Theor Appl Genet (2003) 107:875–883
DOI 10.1007/s00122-003-1313-1

G. Fazio · S. M. Chung · J. E. Staub

Comparative analysis of response to phenotypic and marker-assisted selection for multiple lateral branching in cucumber (*Cucumis sativus* L.)

Received: 8 July 2002 / Accepted: 27 January 2003 / Published online: 14 June 2003

Abstract Yield increase in processing cucumber (*Cucumis sativus* L.) is positively correlated with an increase in number of fruit-bearing branches. Multiple lateral branching (MLB) is a metric trait controlled by at least five effective factors. Breeding efficacy might be improved through marker-assisted selection (MAS) for MLB. Experiments were designed to independently confirm previously determined linkage of molecular markers (L18-2-H19A SNP, CSWTAAA01 SSR, CSWCT13 SSR, W7-2 RAPD and BC-551 RAPD) to MLB, and to determine their utility in MAS. These markers were present in significantly higher frequency than expected (1, presence:3, absence; $p < 0.001$) in BC_2 plants selected based on a high MLB phenotype (BC_2PHE). However, markers that were considered selectively neutral fit the expected segregation of donor parent DNA in BC_2 progeny. Markers linked to MLB were used in MAS of BC_1 and BC_2 plants to produce BC_2MAS, and BC_3MAS progeny. Means for MLB in MAS populations were compared with backcross populations developed through phenotypic selection (BC_2PHE, BC_3PHE) and by random mating where no selection had been applied (BC_2RND, BC_3RND). Statistical analysis showed no significant differences ($p < 0.001$) between means of phenotypic (BC_2PHE = 3.02, BC_3PHE = 3.29) and marker-aided selection (BC_2MAS = 3.12, BC_3MAS = 3.11) for MLB. However, both phenotypic and MAS population means were significantly higher than the random control (BC_2RND = 2.27, BC_3RND = 2.41) for MLB. Thus, given the observed response to selection and the rapid life-cycle of cucumber (4 months), markers linked to MLB when used in MAS will most likely be effective tools in cucumber improvement.

Communicated by H. Nybom

Mention of trade name, proprietary product, or specific equipment does not constitute a guarantee or warranty by the USDA and does not imply its approval to the exclusion of other products that may be suitable

G. Fazio · S. M. Chung · J. E. Staub (✉)
USDA ARS/Department of Horticulture,
University of Wisconsin-Madison, Madison, WI 53706, USA,
e-mail: jestaub@facstaff.wisc.edu
Tel.: 608-262-0028
Fax: 608-262-4743

Present address:
G. Fazio, USDA ARS, Department of Horticultural Sciences,
Cornell University, Geneva, NY 14456, USA

Introduction

The yield (MT/ha) of U.S. processing cucumber [gynoecious (G) or monoecious (M)] reached a plateau in the 1980s, and remained generally unchanged during the 1990s (USDA-NASS 1999). Studies of biomass development in processing cucumbers (Widders and Price 1989) and source-sink relationships (Staub 1989; Staub et al. 1992) suggest that this recent plateau may be associated with net-photosynthetic capacity. Photosynthate resource limitations may explain why fruit developing from the first pollinated flower inhibits the development of subsequent fruits (Denna 1973; Fuller and Leopold 1977). Furthermore, allelic diversity in cucumber suggests that commercial cucumber germplasm has a narrow genetic base (Dijkhuizen et al. 1996; Meglic and Staub 1996; Meglic et al. 1996; Horejsi and Staub 1999). These and other factors (e.g., physiological limitations) may partially explain the yield plateau observed in this cucumber market class.

Serquen et al. (1997a) suggest that few genes (5 to 8) control days to anthesis, sex expression, mainstem length, multiple lateral branching (MLB), and fruit number and weight. All of these traits are considered to be components of yield in cucumber. Path analysis of the correlation between fruit number per plant and several plant traits (e.g., branches per plant, nodes per branch, pistillate nodes and fruit set) in cucumber revealed that MLB had the highest correlation ($r = 0.82$) with number of fruits per plant (Cramer and Wehner 2000). MLB is under the control of at least five effective factors and has a narrow-

sense heritability of about 0.4 to 0.8 depending on growing environment and population structure (Wehner et al. 1987; Serquen et al. 1997a, b). Serquen et al. (1997b) found three QTLs associated with the MLB trait that explained approximately 60% of the observed phenotypic variance. These attributes make MLB a likely candidate for the application of indirect selection procedures to increase yield in processing cucumber.

Standard gynoecious- and monoecious-processing cucumber hybrid varieties (M × M, G × M or G × G hybrids) are indeterminate (*De*), standard-sized leaf (*L*) possess few lateral branches (1 to 3) and display crown-set inhibition (i.e., early non-sequential fruit set). The gynoecious breeding line GY7 (experimental line G421; University of Wisconsin 1997) is a high fruit quality, standard-sized leaf (*LL*), determinate (*de*), non-sequential fruiting germplasm, but does not possess a MLB habit. In contrast, the monoecious elite line H-19 (University of Arkansas-Fayetteville 1993) has little leaves (*ll*), possesses a MLB and sequential fruiting habit, but often bears fruit of poor quality. The combination of these two lines may yield gynoecious, short statured (*de*) genotypes that possess a MLB habit (about 5 to 7 branches). Such genotypes could be sown in machine-harvest operations at relatively high densities (200,000 plants/ha) compared to indeterminate types (50,000 plants/ha), thus allowing for an increase in early, concentrated yield (Staub et al. 1992).

