DOI: 10.1534/genetics.104.040006

DNA Variation at the Invertase Locus *invGE/GF* Is Associated With Tuber Quality Traits in Populations of Potato Breeding Clones

Li Li,* Josef Strahwald,† Hans-Reinhard Hofferbert,‡ Jens Lübeck,† Eckart Tacke,§ Holger Junghans, Jörg Wunder†† and Christiane Gebhardt*,[1]**

**Max-Planck Institut für Züchtungsforschung, 50829 Köln, Germany, †SaKa-Ragis Pflanzenzucht GbR, 24340 Windeby, Germany, ‡Böhm-Nordkartoffel Agrarproduktion GbR, 21337 Lüneburg, Germany, §Bioplant GmbH, 29574 Ebstorf, Germany, **NORIKA GmbH, 18190 Gross Lüsewitz, Germany and ††Istituto Agrario di San Michele all'Adige, 38010 San Michele all'Adige, Italy*

Manuscript received December 16, 2004
Accepted for publication February 28, 2005

ABSTRACT

Starch and sugar content of potato tubers are quantitative traits, which are models for the candidate gene approach for identifying the molecular basis of quantitative trait loci (QTL) in noninbred plants. Starch and sugar content are also important for the quality of processed products such as potato chips and French fries. A high content of the reducing sugars glucose and fructose results in inferior chip quality. Tuber starch content affects nutritional quality. Functional and genetic models suggest that genes encoding invertases control, among other things, tuber sugar content. The *invGE/GF* locus on potato chromosome IX consists of duplicated invertase genes *invGE* and *invGF* and colocalizes with cold-sweetening QTL *Sug9*. DNA variation at *invGE/GF* was analyzed in 188 tetraploid potato cultivars, which have been assessed for chip quality and tuber starch content. Two closely correlated invertase alleles, *invGE-f* and *invGF-d*, were associated with better chip quality in three breeding populations. Allele *invGF-b* was associated with lower tuber starch content. The potato invertase gene *invGE* is orthologous to the tomato invertase gene *Lin5*, which is causal for the fruit-sugar-yield QTL *Brix9-2-5*, suggesting that natural variation of sugar yield in tomato fruits and sugar content of potato tubers is controlled by functional variants of orthologous invertase genes.

MOST agronomic characters of crop plants are quantitative traits, the variation depending on genetic and environmental factors. The molecular basis of the natural variation of agronomic traits is poorly understood. Recently, a single amino acid substitution in an invertase gene of tomato (*Solanum lycopersicum*) was identified to be causal for a quantitative trait locus (QTL) that affects sugar yield of tomato fruits (Fridman *et al.* 2004). In the closely related potato (*Solanum tuberosum*), the tuber content of starch and sugars are quantitative traits, which are important for the nutritional quality of processed food products such as potato chips and French fries. The high frying temperatures cause a nonenzymatic Maillard reaction between free aldehyde groups of the reducing sugars glucose and fructose and free α-amino groups of amino acids and proteins. With increasing amounts of reducing sugars in the raw tubers, chip color changes from light yellow to dark brown, dark-colored chips having inferior culinary quality. The content of glucose and fructose increases in response to cold stress (cold sweetening), when tubers are stored below 10° (Coffin *et al.* 1987; Scheffler *et al.* 1992).

[1]*Corresponding author:* Max-Planck Institute for Plant Breeding Research, Carl von Linné Weg 10, 50829 Köln, Germany.
E-mail: gebhardt@mpiz-koeln.mpg.de

At the functional level, sugar and starch metabolism is one of the best-studied plant processes. Quantitative characters related to carbohydrate metabolism such as tuber starch and sugar content can be used as model traits for exploring the candidate gene approach for identifying the molecular basis of QTL (Pflieger *et al.* 2001). QTL for tuber starch and sugar content or chip color have been mapped in potato (Douches and Freyre 1994; Schäfer-Pregl *et al.* 1998; Menendez *et al.* 2002). Cloned genes functioning in the biosynthesis, degradation, or transport of sugars and starch have been positioned on the potato molecular maps and, on the basis of colocalization with QTL, a number of candidate genes have been identified (Chen *et al.* 2001; Menendez *et al.* 2002). Among others, invertase genes were identified as positional candidates for cold-sweetening QTL.

