 by inter-environment correlation (r_E), ANOVA, and REML for α = 0.05 (one-tailed test) as a function of the number of environments, and test genotypes, and missing data rates.

No. of environments	No. of test genotypes	Method	*P* for missing data rate (%)				
			0	5	10	15	20
2	50	r_E	0.533	0.503	0.518	0.543	0.571
		ANOVA	0.533	0.527	0.507	0.500	0.528
		REML	0.544	0.516	0.502	0.523	0.543
	100	r_E	0.686	0.640	0.606	0.611	0.534
		ANOVA	0.685	0.673	0.653	0.619	0.557
		REML	0.711	0.666	0.643	0.623	0.563
	200	r_E	0.890	0.859	0.837	0.818	0.733
		ANOVA	0.889	0.812	0.705	0.639	0.596
		REML	0.899	0.875	0.858	0.835	0.755
	300	r_E	0.977	0.959	0.953	0.929	0.908
		ANOVA	0.977	0.889	0.775	0.703	0.642
		REML	0.974	0.956	0.947	0.929	0.903
4	50	r_E	0.872	0.849	0.821	0.793	0.733
		ANOVA	0.872	0.886	0.835	0.807	0.741
		REML	0.875	0.859	0.832	0.832	0.763
	100	r_E	0.986	0.976	0.976	0.955	0.913
		ANOVA	0.986	0.982	0.960	0.918	0.858
		REML	0.990	0.984	0.982	0.969	0.947
	200	r_E	≃1	≃1	≃1	0.999	0.998
		ANOVA	≃1	0.999	0.979	0.942	0.888
		REML	≃1	≃1	≃1	≃1	0.999
	300	r_E	≃1	≃1	≃1	≃1	≃1
		ANOVA	≃1	≃1	0.993	0.966	0.917
		REML	≃1	≃1	≃1	≃1	≃1
6	50	r_E	0.962	0.946	0.934	0.916	0.894
		ANOVA	0.964	0.957	0.929	0.897	0.866
		REML	0.969	0.959	0.949	0.939	0.925
	100	r_E	0.999	0.998	0.997	0.994	0.992

Table 4. Contd.

		ANOVA	0.999	0.996	0.976	0.938	0.903
		REML	0.999	0.999	0.998	0.997	0.996
	200	r_E	≃1	≃1	≃1	≃1	≃1
		ANOVA	≃1	≃1	≃1	0.999	0.994
		REML	≃1	≃1	≃1	≃1	≃1
	300	r_E	≃1	≃1	≃1	≃1	≃1
		ANOVA	≃1	≃1	0.999	0.988	0.960
		REML	≃1	≃1	≃1	≃1	≃1
8	50	r_E	0.990	0.988	0.979	0.971	0.958
		ANOVA	0.990	0.985	0.970	0.960	0.929
		REML	0.994	0.992	0.990	0.987	0.977
	100	r_E	≃1	≃1	≃1	≃1	≃1
		ANOVA	≃1	≃1	0.999	0.995	0.986
		REML	≃1	≃1	≃1	≃1	≃1
	200	r_E	≃1	≃1	≃1	≃1	≃1
		ANOVA	≃1	≃1	≃1	0.999	0.995
		REML	≃1	≃1	≃1	≃1	≃1
	300	r_E	≃1	≃1	≃1	≃1	≃1
		ANOVA	≃1	≃1	≃1	≃1	0.999
		REML	≃1	≃1	≃1	≃1	≃1
10	50	r_E	0.996	0.995	0.990	0.988	0.988
		ANOVA	0.997	0.994	0.988	0.983	0.971
		REML	0.996	0.996	0.994	0.992	0.994
	100	r_E	≃1	≃1	≃1	≃1	≃1
		ANOVA	≃1	≃1	≃1	0.999	0.995
		REML	≃1	≃1	≃1	≃1	≃1
	200	r_E	≃1	≃1	≃1	≃1	≃1
		ANOVA	≃1	≃1	≃1	≃1	≃1
		REML	≃1	≃1	≃1	≃1	≃1
	300	r_E	≃1	≃1	≃1	≃1	≃1
		ANOVA	≃1	≃1	≃1	≃1	≃1
		REML	≃1	≃1	≃1	≃1	≃1

are affected by environments and random error. H^2 also represents the repeatability of trait performance in different environments. For highly heritable traits, the performance of genotypes in one environment has a high repeatability in other environments; in other words, the trait performance of the genotypes is strongly correlated between any pair of environments. In contrast, low heritability traits display low correlations of trait performance between any two environments. As such, the inter-environment correlation (r_E) of trait performance should be an indicator of trait heritability. The theoretical derivation confirms that the r_E between two environments corresponds to H^2 (Equation 2). Both simulation and empirical results demonstrate that r_E is an accurate and stable estimate of H^2 on a plot basis.

