Theor Appl Genet (1998) 97: 170–180

**D. Bernacchi · T. Beck-Bunn · D. Emmatty
Y. Eshed · S. Inai · J. Lopez · V. Petiard
H. Sayama · J. Uhlig · D. Zamir · S. Tanksley**

Advanced backcross QTL analysis of tomato. II. Evaluation of near-isogenic lines carrying single-donor introgressions for desirable wild QTL-alleles derived from *Lycopersicon hirsutum* and *L. pimpinellifolium*

Received: 27 October 1997 / Accepted: 25 November 1997

Abstract Improved-processing tomato lines were produced by the molecular breeding strategy of advanced backcross QTL (AB-QTL) analysis. These near-isogenic lines (NILs) contained unique introgressions of wild alleles originating from two donor wild species, *Lycopersicon hirsutum* (LA1777) and *L. pimpinellifolium* (LA1589). Wild alleles targeted for trait improvement were selected on the basis of previously published replicated QTL data obtained from advanced backcross populations for a battery of important agronomic traits. Twenty three NILs were developed for 15 genomic regions which were predicted to contain 25 quantitative trait factors for the improvement of seven agronomic traits: total yield, red yield, soluble solids, brix × red yield, viscosity, fruit color, and fruit firmness. An evaluation of the agronomic performance of the NILs in five locations worldwide revealed that 22 out of the 25 (88%) quantitative factors showed the phenotypic improvement predicted by QTL analysis of the BC_3 populations, as NILs in at least one location. Per-location gains over the elite control ranged from 9% to 59% for brix × red yield; 14% to 33% for fruit color; 17% to 34% for fruit firmness; 6% to 22% for soluble-solids content; 7% to 22% for viscosity; 15% to 48% for red yield, and 20% to 28% for total yield. The inheritance of QTLs, the implementation of the AB-QTL methodology for characterizing unadapted germplasm and the applicability of this method to other crops are discussed.

Key words Molecular breeding · Germplasm · Quantitative traits

Communicated by G. Wenzel

D. Bernacchi · S. Tanksley (✉)
Department of Plant Breeding and Biometry, 252 Emerson Hall, Cornell University, Ithaca, NY 14853-1902, USA
Fax: +1 607 255 6683
E: mail: SDT4@Cornell.edu

T. Beck-Bunn · J. Uhlig
Seminis Vegetable Seeds, 37437 State Highway 16, Woodland, CA 95695, USA

D. Emmatty
Heinz U.S.A., Agricultural Research Department, P.O. Box 57, Stockton, CA 95201-3057, USA

Y. Eshed · D. Zamir
Department of Field Crops, The Faculty of Agriculture, Hebrew University of Jerusalem, Kennedy Lee Building, Room 222, Box 12, Rehovot 76-100, Israel

S. Inai · H. Sayama
Nippon del Monte Corp., Research and Development, 3748 Shimizu-cho, Numata, Gumma 378, Japan

J. Lopez
Nestlé Research and Development Center, Apartado 435, E-06080, Badajoz, Spain

V. Petiard
Centre Recherche Nestlé, 101 Avenue Gustave Eiffel, 37390 Notre Dame D'OE, Tours, France

Introduction

During the last decade molecular marker-based linkage maps have been developed for many crops (reviewed in Paterson 1996). The development of these linkage maps was rapidly followed by studies aimed at identifying discrete factors controlling quantitative traits (QTLs) of agronomic importance. QTL-mapping studies have now been reported in most crops for traits related to yield, quality, disease and insect resistance, and environmental adaptation. The anticipated beneficial impact of QTL mapping was based on the assumption that QTL-alleles detected in segregating populations could be treated as units of Mendelian inheritance and that the incorporation of these alleles into elite breeding lines would result in an enhanced performance. However, in most instances, these assumptions have not yet been experimentally confirmed.

Despite the large amount of QTL-mapping work reported, the impact of QTL mapping on the development of improved varieties has been minimal. This may be due to the fact that QTL-mapping efforts have largely been conducted independently from breeding programs, and in many instances have not involved elite-breeding germplasm. Therefore, the validity of the underlying assumption that markers can aid in the improvement of quantitative traits remains to be demonstrated.

Recently, Tanksley and Nelson (1996) proposed a new molecular breeding strategy based on QTL mapping that integrates the processes of QTL analysis and variety development. This new procedure is referred to as advanced backcross QTL analysis (AB-QTL). The basic aspects of the AB-QTL strategy are described in a companion paper (Bernacchi et al. 1998). This method uses molecular markers to identify beneficial alleles from unadapted germplasm with the potential to improve the agronomic performance of elite cultivated lines. These QTL-alleles are simultaneously transferred into near-isogenic lines (NILs) for replicated testing. Thus, a cycle of AB-QTL breeding (i.e., QTL discovery and NIL development and testing) represents a direct test of the underlying assumption of QTL breeding: namely, that beneficial QTL-alleles identified in segregating populations (BC_2, BC_3 in AB-QTL) will continue to exert their positive effect when transferred to elite lines.

So far, three AB-QTL analysis projects have produced QTL-mapping results for a battery of agronomic traits of interest in processing tomato. The first study used *L. pimpinellifolium* (PM) as donor parent (Tanksley et al. 1996), the second *L. hirsutum* (H) (Bernacchi et al. 1998), and the third *L. peruvianum* (Fulton et al. 1997). Herein, we report on the production and replicated evaluation of a set of NILs developed from two of the studies (*L. pimpinellifolium* and *L. hirsutum*) and we compare the performance of these NILs with that predicted by the previously published QTL analyses. The goal of this research was to test the viability of the QTL-breeding hypothesis in the context of a commercial breeding program. The results from this study are discussed with respect to the feasibility of identifying and transferring beneficial QTL-alleles from unadapted germplasm using a molecular-breeding approach.

Materials and methods

Near-isogenic lines

NILs were produced for two advanced backcross populations, one derived from the cross *L. esculentum* cv E6203 × *L. hirsutum* LA1777 (Bernacchi et al. 1998) and the other from the cross E6203 × *L. pimpinellifolium* accession LA1589 (Tanksley et al. 1996). Hereafter, E6203 is referred to as E, LA1777 as H, and LA1589 as PM. All near-isogenic lines (NILs) contained a single segment of introgressed wild DNA in an otherwise cultivated-tomato E genetic background, as determined by molecular marker analysis.