Invertases (β-d-fructofuranosidase EC3.2.1.26) are ubiquitous enzymes that hydrolyze sucrose into glucose and fructose. Plant invertases occur in several isoforms and play important roles in source/sink relationships, defense responses to pathogens, growth, and other developmental processes (reviewed in Tymowska-Lalanne and Kreis 1998; Roitsch and González 2004). In developing potato plants, invertase functions in the flux of photosynthetic fixed carbon into tuber starch. In ma-

ture, resting tubers, invertase influences the accumulation of reducing sugars in response to cold storage. Invertase enzyme activity in tuber protein extracts is correlated with the amounts of glucose and fructose, which increases at low temperatures and decreases at normal temperatures (Pressey and Shaw 1966). Type and amount of potato invertase gene transcripts are affected by temperature (Zhou *et al.* 1994; Bournay *et al.* 1996). Genetic manipulation of invertase transcript levels in transgenic potato plants has profound effects on starch- and sugar-related phenotypes (Heineke *et al.* 1992; Sonnewald *et al.* 1995; Zrenner *et al.* 1996; Greiner *et al.* 1999). Genes and full-length cDNAs of several different potato invertases have been cloned and characterized (Hedley *et al.* 1993, 1994; Zhou *et al.* 1994; Zrenner *et al.* 1996; Maddison *et al.* 1999).

Molecular mapping in potato of cloned invertase genes identified three loci, *Pain-1* on chromosome III and Inv_{ap}-*a* and Inv_{ap}-*b* on chromosomes X and IX, respectively (Chen *et al.* 2001). The loci on chromosomes IX and X encode putative apoplastic, cell-wall-bound isoforms of acid invertase (Hedley *et al.* 1993, 1994), whereas the locus on chromosome III encodes a putative intracellular, soluble acid invertase (Zhou *et al.* 1994; Zrenner *et al.* 1996). The locus Inv_{ap}-*b* on chromosome IX colocalized with the cold-sweetening QTL *Sug9a*, which was consistently detected across different environments (Menendez *et al.* 2002). Due to high structural collinearity of the potato and tomato genomes (Bonierbale *et al.* 1988), the potato loci Inv_{ap}-*b* and Inv_{ap}-*a* are orthologous to the invertase loci *Lin5/Lin7* and *Lin6/Lin8* on tomato chromosomes 9 and 10, respectively. Each locus contains a pair of tandem duplicated invertase genes. The invertase loci on chromosomes IX and X of potato and tomato are part of a segmental chromosome duplication, which predates speciation (Fridman and Zamir 2003) and is structurally conserved between distantly related plant species, including *Arabidopsis thaliana* (Dominguez *et al.* 2003; Gebhardt *et al.* 2003). The tandem duplicated potato invertase genes *invGE* and *invGF* are most similar to tomato *Lin5* and *Lin7*, respectively, located on tomato chromosome IX (Maddison *et al.* 1999; Fridman and Zamir 2003). One amino acid substitution in the third exon of the *Lin5* gene (ASP^{348} to GLU) was identified as being causal for the QTL *Brix9-2-5* for sugar yield of tomato fruits (Fridman *et al.* 2000, 2004).

Functional studies, map position, and synteny with a cloned tomato-sugar-yield QTL all suggest that alleles of the potato invertase genes encoded at the Inv_{ap}-*b* locus could indeed be causal for the cold-sweetening QTL *Sug9a*. As the resolution of the QTL linkage mapping experiments was low due to limitations in population size and environmental effects on trait assessment, the gene responsible for *Sug9* can be invertase itself or another gene that is linked to the Inv_{ap}-*b* locus. Finding an association between invertase alleles at the Inv_{ap}-*b* locus and the cold-sweetening character in populations of unrelated individuals can validate the role of the candidate locus.