The r_E method has two significant advantages. First, because r_E is a mean correlation coefficient between pairs of environments, its calculation is simple. Second, only the test genotype data is required for calculation of r_E, eliminating the prerequisite for the use of the same control genotypes across all trials and hence permitting flexible field designs. Thus, the r_E method allows joint H^2 estimation over multiple environments for genetic populations that may differ in their MAD2 designs and which may not necessarily include the same control genotypes. The r_E method is equally applicable to trials with the same test genotypes but where different control genotypes are used in the different environments. A practical example was the joint H^2estimation for three flax biparental populations: BM, EV, and SU. These three populations were evaluated in the same six to eight environments using MAD2 designs but with different

Table 5. Correlation of broad-sense heritability ($\hat{H}^2$) estimated by three methods inter-environment correlation (r_E), ANOVA (A), and REML (R) for simulated data sets with missing data rates of 5 to 20%.

No. of environments	No. of test genotypes	$r(r_E, \hat{H}^2$(A))	$r(r_E, \hat{H}^2$(R))	$r(\hat{H}^2$(A), $\hat{H}^2$(R))
2	50 - 300	0.955	0.988	0.945
4	50 - 300	0.984	0.995	0.980
6	50 - 300	0.991	0.997	0.989
8	50 - 300	0.995	0.998	0.993
10	50 - 300	0.996	0.998	0.994
2 - 10	50	0.980	0.998	0.970
2 - 10	100	0.982	0.986	0.979
2 - 10	200	0.984	0.998	0.983
2 - 10	300	0.985	0.999	0.984
2 - 10	50 - 300	0.983	0.995	0.979

$\hat{H}^2$(A): $\hat{H}^2$ on a plot basis estimated by ANOVA; $\hat{H}^2$(R): $\hat{H}^2$ on a plot basis estimated by REML. Simulated data includes data points generated from combinations of environments (2, 4, 6, and 8), test genotypes (50, 100, 200 and 300), and heritability (0.1-0.9 in 0.1 increments) with 500 replicates for each combination.

Table 6. Correlation of broad-sense heritability ($\hat{H}^2$) estimated by three methods inter-environment correlation (r_E), ANOVA (A), and REML (R) for empirical data.

Population	No. of test genotypes	No. of traits	$r(r_E, \hat{H}^2$(A))	$r(r_E, \hat{H}^2$(R))	$r(\hat{H}^2$(A), $\hat{H}^2$(R))
CC	391	22	0.986	0.992	0.995
BM	243	14	0.982	0.982	1.000
EV	90	19	0.948	0.974	0.980
SU	78	11	0.998	0.998	1.000
Total	-	66	0.975	0.985	0.989

CC, Core collection; BM, CDC Bethune/Macbeth; EV, E1747/Viking; SU, SP2047/UGG5-5; $\hat{H}^2$(A): $\hat{H}^2$ on a plot basis estimated by ANOVA; $\hat{H}^2$(R): $\hat{H}^2$ on a plot basis estimated by REML. The estimate for each trait was used as a single data point to calculate the correlation among r_E, $\hat{H}^2$(A), and $\hat{H}^2$(R). No. of traits in each population represent the number of data points used for calculation of correlation between any two methods. Total represents all 66 data points from four individual populations for correlation calculation.

control genotypes. In each MAD2 design, CDC Bethune was used as the main plot control but the subplot controls were the two parents of each of the three biparental populations to improve error control because the two parents shared the genetic background of their offspring. As a consequence of the use of different control genotypes in the three population trials, neither ANOVA- nor REML-based methods could estimate joint H^2 values, but this was achievable with the r_E method using pooled data. The results will be reported separately.

In the ANOVA-based H^2 estimation, the error variance of unreplicated test genotypes is estimated by duplicated control genotypes. This is based on the assumption that control genotypes share the same random error variance with all test genotypes. Theoretically, the total mean square (MS) of the test genotypes will be greater than the error MS in a single trial or the G × E interaction variance in a multi-environment trial. As such, the genetic variance of the test genotypes can be estimated as the total MS minus the error MS of a single trial or by the total variance minus the G × E interaction MS in a multi-environment trial. However, the sampling bias caused by a limited number of control genotypes (typically three in MAD2) may occasionally result in negative genetic variance estimates and failure to correctly estimate genetic parameters, especially when the heritability of a trait is very low. In this case, the r_E method can avoid this potential drawback because the genetic variance in r_E corresponds to the covariance of trait performance between two environments ($COV(y_1, y_2)$) (Equation 2 and Figure 1A). If the $COV(y_1, y_2)$ is less than zero then the genetic variance can be considered zero, whereas negative or null genetic variance obtained by ANOVA or REML might be an incorrect estimate in some cases. In the core collection, for example, both oil and linolenic acid contents have H^2 estimates of zero because their genetic variances were estimated to be zero by both ANOVA- and REML-based methods. This result is

