



A genetic linkage map of sweetpotato [*Ipomoea batatas* (L.) Lam.] based on AFLP markers

Albert Kriegner^{1,*}, Jim Carlos Cervantes², Kornel Burg¹, Robert O.M. Mwangi³ and Dapeng Zhang²

¹Environment and Life Science, Dept. of Life Science/Biotechnology, Austrian Research Centers Seibersdorf

(ARCS), Seibersdorf, A-2444, Austria; ²International Potato Center (CIP), P.O. Box 1558, Lima, Peru;

³Namulonge Agricultural and Animal Production Research Institute (NAARI), National Agricultural Research Organization (NARO), P.O. Box 7083, Kampala, Uganda; *Author for correspondence (e-mail:

albert.kriegner@arcs.ac.at; phone: ++43 50550 3528; fax: ++43 50550 3444)

Received 14 May 2002; accepted in revised form 4 November 2002

Key words: Autopolyploidy, Hexaploid, Polyploidy, Polysomic inheritance, Repulsion linkage

Abstract

Amplified Fragment Length Polymorphism (AFLP) based genetic linkage maps were developed for hexaploid sweetpotato (*Ipomoea batatas* (L.) Lam., $2n = 6x = 90$) using a segregating population derived from a biparental cross between the cultivars 'Tanzania' and 'Bikilamaliya'. A total of 632 ('Tanzania') and 435 ('Bikilamaliya') AFLPs could be ordered in 90 and 80 linkage groups, respectively. Total map lengths were 3655.6 cM and 3011.5 cM, respectively, with an average distance of 5.8 cM between adjacent markers. The genetic linkage analysis was performed in two steps. First a framework map was elaborated from the single dose markers. Interspersed duplex and double-simplex markers were used to detect homologous groups within and corresponding linkage groups among the parental maps. The type of polyploidy (autopolyploidy vs. allopolyploidy) was examined using the ratio of linkage in coupling phase to linkage in repulsion phase and the ratio of non-simplex to simplex markers. Our data support the predominance of polysomic inheritance with some degree of preferential pairing.

Introduction

Sweetpotato (*Ipomoea batatas* (L.) Lam.) is a tropical root crop from the morning glory family Convolvulaceae (Austin 1988). With more than 133 million tons in annual production, sweetpotato ranks as the fifth most important crop in developing countries after rice, wheat, maize, and cassava (International Potato Center (CIP) 1998). Its wide adaptability on marginal land and rich nutritional content provide an enormous potential for preventing malnutrition and enhancing food security in the developing world. In spite of its global importance, the genetic constitution and organization of sweetpotato remain poorly understood. So far little effort has been devoted to the development and application of molecular marker technology for the genetic improvement of sweetpotato.

Sweetpotato is a hexaploid species ($2n = 6x = 90$) (Jones 1965; Magoon et al. 1970).

Cytological and genetic studies of sweetpotato are difficult due to the high number of small chromosomes, its complex self- and cross-incompatibility systems (Jones 1986) and the high level of heterozygosity, consequently, the nature of its polyploidy remains uncertain. A number of hypotheses have been examined in efforts to elucidate the origin of sweetpotato. Both allopolyploidy (Ting and Kehr 1953; Jones 1965; Magoon et al. 1970; Sinha and Sharma 1992) and autopolyploidy (Nishiyama et al. 1975; Shiotani 1987; Ukoskit and Thompson 1997) have been proposed. Shiotani and Kawase (1989) postulated the genome constitution of sweetpotato as ($B_1B_1B_2B_2B_2B_2$), and suggested additional homology between the B_1 and the B_2 genomes, based on the occurrence of frequent formation of tetravalents

and hexavalents. Ukoskit and Thompson (1997) reported a polysomic inheritance in sweetpotato based on the segregation ratio and genetic linkage relationship of RAPD markers. Autopolyploidy is determined by random pairing of homologous chromosomes during meiosis. A highly heterozygous outbreeding population results in a large number of possible allelic combinations at a single locus. In the case of an autohexaploid, 12 different alleles could theoretically segregate independently in a population. This would result in 400 possible genotypic classes in the progeny.

Highly dense genetic linkage maps are powerful tools for the localization and map-based cloning of genes (positional cloning) and marker-assisted breeding. They also provide information for understanding the biological basis of complex traits (Lee 1995) and polyploidy. In recent years, AFLP markers have allowed significant advance in the ability to generate large number of polymorphic DNA bands, which made it feasible to develop linkage map for plants with large genome size. This technique is based on a selective PCR amplification of small restriction fragments (80–400 bp) of genomic DNA, which are used as dominant markers.

In this paper we report the genetic inheritance, segregation and linkage of AFLP markers in two hexaploid sweetpotato varieties ($2n = 6x = 90$). A F1 population was used to construct two separate parental maps integrating simplex and multiplex markers and to analyze the genome constitution of sweetpotato. These are the first reported genetic linkage maps that have substantial genome coverage of sweetpotato. They provide a framework for the tagging of genes and quantitative trait loci (QTL) of economic traits in sweetpotato.

Materials and methods

Plant materials and DNA extraction

The F1-mapping population originated from a pseudo-testcross between two African sweetpotato landraces ('Tanzania' and 'Bikilimaliya'). 'Tanzania' is the most widely grown sweetpotato cultivar in sub-Saharan Africa and is resistant to the sweetpotato virus disease (SPVD) complex in East Africa (Mwanga et al. 2001). 'Bikilimaliya' is susceptible under the same environmental conditions. Linkage mapping was carried out on a subset of 94 randomly selected

plants. Genomic DNA from young leaves of greenhouse-grown sweetpotato plants was extracted using the CTAB method (Murray and Thompson 1980). The quality and quantity of the DNA were evaluated by comparison with a standard weight Lambda DNA (λ -Pst I) by electrophoresis on 1% agarose gels stained with ethidium bromide, and visualized under UV illumination.

Amplified fragment length polymorphism (AFLP) assays

AFLP protocol developed by Vos et al. (1995) was followed with minor modifications. Genomic DNA (1 μ g) was digested with 10 units of EcoRI and 4 units of MseI in 10 mM of Tris H acetate, pH 7.5, 10 mM of MgCl₂, 50 mM potassium acetate, 5 mM of DTT, 5 μ g of BSA at 37 °C for 3 h. After checking for complete digestion EcoRI and MseI adapters (0.1 mM and 1 mM, respectively) were ligated during a 3-hour reaction at 37 °C using 2 mM ATP 1 unit of T₄ DNA ligase and 1 \times RL buffer. After incubation the reaction mix was diluted with T₁₀E_{0.1} to a final volume of 200 μ l. Using the primer notation of Vos et al. (1995) pre-amplification was performed with E00/M00 primer combination (30 ng each) with 5 μ l template DNA in 1 \times PCR buffer (10 mM Tris HCL, pH 8.3, 1.5 mM MgCl₂, 50 mM KCl), 0.2 mM of all four dNTPs, and 0.4 units of Taq polymerase. Pre-amplification was carried out with a Perkin Elmer cyclor under the following conditions: 20 amplification cycles at 92 °C for 60 sec followed by 60 °C for 30 sec and 72 °C for 60 sec. The pre-amplification reaction was diluted in 150 μ l T₁₀E_{0.1} buffer (10 mM Tris and 0.1 mM EDTA, pH 8.3). Five microliters of diluted pre-amplified DNA was selectively amplified using 5 ng γ P³³-labeled EcoRI + 3 primer, 30 ng MseI + 3 primer, 0.4 units of Taq polymerase, 0.2 mM of all four dNTPs and 1 \times PCR puffer. Selective-amplification PCR-cycle profiles were performed as described by Vos et al. (1995). Amplification products were separated on 6% denatured polyacrylamide gels and visualized by autoradiography on X-ray films (Kodak, Tokyo, Japan).