On the other hand, invertase alleles associated with the sweetening trait or any other agronomic character in breeding populations can provide diagnostic markers for the selection of cultivars with improved chip quality. This is an important objective of potato breeding, because cold sweetening causes a dilemma for the potato-processing industry, which favors tuber storage at low temperature to delay sprouting. The phenotypic assessment of chip quality requires several years of testing and sufficient tuber numbers. It can be performed reliably only at more advanced stages of variety development. Diagnostic markers derived from genes that control natural variation of tuber quality traits can be used to screen parental lines for superior alleles. Superior alleles can then be combined by crossing and progeny resulting from such crosses can be selected prior to phenotypic evaluation.

In this article we examine the molecular diversity of the *invGE* and *invGF* genes in populations of tetraploid potato-breeding clones. Molecular variants were identified and tested for association with evaluation data for chip quality, tuber starch, and yield of the same populations. Invertase alleles associated with tuber quality traits were discovered.

MATERIALS AND METHODS

Plant material: Healthy, young leaves of 188 tetraploid breeding clones were collected from field-grown plants, freeze dried, and stored at −20° until use. The breeding clones were selections from different cross combinations. Sixty-three clones (NZ1–NZ63, population BNA) were from the breeding program of Böhm-Nordkartoffel Agrarproduktion OHG, 70 clones (NR1–NR70, population NORIKA) were from NORIKA GmbH, and 55 clones (SR1–SR55, population SARA) were from Saka-Ragis Pflanzenzucht GbR.

Phenotypic data: Data for tuber starch content, tuber yield, starch yield, and chip quality of the three sets of breeding clones were the average values assigned to each clone on the basis of the routine evaluations of breeding materials over years and/or the locations used by each breeding company. The percentage of tuber starch content (w/w) was determined from specific gravity (Von Scheele *et al.* 1937). Tuber yield (deciton per hectare) is expressed as the percentage of standard varieties of different maturity types. Starch yield (deciton per hectare) is the product of yield and the percentage of tuber starch and is also expressed as a percentage relative to the standard varieties. Chip quality was assessed in a frying test by rating the chip color from 1 to 9, where 1 is very bad (very dark chip color) and 9 is very good (very light chip color). Chip quality was examined twice, in autumn after harvest and in spring after tuber storage for several months at 8° or 4°.

DNA extraction: DNA was extracted from 30 to 50 mg freeze-dried leaf material in extraction tubes arranged in the 96-well format (Matrix Technologies, Wehrheim, Germany). Leaves were ground to a fine powder by rapidly shaking the tubes with two 3.5-mm tungsten beads added to each sample. Total genomic DNA was extracted using the DNeasy plant mini kit (QIAGEN, Hilden, Germany) according to the supplier's instructions. DNA quality and quantity were evaluated under

UV illumination after ethidium bromide staining on 1% agarose gels.

Single-strand conformation polymorphism analysis: Single-strand conformation polymorphism analysis (SSCP) is a method for detection of point mutations and insertion/deletion polymorphisms in small DNA fragments amplified by PCR (Orita *et al.* 1989). It is based on differential electrophoretic mobilities of single-stranded nucleic acids in nondenaturing polyacrylamide gels. PCR primers were designed on the basis of the nucleotide sequence of the potato genomic DNA fragment containing the invertase genes *invGE* and *invGF* (accession AJ133765; Maddison *et al.* 1999). PCR was performed in 25 μl of 20 mm Tris-HCl, pH 8.4, 50 mm KCl, 1.5 mm $MgCl_2$, 0.2 mm of each dNTP, 0.5 μm of each primer, containing 1 unit Taq polymerase (Invitrogen, Life Technologies, Karlsruhe, Germany), and 40–50 ng of genomic DNA template. Touchdown PCR was used to increase specificity. The annealing temperature in the initial cycle was set 5° above the optimal annealing temperature. In subsequent cycles the annealing temperature was decreased in steps of 1°/cycle for five cycles and then maintained constantly for 35 cycles. Aliquots of the amplification products were separated on 2% agarose gel and visualized by ethidium bromide staining. PCR products were digested with the 4-bp cutter restriction enzyme *Mse*I. Per 4 μl restricted PCR product, 9 μl dye solution containing 95% formamide, 10 mm NaOH, 0.05% of xylene cyanol, and 0.05% of bromophenol blue was added. After denaturation for 5 min at 95°, a 5-μl sample was loaded onto a vertical 0.5 mm × 28.5 cm × 20 cm 0.5 × mutation detection enhancement gel (Cambrex, Rockland, ME). Electrophoresis was performed in 0.6× TBE buffer at room temperature with constant power (1.5 W) for 16–19 hr. DNA fragments were visualized by silver staining as described by Sanguinetti *et al.* (1994).