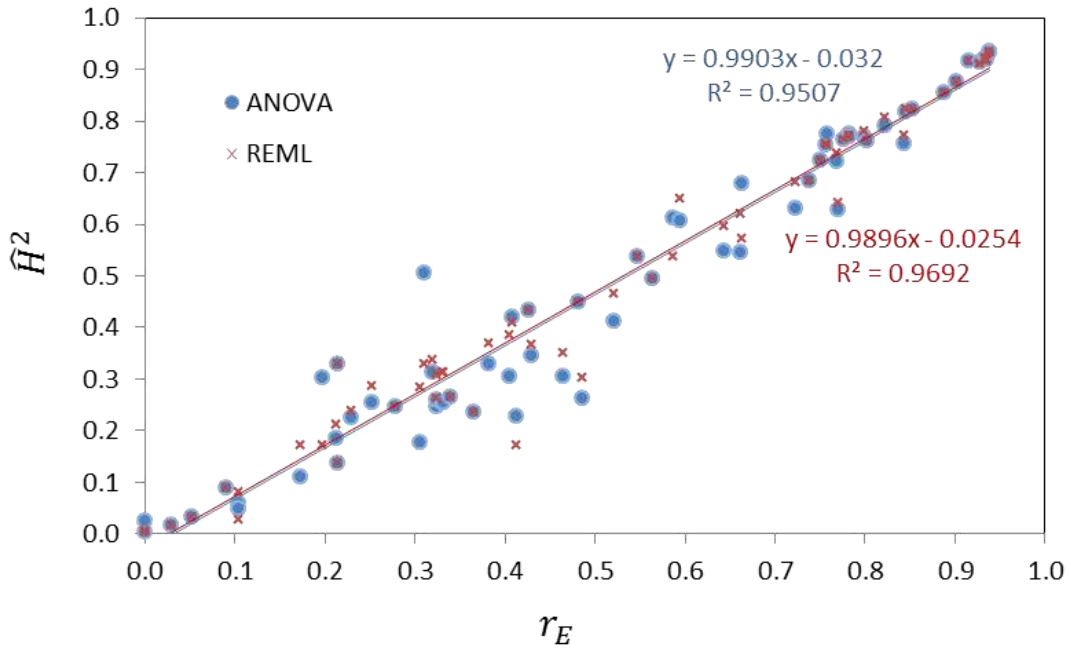


Figure 4. Relationship of inter-environment correlations (r_E) with broad-sense heritability ($\hat{H}^2$) estimated from empirical data by ANOVA and REML. Data points include estimates for r_E and $\hat{H}^2$ from 22 traits in the core collection, 14 in the CDC Bethune/Macbeth, 19 in the E1747/Viking, and 11 in the SP20147/UGG5-5 populations.

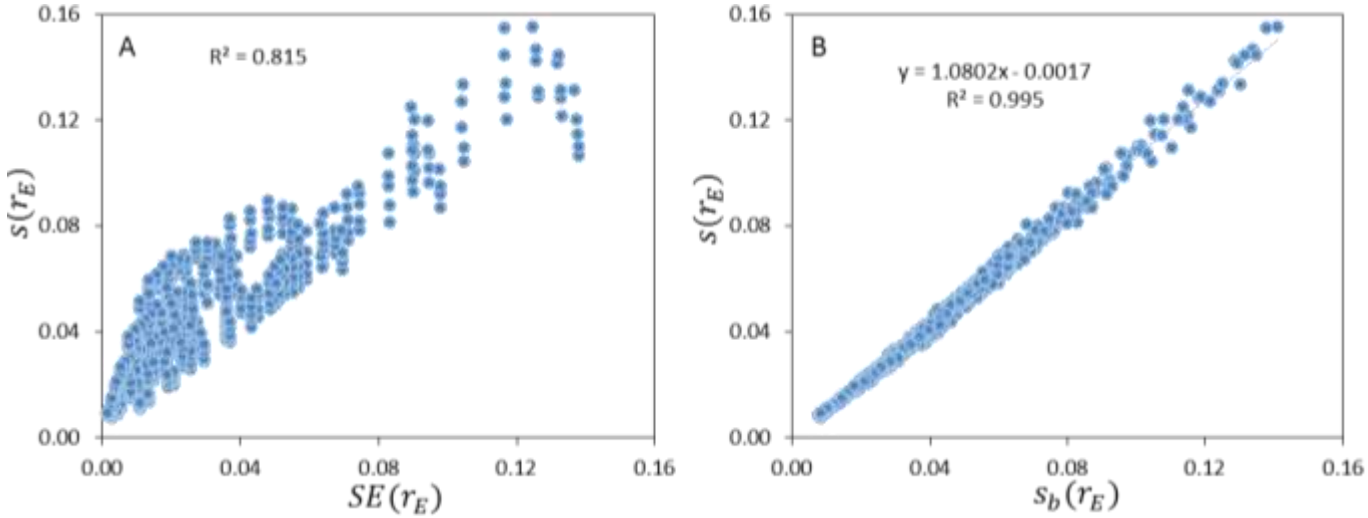


Figure 5. Relationship of standard error (*SE*) of inter-environment correlation (r_E) with *SE* estimates by other methods. (A) $s(r_E)$ vs. *SE* of r_E calculated by Equation 4 ($SE(r_E)$). (B)$s(r_E)$ vs. *SE* estimated by the bootstrap method ($s_b(r_E)$). All data sets including missing and non-missing data) were used.

obviously incorrect because oil and linolenic acid content are traits of moderate to high heritability (You et al., 2016a). However, r_E-based H^2 estimates for the two traits were 0.387 and 0.661, respectively, although these estimates are still smaller than expected.