Specific segments of the wild genome were selected for NIL production on the basis of QTL results (Tanksley et al. 1996; Bernacchi et al. 1998). Targeted regions were selected to contain QTLs for which the wild allele was expected to have a favorable effect on one or more traits. Moreover, priority was given to QTLs that showed favorable effects in several locations or that were detected at higher levels of significance (Tanksley et al. 1996; Bernacchi et al. 1998). As a result, 12 chromosomal regions were selected for H introgressions and five regions for PM introgressions (Table 1, Fig. 1).

The production of NILs for the different targeted donor regions required the identification of single plants that were homozygous for the wild chromosome segment (H/H or PM/PM) of interest and otherwise homozygous E/E at all other unlinked genomic regions. For this purpose bulked BC_3S_1 seed was harvested from BC_3 field plots (Tanksley et al. 1996; Bernacchi et al. 1998). Since the genotypic constitution of all BC_2 progenitor plants was available, from the RFLP genotyping, for QTL analysis (Tanksley et al. 1996; Bernacchi et al. 1998), the chromosomal regions segregating in any given BC_3S_1 family were known. For each targeted region, the initial step was to identify all pedigrees that carried the desired segments of chromosomes from the wild donor at the BC_2 stage. All positive pedigrees were sorted according to the proportion of donor genome they contained, in addition to the targeted segment, using the software package HyperGene (Young and Tanksley 1989). To maximize the probability of identifying NIL plants, BC_3S_1 families were screened that carried the least unwanted wild introgressions, thus requiring analysis with the least number of RFLP markers to select against undesirable, independently segregating fragments. An appropriate number of seedlings was screened to have a 90% chance of recovering at least one plant with the desired genotype. The compound probability of the occurrence of any specific genotype was derived from the probability of a target region being homozygous wild (H/H or PM/PM) (1/8) and the probability of an undesired segregating region being homozygous E/E (5/8 each). Confirmed NIL plants were transplanted to the greenhouse, allowed to self and used to pollinate E plants (to produce hybrid NIL seed, E/H or E/PM, for field testing and confirmation of the putative QTL NIL effect). In cases where the targeted segments were large, several independent NILs carrying overlapping introgressed segments were constructed. This strategy should increase the chance that at least one NIL would contain the targeted QTL (see Fig. 1). It would also allow for variation in the linkage drag around the QTL.

Field evaluation of near-isogenic lines

NIL plots were evaluated for agronomic performance during the summer of 1995 in five locations worldwide: Akko, Israel (IS); Badajoz, Spain (SP); Woodland, California (CAP); Stockton, California (CAH), and Numata, Japan (JP). The cultural conditions of each experimental location are described in Bernacchi et al. (1998) for IS, SP and CAP. CAH was grown in clay loam soil, under drip irrigation and at 25 cm × 150 cm spacing. JP was rainfed with plants spaced at 45 cm. In each location, plots of 30–40 plants were randomized in three blocks (two in JP), with three plots of the recurrent *L. esculentum* E6203 control and one plot per NIL in each block. The performance of the NILs and the control was characterized by measuring 12 traits of agronomic importance. Cover (COV), average fruit weight (AFW), pH (pH), fruit firmness and maturity (FIR, MAT) were evaluated as described in Tanksley et al. (1996). Total yield (YDT), red yield (YDR), soluble solids (SSC), red yield × brix (BYR), percent stem retention (STR), and viscosity (BOS) were measured as described in Bernacchi et al. (1998). Internal fruit color (FC) was rated visually in all locations using a scale from 1 to 5 (increasing redness) by viewing from 10 to 30 randomly selected fruit. In addition, fruit color was evaluated analytically in three

locations: SP and CAH (A/B index), and CAP (Agtron LA/B). The procedures used are as described in Tanksley et al. (1996). Stem retention was measured only in IS and CAP. Viscosity was only measured in SP, CAH and JP. In this last location only one measurement was taken for BOS per NIL, precluding any statistical analysis.

Data analysis

For the NIL evaluation, means and standard deviations were calculated using the software package QGENE (Nelson 1997). Linear contrasts were used to compare the mean of each NIL versus the mean of the control E6203 for each trait within a location. The percent phenotypic difference between the mean of a NIL and the mean of the control E6203 was estimated as $\Delta\% = 100(\text{NIL-E6203})/\text{E6203}$. Phenotypic differences are reported when showing a significance level of $P \leqslant 0.1$. The transfer of QTL-alleles and their associated phenotypic effects is discussed for significance levels of both $P \leqslant 0.1$ and $P \leqslant 0.05$. In cases where targeted factors were represented by more than one NIL, the averaged estimates of $\Delta\%$ across locations or across traits include only the NIL that best manifested the expected phenotypes (Table 2). Across-location and across-trait average QTL effects predicted from BC_2 mapping data were derived from results previously reported (Tanksley et al. 1996; Bernacchi et al. 1998).

Results and discussion

Production of NILs

In total, 23 NILs containing H or PM introgressions were produced, representing quantitative factors for seven traits, positioned at 15 genomic regions and originating from the two different wild donor species. The targeted genomic regions, their putative phenotypic effects and the NILs produced are described in Table 1 and represented in Fig. 1. Some of the targeted introgressions were expected to affect more than one trait. Also, some of the targeted quantitative factors were represented by more than one NIL. Jointly, these