Amplified fragment length polymorphism analysis: AFLP analysis was carried out using essentially the protocol described by Vos *et al.* (1995) with *Eco*RI/*Mse*I and *Hin*dIII/*Mse*I adaptors and seven primer pairs having the three to four arbitrary nucleotide extensions listed in Table 2.

Analysis of genetic similarity and population structure: A total of 166 polymorphic AFLP bands were scored as dominant markers, with presence (1) of a band being dominant over absence (0). Only distinct and well-resolved bands that had at least 10% frequency in the total population were considered. The assumption was made that, for each primer combination, bands of the same apparent size are homologous, representing the same DNA sequence (Bachmann 1997), and are alleles of a single biallelic locus (Lynch and Milligan 1994). For analysis of similarity between the clones, the distance-matrix approach (Weir 1996) was used, calculating the pairwise genetic distances according to the Jaccard algorithm (Jaccard 1908). Dendrograms were created using the neighbor-joining method (Saitou and Nei 1987; Studier and Keppler 1988). To measure the reliability of the branching patterns, the original matrix was bootstrapped 500 times (Felsenstein 1985). Bootstrap values >20% are shown at the corresponding nodes. All calculations were performed using the program TREECON (Version 1.3b; Van De Peer and De Wachter 1997). Principal component analysis (PCA) was done using the program STATISTICA (version 5.5, StatSoft Europe, Hamburg, Germany). The number of extracted factors was limited to two. Population subdivision was analyzed with the STRUCTURE software, version 2.1 (Pritchard *et al.* 2000). The AFLP data of 179 nonidentical genotypes were run with a burn-in length of 50,000, a run length of 50,000, K values from 1 to 8, and adopting the model of dominant loci. Runs were repeated several times using ancestry models of admixture and no admixture, and allele-frequency models of correlated and independent allele frequencies.

Association test: Polymorphic SSCP fragments, generated with primers for the *invGE/GF* locus in the three sets of breeding clones, were scored as present (1) or absent (0). Allele dosage could not be assessed reliably on the SSCP gels in all tetraploid genotypes and was therefore not considered. Unclear SSCP fragments in singular genotypes were declared as missing values. SSCP fragments were tested for association with the nonparametric phenotypic scoring data (chip color) using the Mann-Whitney *U*-test and SPSS Base 10.0 (Chicago). ANOVA and the same software was used to test the markers for association with parametric data (tuber starch content and yield and starch yield). A SSCP marker fragment was considered to be associated with the trait at $P \leq 0.01$.

Cloning of PCR products and DNA sequence analysis of invertase alleles: PCR products of selected genotypes were excised from agarose gels, purified using the QiaEx gel extraction kit (QIAGEN), and cloned in the pGEM-T vector system I (Promega, Madison, WI) following the supplier's instructions. DNA sequences were determined on both strands by custom sequencing by the Automated DNA Isolation and Sequencing (ADIS) group at the Max-Planck Institute for Plant Breeding Research on a PE/Applied Biosystems 377 sequencer using BigDye terminator chemistry. Thirty cloned PCR products were sequenced for each invertase SSCP allele to obtain a consensus sequence. DNA sequence analyses were performed with the Wisconsin Program Package, Version 10.2 (Genetics Computer Group, Madison, WI).