The REML-based method, as an alternative to the more traditional ANOVA techniques, has been widely used for estimation of heritability and genetic correlations (Holland, 2006; Piepho and Möhring, 2011). The advantages compared to ANOVA methods are that REML estimates of variance and covariance components have known asymptotic distributional properties and can

efficiently handle both unbalanced data and complex experimental designs (Meyer, 1985; Holland, 2006). REML's main drawback is that it is much more computationally intensive than ANOVA, a disadvantage that is exacerbated with large data sets. In this study, these two H^2 estimation methods were compared to the r_E method using both simulation and empirical data of multi-environment MAD2 trials. All results corroborate the agreements between the ANOVA- and REML-based methods with $r > 0.99$ in the simulated data sets without missing data, especially in the empirical data of BM and SU populations where their estimates are nearly identical (Table 2). The ANOVA-based H^2estimates are consistent with the r_E estimates when no missing data exist in the data sets. Nevertheless, when missing values occur, the r_E- and REML-based methods show higher power than the ANOVA-based method, confirming that the REML-based method is efficient for tackling unbalanced data. Overall, r_E-, ANOVA-, and REML-based methods can be used for non-missing data sets; the r_E- and REML-based methods are suitable for missing data sets; and the r_E method is versatile for all cases of practical data sets in multi-environment MAD2 trials.

A significance test for r_E estimates requires SE values. Generally, SEs ($s(r_E)$) estimated by simulation with a large sample size (500 in this study) provide a good estimate of the sampling error. Thus, they were used as a relative standard for comparisons. Two potential methods were assessed: the approximate standard error ($SE(r_E)$) of mean simple correlation coefficients (Equation 4) and bootstrap. The non-parametric bootstrap is an effective alternative for determining distribution of an estimator with an unknown probability density, and has been used to estimate standard errors for heritability (Xie and Mosjidis, 1997). Results demonstrate that the bootstrap method outperforms the SE formula of simple correlation coefficient (Equation 4), and provides perfectly consistent SE estimates with $s(r_E)$. However, a systematic difference between $s(r_E)$ and $s_b(r_E)$ was observed; adjustments may be done using the regression equation in Figure 5B, which was constructed using 900 data points generated from different numbers of environments (2, 4, 6, 8, and 10), test genotypes (50, 100, 200 and 300), and missing data rates (0, 5, 10, 15 and 20%) at given heritability values ranging from 0.1 to 0.9. Therefore, the bootstrap method is recommended to estimate the standard error of r_E in significance tests. The bootstrap estimation of *SE* has been implemented in the R pipeline program (http://probes.pw.usda.gov/bioinformatics_tools/MADPipeline/index.html).

To find the effective sample size for estimating H^2 in multi-environment MAD2 trials, the statistical power of three H^2 estimation methods was calculated. As expected, the power is affected by the number of environments and test genotypes, and missing data rates. For traits with H^2 equal to or greater than 0.5, 50 test genotypes at two environments were sufficient to achieve a statistical power over 0.95 (Table 1); for traits with lower heritability (e.g., 0.2), 300 test genotypes at two environments, 100 test genotypes at four environments, or 50 test genotypes at greater than or equal to six environments are required to obtain the same statistical power (Table 1). Increasing the missing data rate decreased the statistical power (Table 4), but an increase of the number of environments and/or test genotypes markedly improved the statistical power of H^2 estimates.

Notably, heritability estimates in the simulation data sets were slightly biased from their true values when true H^2 values were greater than 0.2 (Table 3). This bias was observed in all three H^2 estimation methods, especially when the ANOVA-based method was used for missing data sets. However, the r_E method has less deviation than other two for both missing and non-missing data sets. This bias may be inherently owing to the MAD2's unbalanced feature. Piepho and Möhring (2007) discussed how estimation of broad-sense heritability in unbalanced trials differs from that in the case of a balanced design. The r_E proposed here, as well as the ANOVA- (You et al., 2016b) and REML- (You et al., 2016a) based methods, provide an approximate H^2 estimate for MAD2 trials. This approximation is due to not only unbalanced data but also approximate assumption of independence for adjusted observations from an MAD2 trial.