Selection of primer combinations

The two parental varieties, 'Tanzania' and 'Bikilimaliya', were screened with 240 EcoRI + 3/MseI + 3 primer combinations (Table 3). Those combinations that yielded a high number of polymorphic fragments

Table 1. Expected segregation ratios (presence:absence) for the inheritance of a dominant marker in hexaploid sweetpotato, according to four cytological hypotheses (Jones 1967)

Marker dose	Hypothesis I Autohexaploid (hexasomic)		Hypothesis II and III Tetradiploid (tetra-disomic, tetra- somic, disomic)		Hypothesis IV Allohexaploid (disomic)	
Simplex	Aaaaaa	1:1	Aaaa aa	1:1	Aa aa aa	1:1
			aaaa Aa	1:1		
Duplex	AAaaaa	4:1	AAaa aa	5:1 ²⁾	Aa Aa aa	3:1
			Aaaa Aa	3:1 ³⁾	AA aa aa	1:0
			aaaa AA	1:0 ¹⁾		
Triplex	AAAaaa	19:1	AAAa aa	–	Aa Aa Aa	7:1
			AAaa Aa	11:1	AA Aa aa	1:0
			Aaaa AA	1:0		
Quadruplex	AAAAaa	1:0	AAAA aa	1:0	AA Aa Aa	1:0

¹⁾ disomic inheritance

²⁾ tetrasomic inheritance

³⁾ tetra-disomic inheritance

for each parental line and with a total number of 50–100 fragments were selected for generating AFLP markers in the mapping progeny.

AFLP marker nomenclature and marker scoring

Autoradiographs were scored for the absence (0) or presence (1) of AFLP marker bands, manually and independently by at least two persons. Only those fragments that could be clearly scored were used. Markers ambiguous in a few genotypes were treated as missing data for map construction. Markers that were polymorphic for the offspring population were chosen on the basis of their presence in one parent and absence in the other, or presence in both parents. Each AFLP marker was identified by a primer pair combination using the primer notation of Vos et al. (1995) and a band number or letter as suffix. The polymorphic bands were named serially in descending order of molecular weight. Scored markers were divided into three groups depending on the presence or absence within each parent.

Segregation ratio

The cross was analyzed as a double pseudo-testcross (Grattapaglia and Sederoff 1994). The assessment of marker dosage was done by the expected segregation ratios (presence vs. absence) of AFLP markers in the mapping progeny, in accordance with the allele dosage expected for four alternative cytological hypotheses in sweetpotato (Table 1). Since autopolyploidy

is the most likely genetic configuration in sweetpotato (Nishiyama et al. 1975; Shiotani 1987; Ukoskit and Thompson 1997), strict disomic segregation was ruled out for our linkage analysis. To classify fragments according to their individual frequency (i.e. simplex, duplex, double simplex), acceptance ranges were constructed.

(i) Simplex markers: AFLPs present in one parent and absent in the other were tested for goodness of fit to the 1:1 segregation ratio (presence: absence) in the progeny by a χ^2 test at the 99% confidence level, allowing type I error = type II error = 0.5%. The null hypothesis (H_0) (1:1 segregation) was tested against the alternative hypothesis (3:1 segregation) since all non-simplex ratios would be 3:1 or greater, regardless of whether sweetpotato is an auto-, auto-allo-, or an allopolyploid (Wu et al. 1992). Markers showing a segregation ratio significantly higher than 1:1 are likely to be in two or more copies.

(ii) Duplex markers: Duplex markers were detected under the assumption of hexasomic or tetrasomic inheritance with expected segregation ratios of 4:1, 5:1 respectively. Due to the difficulty to differentiate between the markers fitting into the hexasomic (4:1) or tetrasomic (5:1) segregation groups, a common acceptance region for duplex markers under both types of inheritance was defined, and tested by χ^2 test for goodness of fit at a 10% significance level. Such a significance level is sufficient to distinguish duplex markers from the expected segregation ratios for triplex markers (11:1, 19:1) and simplex markers (1:1) with the progeny size available.

(iii) Double-simplex intercross markers are AFLPs that are present in both parental clones in a single-dose condition. Such markers are expected to segregate in a 3:1 ratio in the population, as tested for $H(0)$ at a 10% significance level, to avoid overlapping with the 11:1 segregation class for duplex-simplex intercross markers.

Estimation of recombination fraction (r) and linkage mapping

Linkage analysis and map construction were performed in two steps: First, two parental framework maps (maternal and paternal) were constructed from simplex markers. Markers were associated at a LOD score of 5. Grouping of markers and map construction was performed by the JoinMap 3.0 (Van Ooijen and Voorrips 2001). Coupling linkage among dominant simplex markers results in a map with 2 n linkage groups (Al-Janabi et al. 1993; Da Silva et al. 1995). Second, duplex and double-simplex markers were interspersed into the fixed order of the parental simplex framework maps. For this purpose, recombination fraction (r) and LOD scores for simplex/duplex, duplex/duplex and simplex/double-simplex marker configurations were calculated assuming (a) hexasomic and (b) tetrasomic inheritance under the assumption of random pairing of homologous chromosomes and absence of double reduction by numerically maximizing the log-likelihood, as follows:

$$L = X_{AB} \log p_{AB}(r) + X_A \log p_A(r) + X_B \log p_B(r) + X_0 \log p_0(r)$$

Where r is the recombination fraction and X_{AB} , X_A , X_B , X_0 are the observed numbers of offspring in each phenotypic class. Phenotype probabilities (p) for the marker pair configurations used are given in Table 2. All configurations are in coupling. LOD scores were calculated for each pair as: $\log_{10}(\text{likelihood for } r = \hat{r}) - \log_{10}(\text{likelihood for } r = 0.5)$. Such pairwise estimates are suitable for input into the JoinMap3.0 program (Van Ooijen and Voorrips 2001). The set of pairwise estimates of r and LOD under the hexasomic model were used to locate duplex and double simplex markers on the framework map. Map positions of interspersed duplex markers were later confirmed using estimates from tetrasomic model. The likelihood surface for calculating linkage among simplex and double-simplex markers is equal for an hexasomic and

tetrasomic model of inheritance. Two linkage groups were declared homologous if they possessed the same duplex markers. Double-simplex intercross markers were used to identify homologous counterparts between the two parental maps.

Estimation of genome coverage

The expected proportion of a genome (E) covered by a linkage map with n markers at random positions (C_n), $E(C_n)$, was estimated according to Bishop et al. (1983) as $E(C_n) = 1 - p_{1,n}$ where $p_{1,n}$ is the probability that a genome is not covered by a linkage map as given by $p_{1,n} = 2r/n + 1[(1 - X/2G)^{n+1} - (1 - X/G)^{n+1} + (1 - rX/G)(1 - X/G)^n]$ where, r is the number of chromosomes, X is the distance between pairs of markers (in cM), and G is the total genome size (in cM).

Assessment of polyploidy type

Detection and quantification of repulsion linkage

The type of chromosome pairing and inheritance was investigated based on the ratio of AFLP markers linked in repulsion phase to markers linked in coupling phase (Wu et al. 1992; Qu and Hancock 2001). Mirror images were created for all simplex marker scores and added to the original marker scores. Recombination fractions (R) [$R = a/n$, where a = number of recombinants in repulsion and n = total number of gametes] were then estimated for all pairs of original/inverted and original/original markers using Join Map 3.0 software. Linkages between original and inverted markers are in repulsion those between original and original markers are in coupling. A 1:1 or 0:1 ratio of repulsion to coupling linkages indicates allopolyploidy or autopolyploidy, respectively.