Backcross breeding is often used to transfer favorable alleles (e.g., MLB) from a donor genotype (e.g., H-19), which usually has poor overall quality, to an elite (high quality) genotype (e.g., G421) (Allard 1960). Gain from selection through backcrossing can be conventionally achieved by using phenotypic data for selection, by marker assisted selection (MAS), or by schemes that employ a combination of both procedures (Melchinger 1990; Stuber 1994; Hospital and Charcosset 1997; Hospital et al. 1997; Ribaut and Hoisington 1998). Possible advantages of MAS via backcrossing during cucumber breeding lies in the fact that sex expression can be chemically manipulated to allow for the completion of three to four life cycles per year. Moreover, MAS can be more efficient than phenotypic selection methods where the target trait is metric, heritability is low and the environment affects trait expression (Hospital et al. 1992; Dudley 1993; Hospital et al. 1997; Knapp 1998).

Selection of determinate MLB genotypes in field or greenhouse environments is not possible because of the lack of apical dominance in MLB genotypes (i.e., the inability to differentiate MLB determinate vs indeterminate genotypes) and the effects of day length and light intensity on trait expression (i.e., the inability to completely classify genotypes differing in size and sex) (Staub et al. 1995). It would be useful to develop a MAS procedure that would allow for the selection of such complex traits in either environment. Therefore, using previously identified and newly developed markers, experiments were designed to compare the effects of marker-assisted and phenotypic selection. Enhancement of selection efficiency for MLB using MAS could result in an increase in the number of selection cycles per year.

Materials and methods

The comparison of efficiency of MAS and phenotypic selection required the development of strategic BC populations, the confirmation of QTLs associated with MLB described by Serquen et al. (1997b), and the analysis of BC populations under a uniform environment. This was accomplished by employing previously identified random amplified polymorphic DNA (RAPD), a sequence-characterized amplified region (SCAR) and simple sequence repeat (SSR) markers (Fig. 1) (Serquen et al. 1997b; Fazio et al. 2002a) for the production and evaluation of backcross progeny originating from MAS, phenotypic field selection and from populations where no selection had been applied.

Phenotypic selection

Germplasm development

Base populations for selection experiments were produced in a greenhouse and in an open-field between 1997 and 2000. In the winter of 1997 the gynoecious (*FF*) determinate cucumber inbred line G421 possessing normal-sized leaves and low lateral branch numbers (between 1 to 2) was crossed with the monoecious (*ff*) indeterminate little-leaf inbred line H-19 possessing high lateral branching (between 6 to 20) to provide F_1 progeny for backcrossing (see Fig. 2). In the spring of 1998, the sex of 15 gynoecious F_1 greenhouse-grown (Arlington, Wisconsin) plants (*Ff*) was altered to produce staminate flowers by aerosol application of 2–3 ml of silver 3-mM thiosulfate at the third leaf stage (Beyer 1976). Converted plants were used to pollinate 200 G421 plants resulting in BC_1 seed, which formed the base population for selection.

Selection cycle 1

Parental lines, 36 F_1 progeny and 2,985 BC_1 progeny were evaluated for MLB number in the summer of 1998 at the University of Wisconsin Agricultural Experiment Station, Hancock, Wisconsin (Typic Udipsament; sandy mixed, mesic). Ten to 36 BC_1 seed of each family (progeny of a single G421 plant) were planted 0.6-m apart on rows spaced 1.5-m apart. Plants were maintained under standard cultural practices.

Data on MLB number were recorded 40 days after sowing when approximately 50% of the BC_1 plants had reached anthesis. MLB number was defined as the number of primary branches (25 cm or longer) present in the first five nodes of the mainstem. The top five percentile (200/4,000) of the population was selected for backcrossing to G421, corresponding to a selection intensity (*i*) of 2.063 and a truncation point of 1.645 standard-deviation units from the mean (Falconer and Mackay 1996). The sex expression of selected gynoecious BC_1 plants (*Ff or FF*) was altered by aerosol application of 6-mM silver thiosulfate as three treatments at 5-day intervals. Two hundred selected plants were then self-pollinated and backcrossed to field-grown G421 plants (recurrent parent) to produce BC_1S_1 and BC_2 seed, respectively. The BC_2 population derived from these backcrosses was designated BC_2PHE.

Selection cycle 2 and evaluation

In the summer of 1999, seed of BC_2PHE (3,233) and BC_1S_1 (1,367) families selected in summer 1998, and parents, F_1 (36), F_2 (400) and BC_1 (397) progeny seed lots were planted at Hancock, Wisconsin. Response to selection in the previous cycle of selection (1998) was evaluated and phenotypic selection was practiced on the