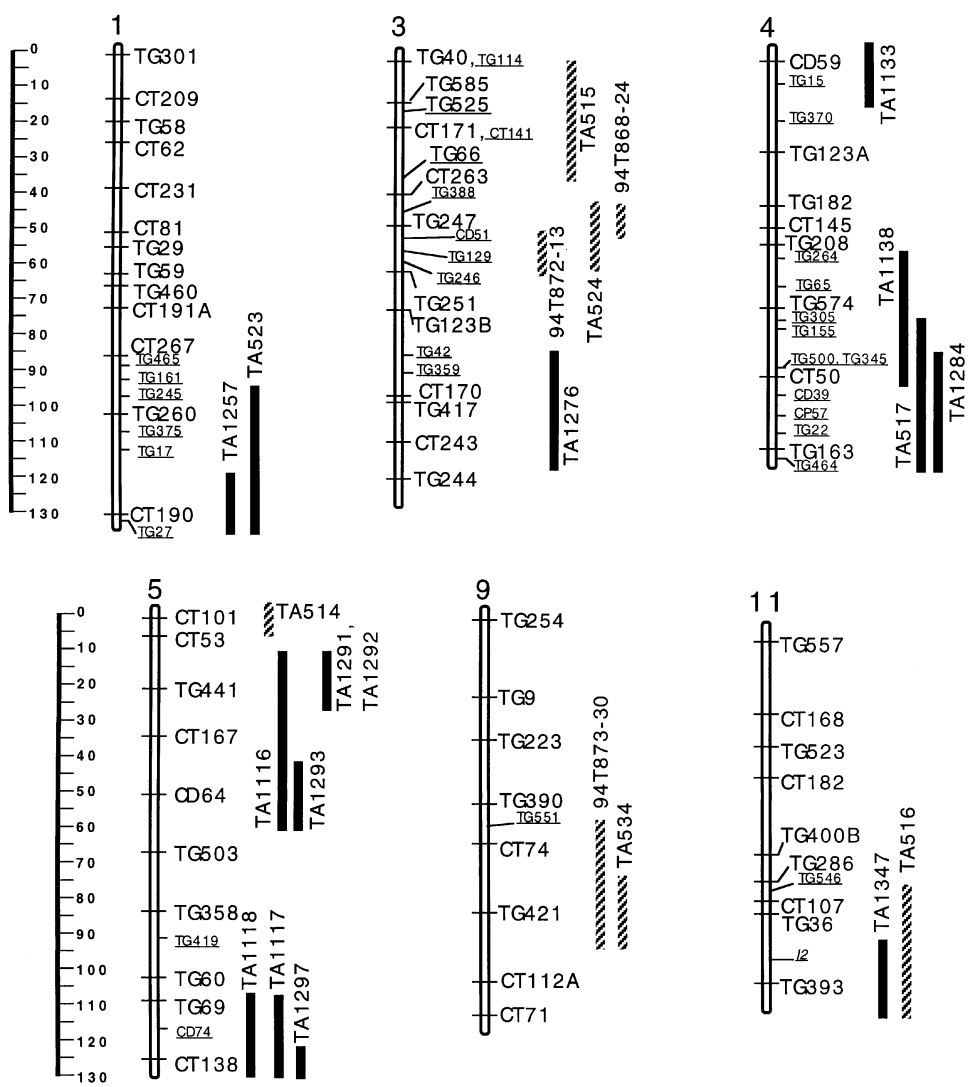

Fig. 1 *L. hirsutum* (H) and *L. pimpinellifolium* (PM) near-isogenic lines (NILs). RFLP linkage map for chromosomes 1, 3, 4, 5, 9 and 11 developed from a BC_1 interspecific *L. esculentum* (E) × *L. hirsutum* H population (Bernacchi et al. 1998). Map distances are in centiMorgans (cM). The introgressed segment of wild DNA for each NIL is represented by the *bars* alongside the chromosomes. H NILs are represented by *black bars*, PM NILs by *striped bars*. The accession number for each NIL is indicated by each bar representing an introgression. Additional *underlined markers* at regions with PM introgressions are from the linkage map derived for a cross of E × PM (Grandillo and Tanksley 1996). *Underlined markers* at regions with H introgressions were used to further characterize the introgressions

NILs were expected to represent 11 QTLs as well as 14 chromosomal regions where quantitative effects had been detected at sub-threshold significance (Tanksley et al. 1996; Bernacchi et al. 1998).

To obtain NILs carrying targeted H introgressions, 2000 BC_3S_1 seedlings were screened resulting in the isolation of 15 homozygous NILs covering ten targeted chromosomal regions. Eight NILs represent the five targeted regions for PM introgressions (Table 1, Fig. 1). The number of seedlings screened to obtain PM NILs was 750. The plants confirming H/H or PM/PM for the target introgressions, and lacking H or PM alleles elsewhere in the genome, were transplanted to the greenhouse to self-pollinate and to be used as the staminate parent in crosses with E to produce NIL-H/E and NIL-PM/E hybrid seed for field testing.

NIL performance

Based on whole-genome marker analysis, all NILs differed from the E6203 control only by the introgressed donor DNA fragment (H or PM); therefore the observable phenotypic differences should be due to the introgressed wild DNA segment. A comparison of the trait distributions of the unselected BC_3, the NIL population and the recurrent control E6203, provides a first indication of the phenotypic gains effected by the targeted wild alleles (Fig. 2). On average, the performance of the NILs was shifted in the expected direction. These gains are more evident when the average performance of NILs selected for specific QTLs are considered (Fig. 2).

The level of significance (P) and the percent phenotypic difference ($\Delta\%$) are reported for each NIL-control contrast within location in Table 3, together with the allelic effect predicted by the BC_3 QTL data. Given the small number of replicates for each NIL in each location (3/location), and the inherently large effect of the environment on yield and other quality measurements, NIL performance was evaluated at two significance levels of $P \leqslant 0.1$ and $P \leqslant 0.05$, despite the higher risk of experiment-wise error.

Proportion of NILs showing improved performance

Overall, contrasts between the performance of each NIL and that of the recurrent parent control, per trait per location, revealed that for 22 out of the 25 (88%) quantitative factors evaluated induced the expected phenotypic improvement in the NILs in at least one location with a significance of $P \leqslant 0.1$ (Table 2). At $P \leqslant 0.1$, and depending on the location, NILs outperformed the elite control for the targeted traits: by 9–59% for brix $\times$ red yield; 14–33% for fruit color; 17–34% for fruit firmness; 6–22% for soluble solids content; 7–22% for viscosity; 15–48% for red yield and 20–28% for total yield (Table 3). For SSC, 4 of the 6 NILs developed showed the predicted effect in at least one location ($P \leqslant 0.1$). For YDR, 2 out of 3 NILs behaved as predicted ($P \leqslant 0.1$). All of the NILs expected to show an improved BOS, YDT, BYR, FC and FIR behaved as predicted, in both direction and magnitude, in at least one location ($P \leqslant 0.1$). If one considers contrasts significant at $P \leqslant 0.05$, 17 out of the 25 (68%) QTL/factors transferred produced the expected phenotype as NILs (Table 2). At a significance of $P \leqslant 0.05$ the same pattern is observed, with only one QTL/factor each for BOS, YDR, BYR, FC and YDT failing to show a significant effect in the NILs.

These results indicate that the majority of the QTLs detected in the BC_3 were not spurious and can be manipulated via marker-assisted selection. Moreover, the QTL-alleles detected from the wild species are likely to be largely additive (versus epistatic) since it is unlikely that epistatic effects could persist after the rest of the wild genome has been replaced by recurrent-parent alleles. That QTLs detected by AB-QTL analysis should be largely non-epistatic was originally predicted from simulations and is most likely due to the donor allele frequencies being very low in advanced backcross populations (Tanksley and Nelson 1996).