In autopolyploids, the observed recombination fractions ($R = a/n$) for markers in repulsion do not represent the exact genetic distance as recombination occurs through crossing over (R_c) [two repulsion-phase linked markers are brought together on one chromosome] and independent assortment (R_i) [two repulsion-phase linked markers segregate into one gamete, but markers are still on two individual homologous chromosomes]. R can accordingly be expressed as $R = R_i + R_c$. R_i is the minimum distance that can be detected between repulsion linked markers. Therefore, the default linkage used for detecting repulsion linkage must be greater than the corresponding R_i for

Table 2. Marker pair configurations and expected phenotypic frequencies used in this study

Hexasomic inheritance			Tetrasomic inheritance		
Marker pair configuration ¹	Phenotype Probabilities		Marker pair configuration	Phenotype probabilities	
Simplex/simplex coupling	AB	$1/2(1 - r)$	Simplex/simplex coupling	equal to hexasomic inheritance	
AB/00/00/00/00 × 00/00/00/00/00/00	A	$1/2r$	AB/00/00/00 × 00/00/00/00		
Simplex/duplex coupling	B	$1/2r$	Simplex/duplex coupling	AB	$1/2 - 1/6r$
AB/0B/00/00/00/00 × 00/00/00/00/00/00	0	$1/2(1 - r)$	AB/0B/00/00 × 00/00/00/00	A	$1/6 r$
Duplex/duplex coupling	AB	$1/2 - 1/5r$	Duplex/duplex coupling	B	$1/3 + 1/6r$
AB/AB/00/00/00/00 × 00/00/00/00/00/00	A	$1/5r$	AB/AB/00/00 × 00/00/00/00	0	$1/6 - 1/6r$
Duplex/duplex coupling	B	$3/10 + 1/5r$	Duplex/duplex coupling	AB	$5/6 - 1/3r + 1/6r^2$
AB/AB/00/00/00/00 × 00/00/00/00/00/00	A	$2/5r - 1/5r^2$	AB/AB/00/00 × 00/00/00/00	A	$1/3r - 1/6r^2$
Simplex/double-simplex coupling	B	$2/5r - 1/5r^2$	Simplex/double-simplex coupling	B	$1/3 - 1/6r^2$
AB/00/00/00/00/00 × A0/00/00/00/00/00	0	$1/5 - 2/5r + 1/5r^2$	AB/00/00/00 × A0/00/00/00	0	$1/6 - 1/3r - 1/6r^2$
Double-simplex/double-simplex coupling	AB	$1/2 - 1/4r$	Double-simplex/double-simplex coupling	equal to hexasomic inheritance	
AB/00/00/00/00/00 × A0/00/00/00/00/00	A	$1/4 + 1/4r$	AB/00/00/00 × A0/00/00/00		
Double-simplex/double-simplex coupling	B	$1/4r$	Double-simplex/double-simplex coupling	equal to hexasomic inheritance	
AB/00/00/00/00/00 × AB/00/00/00/00/00	0	$1/4 - 1/4r$	AB/00/00/00 × AB/00/00/00		
Double-simplex/double-simplex coupling	AB	$(3 - 2r + r^2)/4$	Double-simplex/double-simplex coupling	equal to hexasomic inheritance	
AB/00/00/00/00/00 × AB/00/00/00/00/00	A	$1/2r - 1/4r^2$	AB/00/00/00 × AB/00/00/00		
Double-simplex/double-simplex coupling	B	$1/2r - 1/4r^2$	AB/00/00/00 × AB/00/00/00		
AB/00/00/00/00/00 × AB/00/00/00/00/00	0	$(1 - r)^2/4$	AB/00/00/00 × AB/00/00/00		

¹ Distribution of alleles of two loci in a base chromosome group (chromosomes are separated by "/"). "A": presence of band at locus A, "B": presence of band at locus B, "0": absence of band.

a polyploid (i.e. 0.40, 0.33, 0.0 for autohexaploid, autotetraploid and allopolyploid).

To distinguish disomic from polysomic models of inheritance in sweetpotato the ratio of coupling to repulsion linkage was first analyzed using a default linkage of $R \leq 0.33$, $LOD \geq 4$. Strict allopolyploidy would be clearly established if an equal number of linkages in repulsion to coupling was observed. Conversely, if no repulsion linkage is detected at $R \leq 0.33$ disomic inheritance is highly unlikely. Second, default linkage was lowered to $R \leq 0.4$ and $LOD \geq 2$. If all repulsion linkages are ≥ 0.4 sweetpotato is a true autohexaploid with completely random association of homologous chromosomes. Repulsion linkage detected at $R < 0.4$ indicate some degree of preferential pairing.

Ratio of non-simplex to simplex markers

The proportion of non-simplex to simplex markers was used as a second indication of the type of polyploidy, i.e., autopolyploidy vs. allopolyploidy. This approach has been used to investigate the type of inheritance of *Saccharum officinarum* (Kehrer 1994), *Saccharum robustum* (Al-Janabi et al. 1994), and *Ip-*

omoea batatas (Ukoskit and Thompson 1997). In a hexaploid simplex, duplex and triplex alleles can give rise to polymorphism only (alleles at higher doses produce non-polymorphic markers). The expected frequencies for non-single dose polymorphisms are calculated by summing the probability of transmission of an "absent" allele for duplex and triplex markers under the assumption of either autohexaploidy ($1/5 + 1/20 = 0.25$) or allohexaploidy ($1/4 + 1/8 = 0.375$). In accordance with the assumption of non-preferential chromosome pairing in an autohexaploid, one can expect that 25% of all segregating markers are non-simplex and 75% are simplex. In an allohexaploid, the expected percentage of non-simplex markers would be 37.5% and of simplex markers 62.5% because of the disomic segregation. Three hexasomic (1:1, 4:1, 19:1) ratios of markers (presence to absence) were considered for single-, double- and triple-dose fragments. Observed segregation ratios were tested for goodness-of-fit by a χ^2 test at a 0.05 significance level with the expected segregation ratios.

Results

Selection of primer combinations

A total of 107 out of 240 primer combinations screened were selected and used to generate AFLP markers for the mapping population. Table 3 shows number of scored and mapped AFLP markers for each of 107 EcoRI/MseI primer pairs used for AFLP analysis.

Assessment of marker dosage

AFLPs were grouped according to their presence in the female, male, or in both parents. The segregation ratios observed for 1449 clearly scorable female and male markers are shown in Figure 1. Within these distributions a large group of markers centered around a value of 0.5 of the expected segregation ratio for simplex markers (1:1), and a smaller group around 0.8 where duplex markers (4:1 or 5:1, 3:1) were expected. A notable number of markers fell within the segregation ratios expected for triplex markers (19:1 or 11:1). Our segregation analysis detected 960 simplex markers ($\alpha \leq 0.01$). A total of 269 markers fitted 4:1 or 5:1 segregation ratios ($\alpha \leq 0.10$) that are expected for duplex markers under hexasomic (4:1) or tetrasomic (5:1) inheritance. Quadruplex or higher-dose markers are not expected to result in observable segregation in the offspring, except in the case of random chromatid segregation (double-reduction), which would result in a genotype not showing the marker. Figure 2. shows the segregation ratios observed for each of 540 polymorphic markers present in both parents. A total of 215 markers fitted a 3:1 (presence:absence) segregation ratio, expected for double-simplex markers. These markers were centered around a value of 0.75. The remaining markers fitted segregation ratios significantly higher than 3:1 expected for simplex-duplex, simplex-triplex, double-duplex, duplex-triplex, or double-triplex intercross-markers.

Occurrence of distorted segregation

A total of 173 markers (12.9%) did not fit the expected segregation ratios for a model of hexasomic inheritance at a level of $\alpha \leq 0.05$ (Table 4). Distorted segregation involved both parental clones equally, and it was in the range obtained in eucalyptus species 4–9% (Grattapaglia and Sederoff 1994; Verhaegen

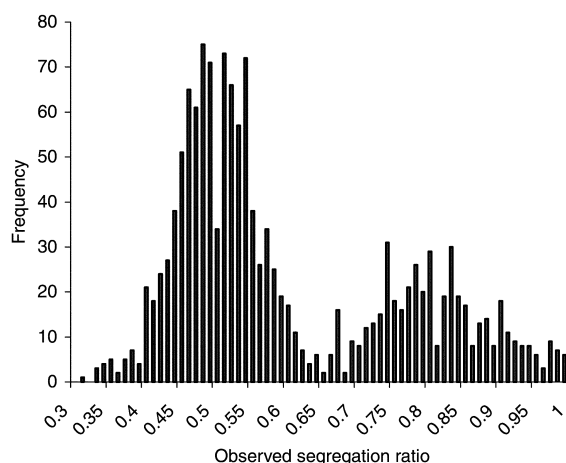


Figure 1. Observed segregation ratios (presence/absence) for 808 'Tanzania' and 641 'Bikilimaliya' AFLP markers.

