

Genetic Mapping of Seed Traits Correlated with Seed Oil Percentage in Watermelon

Geoffrey Meru¹ and Cecilia McGregor

Department of Horticulture, University of Georgia, 1111 Plant Sciences Building, Athens, GA 30602

Additional index words. kernel percentage, seed weight, seed length, seed width

Abstract. Egusi watermelon [*Citrullus lanatus* (Thunb.) Matsum. & Nakai subsp. *mucosospermus* var. *egusi* (C. Jeffrey) Mansf.] is known for its distinctive fleshy-pericarp seed phenotype and high seed oil percentage (SOP). The seed is part of the daily diet in West Africa where it is used in soups and stews or processed for cooking oil. Genetic mapping studies have revealed that most of the variation in SOP between egusi and normal, non-egusi seed is explained by the egusi (*eg*) locus, which is also associated with the unique seed phenotype. However, variation in SOP is also observed within egusi and normal seed types although the basis of this variation remains to be elucidated. A high correlation between kernel percentage (KP) and SOP has been observed in watermelon and other crops, and recent data also suggest an association between seed size and SOP in watermelon. The aim of this study was to elucidate the relationship among SOP, KP, and seed size traits in watermelon and to identify quantitative trait loci (QTL) associated with the latter traits to facilitate marker-assisted selection (MAS) for traits correlated with SOP. KP showed a significant ($\alpha = 0.05$) positive correlation with SOP in both egusi and normal seed types, whereas seed size traits showed significant negative correlations with SOP. QTL associated with KP and seed size traits in normal seed were colocalized with a previously mapped locus for SOP on linkage group (LG) 2, but in egusi seed, a QTL explaining 33% of phenotypic variation in KP was localized on LG 7. The results of this study show that SOP in watermelon is correlated with KP and seed size, but KP is associated with different loci in normal and egusi seed phenotypes.

Watermelon is an important species of the genus *Citrullus* (Robinson and Decker-Walters, 1997) known mainly for its sweet, edible flesh. However, in Asia and West Africa, watermelon seeds are consumed in a wide variety of foods including snacks and soups (Gusmini et al., 2004; Zhang, 1996). The seed of the egusi watermelon is high in oil and is characteristically large in size with a fleshy pericarp but the origin of this phenotype remains uncertain (Gusmini et al., 2004). The egusi seed is prepared for consumption by manually separating the seed-coat from the kernel after which it is eaten raw or ground into powder and added to soups and stews (Achigan-Dako et al., 2008). The seed may also be processed into cooking oil for subsistence use (Akoh and Nwosu, 1992) or used in traditional medicine as an ingredient in sedatives and antiemetics (Achigan-Dako et al., 2008).

In a study comparing SOP of 475 *C. lanatus* genebank accessions, Jarret and Levy (2012) found that the average SOP for egusi seed was higher ($n = 74$; SOP = 35.6%) than normal, non-egusi seed from *C. lanatus* var. *lanatus* ($n = 293$; SOP = 23.2%) and *C. lanatus* var. *citroides* ($n = 108$; SOP = 22.6%). Prothro

et al. (2012a) observed similar results in an F_2 mapping population and mapped the *eg* locus to LG 2 (chromosome 6; Guo et al., 2012). This locus explained 83% of the phenotypic variation for SOP observed between the egusi and normal seed types. However, in both Jarret and Levy (2012) and Prothro et al. (2012a) studies, variation in SOP was also observed within egusi (26% to 43%) and within normal (15% to 29%) seed types, pointing to the involvement of additional genetic factors.

A positive correlation between SOP and KP has been reported in crops such as cotton (*Gossypium hirsutum* and *Gossypium barbadense*) (Song and Zhang, 2007), canola (*Brassica napus*) (Yan et al., 2009), and sunflower (*Helianthus annuus*) (Tang et al., 2006). In watermelon, Jarret and Levy (2012) reported that KP was also highly correlated with SOP ($R^2 = 0.83$) and concluded that it was the thinner hulls of egusi seed relative to normal seed that were responsible for the high SOP. In sunflower (Leon et al., 1995), canola (Yan et al., 2009), and cotton (Song and Zhang, 2007), SOP is a function of kernel oil percentage (KOP), hull components, and KP and the QTL associated with these traits have been identified. However, in sunflower breeding programs, emphasis is put on KP for improvement of SOP in hybrid cultivars (Leon et al., 1995).

