Associating Molecular Markers with Virus Resistance to Classify Sweetpotato Genotypes

M. Mcharo and D. LaBonte
Department of Horticulture, Louisiana State University Agricultural Center, Louisiana Agricultural Experiment Station, Baton Rouge, LA 70803

R.O.M. Mwanga
National Agricultural Research Organization (NARO), Namulonge Agricultural and Animal Production Research Institute (NAARI), P.O. Box 7084, Kampala, Uganda

A. Kriegner
ARC Seibersdorf Research GmbH, Molecular Diagnostics, A-2444 Seibersdorf, Austria

Abstract. Molecular markers linked to resistance to sweetpotato chlorotic stunt closterovirus [SPCSV (genus *Crinivirus*, family Closteroviridae)] and sweetpotato feathery mottle virus [SPFMV (genus *Potyvirus*, family Potyviridae)] were selected using quantitative trait loci (QTL) analysis, discriminant analysis and logistic regression. Eighty-seven F_1 sweetpotato [*Ipomoea batatas* (L.) Lam.] genotypes from a cross of 'Tanzania' and 'Wagabolige' landraces were used to generate DNA marker profiles for this study. Forty-five of the clones were resistant to SPCSV while 37 were resistant to SPFMV. A combination of 232 amplified fragment length polymorphism (AFLP) markers and 37 random amplified polymorphic DNA (RAPD) markers obtained were analyzed to determine the most informative markers. All three statistical procedures revealed that AFLP marker e41m33.a contributed the greatest variation in SPCSV resistance and RAPD marker S13.1130 accounted for most of the variation in SPFMV resistance. The power of discriminant and logistic analyses is that you do not need a parent-progeny population. An evaluation of these two models indicated a classification and prediction accuracy rates of 96% with as few as four markers in a model. Both multivariate techniques identified one important discriminatory marker (e44m41.j) for SPCSV and two markers (e41m37.a and e44m36.d) for SPFMV that were not identified by QTL analysis.

The most widely used approach to germplasm improvement in sweetpotato is mass selection (Jones, 1986). This fundamentally simple technique lends itself well to an allohexaploid crop that possesses few, if any, simply inherited traits (Buteler et al., 1999; Jones, 1986; Zhang et al., 2001). Unfortunately, self-incompatibility prevents quick introgression of desirable traits from landraces or exotics into adapted material. The preponderance of quantitatively inherited traits in sweetpotato (Jones, 1986), environmentally unstable trait expression, and the encumbrance of 2-year breeding cycles underlie recent attempts to use trait-linked DNA molecular markers in breeding programs.

The most compelling reason to develop trait-linked DNA markers in sweetpotato is to effectively breed for virus resistance. Viral diseases have been identified to cause up to 56% to 98% yield reduction in sweetpotato (Gibson et al., 1997; Ngeve, 1990). According to Geddes (1990) the sweetpotato virus disease (SPVD), a result of the co-infection of SPCSV and SPFMV, is the most destructive of the viral diseases within the African continent (Gibson et al., 1997; Karyeija et al., 1998). Hahn et al. (1981) and Mwanga et al. (2002a) estimated high broad-sense heritability for resistance to SPVD, suggesting that breeding for resistance is possible.

Ukoskit and Thompson (1997) and later Kriegner et al. (2003) demonstrated the feasibility of constructing a genetic linkage map for sweetpotato. Mapping was based on the single-dose restriction fragment (SDRF) method described by Wu et al. (1992). Progeny segregating in a 1 present :1 absent marker ratio arise from a simplex x homozygous recessive test cross (e.g., aaaaaa x Aaaaaa). Progeny segregating in 4:1 ratio that generate duplex markers AAaaaa are also used. A map by Kriegner et al. (2003) contained 1067 AFLP markers ordered in ~90 linkage groups with genome coverage suitable for detection of QTL. Mwanga et al. (2002a) used linkage maps to associate DNA RAPD and AFLP markers to SPFMV and SPCSV resistance. Attempts to identify markers associated with SPVD were unsuccessful because few progeny in the defined population were resistant. Despite such efforts, QTL analysis is still a challenge in breeding programs for a crop like the sweetpotato since a complete genetic map is still unavailable.