NILs corresponding to three QTLs/factors did not perform as predicted in any location. These include a sub-threshold factor for YDR associated with H alleles at TG574 (YDR-TG574), and represented by NIL TA1138; the QTL *ssc5.1* which was associated with H alleles linked to TG69 and represented by NILs TA1118, TA1117 and TA1297; and the QTLs *ssc3.1* and *ssc3.2* which were associated with PM alleles in the region TG129-TG388 and represented by the NILs TA524, 94T872-13 and 94T868-24. In these cases, none of the NILs outperformed the control at $P \leqslant 0.1$ for the corresponding trait (Table 3). However, some of these same NILs showed the predicted effects for other QTLs/factors associated with the same introgression. For example, NIL TA1138 showed the predicted effects for BYR and FIR, and NILs TA1118 and TA1117 showed the predicted improvement in FC.

Some of the introgressions in the NILs appeared to reduce the performance of the NILs relative to the control E with respect to the targeted trait. For example, all of the NILs developed for *ssc3.1*/*ssc3.2*, and two of the three NILs developed for *ssc5.1*, had a significantly reduced soluble-solids content relative to the E control in at least one location. Most likely, the inferior performance of NILs is the result of the absence of the putative wild QTL-allele in the NIL due to recombination, or alternatively, that the QTL detection in the BC_3 population was spurious (Type-1 error) (Bernacchi et al. 1998). Another possibility is that the wild QTL-alleles were subjected to novel negative epistatic interactions in the NIL background.

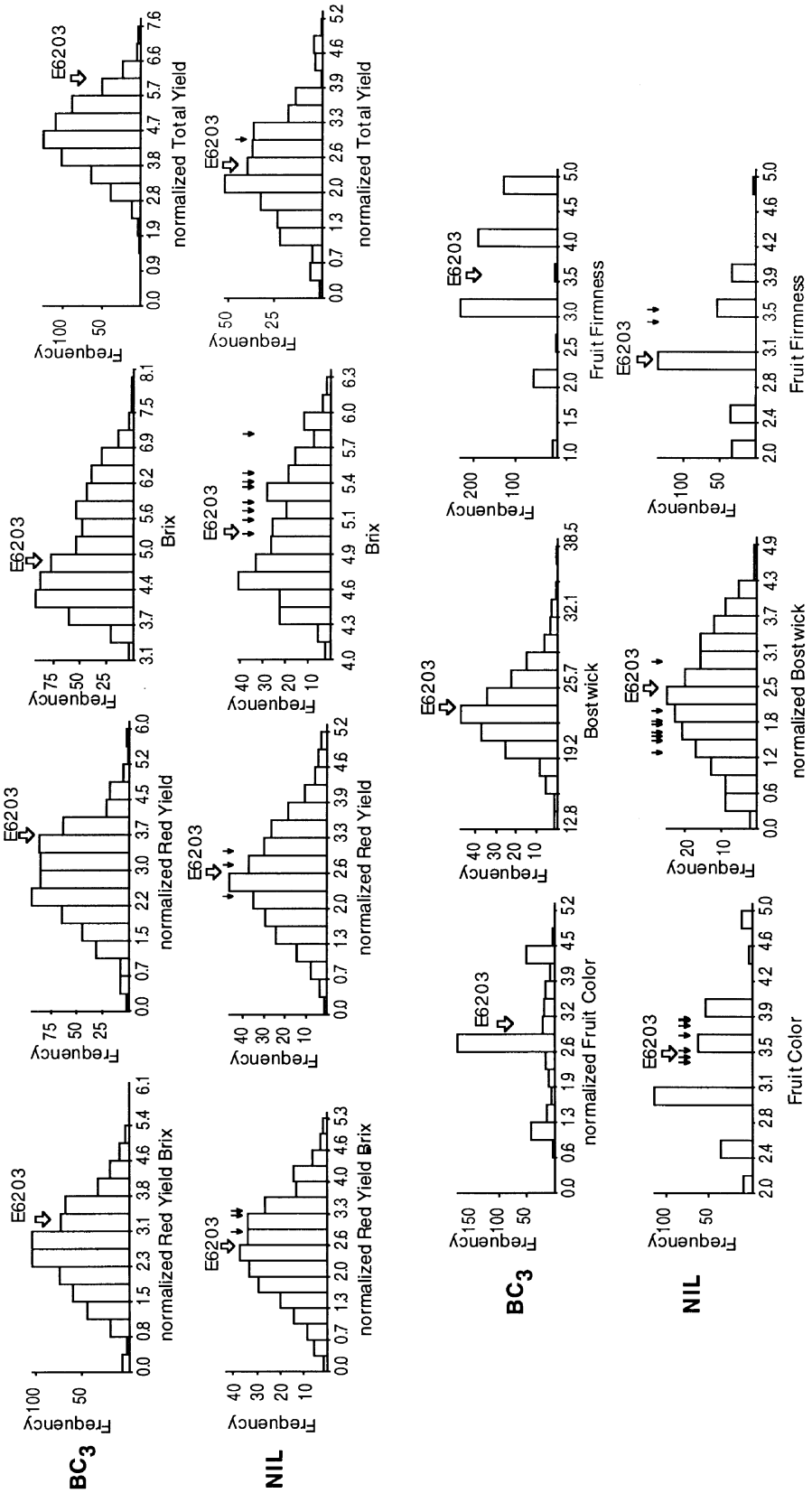

Fig. 2 Frequency distribution for quantitative traits measured in a *L. hirsutum* BC_3 population and *L. hirsutum* NIL plots. Histograms include observation from all locations. Data were normalized where required. The mean level of the recurrent control *L. esculentum* E6203, is indicated by a *large arrow*. In the histograms for NIL plots, *small black arrows* indicate the mean level of the NILs that were selected specifically for each trait (NIL symbols not included)

Ability of BC_3 QTL data to predict NIL performance

The issue of the predictive value of QTL-mapping data is a critical one for AB-QTL analysis and molecular breeding in general. The predictive potential of QTL data can be estimated by determining the relationship between allelic effects predicted from QTL analysis in the BC_3 and the effects of the same alleles on the observed performance of the NILs. This comparison can be made in terms of the magnitude of effects (see next section), the significance level of the effects, and the stability of the effects across locations.