For adjustment of observations in an MAD2 trial, there are four different cases for quantitative traits: (1) significant additive soil variation due to row or column effects (M1); (2) significant non-additive soil variation due to row × column interaction effects (M3); (3) M1+M3, and (4) no additive or non-additive soil variation (You et al., 2013). For case 4, as no data adjustment is required, their estimates of heritability will be unbiased, while for the first three cases, data adjustment are needed and the adjusted data may be correlated to some extent, resulting in biased estimates. In this study, the simulation data was completely independent of each other, but the empirical data of some traits were unnecessarily completely independent due to data adjustment. However, we found that highly similar or consistent results were obtained in both data sets, indicating small effect of defective independence assumption to heritability estimation. In our actual trials, especially those with good quality, most of the traits would not need data adjustment or have minor row or/and column or their interaction effects, the bias due to defective independence would be small and may be disregarded in breeding applications.

Conclusion

A H^2 statistic, r_E, representing an inter-environment correlation of a quantitative trait, was presented for multi-environment MAD2 trials. The r_E method provides a simple approach to approximate H^2 in any multi-

environment MAD2 trial without the limitations of ANOVA- or REML-based methods that require the use of the same control genotypes across trials and/or environments.

Conflicts of Interests

The authors have not declared any conflict of interests.

ACKNOWLEDGEMENTS

This work was conducted as part of the A-base project (No. 1142) funded by Agriculture and Agri-Food Canada, the flax breeding database project funded by the Western Grain Research Foundation (WGRF), and the Total Utilization Flax GENomics (TUFGEN) project funded by Genome Canada and other stakeholders.

REFERENCES

Bates D, Mäechler M, Bolker BM, Walker SC (2015). Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67(1):1-48.

Bowley AL (1928). The standard deviation of the correlation coefficient. J. Am. Stat. Assoc. 23(161):31-34.

Cloutier S, Ragupathy R, Niu Z, Duguid S (2010). SSR-based linkage map of flax (*Linum usitatissimum* L.) and mapping of QTLs underlying fatty acid composition traits. Mol. Breed. 28(4):437-451.

Efron B (1979). Bootstrap methods: another look at the jacknife. Ann. Stat. 7:1-26.

Golparvar AR, Gheisari MM, Naderi D, Mehrabi AM, Hadipanah A, Salehi S (2015). Determination of the best indirect selection criteria in Iranian durum wheat (*Triticum aestivum* L.) genotypes under irrigated and drought stress conditions. Genetika 47:549-558.

Holland JB (2006). Estimating genotypic correlations and their standard errors using multivariate restricted maximum likelihood estimation with SAS Proc MIXED. Crop Sci. 46(2):642-654.

Holland JB, Nyquist WE, Cervantes-Martinez GT (2003). Estimating and interpreting heritability for plant breeding: an update. Plant Breed. Rev. 22:9-112.

Kendall MG, Stuart A (1977). The advanced theory of statistics, Distribution theory. (London: Griffin).

Klein TW (1974). Heritability and genetic correlation: statistical power, population comparisons, and sample size. Behav. Genet. 4(2):171-189.

Kumar S, You FM, Duguid S, Booker H, Rowland G, Cloutier S (2015). QTL for fatty acid composition and yield in linseed (*Linum usitatissimum* L.). Theor. Appl. Genet. 128(5):965-984.

Lin C, Poushinsky G (1983). A modified augmented design for an early stage of plant selection involving a large number of test lines without replication. Biometrics 39(3):553-561.

Lin CS, Poushinsky G (1985). A modified augmented design (type 2) for rectangular plots. Can. J. Plant Sci. 65(3):743-749.

Lin CS, Voldeng HD (1989). Efficiency of type 2 modified augmented designs in soybean variety trials. Agron. J. 81:512-517.

May KW, Kozub GC (1995). Success of a selection program for increasing grain yield of two-row barley lines and evaluation of the modified augmented design (type 2). Can. J. Plant Sci. 75(4):795-799.

May KW, Kozub GC, Schaalje GB (1989). Field evaluation of a modified augmented design (type 2) for screening barley lines. Can. J. Plant Sci. 69(1):9-15.

Meyer K (1985). Maximum likelihood estimation of variance components for a multivariate mixed model with equal design matrices. Biometrics 41(1):153-165.

Piepho HP, Möhring J (2007). Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177:1881-1888.

Piepho HP, Möhring J (2011). On estimation of genotypic correlations and their standard errors by multivariate REML using the MIXED procedure of the SAS system. Crop Sci. 51(6):2449-2454.

Schaalje GB, Lynch DR, Kozub GC (1987). Field evaluation of a modified augmented design for early stage selection involving a large number of test lines without replication. Potato Res. 30:35-45.

Soto-Cerda BJ, Duguid S, Booker H, Rowland G, Diederichsen A, Cloutier S (2014a). Genomic regions underlying agronomic traits in linseed (*Linum usitatissimum* L.) as revealed by association mapping. J. Integr. Plant Biol. 56(1):75-87.