Faced with the complications associated with QTL analysis, this study seeks to determine whether the results of molecular marker and phenotype association for genotype characterization and disease resistance prediction using multivariate statistical procedures like discriminant analysis and logistic regression are comparable to those of QTL analysis. Both permit assemblage of test populations (i.e., resistant and susceptible classes, without regard to genetic relationships).

Discriminant analysis is a multivariate technique that uses multiple predictor variables to associate an individual with a descriptive class (Fisher, 1936). In our genetic studies of sweetpotato, discriminant analysis would involve relating a genotype,

as described by its molecular marker profile, with its disease resistance class or phenotypic group. Using the markers as the multiple variables, the method would develop an optimum classification function that maximizes differences between the resistant and susceptible classes while minimizing differences among members of those two classes. The efficiency of the function is evaluated by cross-validation as described by Lachenbruch and Mickey (1968). Cross-validation is a leave-one-out method that aims to reduce the optimistic bias in calculating classification error-rate. If the assumption of multivariate normality is not met or when no assumptions are valid about the distribution within each phenotypic group, a nonparametric procedure like k-nearest-neighbor (Rosenblatt, 1956) is used for classification. With the k-nearest-neighbor method, a Mahalanobis distance is calculated between group means and this distance is then used to determine an individual's closeness to the mean. All n observations in the dataset are used to calculate distances; however, the observation to be classified is left out from the k-nearest-neighbors of that observation during function development.

Recent work on genetic analysis using discriminant analysis include that of Capdevielle et al. (2000), who associated microsatellites and agronomic traits in rice. Fahima et al. (2002), who investigated microsatellite polymorphism in wheat, also used discriminant analysis in their characterization.

Logistic regression is a statistical technique that is frequently used to associate explanatory variables with a binary outcome (Ostir and Uchida, 2000). When there are more than two response classes, discriminant analysis is more reliable than logistic regression. Models developed using logistic regression have the advantage of summarizing and interpreting the data and hence the researcher can make inferences on factors being investigated. In our studies, the process involved relating molecular markers as the independent factors and presence or absence of viral infection as the dichotomous response. To our knowledge there is no work on logistic regression for molecular marker selection in sweetpotato. However, field disease prediction studies have extensively used logistic regression tools. Mukasa et al. (2003) used logistic regression to determine the odds ratios to detect single- or multiple-virus infections in sweetpotato genotypes.

The specific objectives of this study were to compare marker–trait association using QTL analysis, discriminant analysis, and logistic regression for DNA markers derived from a population of sweetpotato genotypes and to develop and validate genotype classification and resistance class prediction models for SPCSV and SPFMV.

Materials and Methods

Molecular marker detection and phenotypic data acquisition for the population was previously described by Mwanga et al. (2002b). The data consisted of an array of molecular marker and phenotypic information for 87 F_1 genotypes derived from a cross between 'Tanzania' and 'Wagabolige' cultivars. Both 'Tanzania' and 'Wagabolige' are resistant to SPVD, although 'Wagabolige' is more resistant to SPVD than 'Tanzania' under field SPVD inoculum pressure (Mwanga et al., 2002b). Detection and confirmation of virus presence or absence in the genotypes involved using triple-antibody sandwich enzyme-linked immunosorbent assay (TAS-ELISA) test for SPCSV and nitrocellulose membrane enzyme-linked immunosorbent assay (NCM-ELISA) test for SPFMV (International Potato Center, 1990). Resistant genotypes tested negative while susceptible genotypes tested

positive for virus presence (Mwanga et al., 2002b). 'Tanzania' is widely grown in sub-Saharan Africa while 'Wagabolige' is more confined to Uganda. The 87 F_1 genotypes had 269 dominant molecular markers, 232 of them being AFLP fragments and 37 being RAPD fragments. These were two markers less than what Mwanga et al. (2002b) used. Fifty genotypes were classified as resistant and 37 susceptible to SPFMV. Forty-two were classified as resistant and 45 susceptible to SPCSV, respectively. Phenotypic data for SPCSV and SPFMV scores were obtained using ELISA tests and were coded as a one (1) for positive virus infection or zero (0) for negative results.