The role of seed size in SOP has also been studied for several crop species. In sunflower

(Tang et al., 2006), soybean (*Glycine max*) (Panthee et al., 2005) and safflower (*Carthamus tinctorius*) (Yermanos et al., 1967) a negative correlation between seed size and SOP was reported but in jatropha (*Jatropha curcas*), a positive correlation was observed (Kaushik et al., 2007). In watermelon, a seed size QTL (Prothro et al., 2012b) mapped in a similar chromosomal region as a QTL for SOP identified by Prothro et al. (2012a), pointing to a potential role for seed size in SOP.

The direct phenotypic selection for SOP and the indirect selection through correlated traits such as KP and seed size have proved to be viable options for increasing SOP in crops such as sunflower (Tang et al., 2006) and jatropha (Kaushik et al., 2007). However, phenotypic analysis of these traits can only be carried out after seed harvesting at the end of the season. Moreover, the phenotyping of traits such as KP and KOP is destructive. Therefore, the availability of molecular tools for MAS for these traits in watermelon is imperative. Through MAS, selection for SOP can be carried out at the seedling stage, thus saving time. Therefore, the aim of this study was to further elucidate the relationship among seed size traits, KP, and SOP in watermelon through phenotypic correlations and to identify genetic loci associated with the traits to facilitate MAS for higher SOP.

Materials and Methods

Plant materials and genotyping. The F_2 mapping population used by Prothro et al. (2012a) to map loci controlling the egusi seed trait and SOP was used in the current study. Briefly, single nucleotide polymorphism (SNP) genotyping assays were performed on the parents and progeny of the mapping population using an Illumina's GoldenGate SNP array and BeadStudio software (Illumina, San Diego, CA) as described in Sandlin et al. (2012). A genetic map was constructed that included 357 SNP markers spanning 14 LGs with an average gap of 4.2 cM between markers (Sandlin et al., 2012). One hundred F_2 individuals had the normal seed phenotype, whereas 42 had the egusi seed phenotype (Prothro et al., 2012a).

Phenotyping. Phenotyping was carried out on F_3 seeds derived from open-pollinated F_2 individuals. The development of this population is described elsewhere (Prothro et al., 2012a). The KP (Strain II = 44.16%, egusi = 67.31%) was measured as the weight of 15 manually dehulled seed as a percentage of the weight of the intact seed. Seed size was measured as the weight of 100 seeds (100SWT; Strain II = 4.98 g, egusi = 12.95 g), the average seed length (SL; Strain II = 8.01 mm, egusi = 15.01 mm), and seed width (SWD; Strain II = 5.30 mm, egusi = 9.15 mm) of five randomly chosen seeds from each line.

Correlations and QTL detection. Pearson correlations of the phenotypic trait values were calculated using JMP Version 9 (SAS Institute Inc., Cary, NC). The SOP values were previously determined by Prothro et al. (2012a) using nuclear magnetic resonance

Received for publication 13 Feb. 2013. Accepted for publication 14 June 2013.

¹To whom reprint requests should be addressed; e-mail merujeff@uga.edu.