The significance at which QTLs/factors were detected in the BC_3 families (Tanksley et al. 1996; Bernacchi et al. 1997) was a modest predictor of the significance with which the effects were observed in the NILs (Table 4). Most QTLs/factors detected at lower significance in the BC_3 were also detected in the NILs at lower significance (Table 4). Likewise, QTLs/factors detected at high significance levels in the BC_3 showed their effect at higher significance levels in the NIL study. Yet there were several instances in which this relationship did not hold. For example, of the 14 QTLs/factors that were originally detected at subthreshold significance, two were associated with highly significant effects in the NILs (BOS-TG441 and SSC-TG163) (Tables 3, 4). In contrast, *fc4.1* and *vis9.1* are examples of QTLs detected with high significance in the BC_3 study but which resulted in NILs only marginally different from the control (Tables 3, 4).

Nearly half of the QTLs/factors detected with significance in only one location in the BC_3 (Tanksley et al. 1996; Bernacchi et al. 1998) produced the predicted phenotypic improvement in two or more locations as NILs (SSC-TG163, FIR-TG574, FC-TG393 among others) (Table 5). Conversely, several of the QTLs/factors detected with significance in more than one location in the BC_3, produced a detectable improvement over the control at only one location as NILs (*fc5.1*, *ssc3.2*, FC-TG260/CT138) (Table 5). Overall, these results indicate that the degree of conservation of QTLs across locations, as detected in BC_3, was only a moderate predictor of the conservation of NIL performance across locations.

Similarly, there appears to be only a modest relationship between the significance of QTLs/factors detection in the BC_3 and the number of locations at which the corresponding NILs showed significant improvements over the control (Fig. 5). In the BC_3 study ten QTLs/factors were detected at $0.1 \geqslant P > 0.01$, eight at $0.01 \geqslant P > 0.001$, and seven at $0.001 \geqslant P > 0.0001$. Among those detected at the highest level of significance, none showed the predicted effect as NILs in more than two locations. In fact, the only factor to show the predicted effect in all five experimental locations was SSC-TG163 (from H) which was originally detected in the BC_3 in only one of three locations and at subthreshold significance (Bernacchi et al. 1998).

Magnitude of NIL improvement compared with BC_3 predictions

Another way to measure the success (or failure) in transferring and expressing the effect of a selected QTL-allele is given by the comparison of magnitude of the allelic effect predicted by BC_3 QTL-mapping results with the performance of the corresponding NIL (Table 3, Fig. 3). For example, the total yield factor *ydt4.1* was associated with 14% increases in the QTL-mapping study and once transferred to a NIL, induced a 12% yield increase over the E control. Likewise a 20% increase in red yield was predicted for the QTL/factor YDR-CD59. The corresponding NIL was 18% higher yield than the E control. For QTL/factor BOS-TG441, a predicted 12% improvement in viscosity was met by an observed improvement of 22% in the NIL; for QTL/factor SSC-TG163, a predicted improvement of 10% in soluble solids was met by a 12% improvement in the corresponding NIL; for QTL/factor FIR-TG574, a predicted improvement of 17% in firmness corresponded to an observed increase of 25% in the NIL; and for fruit color-QTL/factor *fc5.1* a predicted improvement of 29% was met by an observed improvement of 17% in the corresponding NIL (Table 3, Fig. 3).

In most instances the observed gain in the NILs was less than that predicted from the QTL analysis. Over all traits, 73% of the expected gain was realized in the NILs. Overall, the average genetic gains per trait across locations and across independent QTLs/factors ($P \leqslant 0.1$) for each trait were: 12% for YDT (predicted QTL effect 14%); 12% for YDR (predicted QTL effect 32%); 11% for VIS (predicted QTL effect 13%); 6% for SSC (predicted QTL effect 8%); 24% for FIR (predicted QTL effect 28%); 12% for FC (predicted QTL effect 23%) and 22% for BYR (predicted QTL effect 24%) (Fig. 4). When only contrasts significant at $P \leqslant 0.05$ are considered, these gains per traits become: 33% for YDR, 16% for BOS, 8% for SSC, 28% for FIR, 17% for FC and 42% for BYR (Fig. 4). At this significance level, no NIL constructed for YDT behaved as expected.

Comparison of independent NILs for the same targeted regions

In most cases, there were observable phenotypic differences among NILs selected for the same region of the genome. For example, of two NILs expected to display increased SSC due to an H introgression on chromosome 1 (TA523, TA1257), only the one carrying the longer segment of foreign DNA (TA523) showed the predicted effect (Table 3). Thus, it can be assumed that TA1257 lacks the QTL-allele which must be located in the non-overlapping region between the two lines in the TG17-TG245 region.

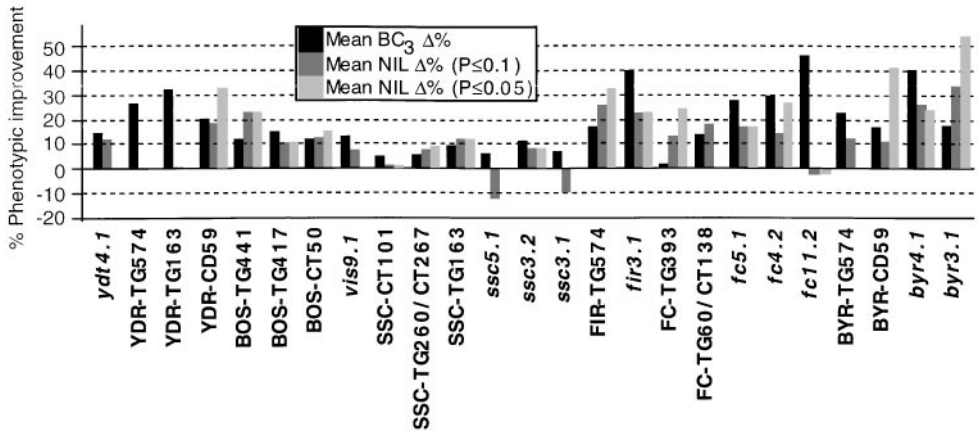

Fig. 3 Mean percent phenotypic improvement (Δ%) of the H allele over the E allele for each trait predicted by BC_3 QTL data (*black bars*) (Tanksley et al. 1996; Bernacchi et al. 1998) compared with the improvement observed in the corresponding NIL (*lighter bars*). Mean NIL Δ% includes observations only from locations in which NILs differed significantly from the control E6203 ($P \leqslant 0.1$). For QTLs/factors that were represented by more than one NIL, only data from the NIL with the best performance is depicted. QTLs/factors are designated by QTL name (in *italics*) or, if the original factor was detected in BC_3 with a sub-threshold significance, the trait symbol and linked the RFLP marker are given. *SSC* = soluble solids; *BOS* = viscosity Bostwick (note that for BOS, a reduction in Bostwick value indicates increased viscosity); *YDT* = total yield; *YDR* = red yield; *BYR* = brix red yield; *FC* = fruit color and *FIR* = fruit firmness