Statistical analyses were done on SPFMV and SPCSV only because only three out of 87 genotypes were resistant to SPVD. These were too few for the statistical procedures to converge and give a solution.

QTL analysis was done using QGENE (Nelson, 1997). Single-point analysis of variance (ANOVA), or regression on a marker (Soller et al., 1976), was conducted to reveal significant ($P < 0.05$) associations between markers and virus resistance. Markers were considered independent variables and virus resistance score as the dependent variable. Proportion of the total phenotypic variance accounted for by the markers was estimated by the coefficient of determination (R^2) obtained from simple linear regression (Nelson, 1997; Soller et al., 1976).

Step discriminant analysis using the STEPDISC procedure (SAS, 2001) was used to select the most informative markers from the original set of markers. The progeny population was divided into two phenotypic groups or classes based on resistance–susceptible reactions, similar to QTL analysis. The forward selection option in STEPDISC was used to select markers to be included in the classification model. As described in SAS (1999), the forward selection process commences with no markers in the model. Entry-significant levels of $P \le 0.03$ and $P \le 0.02$ for SPCSV and SPFMV, respectively, of the chi-square score for entering an effect or marker into the model to achieve at least 95% prediction accuracy were imposed to choose the most discriminating markers (SAS, 1999). At each step the procedure enters the marker that contributes most to the discriminatory power of the model as measured by Wilks' lambda, the likelihood ratio criterion. The process stops when all unselected markers fail to meet the entry condition. The selected markers were then used in a nonparametric discriminant analysis ($k = 1$), DISCRIM option (SAS, 2001), to construct and validate a class prediction function and to predict group membership, resistant or susceptible, of the test. Since the categorical nature of the data could not allow assumption of normal distribution to be made a nonparametric method, the k-nearest-neighbor method (Rosenblatt, 1956) was used to estimate the group-specific densities that produce a classification criterion. The performance of the discriminant criterion was evaluated by group-specific error-count estimates during cross-validation. The error estimator gives the proportion of misclassified observations in each group. Total error, from which we derive percent correct classification, is the weighted mean of error estimates of the two phenotypic groups.

PROC LOGISTIC (SAS, 2001) was used to perform logistic regression with the logit link to select markers that significantly accounted for phenotypic variation, with the forward selection option used for marker selection. With forward selection, PROC LOGISTIC first estimates parameters for effects, markers in our case, forced into the model. These effects are the intercept and the first n explanatory effects in the model; n is zero by default. Next, the procedure computes the score chi-square statistic for

356

J. Amer. Soc. Hort. Sci. 130(3):355–359. 2005.

each effect not in the model and examines the largest of these statistics. If it is significant at a preset probability level, the corresponding effect is added to the model. Once an effect is entered in the model, it is never removed from the model. The process is repeated until none of the remaining effects meet the specified level for entry. For our study entry significant levels of $P \leq 0.05$ and $P \leq 0.02$ for SPCSV and SPFMV to enter marker into model were set. The phenotype was a binary outcome as either resistant or susceptible. The Hosmer and Lemeshow goodness-of-fit test was used to determine model efficiency (Hosmer et al., 1991).