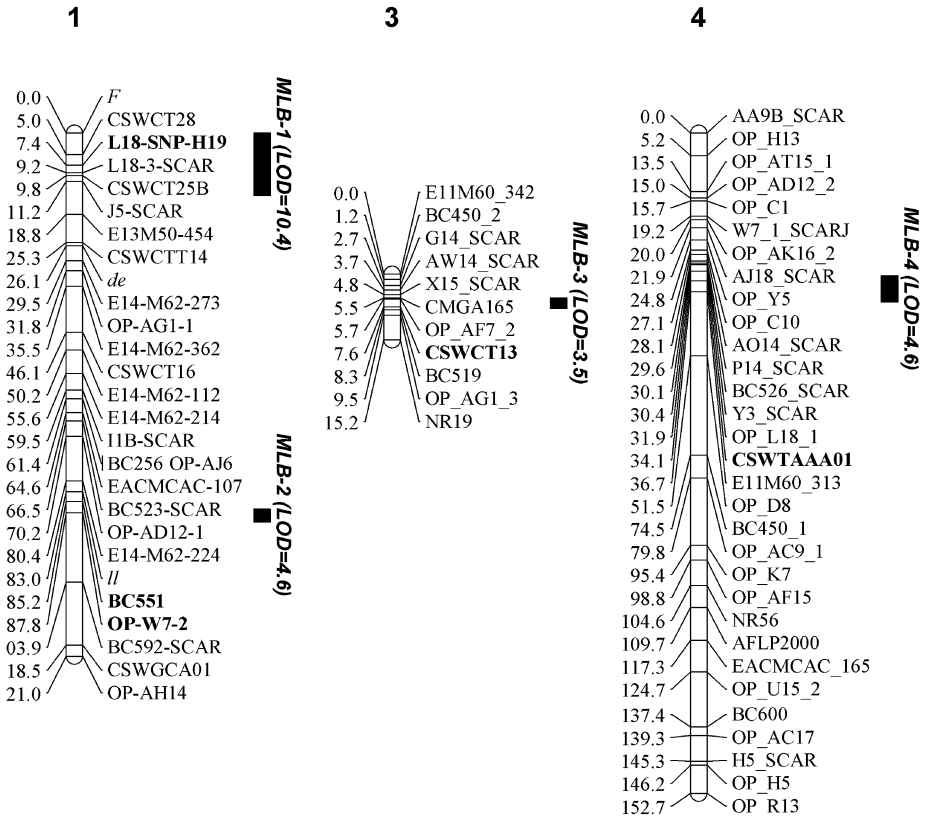

Fig. 1 Linkage Groups 1, 3 and 4 derived from linkage analysis of recombinant inbred lines were derived from the G421 × H-19 line-mating in cucumber (Fazio 2001; Fazio et al. 2002b). Approximate map positions (centimorgans; Kosambi function) of marker loci (L18-2-H19A, CSWTAAA01, OP-W7-2, BC551, CSWCT13) were used in MAS (*bold*) and their associated multiple lateral-branching (MLB) QTLs (*bars*). Correspondence to linkage groups according to Serquen et al. (1997b) and Bradeen et al. (2001) is given by the letters (B, C and D) and roman numerals (I, II, III, IV), respectively

BC_2PHE and BC_1S_1 populations. The BC_1 population was included as a reference point to the previous selection cycle. Parents, F_1, F_2, BC_1, BC_1S_1 and BC_2PHE progeny were arranged in a randomized complete block design with three replications of 12 plants each (a total of 36 plants per family). Seed was planted 0.45-cm apart within rows and positioned 1.5-m apart. Data collection, sex conversion and selection of BC_2PHE plants were performed as in 1998. Similar selection intensity was applied to the BC_2PHE population as was applied to the BC_1 population (1998) (i.e., i = 2.063). Selections were then backcrossed to G421 to produce the BC_3PHE generation. Concurrently, monoecious high MLB BC_1S_1 and BC_2PHE plants (MLB $\geq$ 5) were self-pollinated to provide BC_1S_2 and BC_2S_1 progeny, respectively.

QTL confirmation

Germplasm

Empirical validation of markers previously associated with the MLB trait (Serquen et al. 1997b) was undertaken by testing the frequency of such markers in BC_2PHE plant selections (i.e., two cycles of selection). These selections then became parents of the BC_3PHE population. This required the development, and analysis of BC_2PHE individuals to determine if alleles of markers linked to MLB, were in higher frequency. For this purpose, DNA was extracted from a randomly sampled subset of 75 individuals in the upper 5% distribution (tail) of MLB in the BC_2PHE population. These individuals had undergone two cycles of phenotypic selection for MLB, and had a minimum of five lateral branches. Selected BC_2PHE individuals were the same as those backcrossed to G421 to generate the BC_3PHE population.

DNA markers

Allelic frequencies were determined in selected BC_2PHE individuals at two RAPD loci (OP-W7-2 RAPD, BC 551 RAPD), two SSR loci (CSWCT13, CSWTAAA01) and one single-nucleotide polymorphism (SNP) locus (L18-2/H19A) (Fazio 2001; Fazio et al. 2002a). These marker loci were found to be associated with the MLB character (Serquen et al. 1997b; Fazio et al. 2002b). For instance, the RAPD markers and the original L18-2 band leading to the development of the SNP marker were identified by Serquen et al. (1997b), and explained approximately 40% of the observed phenotypic variation for multiple lateral branching in F_3 families derived from a G421 × H-19 mating. For certain QTLs (i.e., L18-2, CSWTAAA01 and CSWCT13) only a single marker (instead of flanking markers) was available for MAS since attempts to convert other linked markers to SCARs or SNPs were unsuccessful (Fazio et al. 2002a).

Plants were also genotyped using three randomly chosen unlinked markers that were not associated with MLB. These markers were used as positive controls for the expected segregation of markers in BC_2 individuals, and included two RAPD markers (W7-1 RAPD, OP-H5 RAPD) and a SCAR marker ($N6_{800}$) developed from N6 RAPD (Serquen et al. 1997b; Fazio 2001; Fazio et al. 2002a). The expectation of observing an individual being heterozygous (Aa) for the non-recurrent parental allele (A) in a random BC_2 population is one (Aa) in four individuals (25%, a 1:3 ratio). A one-tailed chi-square test (one degree of freedom) was performed on the markers examined to determine their fit to the expected frequency.

Marker-assisted selection

Germplasm development

During the fall of 1999 and winter 2000, remnant seed of the BC_1 and BC_2PHE populations were planted for the development of