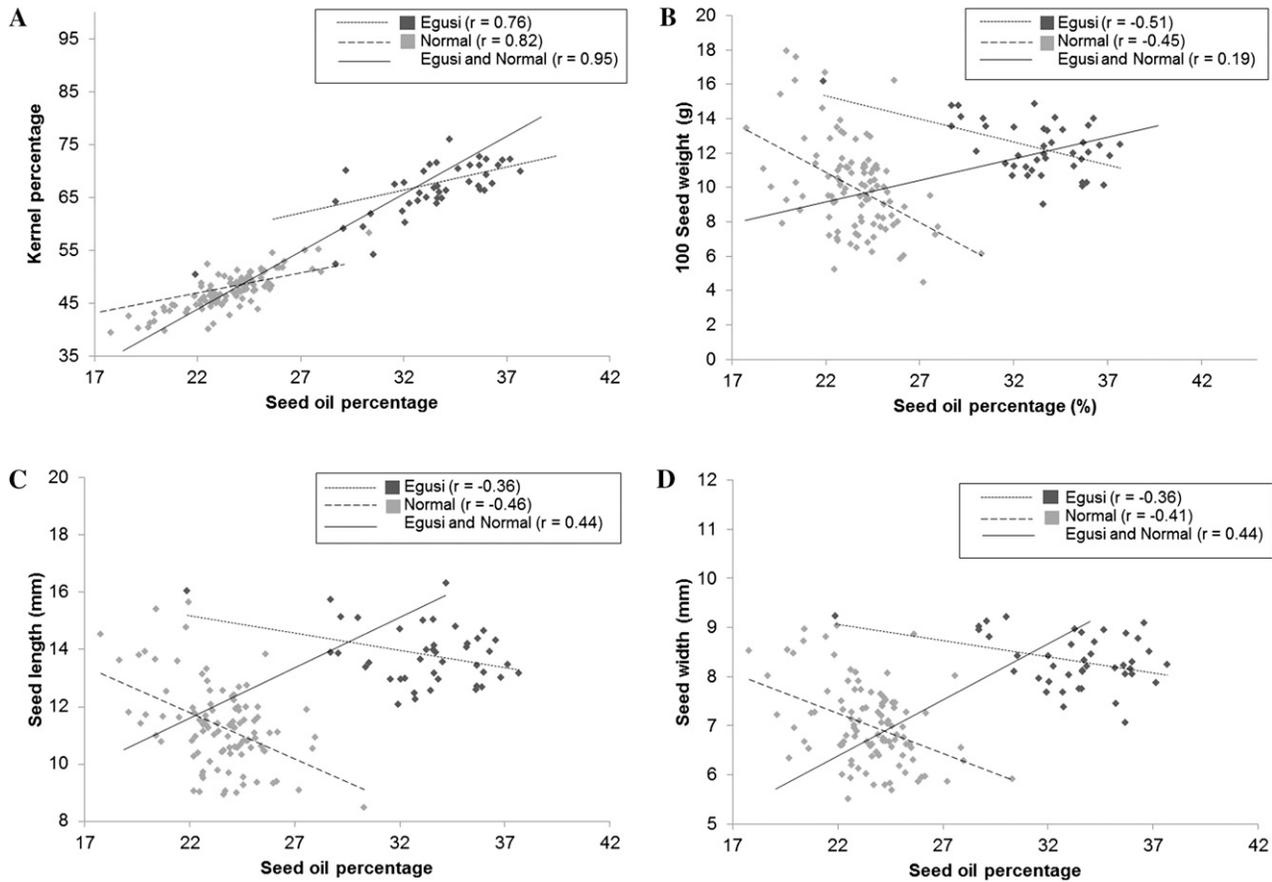


Fig. 1. Scatterplots and correlations for (A) kernel percentage, (B) 100 seed weight, (C) seed length, and (D) seed width in the normal and egusi seed types and the two seed types combined in the Strain II \times egusi F_2 population.