Results and Discussion

SWEETPOTATO CHLOROTIC STUNT VIRUS. All three statistical procedures selected two common markers (e41m33.a and e38m36.u) that had a significant association with resistance to SPCSV (Tables 1, 2, and 3). These results are in concordance with those of Mwanga et al. (2002b). Marker e38m36.u was selected by discriminant and logistic but not QTL procedures as one of the informative markers. Marker e41m33.a was the most informative, regardless of the method of selection used, accounting for 66.75% of the variation in SPCSV resistance according to QTL analysis. Our results suggest the possibility of using these multivariate methods for identifying powerful QTLs. With the six markers selected by discriminant analysis to create a classification model, we achieved 96% correct classification for SPCSV. Discriminant analysis misclassified one genotype out of 42 that were phenotypically classified as being resistant. This model also misclassified into the resistant class three genotypes out of 45 that were visually described as susceptible. The model developed using logistic regression fitted the data well according to the Hosmer and Lemeshow (Hosmer et al., 1991) goodness-of-fit test ($\chi^2 = 0.082$; df = 3; $P = 0.9939$). Table 4 shows observed vs. expected class values for SPCSV and SPFMV.

SWEETPOTATO FEATHERY MOTTLE VIRUS. Marker S13.1130 was identified by all three selection procedures and accounted for 64.62% of the variation in SPFMV resistance (Tables 1, 2, and 3) according to QTL analysis. Marker e39m32.f was the second most powerful marker selected by QTL analysis but neither by discriminant nor logistic analyses. Markers e41m37.a and e44m36.d were selected as being informative by discriminant and logistic analyses only (Tables 2 and 3) but in differing order of importance. The discriminant analysis model constructed using the four selected markers achieved a 96% correct classification rate. Four clones out of 50 that were phenotypically classified as being resistant were classified by discriminant analysis as being susceptible. These are the clones that discriminant analysis misclassified and they give us the total error rate of the model developed. This model achieved a highly accurate classification rate and with only four variables was a desirable result because we aim to have as few markers as possible to create an efficient class prediction model. Results of Hosmer and Lemeshow model evaluation indicate that the logistic model fitted the data well ($\chi^2 = 0.071$; df = 2; $P = 0.9951$) and the proximity of expected values to observed values (Table 4) confirm this.

QTL analysis, while providing useful results, is limited by the fact that it can only be used in a population of individuals that are closely related. Furthermore, results from QTL analysis have in the past been inapplicable in other populations. Lubberstedt et al. (1998) in their work on maize reported that QTL results were not consistent among crosses within the flint heterotic pool, suggesting that prior to marker-assisted selection, QTL mapping

Table 1. QTL analysis for DNA markers associated with ≥10% resistance to sweetpotato chlorotic stunt closterovirus (SPCSV) and sweetpotato feathery mottle virus (SPFMV) in sweetpotato.

Disease	Marker	F	R^2	P
SPCSV	e41m33.a	170.66	0.67	<0.0001
	e40m34.c	111.16	0.57	<0.0001
	e38m36.u	19.24	0.18	<0.0001
	e36m49.a	14.42	0.14	0.0003
	e35m49.d	12.58	0.13	0.0006
	R9.650	9.41	0.10	0.0029
	e39m40.e	9.23	0.10	0.0032
SPFMV	S13.1130	155.28	0.65	<0.0001
	e39m32.f	52.49	0.38	<0.0001
	e36m59.a	10.53	0.11	0.0017
	e33m59.a	9.76	0.10	0.0024

Table 2. STEPDISC marker selection for DNA markers associated with ≥10% resistance to sweetpotato chlorotic stunt closterovirus (SPCSV) and sweetpotato feathery mottle virus (SPFMV) in sweetpotato.

Disease	Marker	Wilk's lamda	P
SPCSV	e41m33.a	0.33	<0.001
	e38m36.u	0.28	<0.001
	K2.650	0.26	<0.001
	AL3.1300	0.24	<0.001
	e44m.41.j	0.23	<0.001
	e39m.41f	0.22	<0.001
SPFMV	S13.1130	0.34	<0.001
	e41m37.a	0.31	<0.001
	e40m36.d	0.29	<0.001
	e44m36.d	0.26	<0.001

Table 3. Logistic regression selection for DNA markers associated with ≥10% resistance to sweetpotato chlorotic stunt closterovirus (SPCSV) and sweetpotato feathery mottle virus (SPFMV) in sweetpotato.