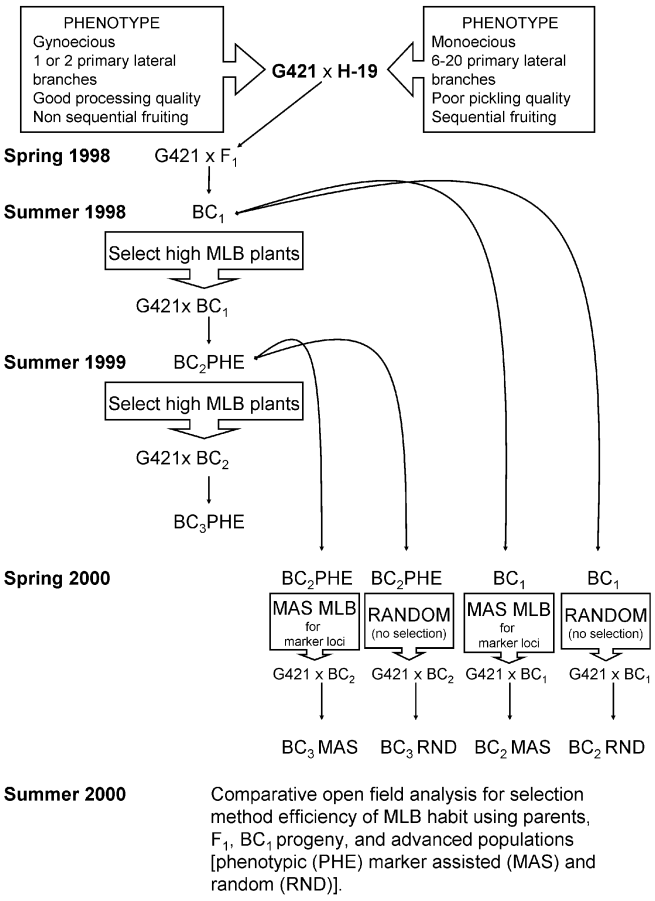

Fig. 2 Population development for comparative analysis of selection method efficiency of multiple lateral branching in cucumber. The text in *boxes* outline the type of selection applied to a population. Progeny in two base backcross populations (BC_1 and BC_2PHE) were selected using three selection schemes: marker-assisted (MAS), phenotypic (PHE), and random (RND)

populations for MAS. Nine hundred and sixty BC_1 and 960 BC_2PHE seedlings were grown in a greenhouse (Madison, Wisconsin) in 4 × 12 plastic flats (Com-Pack, Minneapolis, Minn.) filled with Fafard Pro-Mix #2 (Conrad Fafard Inc., Agawan, Mass.).

Genotyping and selection

One leaf (<1 cm^2) from each plant in each population was collected from 2-week-old seedlings and introduced in a unique well-position of a 0.5-ml polypropylene 96-well microtiter plate (Fisher Scientific, Pittsburgh, Pa.). The tissue was then immediately lyophilized in preparation for DNA extraction using the Aquapure genomic DNA extraction Kit (Bio-Rad Laboratories, Hercules, Calif.) methodology. Marker loci W7-2, CSWTAAA01, CSWCT13, BC-551 and L18-2/H19A were used to genotype BC_1 (960) and BC_2PHE (960) progeny. The PCR reactions (five markers × 960 plants × two populations) were performed in 15-μl volumes of a uniform reaction mixture [3 mM of MgCl, 0.2 mM of dNTPs, 15–20 ng of DNA, 0.4 μM of each primer, *Taq* polymerase and commercial buffer (Promega, Madison, Wis.)] incorporating 10 μl of light-weight mineral oil overlay (Fisher Scientific, Pittsburg, Pa.). Thermal cycling conditions for each primer were different for each marker locus and followed published protocols for those loci (Serquen et al. 1997b; Fazio et al. 2002a). The PCR reactions were resolved in 1.6% agarose (Life Technologies, Gaithersburg, Md.) by electrophoresis and staining with ethidium bromide (0.5 μg/ml). Gels were visualized with the Dark Reader trans-illuminator (Clare Chemical Inc., Denver, Colo.) and a closed circuit digital (CCD) camera.

BC_1 and BC_2PHE plants possessing high parent (H-19) alleles for the MLB number of each marker were then transplanted into 20 cm-diameter pots. Selected plants originating from the BC_2PHE population possessed varying numbers of high-parent alleles. Thus, populations derived from individuals lacking one of the five high-parent alleles for markers CSWCT13, CSWTAAA01 and L18-2-H19A were designated BC_3MAS-A (eight individuals), BC_3MAS-B (nine individuals) and BC_3MAS-C (five individuals), respectively. A population designated as BC_3MAS-D was derived from 21 selected BC_2PHE plants possessing all five high-parent alleles. Selected seedlings were gynoecious, and therefore they were treated as previously stated with silver thiosulfate for sex conversion to staminate flowers for further backcrossing to G421 to produce BC_nMAS families (i.e., BC_2MAS and BC_3MAS progeny were produced from BC_1 and BC_2PHE selected plants, respectively). DNA was extracted from each selected donor plant, and its allelic constitution at each selectable MLB marker locus was confirmed by PCR assay.

Randomly advanced control population

Backcross populations were random-mated to produce control progeny (BC RND) simulating lack of selection. Fifty randomly selected seed from the BC_1 and BC_2PHE were germinated and grown in a greenhouse (Arlington, Wis.), and then backcrossed to the recurrent parent (G421) to obtain BC_2RND and BC_3RND populations, respectively (see Fig. 2).

Experimental design and field evaluation

In the summer of 2000, 39 seeds each of 39 BC_1, 24 BC_1S_1, 25 BC_1S_2, 44 BC_2MAS, 38 BC_2PHE, 37 BC_2RND, 41 BC_3MAS, 59 BC_3PHE, and 31 BC_3RND families (experimental units), along with original parents (G421, H-19) and their F_1 progeny, were planted in a randomized complete block design. Experimental units (360) were replicated in three blocks (i.e., 1,080 plots) to yield 14,040 observational units (plants). Seeds of each entry were planted 40-cm apart in 5.2-m long plots (13 seeds) on 1.5-m centered rows at Hancock Wis. Standard cultural practices for Wisconsin were used. The number of lateral branches was counted on all emerged plants 55 days post-planting when approximately 85% of the plots had reached anthesis.

Statistical analysis

Data on MLB number were analyzed using a mixed-models procedure (PROC Mixed) employed by SAS (Littell et al. 1996). While population types were considered fixed effects, families nested within populations, blocks and interactions were considered as random effects. Least square means and *t*-test probabilities (p = 0.01) of mean differences were calculated for all treatments.