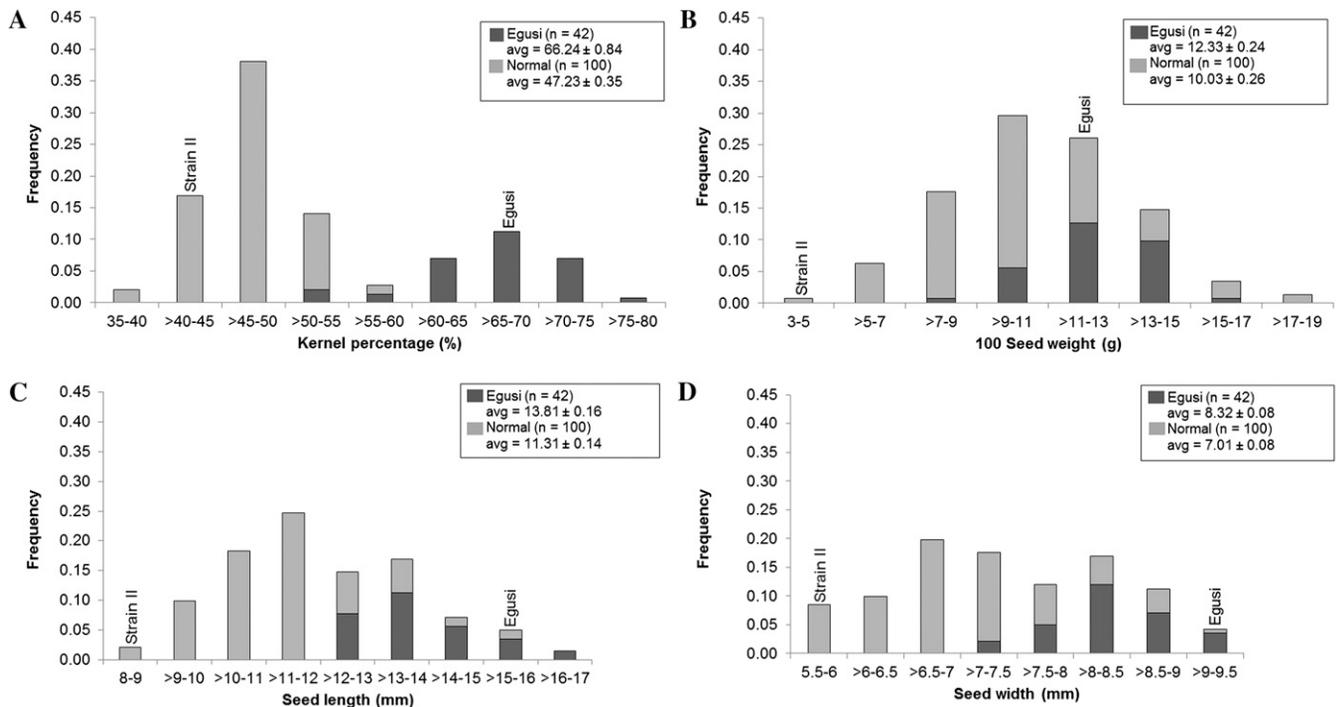


Fig. 2. Frequency distribution for (A) kernel percentage, (B) 100 seed weight, (C) seed length, and (D) seed width in the Strain II \times egusi F_2 population as well as the parental phenotypes.

(MiniSpec MQ20, NMR Analyzer; Bruker Optics Inc., Billerica, MA).

The KP was arcsine square root transformed before QTL analysis because it was

expressed as a proportion of intact seed weight (Wills et al., 2010). QTL were detected by composite interval mapping (CIM) performed with a 10-cM window in WinQTL

Cartographer Version 2.5 (Wang et al., 2011). The standard model (Model 6) with a walk speed of 1 cM was used for CIM analysis. Statistical significance of a QTL was determined

by likelihood odds thresholds set by 1000 permutations ($\alpha = 0.05$) (Churchill and Doerge, 1994). All LGs and QTL were visualized using MapChart 2.2 (Voorrips, 2002).

Results and Discussion

Phenotypic analysis of traits. When using the entire F_2 population ($n = 142$), positive correlations were observed among KP, seed size traits, and SOP (Fig. 1). However, because we were interested in the factors affecting SOP within the two seed types, the data were split and analyzed separately for the lines with egusi ($n = 42$) and normal seed ($n = 100$). Splitting the population this way eliminated the masking effect of the *eg* locus ($R^2 = 83\%$) already described by Prothro et al. (2012a).

KP ranged from 39.69% to 58.24% in normal seed and from 50.40% to 75.85% in egusi seed and followed a distribution (Fig. 2A) similar to that of SOP observed by Prothro et al. (2012a). The 100SWT for the normal seed type ranged from 4.48 to 16.70 g, whereas values for the egusi seed type ranged from 9.00 to 16.16 g (Fig. 2B). Values for SL ranged from 8.47 to 15.65 mm in normal seed and from 12.08 to 16.04 mm in egusi seed

(Fig. 2C), whereas SWD values ranged from 5.51 to 9.12 mm in normal seed and from 7.05 to 9.23 mm in egusi seed (Fig. 2D). In both egusi and normal seed types, significant ($\alpha = 0.05$) positive correlations between KP and SOP were observed supporting the results of Jarret and Levy (2012) who also found a positive correlation ($R^2 = 0.83$) between KP and SOP in egusi watermelon accessions. Significant negative correlations were observed between seed size traits and SOP in both seed types (Table 1; Fig. 1). Tang et al. (2006) reported similar results in sunflower, where seed oil content was positively correlated with KP but negatively correlated with 100SWT, SL, and SWD. These results are expected because breeding for improved seed oil content in sunflower has for many years focused on indirect selection for small seeds (Tang et al., 2006) and seeds with thinner hulls (Leon et al., 1995). Similar progress in breeding for increased SOP may be achieved for watermelon through the selection of small-sized seeds with thinner hulls in both egusi and normal seed types. Another way to increase watermelon seed oil content would be by increasing the total oil content per fruit, which would be dependent on the seed yield per fruit. On the other hand, seed yield per fruit is a function of seed size and seed number (Nerson and Paris, 2000). Increased seed yield ($\text{kg}\cdot\text{ha}^{-1}$) in the edible seed watermelon types has been achieved by breeding for increased seed width (Zhang, 1996). Nerson and Paris (2000) found no relationship between fruit weight and seed weight in watermelon but noted that very small fruits (less than 500 g) had smaller seeds in some cultivars. Therefore, fruit size may not be an important factor in determining seed yield, hence not a crucial determinant of total fruit oil content.