Disease	Marker selected	χ^2 score	P
SPCSV	e41m33.a	58.49	<0.0001
	e38m36.u	11.36	0.0007
	H14.700	7.44	0.0064
	e44m41.j	6.61	0.0101
	e40m41.c	5.71	0.0169
	B7.1130	5.47	0.0194
	e32m60.g	5.31	0.0211
SPFMV	S13.1130	57.38	<0.0001
	e40m38.b	8.90	0.0029
	e44m36.d	8.56	0.0034
	e41m37.a	7.56	0.0060
	e39m45.c	6.98	0.0083
	e39m39.e	6.00	0.0143

Table 4. Partition for the Hosmer and Lemeshow test for sweetpotato chlorotic stunt closterovirus (SPCSV) and sweetpotato feathery mottle virus (SPFMV) in sweetpotato.

Disease	Observed clones (no.)	Expected clones (no.)	Disease resistance group
SPCSV	45	44.98	1 (present)
SPCSV	42	42.02	0 (absent)
SPFMV	37	36.98	1 (present)
SPFMV	50	50.02	0 (absent)

must be performed separately for each population. In a review of QTL methods in plants, Kearsey and Farquhar (1998) showed that analytical methods locate QTL with poor precision [10–30 centimorgans (cM)], unless the heritability of an individual QTL is high. They further stated that this unreliability of QTL location may suggest false candidate genes. In his review of QTL analysis, Young (1999) states that phenotypic mis-scoring of even a few individuals can totally confound QTL discovery and placement. He also suggests that nonparametric categorical data may not be as amenable to QTL analysis as normally distributed quantitative data, thus underscoring the difficulties associated with using QTL for marker-assisted selection.

Mwanga et al. (2002b) reported studies combining both quantitative genetics and molecular marker investigations in an attempt to describe inheritance of resistance to SPCSV, SPFMV, and SPVD. However, these workers were faced with the undesirable difficulty of constructing genetic maps using null alleles and having insufficient plant numbers in the SPVD resistant category. In their QTL studies on molecular markers, Mwanga et al. (2002b) found that marker e41m33.a was responsible for 70% of the variation in SPCSV resistance. They also found that marker S13.1130 accounted for 72% of the resistance to SPFMV. Their conclusions were based on single-point ANOVA and multiple-regression analyses with two markers in the model. Unfortunately, a lack of progeny in the SPVD resistant class prevented disclosure of markers linked to SPVD resistance.

We attribute the slight differences between our QTL results and those of Mwanga et al. (2002b) to the fact that the data set that we used contained two markers less. Nonetheless our conclusions concur with theirs and do confirm the power of markers e41m33.a and S13.1130 to influence resistance to SPCSV and SPFMV, respectively.

Discriminant and logistic regression analyses have greater commonality with each other than with QTL analysis as far as markers selected is concerned possibly because the former two are both multivariate techniques. These multivariate techniques determine the power of a marker while other markers are still in the model hence giving a more accurate comparative analysis compared to the single point ANOVA for QTL analysis. Single point ANOVA is also referred to as regression on a marker and it assumes the presence of a single QTL on the genome (Broman, 2001). The ANOVA calculates the variance at marker loci without considering the effect of the other markers when they act together, hence the value obtained may not be the true effect of the QTL. In his study on genetic mapping of agronomic traits from the interspecific cross of *Oryza sativa* L. and *O. glaberrima* Steud., Aluko (2003) concluded that discriminant analysis was superior to interval mapping in selecting markers that could be used in grouping doubled haploid lines of rice into predefined groups.