Results

Response to phenotypic selection

Comparative mean analysis of BC_1, BC_2PHE, BC_1S_1 and F_2 populations, as well as parents and F_1 progeny in the summer of 1999, was used to ascertain whether there had been a response to phenotypic selection in the backcross populations examined (Fig. 2). Although 200 plants had

Table 1 Means and standard errors for multiple lateral branching of cucumber populations developed by open-field phenotypic selection during 1998 and 1999

Year	Generation	N	Mean	SE	SD	N selected[a]	Mean of selected[a]
1998	BC_1	2,985	2.23	0.02	1.20	126	4.9
1999	G421	150	1.75	0.06	0.84	n.a.	n.a.
1999	H19	36	7.9	0.26	1.60	n.a.	n.a.
1999	F_1	36	4.61	0.16	0.98	n.a.	n.a.
1999	BC_1	397	2.21	0.06	1.14	n.a.	n.a.
1999	BC_2PHE	3,233	2.60	0.02	1.21	167	5.6
1999	BC_1S_1	1,367	3.38	0.04	1.62	n.a.	n.a.
1999	F_2	400	3.58	0.07	1.38	n.a.	n.a.

[a] Plants with a high multiple lateral branching phenotype selected for generation advancement

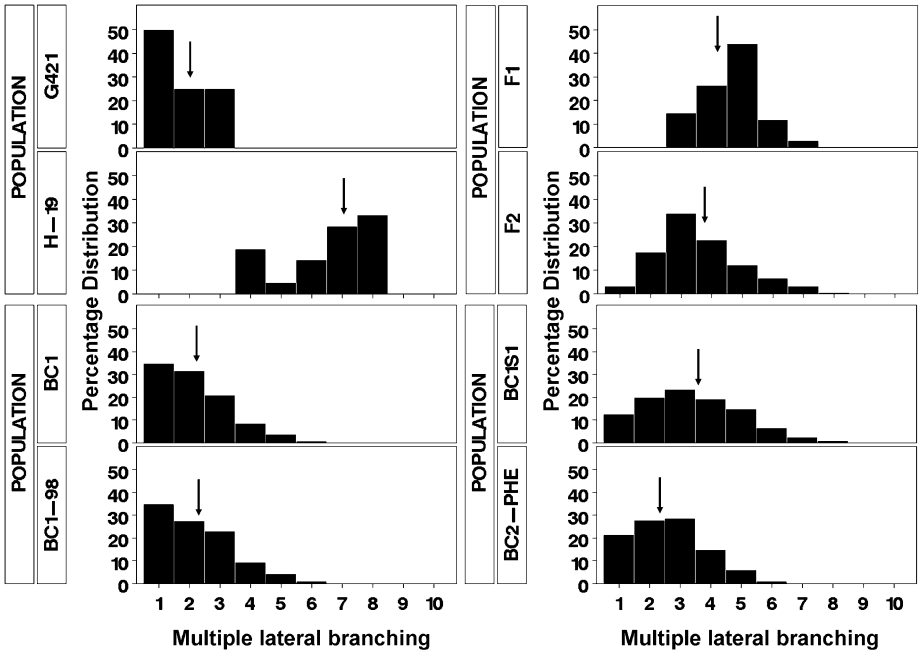

Fig. 3 The percentage distribution of lateral branch number in parental, F_1 and segregating populations evaluated in an open-field (Hancock, Wis.) in the summer of 1999. Population means are indicated by *arrows*

been backcrossed to G421 during Cycle 1 (1998), seed was recovered from only 126 backcrosses (64% of the BC_1 population) because of unsuccessful pollinations and the formation of parthenocarpic fruits. The mean of selected BC_1 plants was 4.90. The selection intensity was 1.06 (adjusted to 1/2 of 2.116 because no selection was applied to the female recurrent parent), and the expected response from selection was 0.6 units. The realized gain from selection based on the difference of the generation means (BC_1 = 2.21 vs BC_2PHE = 2.60) was approximately 0.4 units. The base BC_1 population mean (2.23 ± 0.02), as well as parents and F_1 progeny means, were similar across years (1998 and 1999; Table 1). The distributions of the number of MLB differed between BC_1 (mean = 2.2; range = 1 to 7), BC_2PHE (mean = 2.6; range = 1 to 7) and BC_1S_1 (mean = 3.4; range = 1 to 10), and F_2 (mean = 3.6; range = 1 to 8) populations (Fig. 3). All backcross progeny were gynoecious, reflecting the dominant nature of the *F* locus derived from the recurrent parent. Monoecious plants were recovered in BC_1S_1, BC_1S_2, and BC_2S_1 progeny in high frequency, possibly because of the linkage between the *f* allele and *MLB-1*.

QTL confirmation

One-tailed chi-square tests (*df* = 1) revealed that all markers previously identified by Serquen et al. (1997b) as associated to MLB were present in significantly higher frequency than the expected 1:3 (donor:recurrent parent allele) (p = 0.001) in the selected 75 BC_2PHE plants (Table 2). However, markers not associated with MLB and presumably "neutral" in their effect on MLB did fit the expected 1:3 (25%) frequency. The markers exhibiting the largest deviation from the expected frequency were L18-2/H19A, W7-2, and BC551.

Marker-assisted selection

No significant difference was detected in the mean number of lateral branches between populations derived by phenotypic selection (BC_2PHE, BC_3PHE) and marker-assisted selection (BC_2MAS, BC_3MAS) (Table 3). Population means derived from either selection method were, however, significantly higher ($p < 0.0001$) than populations where no selection had been applied (BC_2RND, BC_3RND). Moreover, the percent of plants with five or

Table 2 Chi-square tests (p = 0.001) for segregation of markers associated with QTLs for multiple lateral branching (MLB) in cucumber and markers not associated (neutral) with MLB in BC_2PHE plants

Marker[a]		Locus type	LG[b]	N	Expected[c]	Observed	χ^2	R^2 (%)[d]
Name	Type							
CSWCT13	SSR	QTL	3	75	19	39	21.87	3.5
OP W7-2	RAPD	QTL	1	75	19	47	42.56	19.2
BC551	RAPD	QTL	1	74	19	48	45.63	19.2
CSWTAAA01	SSR	QTL	4	75	19	44	34.00	3.2
L18-2/H19A	SNP	QTL	1	75	19	58	82.16	10.4
OP W7-1	RAPD	Neutral	6	75	19	18	0.03	n.a.
N6	SCAR	Neutral	6	74	19	21	0.27	n.a.
OP H5	RAPD	Neutral	4	75	19	20	0.08	n.a.