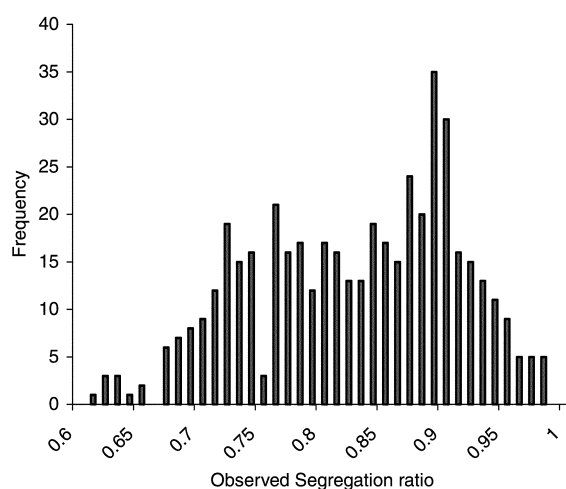


Figure 2. Observed segregation ratios (presence/absence) for 540 AFLP markers present in both parents

and Plomion 1996), pine species 14–15% (Kubisiak et al. 1995), oak 18% (Barreneche et al. 1998), and potato 27% (Gebhardt et al. 1994). No clustering of

Table 4. Segregation analysis of AFLP markers based on hexasomic chromosome pairing in 'Tanzania' and 'Bikilimaliya' sweet-potato cultivars (Chi-square test $\alpha \leq 0.05$).

Marker type	AFLP marker					
	Tanzania		Bikilimaliya		Total	
	Number	%	Number	%	Number	%
Simplex	519	64.2	410	63.9	929	63.9
Duplex	161	20.0	120	18.7	281	18.7
Triples	38	4.7	28	4.3	66	4.5
Distorted	90	11.1	83	12.9	173	12.9
Total	808	100.0	641	100.0	1445	100.0