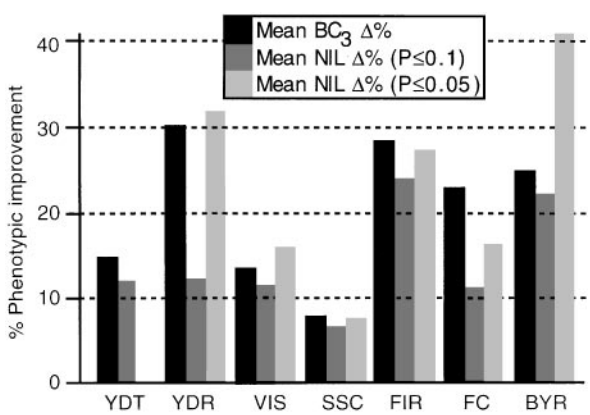

Fig. 4 Average percentage phenotypic improvement (Δ%) per trait predicted by BC_3 QTL data (Tanksley et al. 1996; Bernacchi et al. 1998) and the observed improvement associated with NILs. *Dark bars* are BC_3 Δ% and *light bars* are NIL Δ%. Means include only differences significant at $P \leqslant 0.1$ for the NIL showing the best performance for the targeted trait. *YDT* = total yield; *YDR* = red yield; *SSC* = soluble solids; *BYR* = brix × red yield; *FC* = fruit color; *FIR* = firmness and *VIS* = viscosity

A similar situation was observed with NILs TA517 and TA1248 which carry H alleles for increased fruit color (*fc*4.2) on chromosome 4 (Table 3). In this case, the NIL with the longer H segment, TA517, displayed the predicted effect in two locations (CAP and JP) whereas the NIL with the shorter introgression, TA1248, did not differ significantly from the control in any location. Again, these lead us to believe that the QTL lies in the non-overlapping region between the introgressions of the two NILs (TG155–TG305). On the other hand, these same two NILs (TA517 and TA1248) were also expected to display improved soluble solids, and both the longer introgression (TA517) and the short introgression (TA1248) showed the predicted effect in five and three locations, respectively. These results indicate that the gene(s) causing increased solids may be present in both NILs in the TG500– TG464 region. This interpretation implies that there are different genes for fruit color and soluble solids on the introgressed segment. Also in this region of chromosome 4, NILs TA517 and TA1138 contained overlapping introgressions of H in the region between markers TG155 and CT50, and both showed a similar improvement of BYR over the elite control E indicating that the BYR QTL may be more precisely mapped to this same interval.

The inverse situation occurred with NIL TA534 and NIL 94T873-30 which were developed for PM alleles associated with the viscosity QTL *vis9.1*. In this case the NIL carrying the shorter introgression showed the predicted effect whereas the NIL with the longer introgression did not (Table 3). The most likely explanation in this case is that the longer introgression does not contain the QTL-alleles due to recombination and that the QTL would in fact lie beyond the distal common marker (TG421). Alternatively it is possible that the QTL-alleles are also present in the longer introgression but their effects are masked by deleterious linked linkage drag absent in the shorter introgression.

Another case in which the effect of a PM QTL was confirmed, and the factor more precisely mapped, was for BYR QTL *byr3.1*. In this case the NIL with the longer introgression (TA524) and one overlapping in one extreme (94T872-13) showed the predicted effect while the NIL with overlapping introgressions towards the other distal extreme (94T868-24) did not. These results indicate that *byr3.1* most likely lies in the CD51-TG129 region.

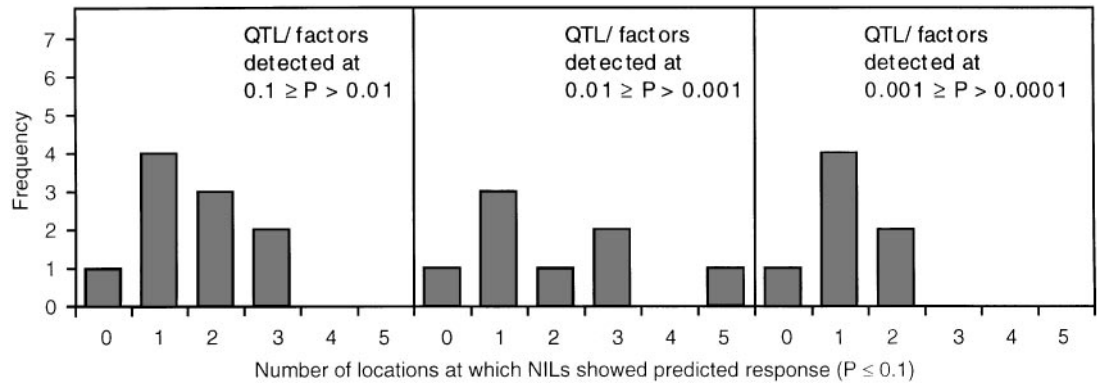

Fig. 5 Distribution of QTLs/factors transferred according to the number of locations in which NILs gave the predicted effect ($P \leqslant 0.1$) and grouped by the level of significance detection in the BC_3 population (Tanksley et al. 1996; Bernacchi et al. 1998)

The differential behavior of NILs TA1116, TA1291, TA1292 and TA1293, developed to contain H alleles for improved fruit color and viscosity on chromosome 5, also allowed us to narrow down the position of the QTL-alleles. Since TA1116, with the longer introgression, and TA1293 overlapping with it on the distal side (opposite TA1291 and TA1292) showed the predicted FC and VIS improvements, while TA1291 and TA1292 completely failed to display improvements for either trait, it can be assumed that the H QTL-alleles lie in the overlapping region between TA1116 and TA1293 (CD64). Also on chromosome 5, NILs TA1117, TA1118 and TA1297 all failed to show the expected improvements in SSC indicating, as mentioned, possible absence of the QTL-alleles or spurious QTL detection. Yet the two lines with the long introgression (NILs TA1117 and TA1118) showed the expected improvements in FC while the one with the short introgression (TA1297) did not. These facts narrow the position of the FC factor to the region TG69–CD74.