[a] SSR = simple sequence repeat, RAPD = random amplified polymorphic DNA, SCAR = sequence characterized amplified region, SNP = single nucleotide polymorphism
[b] LG = Linkage group according to Fazio 2001 and 2002b
[c] Expected number of individuals being heterozygous (Aa) for the non-recurrent dominant parental allele (A) in a random BC_2 population (25%)
[d] Percent of the observed phenotypic variation averaged over locations (n.a. = not applicable) (Fazio et al. 2002b)

Table 3 Comparison of mean lateral branch number between cucumber populations derived from identical base populations (α = 0.01)

Base population	Comparison[a] pop. (1) vs pop. (2)	LS mean difference	SE of mean difference	P value[b]
BC_1	BC_2MAS vs. BC_2PHE	0.11	0.11	0.3295
BC_1	BC_2MAS vs. BC_2RND	0.85	0.11	<0.0001
BC_1	BC_2PHE vs. BC_2RND	0.74	0.11	<0.0001
BC_2PHE	BC_3MAS-D vs. BC_3PHE	–0.18	0.15	0.2106
BC_2PHE	BC_3MAS-D vs. BC_3RND	0.70	0.16	<0.0001
BC_2PHE	BC_3PHE vs. BC_3RND	0.89	0.11	<0.0001

[a] MAS = marker-assisted selection, PHE = phenotypic selection and RND = random control
[b] $Pr > |t|$ for H_0: LSMean(1) = LSMean(2)

Table 4 Descriptive statistics of the mean number of lateral branches of cucumber populations evaluated in an open field in Wisconsin (2000)

Population[a]	N	Mean	Standard deviaton	Standard error	Variance
F_1	44	4.68	0.88	0.36	0.78
G421	52	1.50	0.64	0.36	0.41
H-19	45	7.60	1.35	0.36	1.83
BC_1	1,000	2.96	1.13	0.11	1.27
BC_1S_1 PHE	532	4.37	1.60	0.13	2.58
BC_1S_2 PHE	638	5.40	1.57	0.13	2.46
BC_2MAS	1,265	3.12	1.26	0.11	1.59
BC_2PHE	957	3.02	1.17	0.11	1.37
BC_2RND	1,085	2.27	1.08	0.11	1.18
BC_3MAS-A	238	2.83	1.16	0.19	1.36
BC_3MAS-B	370	2.95	1.10	0.16	1.21
BC_3MAS-C	171	2.63	1.05	0.21	1.10
BC_3MAS-D	400	3.11	1.23	0.15	1.53
BC_3PHE	1,385	3.29	1.22	0.11	1.50
BC_3RND	955	2.40	1.11	0.12	1.25

[a] PHE = phenotypic selection, MAS = marker-assisted selection and RND = random mated (no selection)

more lateral branches in the BC_2PHE (10%) and BC_2MAS (14%) populations were higher than those of their random control (3%). This tendency was also evident among BC_3 populations [BC_3PHE (16%), BC_3MAS (14%) and BC_3RND (3.6%)]. The mean lateral branch number of F_1 progeny (4.68) approached the mid-parent value (4.55) (Table 4). Thus, empirical analysis of parental and F_1 progeny distributions of MLB confirms its additive nature as described by Serquen et al. (1997b).

Discussion

Backcross populations (BC_1, BC_2PHE) were developed to produce equivalent BC generations through MAS, phenotypic selection, and relaxed selection. These populations were then compared to determine if MAS could be effective as an indirect selection tool in populations developed from matings between elite parents [G421 (G, *de*) × H-19 (M, *De*)]. The selection experiments described herein revealed that the MLB mean of populations derived by either phenotypic or marker-assisted selection

was always less than the mid-parent (F_1) value. The value of the F_1 progeny represents the maximum achievable value for an additive trait like MLB in a backcross procedure (the low parent is the recurrent parent; Serquen et al. 1997b). However, as a result of an increase in homozygous loci that condition MLB, this mean value was predictably exceeded in BC_1S_1 and BC_1S_2 families derived from selected high MLB BC_1 and BC_1S_1 plants, respectively. The mean MLB of any BC_1S_2 family, however, was never equal to or higher than the mean of the high lateral-branched donor parent (H-19) (Table 4). The fruit and several economically important plant characteristics (i.e., *de, l*, and *F*) of the BC_3 populations (about 87.5% of G421) were predictably similar to G421 by observation, indicating that high MLB BC_3 progeny possessed high quality attributes and were amenable to further refinement.

Reliable estimates of QTL positions and effects are a requirement for the effective application of MAS (Beavis 1998; Melchinger et al. 1998; Stuber et al. 1999). This requires the discovery and validation of such estimates. The experiments conducted sought validation of marker-QTL association by monitoring increases in the allelic frequency of high-donor parent alleles for MLB in BC PHE populations developed during phenotypic selection for MLB. These populations were used to test the H_o that, under phenotypic selection, marker alleles linked to MLB would not be present in higher frequency in selected individuals (BC_2PHE) than in "neutral" marker alleles segregating randomly. Markers previously identified by Serquen et al. (1997b) as linked to MLB exhibited a significant deviation from the expected ratio 1:3 (donor:recurrent parent alleles) in a BC_2 population. However, presumably "neutral" markers not associated with MLB fit this expectation (Table 2). Theoretically, deviation from this ratio could be the result of non-random effects such as selection, random genetic drift, or spurious sampling error. The relative size of the BC_2PHE population (derived from 126 BC_1 plants) most likely mitigates the potential of random drift as a source of deviation from the expected outcome.