KP was significantly negatively correlated with SWD and 100SWT in both seed types but only to SL in normal seed (Table 1) suggesting that larger seeds have a higher hull:kernel ratio than smaller seeds. As expected, there were significant positive correlations among the seed size traits (100SWT, SL, and SWD) in this study for both seed

types, confirming the results of Prothro et al. (2012b). The high correlation (0.89 to 0.92) among seed size traits (100SWT, SL, and SWD) in the normal seed type suggests that any of these traits may be used as a measure of seed size (Poole et al., 1941).

KOP contributes to SOP and would have been a useful and more direct phenotype to measure for watermelon because seeds are usually dehulled before use. However, there was not enough seed available from each of the $F_{2:3}$ lines to measure KOP in the current study.

Detection of QTL. A total of six QTL were detected on LG 2, 6, and 7 in normal and egusi seed phenotypes (Table 2; Fig. 3). QTL for KP, 100SWT, SL, and SWD in normal seed types were colocalized on LG 2 and overlapped a QTL for SOP ($R^2 = 35.70\%$) identified for normal seed by Prothro et al. (2012a). A similar colocalization of QTL associated with SOP, KP, and seed size traits has also been observed in sunflower (Bert et al., 2003; Leon et al., 1995; Tang et al., 2006). Colocalization of QTL could indicate linkage of multiple QTL or pleiotropy (Song and Zhang, 2007) and fine mapping would be required to distinguish between the two. The QTL for seed size detected on LG 2 in normal seed type is in the same region as QTL for seed size described by Prothro et al. (2012b), confirming that this region is associated with seed size in watermelon populations from diverse genetic backgrounds. A novel QTL ($R^2 = 7.90\%$) for SL was detected on LG 6 in normal seed type (Fig. 2).

No genomic regions were associated with seed size in egusi seed phenotypes, perhaps as a result of the small sample size ($n = 42$). Population size is known to affect the ability to detect QTL, particularly those of minor effects. While studying resistance to stripe rust in barley using a double haploid population, Vales et al. (2005) demonstrated that in a small population size ($n = 50$), only major QTL could be detected, whereas both the major and minor QTL could be detected when a large population size ($n = > 300$) was used. It is therefore not clear from our results whether seed size in egusi seed is not associated with

Table 1. Pearson correlations for seed oil percentage (SOP; Prothro et al., 2012a), kernel percentage (KP), 100 seed weight (100SWT), seed length (SL), and seed width (SWD) for individuals with (A) normal seed phenotype and (B) egusi seed phenotype in the Strain II (PI 279261) \times egusi (PI 560023) F_2 population.

(A)	SOP	KP	100SWT	SL
KP	0.82*			
100SWT	-0.45*	-0.47*		
SL	-0.46*	-0.47*	0.90*	
SWD	-0.41*	-0.45*	0.89*	0.92*
(B)	SOP	KP	100SWT	SL
KP	0.76*			
100SWT	-0.51*	-0.47*		
SL	-0.36*	-0.01	0.57*	
SWD	-0.36*	-0.32*	0.68*	0.64*

*Significant at $P < 0.05$.

Table 2. Chromosomal regions associated with kernel percentage (KP), 100 seed weight (100SWT), seed length (SL), and seed width (SWD) for individuals with the normal and egusi seed phenotype in the Strain II (PI 279261) \times egusi (PI 560023) F_2 population.