Soto-Cerda BJ, Duguid S, Booker H, Rowland G, Diederichsen A, Cloutier S (2014b). Association mapping of seed quality traits using the Canadian flax (*Linum usitatissimum* L.) core collection. Theor. Appl. Genet. 127(4):881-896.

Xie C, Mosjidis JA (1997). Influence of sample size on precision of heritability and expected selection response in red clover. Plant Breed. 116:83-88.

You FM, Duguid SD, Thambugala D, Cloutier S (2013). Statistical analysis and field evaluation of the type 2 modified augmented design (MAD) in phenotyping of flax (*Linum usitatissimum*) germplasms in multiple environments. Aust. J. Crop Sci. 7(11):1789-1800.

You FM, Booker MH, Duguid DS, Jia G, Cloutier S (2016a). Accuracy of genomic selection in biparental populations of flax (*Linum usitatissimum* L.). Crop J. 4(4):290-303.

You FM, Song Q, Jia G, Cheng Y, Duguid S, Booker H, Cloutier S (2016b). Estimation of genetic parameters and their sampling variances for quantitative traits in the type 2 modified augmented design. Crop J. 4(2):107-118.

Supplementary data

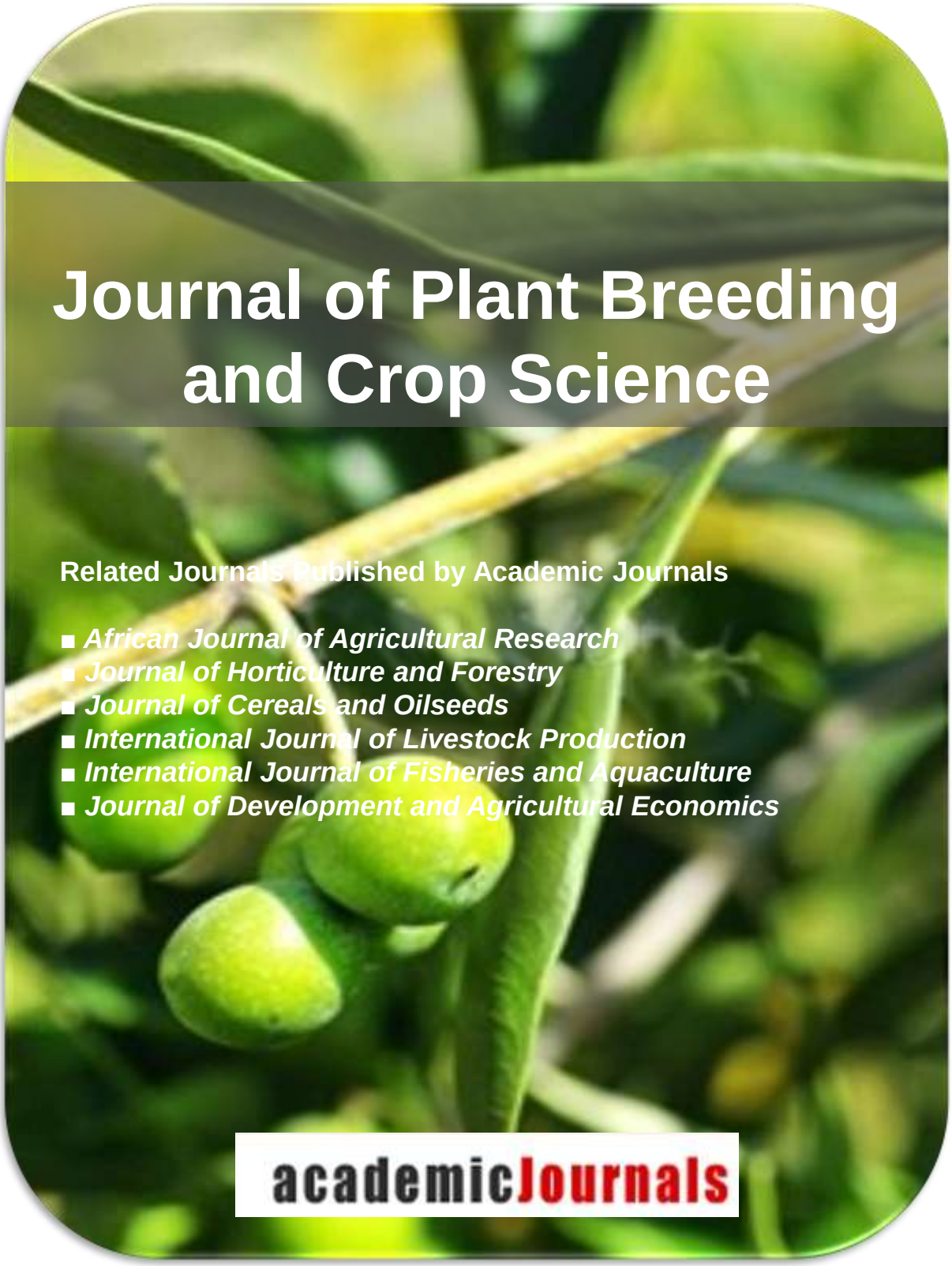
Table S1. Broad-sense heritability estimates ($\widehat{H}^2$) determined by inter-environment correlation (r_E), ANOVA (A), and REML (R), and their standard error (SE) for traits for three flax biparental populations (BM, EV, and SU) and the flax core collection (CC).