Mapping of invertase alleles: Segregation of invertase alleles *invGF-4c* and *invGF-4e* was scored in the mapping populations BC916[2] (Gebhardt *et al.* 1991) and "Erwinia" (Zimnoch-Guzowska *et al.* 2000), respectively, by SSCP analysis of the InvGF-4 amplicons in parents and progeny. SSCP fragments were scored as present or absent and were mapped relative to the available molecular maps of these populations as described (Gebhardt *et al.* 1991; Zimnoch-Guzowska *et al.* 2000).

Allele-specific marker assays: Invertase alleles *invGF-4d* and *invGF-4b* were specifically amplified using the primer sequences InvGF-4d and InvGF-4b shown in Table 3. The PCR conditions were 40 ng genomic DNA amplified in a total volume of 25 μl of 20 mm Tris-HCl, pH 8.4, 50 mm KCl, 1.5 mm $MgCl_2$, 0.2 mm of each dNTP, 0.5 μm of each primer, and 1 unit Taq polymerase. Reaction conditions were initial denaturation for 3 min at 94°, followed by a touchdown protocol consisting of a 60-sec denaturation at 94°, a 60-sec annealing, starting at the temperature 5° above the optimal annealing temperature (Table 3), stepwise decreasing by 1° in cycles 1–5, and remaining constant at optimal annealing temperature in the subsequent 30 cycles, a 60-sec extension at 72°, and ending with a final extension at 72° for 5 min. PCR products were separated on 2% agarose gels and were stained with ethidium bromide.

RESULTS

Tuber quality traits: Evaluation data for chip quality in autumn after harvest and in spring after low-temperature storage for tuber yield and starch content and for starch yield are shown in Table 1. The traits, in particular chip quality, were evaluated at the three breeding stations under similar but nonidentical conditions. Trait means and ranges were comparable among the sets of breeding clones, except that the mean tuber starch content was lower in the NORIKA clones (TSC_N) and chip quality after cold storage had a higher mean and a narrower range in the SARA clones (CQS_S). As expected, means for chip quality decreased after low-temperature storage (CQS) when compared to the autumn scores (CQA) due to the accumulation of reducing sugars.

TABLE 1

Traits evaluated, trait codes, and statistical parameters for phenotypic data of tuber quality in populations BNA, NORIKA, and SARA

Population	Trait	Trait code	N^a	Data type	Population mean (standard deviation)	Range
BNA	Chip quality autumn	CQA_B	63	Score from 1 to 9	4.9 (2.4)	1.0–9.0
	Chip quality spring (storage at 8°)	CQS_B	63	Score from 1 to 9	3.9 (2.6)	1.0–8.3
	Tuber starch content	TSC_B	63	%	19.9 (2.5)	15.3–25.5
	Relative tuber yield	RTY_B	60	%	99.3 (11.6)	71.0–127.9
	Relative starch yield	RSY_B	60	%	107.6 (15.2)	74.9–147.2
NORIKA	Chip quality autumn	CQA_N	70	Score from 1 to 9	5.6 (1.9)	1.0–8.0
	Chip quality spring (storage at 4°)	CQS_N	70	Score from 1 to 9	3.6 (2.0)	1.0–8.0
	Tuber starch content	TSC_N	70	%	15.5 (2.5)	10.6–20.6
	Relative tuber yield	RTY_N	70	%	101.5 (14.6)	56.0–132.0
	Relative starch yield	RSY_N	70	%	107.1 (16.2)	77.0–148.0
SARA	Chip quality autumn	CQA_S	28	Score from 1 to 9	6.3 (0.8)	4.8–8.1
	Chip quality spring (storage at 4°)	CQS_S	41	Score from 1 to 9	4.7 (1.2)	1.5–6.5
	Tuber starch content	TSC_S	55	%	19.3 (1.3)	16.7–22.0
	Relative tuber yield	RTY_S	55	%	100.8 (9.0)	88.0–123.0
	Relative starch yield	RSY_S	55	%	100.4 (10.2)	80.0–138.0

[a] Number of breeding clones for which data were available.