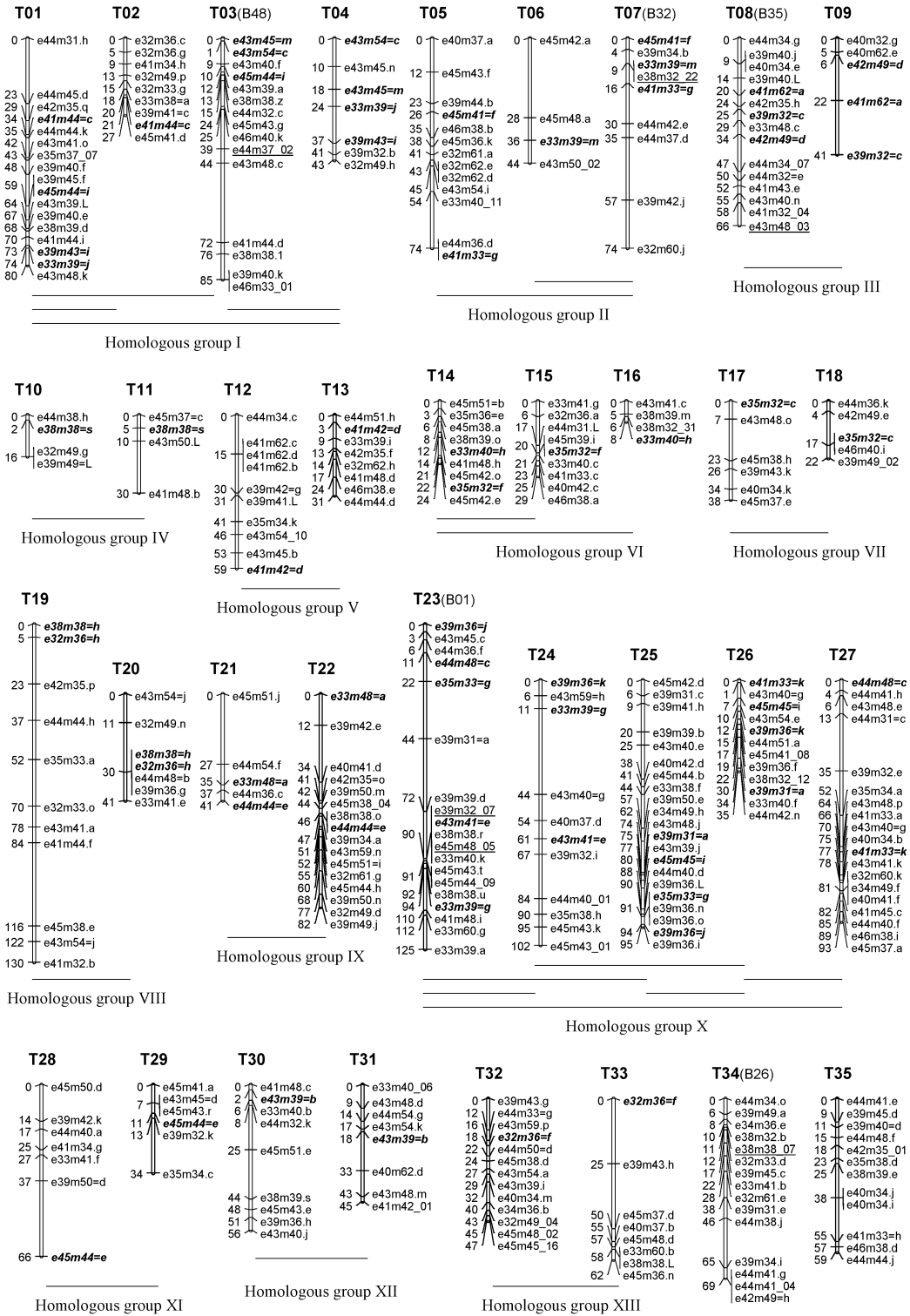


Figure 3. Genetic linkage map of Tanzania.*

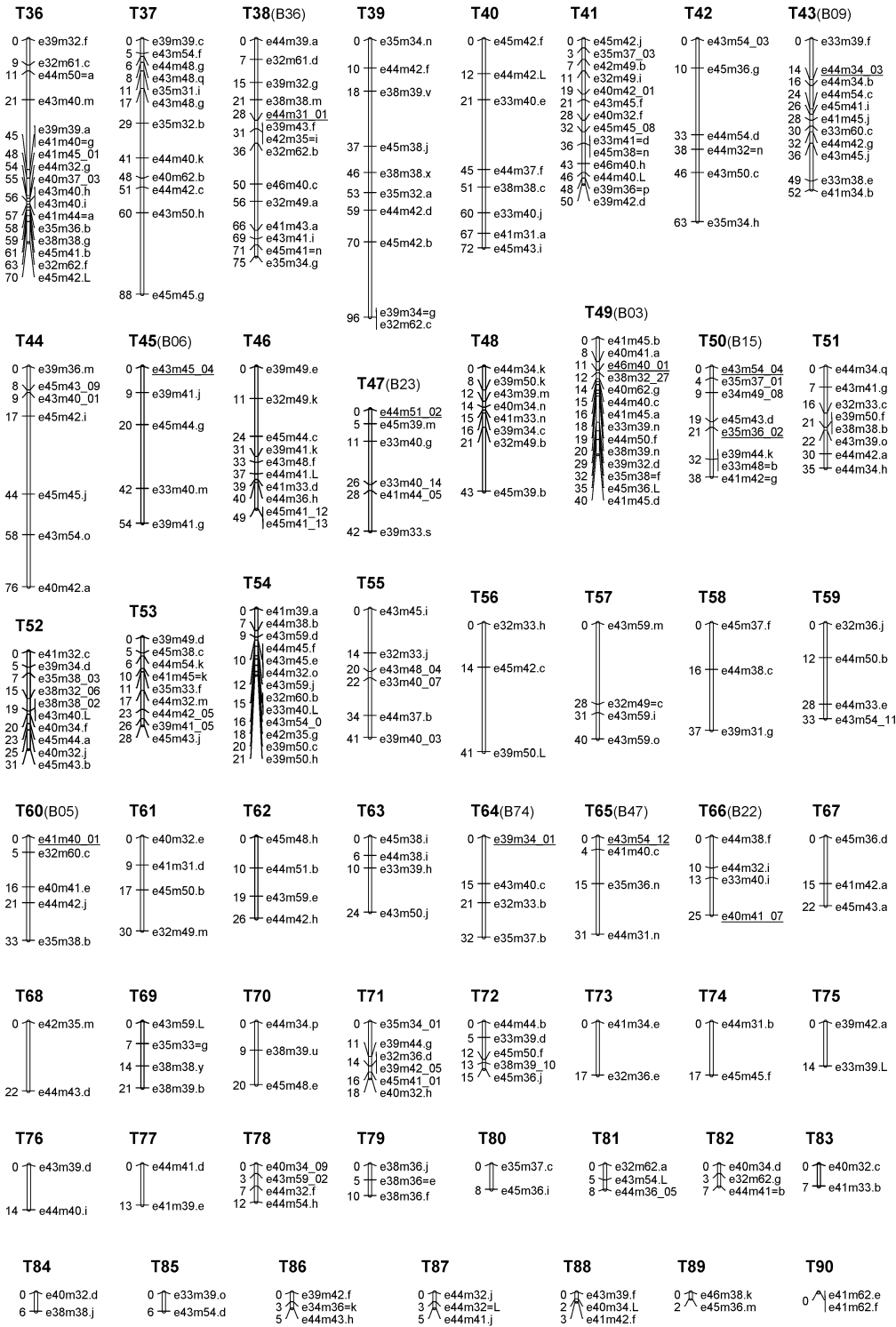


Figure 3. Continued

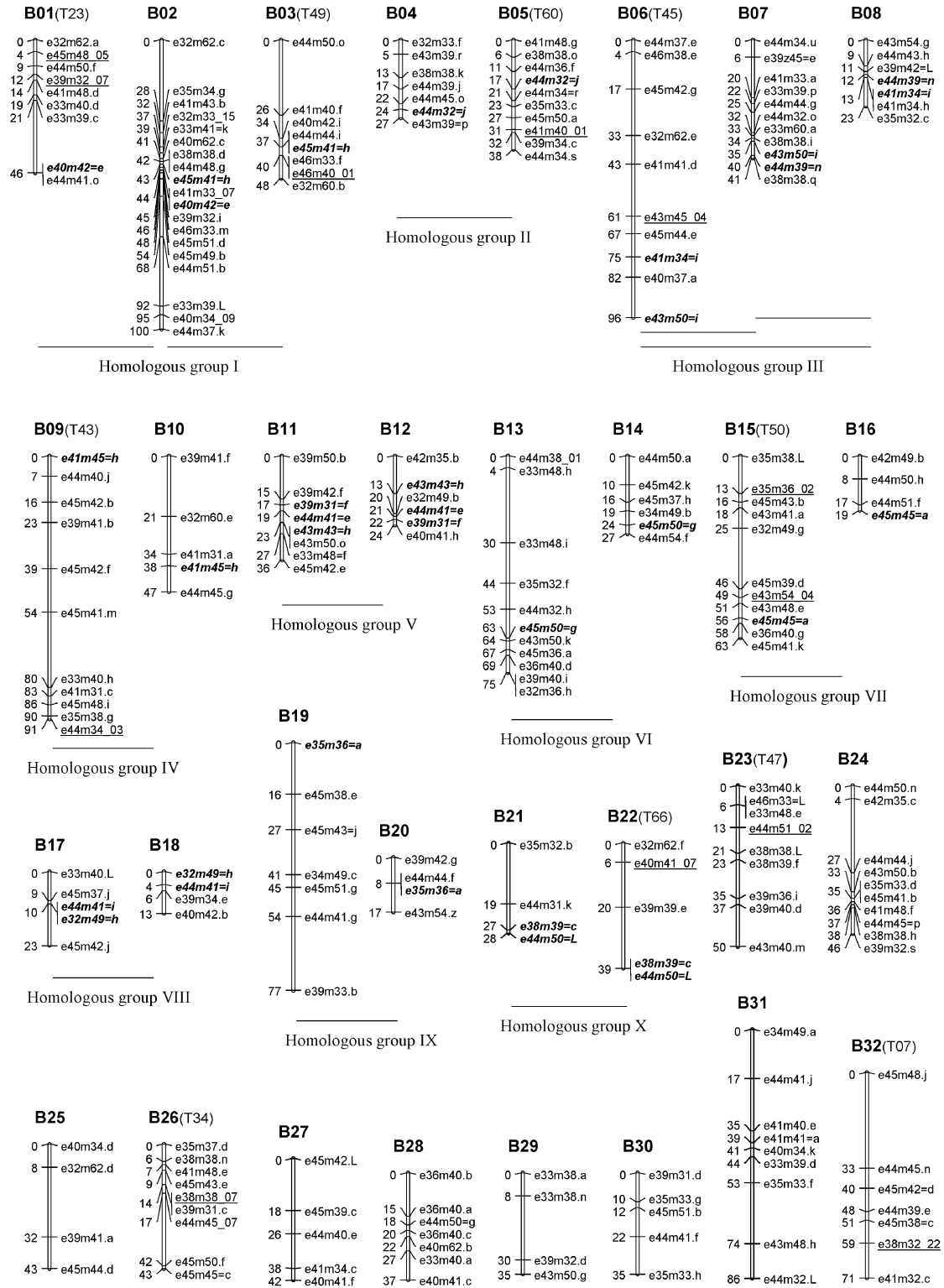


Figure 4. Genetic linkage map of Bikilimaliya

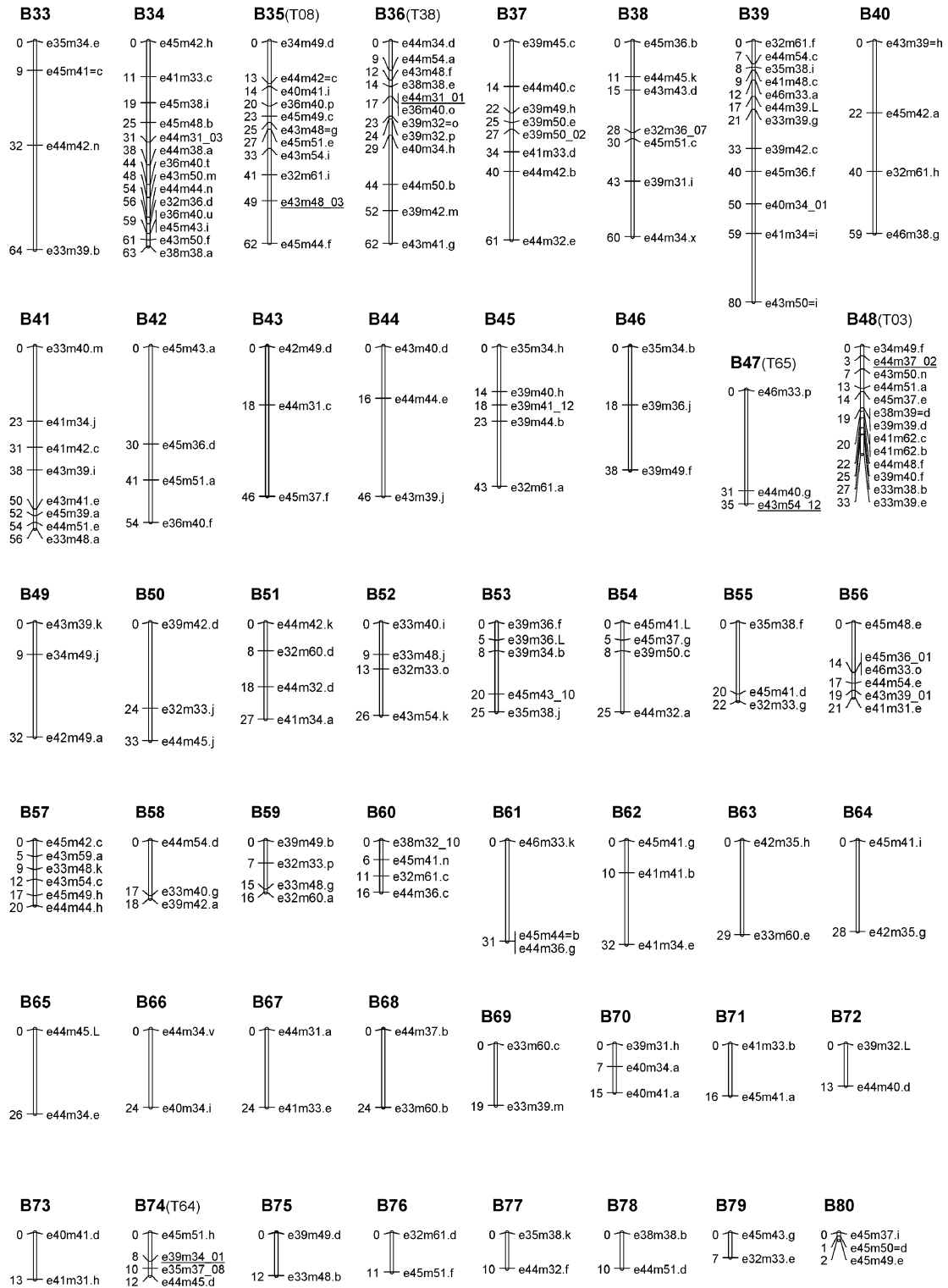


Figure 4. Continued

distorted simplex markers ($0.05 > p > 0.01$) on the map was obtained. Segregation distortion may be caused by various processes, including gametic selection or specific chromosome pairing factors, associations between heterozygosity and plant vigor, as was observed in alfalfa (Brummer et al. 1993), or natural selection toward one parental type during recombinant inbred line development (Wang et al. 1994). Calculation of percentage of distorted markers was done under the hexasomic model which might simplify the real type of segregation occurring in sweetpotato and therefore overestimate marker distortion. The occurrence of non-viable individuals homozygous for sublethal loci that are identical by descent would result in the absence of one class of genotypes (zygotic selection), and thereby could produce segregation distortion.