Trait	LG ^z	Chr ^y	Position (cM)	Closest marker ^x	Marker position on Chr (bp) ^y	LOD ^w	R^2 (%) ^v	Additive effect ^u	Dominance effect ^u	LOD-1 support interval (cM)	LOD-1 support interval (cM)	
Normal seed												
KP	2	6	36.6	NW0250854	4,581,628–4,581,749	5.6	22.3	0.02	-0.01	11.8	48.6	
100SWT	2	6	38.6	NW0250854	4,581,628–4,581,749	17.4	62.9	-3.73	-1.01	33.5	43.8	
SL	2	6	33.6	NW0250854	4,581,628–4,581,749	23.1	60.8	-1.74	-0.34	29.6	42.1	
		6	5	88.0	NW0248236	8,808,727–8,808,847	4.2	7.9	-0.80	-0.18	83.4	89.0
SWD	2	6	36.6	NW0250854	4,581,628–4,581,749	13.6	34.4	-0.88	-0.22	32.1	41.3	
Egusi seed												
KP	7	1	102.21	NW0249392	20,067,660–20,067,779	5.83	33.8	1.25	-6.83	101.4	107.8	

^zLinkage group in Sandlin et al. (2012).

^yChromosome and marker position in the draft watermelon genome sequence (Guo et al., 2012; Ren et al., 2012).

^xMarker sequence information available in Sandlin et al. (2012).

^w Log_{10} likelihood ratio.

^vPhenotypic variation explained.

^uNegative values indicate that the effect is contributed by the allele from the egusi parent (PI 560023). For KP, the results are for the arcsine square root-transformed data.

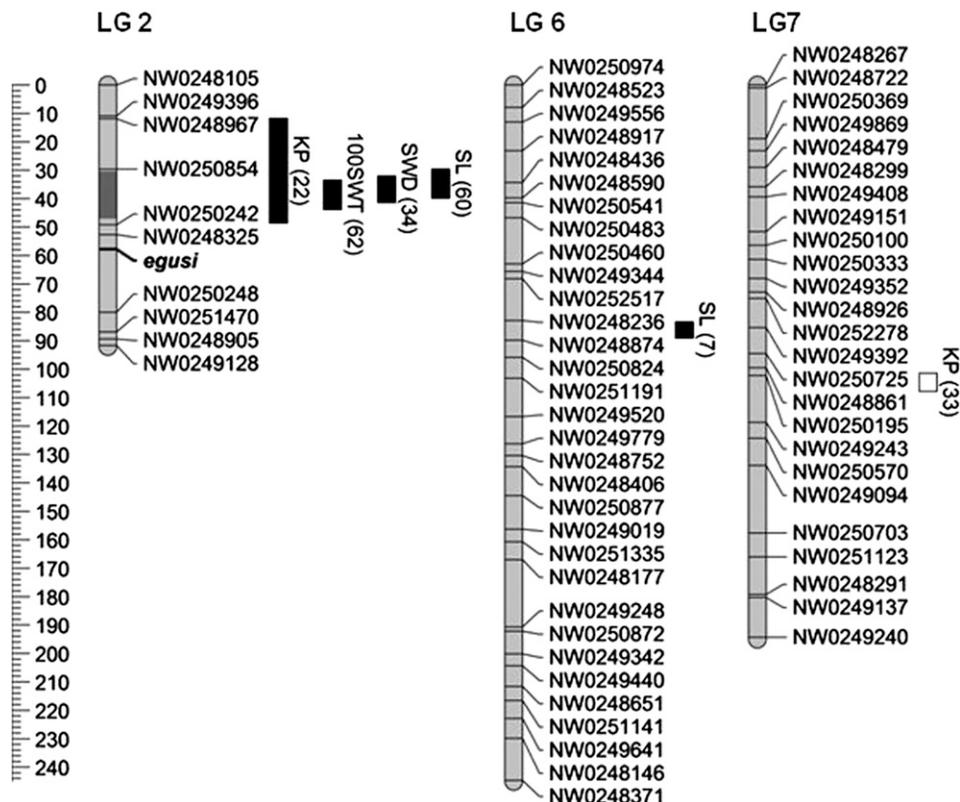


Fig. 3. Quantitative trait loci (length of the bar is equal to 1 LOD support interval) identified by composite interval mapping for kernel percentage (KP), 100 seed weight (100SWT), seed length (SL), and seed width (SWD) in the normal (black fill) and egusi (unfilled bars) seeds in the Strain II (PI 279261) × egusi (PI 560023) mapping population. Values for the phenotypic variation (R^2) for the individual QTL are given in parentheses. The location of the SOP QTL in normal seeds (shaded area on LG 2) and the egusi locus previously identified by Prothro et al. (2012a) on LG is indicated. Figure generated using MapChart Version 2.2 (Voorrips, 2002). LOD = likelihood odds; QTL = quantitative trait loci; SOP = seed oil percentage; LG = linkage group.

the QTL on LG 2 or whether we are unable to detect the effect resulting from the small population size.