Population	Trait	MDR (%)	r_E ± SE	$\widehat{H}^2(A)$ ± SE	$\widehat{H}^2(R)$ ± SE
CC	Bolls (m^{-2})	6.07	0.319 ± 0.051	0.313 ± 0.022	0.337 ± 0.023
CC	Cellulose content (%)	2.23	0.306 ± 0.046	0.178 ± 0.014	0.284 ± 0.020
CC	Cell wall content (%)	2.23	0.173 ± 0.050	0.111 ± 0.012	0.170 ± 0.017
CC	Fiber content (%)	2.23	0.331 ± 0.052	0.253 ± 0.018	0.312 ± 0.021
CC	Days to 5% flowering	2.05	0.643 ± 0.043	0.548 ± 0.020	0.595 ± 0.020
CC	Days to 95% flowering	2.40	0.662 ± 0.038	0.546 ± 0.020	0.619 ± 0.019
CC	Plant height (cm)	2.05	0.594 ± 0.033	0.608 ± 0.019	0.650 ± 0.019
CC	Iodine value	3.11	0.800 ± 0.026	0.769 ± 0.014	0.780 ± 0.014
CC	Lignin content (%)	2.23	0.324 ± 0.051	0.247 ± 0.018	0.307 ± 0.021
CC	Linoleic content (%)	3.11	0.934 ± 0.050	0.919 ± 0.006	0.921 ± 0.006
CC	Lodging	1.98	0.104 ± 0.048	0.060 ± 0.010	0.081 ± 0.014
CC	Days to maturity	2.23	0.212 ± 0.055	0.185 ± 0.017	0.212 ± 0.019
CC	Mildew score	5.53	0.522 ± 0.038	0.411 ± 0.028	0.466 ± 0.031
CC	Oleic content (%)	3.11	0.768 ± 0.028	0.722 ± 0.016	0.736 ± 0.016
CC	Palmitic content (%)	3.11	0.822 ± 0.027	0.791 ± 0.013	0.806 ± 0.012
CC	Pasmo score	3.02	0.253 ± 0.042	0.256 ± 0.022	0.288 ± 0.025
CC	Plant branching score	9.85	0.000 ± 0.053	0.004 ± 0.013	0.006 ± 0.018
CC	Shive content (%)	2.23	0.332 ± 0.052	0.254 ± 0.018	0.312 ± 0.021
CC	Protein content (%)	3.46	0.723 ± 0.061	0.631 ± 0.021	0.680 ± 0.020
CC	Stearic content (%)	3.11	0.845 ± 0.023	0.817 ± 0.012	0.822 ± 0.011
CC	Thousand seed weight (g)	0.38	0.770 ± 0.028	0.629 ± 0.030	0.641 ± 0.030
CC	Seed yield ($T\ H^{-1}$)	2.84	0.405 ± 0.044	0.306 ± 0.020	0.386 ± 0.022
BM	Cell wall content (%)	0.29	0.091 ± 0.064	0.089 ± 0.020	0.089 ± 0.020
BM	Iodine value	0.46	0.783 ± 0.032	0.769 ± 0.018	0.769 ± 0.018
BM	Linoleic content (%)	0.46	0.756 ± 0.030	0.755 ± 0.019	0.755 ± 0.019
BM	Linolenic content (%)	0.46	0.783 ± 0.032	0.774 ± 0.018	0.773 ± 0.018
BM	Days to maturity	0.00	0.427 ± 0.050	0.432 ± 0.031	0.432 ± 0.031
BM	Oil content (%)	0.36	0.564 ± 0.047	0.494 ± 0.027	0.494 ± 0.027
BM	Oleic content (%)	0.46	0.777 ± 0.029	0.764 ± 0.018	0.764 ± 0.018
BM	Palmitic content (%)	0.46	0.803 ± 0.029	0.763 ± 0.018	0.762 ± 0.018
BM	Seeds per boll	0.55	0.366 ± 0.056	0.234 ± 0.028	0.235 ± 0.028
BM	Protein content (%)	0.00	0.215 ± 0.061	0.328 ± 0.029	0.328 ± 0.029
BM	Stearic content (%)	0.46	0.852 ± 0.028	0.822 ± 0.015	0.823 ± 0.014
BM	Straw weight (g)	0.51	0.340 ± 0.055	0.266 ± 0.035	0.266 ± 0.035
BM	Thousand seed weight (g)	0.55	0.324 ± 0.071	0.261 ± 0.028	0.261 ± 0.028
BM	Seed yield ($T\ H^{-1}$)	0.26	0.029 ± 0.061	0.016 ± 0.013	0.016 ± 0.013
EV	Cellulose content (%)	0.00	0.383 ± 0.087	0.329 ± 0.045	0.368 ± 0.043
EV	Cell wall content (%)	0.00	0.052 ± 0.099	0.034 ± 0.027	0.033 ± 0.028
EV	Fiber content (%)	0.00	0.429 ± 0.084	0.344 ± 0.045	0.368 ± 0.048
EV	Days to flowering	0.00	0.000 ± 0.083	0.024 ± 0.089	0.004 ± 0.115
EV	Plant height (cm)	0.00	0.310 ± 0.090	0.506 ± 0.047	0.330 ± 0.078
EV	Iodine value	0.00	0.916 ± 0.031	0.916 ± 0.012	0.917 ± 0.013
EV	Lignin content (%)	0.00	0.465 ± 0.086	0.304 ± 0.048	0.352 ± 0.050
EV	Linoleic content (%)	0.00	0.928 ± 0.029	0.913 ± 0.013	0.912 ± 0.013
EV	Linolenic content (%)	0.00	0.936 ± 0.026	0.918 ± 0.012	0.918 ± 0.012
EV	Days to maturity	0.00	0.230 ± 0.105	0.226 ± 0.045	0.237 ± 0.048
EV	Oil content (%)	0.00	0.587 ± 0.072	0.612 ± 0.041	0.536 ± 0.055
EV	Oleic content (%)	0.00	0.408 ± 0.092	0.419 ± 0.046	0.409 ± 0.050