A sweetpotato framework map constructed using simplex markers

Linkage analysis of 539 maternal and 421 paternal simplex markers resulted in maps of 90 and 80 linkage groups for 'Tanzania' and 'Bikilimaliya' respectively. The grouping was consistent for LOD scores ranging from 4 to 6, suggesting that the linkage groups detected were highly reproducible and should represent the corresponding physical chromosomes. Markers that could not be placed on the map during "round one" or "round two" of the Join Map-mapping procedure were omitted from the map. A total of 54 female and 58 male simplex markers did not meet the grouping or ordering criteria and remained unlinked.

* Figures 3 and 4. Maternal ('Tanzania') and paternal ('Bikilimaliya') linkage maps are based on AFLP markers obtained on a progeny generated by a pseudo-testcross among the varieties 'Tanzania' and 'Bikilimaliya'. The two hexaploid maps comprise 485 and 363 simplex (indicated by a "."), 76 and 42 duplex ("=") and 70 and 31 double-simplex ("_") markers distributed over 90 and 80 linkage groups, respectively. Linkage groups are named at their top; a second name in parenthesis indicates the corresponding linkage group on the 'Tanzania' or 'Bikilimaliya' map. These homologous female and male linkage groups contain the same double-simplex marker (markers underlined). Fifty chromosomal connections were provided by duplex markers that mapped to two linkage groups (markers in italic and bold). Connections among homologous linkage groups are graphically indicated below the linkage groups. Cumulative Haldane map distances are indicated at the left side of each marker. Each marker locus is identified by the corresponding combination of primers (primer notation of Vos et al. (1995)) used for its generation. A letter or a number at the end of each marker identifies the individual fragments amplified by the same primer combination.

The linkage groups obtained were numbered at random until they may be ascribed to individual chromosome karyotypes.

Mapping of duplex markers

The absolute differences of pairwise estimates (of r and LOD scores) under hexasomic and tetrasomic model of inheritance were low. As few as seven duplex markers were placed at different sites on the simplex map when linkage was analyzed under the hexasomic and tetrasomic model of inheritance. Such markers were removed from the map. The remaining 118 mapped duplex markers were either placed on one or two linkage groups, thus providing an association between homologous chromosomes. Fifty chromosomal connections were provided by duplex markers (duplex markers that mapped to two linkage groups), thus allowing to order 33 female and 22 male linkage groups into 13 and 10 homologous groups respectively (Figures 3 and 4).

Mapping of double simplex intercross markers

A total of 15 corresponding male and female linkage groups could be aligned by interspersed double-simplex intercross markers (Figures 3 and 4). The detection of homology between the two parental maps is an important criterion for consistency of the maps.

Genome length

The final maps of 'Tanzania' (Figure 3) and 'Bikilimaliya' (Figure 4) comprised a total of 632 and 435 markers covering a cumulative map length of 3655.6 cM and 3011.5 cM, respectively. A linkage group had on average seven markers with a maximum of 21 markers detected for the largest group. The maternal map consisted of 90 linkage groups (71 groups with 4 or more markers) compared to the paternal map with 80 linkage groups (56 groups with 4 or more markers). The largest female and male linkage groups were 129.6 cM and 99.5 cM, respectively. The average spacing between marker loci was 5.8 cM with only four intervals exceeding 30 cM. The percentage of mapped simplex, duplex and double-simplex intercross markers were 90%, 54% and 30%, respectively. Based on an estimated nuclear genome size of sweetpotato of approximately 3.03×10^9 basepairs/2C nuclei (Krieger 2001) every cM equals approximately

Table 5. Chi-square analysis ($\alpha \leq 0.001$) for polyploidy type in sweetpotato based on the percentages of non-simplex vs. simplex markers supported a autopolyploid nature of sweetpotato.

Marker type	Observed number of markers		Expected number of markers			
	Number	%	Autohexaploid		Allohexaploid	
			Number	%	Number	%
Simplex	929	72.8	957	75.0	798	62.5
Non-simplex (duplex, triplex)	347	27.2	319	25.0	478	37.5
Total	1276	100.0	1276	100.0	1276	100.0
χ^2			p = 0.07 ^{ns}		P = 0.000 ^{***}	

^{ns} and ^{**} Non-significant and significant at 0.001 level for H_0 : Observed ratio – expected ratio = 0

832 Kbp (3.03×10^9 divided by the 3655.6 cM of the 'Tanzania' map).

Assessment of genome coverage by the parental maps

The expected proportion of genome coverage was estimated using 485 and 363 simplex markers for the female and male maps, respectively. Only framework markers were used in this procedure to avoid an over-estimation of genome coverage. Under the assumption of random marker distribution, any new female or male marker had a 89% and 86% probability of being within 20 cM of any existing marker in the female and male map (Bishop et al. 1983). Since the 'Tanzania' and 'Bikilimaliya' linkage groups might cover non-overlapping regions, it can be supposed that the major portion of the entire sweetpotato genome has already been covered by these maps.

Estimation of ploidy type

Detection and quantification of repulsion linkage

The presence and absence data for 539 'Tanzania' and 421 'Bikilimaliya' simplex markers were inverted (re-coded) and added to the original markers. Recombination fractions were estimated for $(539 \times 538) = 289982$ female and $(421 \times 420) = 176820$ male pairs of original/inverted and original/original markers.

A total of 2800 coupling linkages, and no repulsion linkages were detected at a default linkage of $R \leq 0.33$, $LOD \geq 4$. This result would be highly unlikely if sweetpotato was an allopolyploid with strict disomic inheritance and supports prevalence of polysomic inheritance. A total of 243 repulsion linkages were detected when stringency of detection was lowered to a LOD of 2 and a maximum recombination

fraction of 0.4. This indicates some degree of preferential pairing.

Ploidy estimation based on the percentage of simplex vs non-simplex markers

The number of simplex, duplex, and triplex markers was determined by segregation analysis of 808 and 641 'Tanzania' and 'Bikilimaliya' AFLP markers, respectively. The polyploidy type was tested using the combined data sets from both parents comparing the proportion of simplex markers to non-simplex (duplex and triplex) markers. The present study showed 72.8% simplex and 27.2% non-simplex markers (Table 5), and supported an autopolyploid nature of sweetpotato.

Discussion

Genetic linkage maps were constructed for sweetpotato, one of the most important food crops in developing countries. In selecting a mapping population we considered the capacity to differentiate between 1:1 and 3:1 or larger segregation ratios with a high level of certainty and to detect linkages in a range that is useful for future genetic and breeding studies. Female and male meioses leading to gamete formation are independent events because they occur on different sporophytes, and the combination of gametes to constitute any individual offspring is a random process. Therefore, each individual parent can be regarded as independent data. In this manner, for each of the two parents an independent map was constructed. All offspring plants were found to contain AFLP markers from one or the other parental clone, and no markers were observed to be absent for both parents. This finding underlined the high quality and good score-

ability of the markers used for construction of the linkage maps.

Mapping in sweetpotato reveals several difficulties. The estimation of marker dosage was complicated because the exact type of chromosome inheritance remains uncertain. At least, two different inheritance models (hexasomic, tetra-disomic) had to be considered. Type of ploidy will be discussed later. However, an accurate estimation of marker dosage is critical for subsequent analysis. Thus, only three categories of markers were analyzed. Among these, simplex and double-simplex markers have equal expected segregation ratios and linkage likelihood surfaces regardless of the true type of inheritance. For duplex markers, segregation ratios and likelihood surfaces differ slightly between both models of inheritance. To overcome these differences, an alternative approach, including the assumptions for both hexasomic and tetrasomic models, was used.