A QTL for KP was detected on LG 7 ($R^2 = 33.80\%$) in the egusi seed type. Once again, it is not clear whether the small population size is preventing us from detecting the KP QTL on LG 2 or whether that region does not play a role in KP in egusi seed.

The QTL detected for KP only accounted for a relatively small proportion of the observed phenotypic variation. QTL for KP explained only 22.30% and 33.80% of the observed phenotypic variation in normal and egusi seed, respectively. QTL for 100SWT, SL, and SWD in normal seed explained 62.90%, 60.80%, and 34.40% of the observed phenotypic variation, respectively. These results suggest additional genetic factors, not identified in the present study, may be contributing to the seed traits; thus, further research is required to elucidate this matter.

This is the first report in watermelon to establish the relationship between seed size and SOP and to identify QTL associated with KP in the normal and egusi seed types. Although selection for the egusi seed phenotype alone is a good indication of high SOP, the current results indicate that further enhancement of SOP in watermelon may be achieved through indirect selection of several phenotypically correlated seed traits as has been

achieved for other crops. However, caution is to be exercised when applying MAS for KP in the egusi and normal seed types because this trait seems to be associated with different loci in the two seed types.

Literature Cited

- Achigan-Dako, E.G., R. Fagbemissi, H.T. Avohou, R.S. Vodouhe, O. Coulibaly, and A. Ahanchede. 2008. Importance and practices of egusi crops [*Citrullus Lanatus* (Thunb.) Matsum. & Nakai, *Cucumeropsis mannii* Naudin and *Lagenaria siceraria* (Molina) Standl. cv. 'Aklamkpa'] in sociolinguistic areas in Benin. *Biotechnol. Agron. Soc. Environ.* 12:393–403.
- Akoh, C.C. and C.V. Nwosu. 1992. Fatty acid composition of melon seed oil lipids and phospholipids. *J. Amer. Oil Chem. Soc.* 69:314–316.
- Bert, P.F., I. Jouan, D. Tourevieille, F. Serre, J. Philippon, P. Nicolas, and F. Vear. 2003. Comparative genetic analysis of quantitative traits in sunflower (*Helianthus annuus* L.). 2. Characterisation of QTL involved in developmental and agronomic traits. *Theor. Appl. Genet.* 107: 181–189.
- Churchill, G.A. and R.W. Doerge. 1994. Empirical threshold values for quantitative trait mapping. *Genet.* 138:963–971.
- Guo, S., J. Zhang, H. Sun, J. Salse, W.J. Lucas, H. Zhang, Y. Zheng, L. Mao, Y. Ren, Z. Wang, J. Min, X. Guo, F. Murat, B.K. Ham, Z. Zhang, S. Gao, M. Huang, Y. Xu, S. Zhong, A. Bombarely, L.A. Mueller, H. Zhao, H. He, Y. Zhang, Z. Zhang, S. Huang, T. Tan, E. Pang, K. Lin, Q. Hu, H. Kuang, P. Ni, B. Wang, J. Liu, Q. Kou, W. Hou, X. Zou, J. Jiang, G. Gong, K. Klee, H. Schoof, Y. Huang, X. Hu, S. Dong, D. Liang, J. Wang, K. Wu, Y. Xia, X. Zhao, Z. Zheng, M. Xing, X. Liang, B. Huang, T. Lv, J. Wang, Y. Yin, H. Yi, R. Li, M. Wu, A. Levi, X. Zhang, J.J. Giovannoni, J. Wang, Y. Li, Z. Fei, and Y. Xu. 2012. The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nature Genet.* 45:51–58.
- Gusmini, G., T. Wehner, and R. Jarret. 2004. Inheritance of egusi seed type in watermelon. *J. Hered.* 95:268–270.
- Jarret, R. and I. Levy. 2012. Oil and fatty acid contents in seed of *Citrullus lanatus* Schrad. *J. Agr. Food Chem.* 60:5199–5204.
- Kaushik, N., K. Kumar, S. Kumar, and S. Roy. 2007. Genetic variability and divergence studies in seed traits and oil content of *Jatropha (Jatropha curcas* L.) accessions. *Biomass Bioenergy* 31:497–502.
- Leon, A., M. Lee, G. Rufener, S. Berry, and R. Mowers. 1995. Use of RFLP markers for genetic linkage analysis of oil percentage in sunflower seed. *Crop Sci.* 35:558–564.
- Nerson, H. and H.S. Paris. 2000. Relationship between fruit size and seed size in cucurbits. *Cucurbit Genet. Coop. Rpt.* 23:64–67.
- Panthee, D., V. Pantalone, D. West, A. Saxton, and C. Sams. 2005. Quantitative trait loci for seed protein and oil concentration, and seed size in soybean. *Crop Sci.* 45:2015–2022.
- Poole, C.F., P.C. Grimball, and D.R. Porter. 1941. Inheritance of seed characters in watermelon. *J. Agr. Res.* 63:433–456.