We suggest that logistic regression is an increasingly valuable tool in molecular marker selection protocols in a sweetpotato breeding program. Investigators in the medical sciences have exploited the power of logistic regression to select important markers. In a study involving boars (*Sus scrofa* L.), Thurston et al. (2002) identified 16 candidate genetic markers ($P \leq 0.005$) by comparing the AFLP profile with semen freezability trait using logistic regression analysis. These findings support the hypothesis that there is a genetic basis for variation in postthaw semen quality between individuals, and that AFLP technology may be able to identify molecular markers linked to genes influencing this variation. Dunsmuir et al. (2000) used logistic regression

modeling to identify three immunohistochemistry markers that were statistically significant predictors of the metastatic status (M-stage, bone metastasis vs. no bone metastasis) of prostate cancer in men.

Results from our work suggest that it is possible to use logistic and discriminant multivariate techniques to select powerful markers that may be useful to breeders. While multivariate techniques may not replace QTL analysis, the flexibility of assembling test populations, regardless of relatedness, is particularly useful when progeny from a controlled cross rarely arise as resistant (e.g., SPVD progeny). Our results demonstrated that multivariate techniques can achieve results that are comparable to those of QTL analysis, even in the absence of a genetic map, and identify markers not disclosed through QTL analysis.

Literature Cited

Aluko, G.K. 2003. Genetic mapping of agronomic traits from the interspecific cross of *Oryza sativa* L. and *Oryza glaberrima* Steud. PhD Diss., Louisiana State Univ., Baton Rouge. ETD-1110103–143019.

Broman, K. 2001. Review of statistical methods for QTL mapping in experimental crosses. Lab Animal 30:44–52.

Buteler, M.I., R.L. Jarret, and D.R. La Bonte. 1999. Sequence characterization of microsatellites in diploid and polyploid *Ipomoeas*. Theor. Appl. Genet. 99:123–132.

Capdevielle, F.M., G.K. Aluko, M. Balzarini, and J.H. Oard. 2000. Application of molecular markers and discriminant analysis to identify rice lines with contrasting phenotypes for agronomic traits, p. 216. In: G.S. Khush, D.S. Brar and B. Hardy (eds.). Proc. Fourth Intl. Rice Genet. Symp., Intl. Rice Res. Inst., Los Banos, Philippines. (Abstr.)

Dunsmuir, W.D., C.E. Gillett, L.C. Meyer, M.P. Young, C. Corbishley, R.A. Eeles, and R.S. Kirby. 2000. Molecular markers for predicting prostate cancer stage and survival. BJU Intl. 86:869–878.

Fahima, T., M.S. Roder, K. Wendehake, V.M. Kirzhner, and E. Nevo. 2002. Microsatellite polymorphism in natural populations of wild emmer wheat, *Triticum dicoccoides*, in Israel. Theor. Appl. Genet. 104:17–29.

Fisher, R.A. 1936. The use of multiple measurements in taxonomic problems. Ann. Eugenics 7:179–188.

Geddes, A.M.W. 1990. The relative importance of crop pests in sub-Saharan Africa. Natural Resources Inst. Bul. No. 36. NRI, Kent, U.K.

Gibson, R.W., R.O.M. Mwanga, S. Kasule,, I. Mpembe, and E.E. Carey. 1997. Apparent absence of viruses in most symptomless field grown sweetpotato in Uganda. Ann. Appl. Biol. 130:481–490.

Hahn, S.K., E.R. Terry and K. Leuschener. 1981. Resistance of sweetpotato to virus complex. HortScience 16:535–537.

Hosmer, D.W., S. Taber, and S. Lemeshow. 1991. The importance of assessing the fit of logistic regression models: A case study. Amer. J. Public Health 81(12):1630–1635.

International Potato Center. 1990. Annual report for 1990. Intl. Potato Ctr., Lima, Peru.

Jones, A. 1986. Sweet potato heritability estimates and their use in breeding. HortScience 21:14–17.

Karyeija, F., R.W. Gibson, and J.P.T Valkonen. 1998. Resistance to sweetpotato virus disease (SPVD) in wild East African *Ipomoea*. Ann. Appl. Biol. 133:39–44.