Our genetic linkage map comprises two different levels of confidence. The simplex marker framework, on the one hand, represents the precisely located core of the genome maps. The high confidence level of the local order in the framework maps ($\text{LOD} \geq 5$), arises from the high information content of the simplex markers compared to the limited information content of multiplex markers. If a genome is scanned for quantitative trait loci (Tanksley et al. 1992), analysis is mostly restricted to framework loci. With an effective size of the mapping population of 91 individuals, linkages of simplex markers at 25 cM are expected to be associated with a high confidence level. A LOD score of 5 reduces the rate of false positives to 10^{-5} , i.e. less than 1 of 100,000 linkages detected are erroneous. The LOD scores used by Join Map 3.0 are based on the χ^2 test for independence of segregation compared to LOD scores recommended by Lander et al. (1987), which can result in spurious linkage of markers with segregation distortion. The test of independence is not affected by distorted segregation ratios.

Accessory markers (i.e. duplex, double simplex markers), on the other hand, are less precisely located, but provide useful information on the pairing behavior of homologous chromosomes and the consistency of parental maps. In order to minimize the number of redundant associations of linkage groups, only the most informative configurations (duplex/duplex coupling, simplex/duplex coupling) (Meyer et al. 1998) were analyzed, and linkage was declared at a LOD score of 5. The maximum number of linkage

groups to which a multiplex marker should align is determined by its dosage. The fact that no duplex marker aligned to more than two simplex linkage groups demonstrates the high quality of detected chromosomal bridges. In a simulation study, Ripol et al. (1999) estimated that the number of chromosomal connections necessary to align all 64 sugarcane chromosomes was about 200. In their linkage map, duplex markers provided 30 connections and 41 additional connections were established from highly polymorphic simplex RFLP markers. The number of connections resolved by duplex markers was comparable in our study, yet the sweetpotato linkage map could not be resolved into 15 homologous groups.

Interspersed double simplex markers identified 15 associations between two parental maps. The consistency of identified chromosomal bridges should be reconfirmed since there are sources of potential error due to the occurrence of non-homologous fragments (Dowling et al. 1996), or statistical error (Plomion et al. 1995). A complete merging of the two maps into one comprehensive map would be possible if enough fully informative co-dominant markers were available.

Sweetpotato chromosomes are extremely small and differ considerably in size (Sinha and Sharma 1992), which may be partially responsible for the differential distribution of markers to the different linkage groups observed in our study. In tomato, Tanksley et al. (1992) observed a high correlation ($r = 0.9$) among the size of mitotic metaphase chromosomes and the number of markers per chromosome. The AT-rich target sequences of restriction enzymes EcoRI and MseI used for generation of AFLP preferentially target the AT – rich centromere flanking regions (Alonso-Blanco et al. 1998; Vuylsteke et al. 1999; Bennetzen and Freeling 1993) and therefore constitute another explanation for clustering of markers. The existing linkage maps are being used as a framework for the mapping of genes and quantitative trait loci that are involved in the resistance to the sweetpotato virus complex and nematodes. Based on the linkage information, a subset of well distributed AFLP fragments will be selected for phylogenetic studies. Existing linkage groups will be chromosomally allocated, and oriented in relation to the centromeres and telomeres. Inheritance, as determined by chromosome pairing, is of major interest and importance for understanding the evolution of polyploids. This emphasizes the importance of integrating data from cytological, morphological, ecological, and ge-

netic studies. In the present study, marker segregation and repulsion linkage were analyzed to distinguish between disomic and polysomic models of inheritance. Detection of linkage in repulsion is determined by the type of meiotic chromosome pairing and assortment in each parent. Allopolyploidy with strict disomic inheritance is functionally equivalent to diploidy where likelihood functions for repulsion linkage are direct transformations of those for coupling, with r replaced by $(1-r)$ (except for dominant markers in an F_2 population). Because likelihood surfaces are symmetric about 0.5, the log likelihood ratio test statistics are the same for both phases and the recombination fractions can be transformed directly for the two phases. In this way, markers located on both homologous chromosomes can be included into one linkage group. Alternatively, repulsion linkage can also be detected as coupling linkage among original and inverted markers.

In autopolyploids the likelihood surfaces are different for the linkage in coupling and repulsion phases, and recombination fractions cannot be directly transformed for the two phases. Wu et al. (1992) described likelihood estimators for repulsion recombination fractions in autopolyploids: $r_2 = [(h-1)a - 0.5(h-2)n]/n$ where h = number of homologous chromosomes, a = observed number of recombinants in repulsion phase. An exact estimation of genetic distance in repulsion requires a known degree of preferential pairing. Qu and Hancock (2001) described a method for the detection of preferential pairing in polyploids using smallest observed repulsion linkages from a large number of marker pairs. In autopolyploids, recombination occurs through crossing over (R_c) and independent assortment (R_i) according to $R = R_i + R_c$. When a large number of marker linkages in repulsion are evaluated, several pairs of markers with a genetic distance of 0 ($R_c = 0$) may be found. Therefore the lowest detected R value can be used as R_i for the polyploid. For polyploids with complex genome structure such as sweetpotato this approach is not very efficient as it only detects the highest degree of preferential pairing occurring in a genome. Varying degrees of preferential pairing over a genome cannot be resolved. This limitations could be overcome by using co-dominant markers. For any pair of homologous simplex markers (simplex markers that belong to the same locus) the exact R_i can be directly obtained from the repulsion recombination fraction. This is because recombination for such marker pairs only arises from independent assortment

and not from crossing over as their genetic distance is 0 (i.e., $R = R_i$; $R_c = 0$). Therefore, R can be directly used as R_i . For a fully informative co-dominant marker loci in a hexaploid (all 6 marker alleles are known) a total of 15 different R_i values can be estimated. The individual R_i values quantify the pairing affinities of the chromosomes carrying the markers. The smaller the R_i the stronger is the preferential pairing of two chromosomes. In case of complete random pairing of homologous chromosomes all 15 R_i would be 0.40. We suggest that a small number of fully informative co-dominant loci would give a good indication of the type of meiotic chromosome pairing and inheritance in sweetpotato.

The following conclusions can be made for polyploidy type of sweetpotato:

(1) The observed ratio of non-simplex (27.2%) to simplex markers (72.8%) generally supports autopolyploidy. The analysis of repulsion linkage further confirmed the prevalence of polysomic inheritance. Also, partial preferential pairing between certain chromosomes could not be excluded, as a few repulsion linkages were detected at a minimum LOD of 2 and a maximum recombination fraction of 0.4. (2) Strict allopolyploidy with disomic inheritance can be excluded as no linkage in repulsion was detected at a LOD of 4 and a maximum recombination fraction of 0.33. The proposed status of ploidy for the sweetpotato varieties Tanzania and Bikilimaliya concurs with studies on pairing behavior and inheritance of RAPD markers (Ukoskit and Thompson 1997), inheritance of genes such as β -amylase (Kumagai et al. 1990), and cytological studies of sweetpotato chromosomes (Shiotani and Kawase 1989).

Acknowledgements

We thank Dr Z-B. Zeng for the comments and useful discussions. We also thank Doris Carbajulca and Geneveva Rossell and Oscar Hurtado for helping with the marker scoring.

References

- Al-Janabi S.M., Honeycutt R.J., McClelland M. and Sobral B.W.S. 1993. A genetic linkage map of *Saccharum spontaneum* (L.) 'SES 208'. *Genetics* 134: 1249–1260.