Response to MAS of a quantitative trait such as MLB is dependent on the magnitude of QTL effects, the distance between that marker, and a specific QTL and environmental variance (Dudley 1993; Knapp 1998). Selection experiments allowed for the ranking of markers according to their importance in response to selection for use in MAS. The high proportion of progeny possessing the L18-2 and W7-2 favorable allele in selected individuals (Table 2) agreed with the findings by Serquen et al. (1997b). The QTLs associated with these markers explained more than 50% of the observed phenotypic variation in QTL experiments that used F_2:F_3 families derived from G421 and H-19. The relative importance of L18-2 was also defined by the comparatively high MLB mean of the BC_3MAS-D population (3.11), which possessed all five markers used in MAS. In contrast, the mean MLB in BC_3MAS-C individuals was 2.63, and lacked the high parent L18-2 allele. This is an indication that selection could contribute to the observed deviation from the expected 1:3 ratio.

Genetic linkage mapping using 171 recombinant inbred lines detected a loose linkage (12.5 cM) between L18-2 and the determinate gene (*de*) (Fazio 2001; Fazio et al. 2002b). Composite interval mapping of the MLB trait revealed that a QTL with relatively large effect (R^2 = 11.5%) was tightly linked to the determinate gene (present in G421 and absent in H-19). This QTL region has a pleiotropic effect on lateral branching, sex expression, and mainstem length (Serquen et al. 1997a; Fazio 2001). Although *de* conditions the termination of the growing point at the terminal whorl, the length of lateral branches is quantitatively inherited, and determinate plants can vary dramatically (Staub et al. 1995; Serquen et al. 1997b). Moreover, the identification of determinate, multiple lateral individuals is difficult because of the lack of apical dominance in multiple-branched genotypes. The inability to isolate the determinate character in high MLB genotypes during phenotypic selection is most likely due to the documented pleiotropic effects of QTLs and epistatic interactions of their alleles. Several codominant markers mapped to the QTL region around *de* (Fazio 2001; Fazio et al. 2002b) may prove useful in MAS for the development of high yielding MLB, *de* genotypes. This task would be accomplished by high resolution mapping and screening large populations for crossover events in the *de* region.

Studies relating the efficiency of MAS to the improvement of quantitative traits are scarce. Romagosa et al. (1999) compared MAS to phenotypic selection and to MAS-based tandem selection (MAS followed by phenotypic selection) for grain yield in doubled-haploid lines (DHL) in barley. MAS for the QTL that had the largest effect was as effective as phenotypic selection. Tandem selection was at least as effective as MAS and phenotypic selection. A similar study compared sweet-corn progeny produced by MAS and phenotypic selection for seed emergence and eating quality (Yousef and Juvik 2001). MAS resulted in higher gains than phenotypic selection in 38% of the paired family comparisons versus 4% where phenotypic selection had higher gains. Moreover, MAS and phenotypic selection were on average 10.9% and 6.9% higher, respectively, than randomly selected controls for the traits examined.

Marker-assisted backcrossing has been proposed as the most effective strategy for transfering a small number of QTL alleles from the same parent (Hospital et al. 1997). Shen et al. (2001) used MAS in rice (*O. sativa* L.) during backcrossing of five QTLs conditioning rooting depth to obtain a positive response to selection. Such progress during phenotypic selection would have been difficult because of the problems associated with the evaluation of root phenotypes. Likewise, Causse et al. (2001) conducted experiments with marker-assisted backcrossing to transfer five QTLs for organoleptic quality in elite tomato lines utilizing a selection scheme according to Hospital et al. (1997). They reported that only a relatively small number of plants (5 of 267) were further backcrossed after the first

generation of foreground MAS, suggesting the use of an extremely stringent selection protocol based solely on molecular QTL analysis.

The QTL-MLB associations reported by Serquen et al. (1997b) were confirmed in this study. Their studies demonstrated that L18-2 and W7-2 explained about 40% and about 14% of the phenotypic variation, respectively. A more recent analysis using RILs derived from the same population used by Serquen et al. (1997b) indicates that MLB is controlled by relatively few loci (4) with major effects, and an additional number (6) with smaller but significant effects (Fazio 2001; Fazio et al. 2002b). The present study attempted the introgression of marker-associated QTL regions with large effects on MLB (i.e., L18-2, W7-2, BC551, CSWTAAA01, and CSWCT13) under greenhouse conditions. The data presented indicate that these markers will have utility for increasing the efficiency of selection for MLB (Table 3). MAS allows for an increase of the number of selection cycles per year when compared to open-field phenotypic selection. Moreover, MAS can identify individuals with allelic arrays that can be customized during greenhouse selection leading to an increase in preferred genotypes for transplanting to an open-field. These "selected" plants can then be selected for other desirable traits.

The ultimate intent of selection was to retain BC populations large enough (>100 selections, 5% selection intensity) to maintain allelic variability at unselected loci for further evaluation of MLB and other plant architecture (e.g., plant stature) and quality (e.g., fruit length and diameter) traits. MAS was effective in this regard since not only was the number of selected BC_2 and BC_3 plants possessing five or more lateral branches similar to that resulting from phenotypic selection, but the number of individuals selected was large enough to allow adequate phenotypic variation for continued selection of other traits. Although this study shows that MAS of QTLs can work, the utility of MAS in a breeding program needs to be evaluated on a case by case basis, taking into consideration important economic factors such as investment costs, profit margin, market share, and opportunity costs. These factors were not considered in this study, but are of consequence for commercial implementation of MAS technologies.

Acknowledgements This research was partially supported by grant No. IS-2708-96 from the U.S.-Israel Binational Agricultural Research and Development (BARD) Fund.