Table S1. Contd.

EV	Palmitic content (%)	0.00	0.758 ± 0.048	0.775 ± 0.029	0.754 ± 0.036
EV	Protein content (%)	0.00	0.664 ± 0.056	0.680 ± 0.040	0.573 ± 0.059
EV	Shive content (%)	0.00	0.486 ± 0.086	0.262 ± 0.043	0.301 ± 0.046
EV	Stearic content (%)	0.00	0.843 ± 0.042	0.756 ± 0.030	0.773 ± 0.026
EV	Straw weight (g)	0.00	0.198 ± 0.085	0.302 ± 0.086	0.171 ± 0.086
EV	Seed yield (T H^{-1})	0.00	0.414 ± 0.098	0.227 ± 0.041	0.172 ± 0.053
SU	Plant height (cm)	0.00	0.278 ± 0.101	0.245 ± 0.055	0.245 ± 0.055
SU	Iodine value	0.00	0.934 ± 0.033	0.921 ± 0.013	0.921 ± 0.013
SU	Linoleic content (%)	0.00	0.939 ± 0.031	0.935 ± 0.011	0.935 ± 0.011
SU	Linolenic content (%)	0.00	0.939 ± 0.033	0.932 ± 0.012	0.932 ± 0.012
SU	Days to maturity	0.00	0.547 ± 0.078	0.537 ± 0.054	0.537 ± 0.054
SU	Oil content (%)	0.00	0.738 ± 0.049	0.685 ± 0.042	0.685 ± 0.042
SU	Oleic content (%)	0.00	0.752 ± 0.052	0.724 ± 0.039	0.724 ± 0.039
SU	Palmitic content (%)	0.00	0.902 ± 0.020	0.878 ± 0.020	0.878 ± 0.020
SU	Protein content (%)	0.00	0.481 ± 0.089	0.450 ± 0.054	0.450 ± 0.054
SU	Stearic content (%)	0.00	0.888 ± 0.031	0.854 ± 0.023	0.854 ± 0.023
SU	Seed yield (T H^{-1})	0.00	0.214 ± 0.105	0.136 ± 0.044	0.136 ± 0.044

MDR, Missing data rate; BM, CDC Bethune/Macbeth; EV, E1747/Viking; SU: SP2047/UGG5-5; CC, core collection. The population sizes of BM, EV, SU, and CC are 243, 86, 70, and 391, respectively. $\widehat{H}^2$(A): $\widehat{H}^2$ on a plot basis estimated by ANOVA; $\widehat{H}^2$(R): $\widehat{H}^2$on a plot basis estimated by REML. The standard error (SE) for $r_E$was estimated using the bootstrap method and the SE for $\widehat{H}^2$(A) and $\widehat{H}^2$(R) was calculated based on the Delta method implemented in You et al. (2016b).



Journal of Plant Breeding and Crop Science

Related Journals Published by Academic Journals

- *African Journal of Agricultural Research*
- *Journal of Horticulture and Forestry*
- *Journal of Cereals and Oilseeds*
- *International Journal of Livestock Production*
- *International Journal of Fisheries and Aquaculture*
- *Journal of Development and Agricultural Economics*

academicJournals