- Al-Janabi S.M., Honeycutt R.J. and Sobral B.W.S. 1994. Chromosome assortment in *Saccharum*. *Theor. Appl. Genet.* 89: 959–963.
- Alonso-Blanco C., Peeters A.J.M., Koorneef M., Lister C., Dean C., van den Bosh N. et al. 1998. Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a Ler/Cvi recombinant inbred line population. *Plant J.* 14: 259–271.
- Austin D.F. 1988. The taxonomy, evolution and genetic diversity of sweet potato and related wild species. In: Gregory P. (ed.), *Exploration Maintenance and Utilization of Sweetpotato Genetic Resources*. International Potato Center (CIP), Lima, Peru, pp. 27–59.
- Barreneche T., Bodenes C., Lexer C., Trontin J.F., Fluch S., Streiff R. et al. 1998. A genetic linkage map of *Quercus robur* L. (pedunculate oak) based on RAPD, SCAR, microsatellite, minisatellite, isoenzyme and 5S rDNA markers. *Theor. Appl. Genet.* 97: 1090–1103.
- Bennetzen J.L. and Freeling M. 1993. Grasses as a single genetic system: genome composition, co-linearity and compatibility. *Trends Genet.* 9: 259–261.
- Bishop D.T., Cannings C., Skolnick M. and Williamson J.A. 1983. The number of polymorphic clones required to map the human genome. In: Weir B.S. (ed.), *Statistical analysis of DNA sequence data*. Marcel Dekker, New York, pp. 181–200.
- Brummer F.C., Fecht C.S., Mc Coy I.J., Kidwell K.K., Osborn T.C., Kiss G.B. et al. 1993. Molecular maps of alfalfa. In: Vasil I. and Philips R. (eds), *DNA-based markers in plants*. Kluwer Academic Press, New York.
- Da Silva J.A.G., Honeycutt R.J., Burnquist W., Al-Janabi S.M., Sorrells M.E., Tanksley S.D. et al. 1995. *Saccharum spontaneum* L. SES 208 genetic linkage map combining RFLP – and PCR-based markers. *Mol. Breed.* 1: 165–179.
- Dowling T.E., Moritz C., Paimer J.D. and Rieseberg L.H. 1996. Nucleic acids III: analysis of fragments and restriction sites. In: Hills D.M., Moritz C. and Mable B.K. (eds), *Molecular Systematics*. Sinauer Assoc. Publ., Sunderland, Mass, USA, pp. 249–320.
- Gebhardt C., Ritter E. and Salamini F. 1994. RFLP map of the potato. In: Phillips R.L. and Vasil I.K. (eds), *DNA-based markers in plants*. Kluwer, Dordrecht/Boston, pp. 271–285.
- Grattapaglia D. and Sederoff R. 1994. Genetic Linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using pseudotestcross mapping strategy and RAPD markers. *Genetics* 137: 1121–1137.
- International Potato Center (CIP) 1998. *Sweetpotato Facts*. International Potato Center, Lima, Peru.
- Jones A. 1965. Cytological observations and fertility measurements of sweetpotato (*Ipomoea batatas* (L.) Lam.). *Proc. Amer. Soc. Hort. Sci.* 86: 527–537.
- Jones A. 1967. Theoretical segregation ratios of qualitatively inherited characters for hexaploid sweetpotato (*Ipomoea batatas*). *Bul. No. 1368*. USDA Tech.
- Jones A. 1986. Sweetpotato heritability estimates and their use in breeding. *HortScience* 21: 14–17.
- Kehrer R.L. 1994. A RAPD analysis of the segregation patterns in the cross of *Saccharum officinarum* (La Purple) with *Saccharum robustum* (Molokai 5829). MSc Dissertation, Brigham Young University, Salt Lake City, Utah, USA.
- Kriegner A. 2001. Genetic linkage mapping, determination of ploidy type, and genetic variability of hexaploid sweetpotato (*Ipomoea batatas* (L.) Lam.) – 115 Bl. University of Bodenkultur, Vienna, Austria, (Dissertation).
- Kubisiak T.L., Nelson C.D., Nance W.L. and Stine M. 1995. RAPD linkage mapping in a longleaf pine x slash pine F1 family. *Theor. Appl. Genet.* 90: 1119–1127.
- Kumagai T., Umemura Y., Baba T. and Iwanaga M. 1990. The inheritance of beta-amylase null in storage roots of sweetpotato, *Ipomoea batatas* (L.) Lam. *Theor. Appl. Genet.* 79: 369–376.
- Lander E.S., Green P., Abrahamson J., Barlow A., Daly M.J., Lincoln S.E. et al. 1987. Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1: 174–181.
- Lee M. 1995. DNA markers and plant breeding programs. *Adv. Agron.* 55: 265–344.
- Magoon M.L., Krishnar R. and Vijaya Bai K. 1970. Cytological evidences on the origin of sweetpotato. *Theor. Appl. Genet.* 40: 360–366.
- Meyer R.C., Milbourne D., Hackett C.A., Bradshaw J.E., McNichol J.W. and Waugh R. 1998. Linkage analysis in tetraploid potato and association of markers with quantitative resistance to late blight (*Phytophthora infestans*). *Mol. Gen. Genet.* 259: 150–160.
- Murray M. and Thompson W.F. 1980. Rapid isolation of high-molecular-weight plant DNA. *Nucleic Acids Res.* 8: 4321–4325.
- Mwanga R.O.M., Odongo B., Ocitti p'Obwoya C., Gibson R.W., Smit N.E.J.M. and Carey E.E. 2001. Release of five sweetpotato cultivars in Uganda. *Hort. Science* (in press).
- Nishiyama I., Niyazaki T. and Sakamoto S. 1975. Evolutionary autopolyploidy in sweet potato and its progenitors. *Euphytica* 24: 197–208.
- Plomion C., Bahrman N., Durel C.E. and O'Malley D.M. 1995. Genomic mapping of *Pinus pinaster* (maritime pine) using RAPD and protein markers. *Hered.* 74: 661–668.
- Qu L. and Hancock J.F. 2001. Detecting and mapping of repulsion-phase linkage in polyploids with polysomic inheritance. *Theor. Appl. Genet.* 103: 136–143.
- Ripol M.I., Churchill G.A., Da Silva J.A.G. and Sorrells M. 1999. Statistical aspects of genetic mapping in autopolyploids. *Gene* 235: 31–41.
- Shiotani I. 1987. Genomic structure and the gene flow in sweetpotato and related species. In: Gregory P. (ed.), *Exploration, Maintenance, and Utilization of Sweetpotato Genetic Resources*. Rep 1st Sweetpotato Planning Conf. CIP, Lima, Peru, pp. 61–73.
- Shiotani I. and Kawase T. 1989. Genomic structure of the sweetpotato and hexaploids in *Ipomoea trifida* (HBK) DON. *Japan J. Breed.* 39: 57–66.
- Sinha S. and Sharma S.N. 1992. Taxonomic significance of karyomorphology in *Ipomoea* spp. *Jap. Mendel Soc.*
- Tanksley S.D., Ganai M.W., Prince J.P., De Vincente M.C., Bonierbale M.W., Broun P. et al. 1992. High-density molecular linkage maps of the tomato and potato genomes. *Genetics* 132: 1141–1160.
- Ting Y.C. and Kehr A.E. 1953. Meiotic studies in sweetpotato (*Ipomoea batatas* (L.) Lam.). *J. Hered.* 44: 207–211.
- Ukoskit K. and Thompson P.G. 1997. Autopolyploidy versus allopolyploidy and low-density randomly amplified polymorphic DNA linkage maps of sweetpotato. *J. Amer. Soc. Hort. Sci.* 122: 822–828.

- Van Ooijen J.W. and Voorrips R.E. 2001. Join Map® 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, The Netherlands.
- Verhaegen D. and Plomion C. 1996. Genetic mapping in *Eucalyptus urophylla* and *Eucalyptus grandis* using RAPD markers. *Genome* 39: 1051–1061.
- Vos P., Hogers R., Bleeker M., Reijans M., Van de Lee T., Hornes M. et al. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 23: 4407–4414.
- Vuylsteke M., Mank R., Antonise R., Bastiaans E., Senior M.L., Stuber C.W. et al. 1999. Two high-density AFLP linkage maps of *Zea mays* L.: analysis of distribution of AFLP markers. *Theor. Appl. Genet.* 99: 921–935.
- Wang G.L., Mackill D.J., Bonmann J.M., McCouch S.R., Champoux M.C. and Nelson R.J. 1994. RFLP mapping of genes conferring complete and partial resistance to blast in a durably resistant rice cultivar. *Genetics* 136: 1421–1434.
- Wu K.K., Burnquist W., Sorrells M.E., Tew T.L., Moore P.H. and Tanksley S.D. 1992. The detection and estimation of linkage in polyploids using single-dose restriction fragments. *Theor. Appl. Genet.* 83: 294–300.