References

Allard RW (1960) Principles of plant breeding. Wiley, New York

Beavis WD (1998) QTL analyses: power, precision, and accuracy. In: Paterson AH (ed) Molecular dissection of complex traits. CRC Press, Boca Raton pp 145–162

Beyer EMJ (1976) Silver ion: a potent anti-ethylene agent in cucumber and tomato. HortScience 11:195–196

Bradeen JM, Staub JE, Wye C, Antonise R, Peleman J (2001) Towards an expanded and integrated linkage map of cucumber (*Cucumis sativus* L.). Genome 44:111–119

Causse M, Lecomte L, Baffert N, Duffe P, Hospital F (2001) Marker-assisted selection for the transfer of QTLs controlling fruit quality traits into tomato elite lines. Acta Hort 546:557–564

Cramer CS, Wehner TC (2000) Path analysis of the correlation between fruit number and plant traits of cucumber populations. HortScience 35:708–711

Denna DW (1973) Effects on genetic parthenocarpy and gynoecious flowering habit on fruit production and growth in cucumber, *Cucumis sativus*. J Am Soc Hort Sci 98:602–604

Dijkhuizen A, Kennard WC, Havey MJ, Staub JE (1996) RFLP variation and genetic relationships in cultivated cucumber. Euphytica 90:79–87

Dudley JW (1993) Molecular markers in plant improvement: manipulation of genes affecting quantitative traits. Crop Sci 33:660–668

Falconer DS, Mackay FC (1996) Introduction to quantitative genetics. Longman Ltd, Harlow Essex, England

Fazio G (2001) Comparative study of marker-assisted and phenotypic selection and genetic analysis of yield components in cucumber. Dissertation, University of Wisconsin-Madison

Fazio G, Staub JE, Chung SM (2002a) Development and characterization of PCR markers in cucumber (*Cucumis sativus* L.). J Am Soc Hort Sci 127:545–557

Fazio G, Staub JE, Stevens MR (2002b) Genetic mapping and QTL analysis of horticultural traits in cucumber (*Cucumis sativus* L.) using recombinant inbred lines. Theor Appl Genet (in press)

Fuller GL, Leopold CA (1977) The role of nucleic acid synthesis in cucumber fruit set. J Am Soc Hort Sci 102:384–388

Horejsi T, Staub JE (1999) Genetic variation in cucumber (*Cucumis sativus* L.) as assessed by random amplified polymorphic DNA. Genet Res Crop Evol 46:337–350

Hospital F, Charcosset A (1997) Marker-assisted introgression of quantitative trait loci. Genetics 147:1469–1485

Hospital F, Chevalet C, Mulsant P (1992) Using markers in gene introgression breeding programs. Genetics 132:1199–1210

Hospital F, Moreau L, Lacoudre F, Charcosset A, Gallais A (1997) More on the efficiency of marker-assisted selection. Theor Appl Genet 95:1181–1189

Knapp SJ (1998) Marker-assisted selection as a strategy for increasing the probability of selecting superior genotypes. Crop Sci 38:1164–1174

Littell RC, Milliken GA, Stroup WW, Wolfinger RD (1996) SAS System for Mixed Models. SAS Istitute Inc, Cary, North Carolina

Meglic V, Staub JE (1996) Genetic diversity in cucumber (*Cucumis sativus* L.). II. An evaluation of selected cultivars released between 1846 and 1978. Genet Res Crop Evol 43:547–558

Meglic V, Serquen F, Staub JE (1996) Genetic diversity in cucumber (*Cucumis sativus* L.). I. A re-evaluation of the U.S. germplasm collection. Genet Res Crop Evol 43:533–546

Melchinger AE (1990) Use of molecular markers in breeding for oligogenic disease resistance. Plant Breed 104:1–19

Melchinger AE, Utz HF, Schon CC (1998) Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149:383–403

Ribaut JM, Hoisington D (1998) Marker-assisted selection: new tools and strategies. Trends Plant Sci 3:236–239

Romagosa I, Han F, Ullrich SE, Hayes PM, Wesenberg DM (1999) Verification of yield QTLs through realized molecular marker-assisted selection responses in a barley cross. Mol Breed 5:143–152

Serquen FC, Bacher J, Staub JE (1997a) Genetic analysis of yield components in cucumber at low plant density. J Am Soc Hort Sci 122:522–528

Serquen FC, Bacher J, Staub JE (1997b) Mapping and QTL analysis of horticultural traits in a narrow cross in cucumber

(*Cucumis sativus* L.) using random-amplified polymorphic DNA markers. Mol Breed 3:257–268
Shen L, Courtois B, McNally KL, Robin S, Li Z (2001) Evaluation of near-isogenic lines of rice introgressed with QTLs for root depth through marker-aided selection. Theor Appl Genet 103:75–83
Staub JE (1989) Source-sink relationships in cucumber. Cucurbit Genet Coop Rep 12:11–14
Staub JE, Knerr LD, Hopen HJ (1992) Plant density and herbicides affect cucumber productivity. J Am Soc Hort Sci 117:48–53
Staub JE, Bacher J, Crubaugh L (1995) Problems associated with the selection of determinate cucumber (*Cucumis sativus* L.) plant types in a multiple lateral background. Cucurbit Genet Coop Rep 18:7–9
Stuber CW (1994) Breeding multigenic traits. Adv Cell Mol Biol Plant 1:97–115
Stuber CW, Polacco M, Senior ML (1999) Synergy of empirical breeding, marker-assisted selection, and genomics to increase crop yield potential. Crop Sci 39:1571–1583
University of Arkansas-Fayetteville (1993) H-19 Cucumber Plant Variety Protection Certificate PVP no. 8900073. United States Dept of Agriculture, Agricultural marketing service Washington D.C., USA
USDA-NASS (1999) Statistical Bulletin 946c
Wehner TC, Staub JE, Peterson CE (1987) Inheritance of littleleaf and multi-branched plant habit in cucumber. Cucurbit Genet Coop Rep 10:33–34
Widders IE, Price HC (1989) Effects of plant density on growth and biomass partitioning in pickling cucumbers. J Am Soc Hort Sci 114:751–755
Yousef GG, Juvik JA (2001) Comparison of phenotypic and marker-assisted selection for quantitative traits in sweet corn. Crop Sci 41:645–655