- Prothro, J., K. Sandlin, R. Gill, E. Bachlava, V. White, S. Knapp, and C. McGregor. 2012a. Mapping of the egusi seed trait locus (*eg*) and quantitative trait loci associated with seed oil percentage in watermelon. *J. Amer. Soc. Hort. Sci.* 137:311–315.
- Prothro, J., K. Sandlin, H. Abdel-Haleem, E. Bachlava, V. White, S. Knapp, and C. McGregor. 2012b. Main and epistatic quantitative trait loci associated with seed size in watermelon. *J. Amer. Soc. Hort. Sci.* 137:452–457.
- Ren, Y., H. Zhao, Q. Kou, J. Jiang, S. Guo, H. Zhang, W. Hou, X. Zou, H. Sun, G. Gong, A. Levi, and Y. Xu. 2012. A high resolution genetic map anchoring scaffolds of the sequenced watermelon genome. *PLoS One* 7:e29453.
- Robinson, R.W. and D. Decker-Walters. 1997. *Cucurbits*. CAB International Publishing, Wallingford, UK.
- Sandlin, K., J. Prothro, A. Heesacker, N. Khalilian, R. Okashah, W. Xiang, E. Bachlava, D. Caldwell, D. Seymour, V. White, E. Chan, G. Tolla, C. White, D. Safran, E. Graham, S. Knapp, and C. McGregor. 2012. Comparative mapping in watermelon [*Citrullus lanatus* (Thunb.) Matsum. et Nakai]. *Theor. Appl. Genet.* 125:1603–1618.
- Song, X.L. and T.Z. Zhang. 2007. Identification of quantitative trait loci controlling seed physical and nutrient traits in cotton. *Seed Sci. Res.* 17:243–252.
- Tang, S., A. Leon, W.C. Bridges, and S.J. Knapp. 2006. Quantitative trait loci for genetically correlated seed traits are tightly linked to branching and pericarp pigment loci in sunflower. *Crop Sci.* 46:721–734.
- Vales, M., C. Schön, F. Capettini, X. Chen, A. Corey, D.E. Mather, C.C. Mundt, K.L. Richardson, J.S. Sandoval-Islas, H.F. Utz, and P.F. Hayes. 2005. Effect of population size on the estimation of QTL: A test using resistance to barley stripe rust. *Theor. Appl. Genet.* 111:1260–1270.
- Voorrips, R.E. 2002. MapChart: Software for the graphical presentation of linkage maps and QTLs. *J. Hered.* 93:77–78.
- Wang, S., C.J. Basten, and Z.B. Zeng. 2011. *Windows QTL Cartographer 2.5*. Department of Statistics, North Carolina State University, Raleigh, NC.
- Wills, D.M., H. Abdel-Haleem, S.J. Knapp, and J.M. Burke. 2010. Genetic architecture of novel traits in the Hopi sunflower. *J. Hered.* 101:727–736.
- Yan, X., J. Li, F. Fu, M. Jin, L. Chen, and L. Liu. 2009. Co-location of seed oil content, seed hull content and seed coat color QTL in three different environments in *Brassica napus* L. *Euphytica* 170:355–364.
- Yermanos, D., S. Hemstreet, and M. Garber. 1967. Inheritance of quality and quantity of seed-oil in safflower (*Carthamus tinctorius* L.). *Crop Sci.* 7:417–422.
- Zhang, J. 1996. Breeding and production of watermelon for edible seed in China. *Cucurbit Genet. Coop. Rpt.* 19:66–67.