Interestingly, the independent H and PM QTL studies had both identified the bottom of chromosome 11 as associated with improved color and one NIL from each donor was available for testing (TA1347 from H and TA516 from PM). TA516 which carried a longer introgression of PM alleles showed the predicted improvement in FC in most locations. However, TA1347, carrying a shorter introgression from H, showed an improved FC in only one location and a reduced FC in another, precluding the possibility of fine mapping the effect or confirming the orthology of the H and PM fruit-color loci.

Variation in NIL performance across locations

The performance of the different NILs, relative to the recurrent parent control, was variable across environments. Forty percent of the NILs showed the predicted phenotypic improvement over the control, with statistical significance ($P \leqslant 0.1$) , in the same location in which a particular QTL/factor was originally detected in the BC_3 population (Table 3). Of the QTLs/factors that produced the predicted effect in at least one location (23), the targeted QTL-allele SSC-TG163, present in TA517, was the only one that produced the expected improvement in all locations tested (5), while 9 out of 23 target wild alleles produced the predicted improvement in only a single location. A factor that may have conditioned these results is the relatively limited number of replications used for the NIL plots at each location (three plots per NIL versus nine plots for the control). Eight QTL-factors (*ydt4.1*, YDR-CD59, BYR-CD59, YDR-TG163, *fc4.2*, *fc11.2*, SSC-CT101 and FC-TG393) out of 23 that showed the predicted effect as NILs also appeared to affect the trait adversely in one of the locations (Table 3). For example, YDR-CD59 which significantly increased YDR in CAP, CAH and JP also significantly reduced YDR in IS (Table 3). Since locations differ greatly in environmental and cultural conditions, it is likely that some QTL-alleles interact with the environment resulting in the variable performance of the NILs across locations.

Expression of phenotypes in NILs not predicted from BC_3 data

Without exception, all of the evaluated NILs showed significant deviations from the E control for traits other than those putatively associated with the introgressed segments on the basis of the BC_3 QTL results (Table 6) (Bernacchi et al. 1998). The unpredicted effects from the introgressed fragments were either horiculturally desirable or deleterious with a similar frequency. For example, NIL TA523, in addition to showing the predicted increases in soluble solids, also showed improved performance for FC (IS), BYR (CAH and JP) and FIR (CAH), though it showed significantly reduced AFW in three locations (CAH, CAP and SP). The chromosome-5 introgression tested in NIL TA1116 (TG441-CT167-CD64), which showed improved FC and VIS as predicted, also showed improved fruit

firmness in three locations (CAH, IS and JP). Several NILs (TA517 in four locations; TA1133, TA523, TA1276, TA515; TA516 and TA1248 in three locations; TA1138, TA1116 and TA524 in two locations) showed reduced average fruit weight despite the fact that no QTLs for fruit weight were detected for those regions in the BC_3 (Bernacchi et al. 1998). In most cases, the direction in which a trait was affected by the introgression was constant across locations (Table 6). For example, the chromosome-4 introgression present in NIL TA1138 (TG574) was associated with improved BYR and FIR, as predicted, and in addition showed improved viscosity in three locations (CAH, CAP and JP). When BC_3 QTL results were reviewed in the light of the NIL results, it was found that for marker CT50, adjacent to TG574, the H allele was also marginally associated with improved viscosity ($P \leqslant 0.1$).

One likely explanation for the expression of phenotypes in the NILs not predicted by the BC_3 data is the overall higher genetic variance in the BC_3 which reduces the power of detecting QTL effects compared with NILs. Epistatic interactions may also play a role. Multiple, segregating H alleles in the BC_3 population may have masked the effects of particular alleles which may become evident in the NILs. Eshed and Zamir (1996) demonstrated that epistatic interactions between unadapted alleles in an elite genetic background result in a less-than-additive expression of the respective alleles additive effects. By extension, the authors propose that this diminishing additivity of QTL effects is greater when a larger number of interacting unadapted alleles segregate in a given genotype. Since even advanced backcross QTL-mapping populations segregate for several independent chromosomal regions, certain QTLs with favorable effects may be masked by more predominant unfavorable wild alleles. Alternatively, this unpredicted response may originate from the onset of novel epistatic interactions between the introgressed foreign alleles and the recurrent-parent genetic background. This would be the case if genes that are not rate limiting for a particular complex metabolic pathway in the genome of *L. hirsutum*, and therefore are not detected by QTL mapping, become rate limiting once the rest of the genome is replaced by *L. esculentum* alleles, as occurs in the NILs.

Potential for pyramiding novel QTLs

The different NILs produced by AB-QTL breeding can be easily intercrossed to pyramid different QTL alleles for the same or different traits to effect an even greater potential improvement. In effect, a library of wild-species introgressions having positive effects on different traits can be produced and distributed for use by breeding programs and scientists. Since these QTLs are linked to molecular markers, MAS can be readily applied. Further, NILs are ideal experimental tools for studies trying to determine the physiological basis of the differences observed between NILs. Also the NILs provide the starting point for the fine mapping, and ultimately the cloning, of the genes responsible for the QTLs controlling traits of economic importance (Alpert and Tanksley 1996). The AB-QTL strategy can be easily implemented in other crops for which DNA markers and cross-compatible unadapted germplasms are available. Currently, AB-QTL studies are also under way in rice (Xiao et al. 1997) and maize (McCouch, personal communication).

Advanced backcross breeding compared with traditional breeding

The comparison of the AB methodology with traditional breeding approaches can be discussed in terms of the potential genetic gains per unit time. The historical genetic gains for yield in processing tomatoes is estimated to be 1% per year (Grandillo et al. submitted). In comparison, after the 4 years required for the implementation of the first cycle of AB-QTL breeding using *L. hirsutum* as the donor parent, the gains obtained for brix $\times$ red yield ranged from 14% to 33% resulting in yearly gains of between 3.5% to 8%. The likelihood that similar genetic gains could be obtained using the same exotic sources and conducting the breeding exclusively on the basis of phenotypic selection is low. Even though superior genotypes may occur for some traits in segregating populations, numerous backcrosses would be required to recover the majority of the recurrent parent genome without the assistance of markers (Young and Tanksley 1989). Also, the relatively low heritability of quantitative traits makes it difficult to select for useful alleles from exotic germplasm based only on phenotypic selection. It is likely that many low-frequency, beneficial exotic alleles would go undetected in early generations due to the effect of major adverse exotic alleles (Tanksley and Nelson 1996). However, as shown here, phenotypic selection can be used effectively to remove exotic alleles with strong undesirable effects while beneficial QTL alleles can be detected and manipulated with markers.