Kearsey, M.J. and A.G.L. Farquhar. 1998. QTL analysis in plants; where are we now? Heredity 80:137–142.

Kriegner, A., J.C. Cervantes, K. Burg, R.O.M. Mwanga, and D. Zhang. 2003. A genetic linkage map of sweetpotato [*Ipomoea batatas* (L.) Lam.] based on AFLP markers. Mol. Breeding 11:169–185.

Lachenbruch, P.A. and M.A. Mickey. 1968. Estimation of error rates in discriminant analysis. Technometrics 10:1–10.

Lubberstedt T., A.E. Melchinger, S. Fahr, D. Klein, A. Dally, and P. Westhoff. 1998. QTL mapping in testcrosses of flint lines of maize: III. Comparison across populations for forage traits. Crop Sci. 38:1278–1289.

358

J. Amer. Soc. Hort. Sci. 130(3):355–359. 2005.

Mukasa, S.B., P.R. Rubaihayo, and J.P.T. Valkonen. 2003. Incidence of viruses and viruslike diseases of sweetpotato in Uganda. Plant Dis. 87:329–335.

Mwanga, R.O.M., G.C. Yencho, and J.W. Moyer. 2002a. Diallel analysis of sweetpotatoes for resistance to sweetpotato virus disease. Euphytica 128:237–248.

Mwanga, R.O.M., A. Kreigner, J.C. Cervantes-Flores, D.P. Zhang, J.W. Moyer, and G.C Yencho. 2002b. Resistance to sweetpotato chlorotic stunt virus and sweetpotato feathery mottle virus is mediated by two separate recessive genes in sweetpotato. J. Amer. Soc. Hort. Sci. 127:798–806.

Nelson, J.C. 1997. QGENE: Software for marker-based genomic analysis and breeding. Mol. Breeding 3:239–245.

Ngeve, J.M. 1990. Yield stability and yield depression in sweetpotato cultivars susceptible to the sweetpotato virus disease. J. Hort. Sci. 65:225–230.

Ostir, G.V. and T. Uchida. 2000. Logistic regression: A non-technical review. Amer. J. Phys. Med. Rehab. 79:565–572.

Rosenblatt, M. 1956. Remarks on some nonparametric estimates of a density function. Ann. Math. Stat. 27:832–837.

SAS Institute Inc. 1999. SAS/STAT User's guide. SAS Inst., Cary, N.C.

SAS Institute. 2001. The SAS system for windows v8. SAS Inst., Cary, N.C.

Soller, M., A. Genizi, and T. Brody. 1976. On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. Theor. Appl. Genet. 47:35–39.

Thurston L.M., K. Siggins, A.J. Mileham, P.F. Watson, and W.V. Holt. 2002. Identification of amplified restriction fragment length polymorphism markers linked to genes controlling boar sperm viability following cryopreservation. Biol. Reprod. 66:545–554.

Ukoskit, K. and P.G. Thompson. 1997. Autoploidy vs. allopolyploidy and low-density randomly amplified polymorphic DNA linkage maps of sweetpotato. J. Amer. Soc. Hort. Sci. 122:822–828.

Wu, K.K., W. Burnquish, M.E. Sorrells, T.L. Pew, P.H. Moore, and S.D. Tanksley. 1992. The detection and estimation of linkage in polyploids using single-dose restriction fragments. Theor. Appl. Genet. 83:294–300.

Young, N. D. 1999. A cautiously optimistic vision for marker-assisted breeding. Mol. Breeding 5:505–510.

Zhang, D.P., D. Carbajulca, L. Ojeda, G. Rossel, S. Milla, and C. Herrera, and M. Ghislain. 2001. Microsatellite analysis of genetic diversity in sweetpotato varieties from Latin America, p.295–301. In: Program report 1999–2000, Intl. Potato Ctr., Lima, Peru.