Long-term effect of enriching the germplasm base of crops

In addition to resulting in the identification of QTL-alleles capable of improving the performance of elite cultivars, the advanced backcross strategy results in population improvement and germplasm enhancement. This approach addresses one of the limitations of traditional breeding strategies which is that it normally only utilizes genetically similar elite stocks and does

not favor the exploitation of exotic germplasm due to the unfavorable effects often associated with it.

The novel introgressions identified in this report, once incorporated into elite lines, will result in an enrichment of the genetic base of cultivated tomatoes. These novel alleles can most likely be combined with the genetic variation already present in elite tomato germplasm resulting in even greater genetic gains. Further, this ability of AB-QTL analysis to selectively enrich the base of crops should provide an effective way to improve the environmental adaptation of basic breeding populations, by selecting unadapted donors originating in the targeted environments and identifying the QTL-alleles that, without pleiotropic or tightly linked adverse effects, improve the performance in the targeted environment.

Conclusions

The advanced backcross method proposed by Tanksley and Nelson (1996) is predicated upon the assumption that QTLs detected in advanced backcross populations will continue to exert their positive effects when isolated in near-isogenic lines. Results from the current study suggest this is likely to be the case in most instances since 88% of the *L. pimpinellifolium* and *L. hirsutum* QTLs/factors detected in advanced backcross populations resulted in NILs with superior performance in at least one location. Until now germplasm curators and breeders have measured the potential of exotic germplasm almost exclusively on the basis of its phenotype, yet the results from the current study, and others, indicate that unadapted germplasm has genetic potential for the improvement of quantitative traits of agronomic importance that cannot be predicted from its phenotype (Ragot et al. 1995; Eshed and Zamir 1996; Tanksley et al. 1996; Xiao et al. 1997). The AB-QTL analysis method represents one way in which valuable wild alleles can be unmasked and transferred into elite cultivars to effect superior performance. This process not only results in improved elite varieties, but also in a general enrichment of cultivated germplasm.

However, the significance with which QTLs are detected in the BC_3 is only a modest predictor of the significance with which the effects are realized in the NILs. This is probably due in large part to the large variance encountered when estimating allelic effects in both the BC_3 mapping stage and the NIL evaluation stage due to limited replication. Also, independent NILs constructed for the same target region of the donor genome frequently have significantly different phenotypes and often only one was superior to the cultivated control with respect to the targeted trait. These differences may be due to linkage-drag effects resulting from the varying amounts of foreign DNA present in each NIL. Most NILs also demonstrated significant phenotypic differences from the control (both beneficial and detrimental) not predicted by BC_3 mapping data. This is most likely due to linkage drag, pleiotropic effects or epistatic effects that could not be detected in the BC_3 mapping due to the higher overall level of variance as compared to the NILs. These observations emphasize the need to construct more than one NIL for each target region, in order to maximize the chance of obtaining lines with superior agronomic potential.

This work demonstrates that wild introgressions can improve yield and quality traits over the near-isogenic inbred control, yet it remains to be seen what response these introgressions produce when evaluated in hybrid combinations and if these can surpass the leading processing hybrids.

Acknowledgements This work was supported in part by grants from the National Research Initiative Cooperative Grants Program, USDA Plant Genome Program (No. 58-1908-5-001), and by the Binational Agricultural Research and Development Fund (No. US-2427-94) to S.T. and D.Z. D.B. was supported in part by a Fulbright Scholarship. Thanks to C. R. Brown and K. D. Livingstone for reviewing this manuscript.

References

Alpert K, Tanksley SD (1996) High-resolution mapping and isolation of a yeast artificial chromosome contig containing *fw2.2*: a major fruit-weight quantitative trait locus in tomato. Proc Natl Acad Sci USA 93:15503–15507

Bernacchi D, Tanksley SD (1997) An interspecific backcross of *Lycopersicon esculentum* × *L. hirsutum*: linkage analysis and a QTL study of sexual compatibility factors and floral traits. Genetics (in press)

Bernacchi D, Beck-Bunn T, Eshed Y, Lopez J, Petiard V, Uhlig J, Zamir D, Tanksley SD (1998) Advanced backcross QTL analysis in tomato. I. identification of QTLs for traits of agronomic importance from *Lycopersicon hirsutum*. Theor Appl Genet (in press)

Eshed Y, Zamir D (1996) Less-than-additive interactions of quantitative trait loci in tomato. Genetics 143:1807–1817

Fulton TM, Beck-Bunn T, Emmatty D, Eshed Y, Lopez J, Petiard V, Uhlig J, Zamir D, Tanksley SD (1997) QTL analysis in an advanced backcross of *Lycopersicon peruvianum* to the cultivated tomato and comparisons with QTLs found in other wild species. Theor Appl Genet 95:881–894

Grandillo S, Tanksley SD (1996) QTL analysis of horticultural traits differentiating the cultivated tomato from the closely related species *Lycopersicon pimpinellifolium*. Theor Appl Genet 92:935–951

Nelson CJ (1997) QGENE: software for marker-based genomic analysis and breeding. Mol Breed 3:229–235

Paterson AH (1996) Genome mapping in plants. R.G. Landers Co., Austin, Texas

Ragot M, Sisco PH, Hoisington DA, Stuber CW (1995) Molecular-marker-mediated characterization of favorable exotic alleles at quantitative trait loci in maize. Crop Sci 35:1306–1315

Tanksley SD, Nelson JC (1996) Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor Appl Genet 92:191–203

Tanksley SD, Grandillo S, Fulton TM, Zamir D, Eshed Y, Petiard V, Lopez J, Beck-Bunn T (1996) Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. Theor Appl Genet 92 : 213–224

Xiao J, Grandillo S, Ahn SN, McCouch SR, Tanksley SD (1997) Genes from the wild boost rice yields. Nature 384 : 223–224

Young ND, Tanksley SD (1989) Graphical-based whole-genome selection using RFLP. In: Helentjaris T, Burr B (eds) Development and application of molecular markers to problems in plant genetics. Current communications in molecular biology. Cold Spring Harbor Laboratory. Cold Spring Harbor, New York, pp 122–125