Population structure: The three sets of breeding clones (188 genotypes total) were scored for 166 AFLP bands generated with seven different primer combinations (Table 2). Principal component analysis of the AFLP data showed that members of all three populations were intermixed, with the NORIKA and BNA genotypes only partially overlapping with each other and with the SARA genotypes (Figure 1). No distinct clusters were observed either within or between populations. Phenograms were constructed for each population separately and with all 188 genotypes (supplementary Figure 1 at http://www.genetics.org/supplemental/). Nine genotype pairs were closely related (bootstrap support values 100%) within or between populations. They differed for up to two to three AFLP bands. Otherwise, all genotypes were clearly separated by long branches of similar length. Bootstrap support values for the branching pattern at the base of the phenograms were low. For the association tests, one genotype of the nine highly similar pairs was removed from the data set. Analyzing the AFLP data of the reduced number of 179 genotypes multiple times with the *structure* software (Pritchard *et al.* 2000) under various models and parameter settings did not provide evidence for the existence of population substructure.

Molecular diversity at the *invGE/GF* locus: Eight primer pairs (Table 3) were designed to amplify by PCR specific parts of the 9-kbp genomic region of the *invGE/GF* locus (Maddison *et al.* 1999; GenBank accession AJ133765). The structure of the locus and the positions of the amplicons are shown in Figure 2. Seven primer pairs

TABLE 2

AFLP primer combinations used and AFLP bands scored

Primer combination	No. of bands scored	AFLP marker no.
C1: H + ACA/M + CTG	23	V1–V23
C2: E + ACT/M + CAG	27	V24–V50
C3: H + ACT/M + CAT	25	V51–V75
C4: E + ACG/M + CAG	32	V76–V107
C5: E + ACG/M + CTA	29	V108–V136
C6: E + ACT/M + CTGT	13	V137–V149
C7: E + AGG/M + CAGA	17	V150–V166
Total:	166	

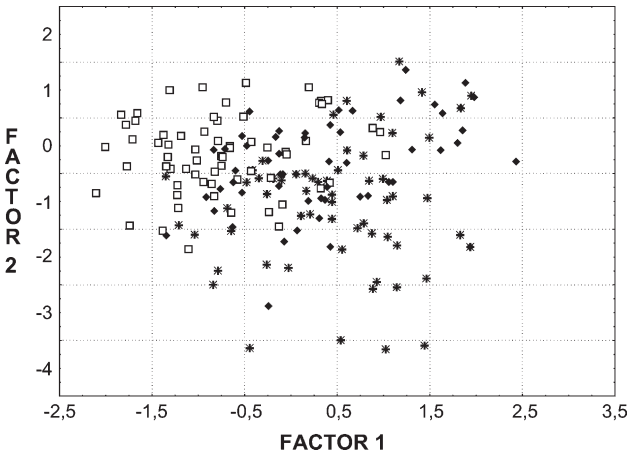

Figure 1.—PCA of 188 tetraploid potato breeding clones based on 166 AFLP markers. (*) BNA, (□) NORIKA, (◆) SARA.

TABLE 3

PCR primers for specific amplification of fragments of invertase genes *invGE* and *invGF*

Primer name	Position[a]	Forward primer sequence 5′–3′	Reverse primer sequence 5′–3′	T_a	Expected size[a] of PCR product in base pairs
InvGE-1	3553–4138	ctc agc atc aca ggt ttt aac	tca tta caa cta att caa ttg	56°	585
InvGE-2	789–2715	gaa aag ctc ttc tct ttg ggg t	ccg gac caa gca cca tat tt	56°	1926
InvGE-3	3963–4566	ggg ttc gac tat cca agg tg	cag cac caa aac tct cca cta c	58°	603
InvGF-4	6481–7056	gtt ggg ctt tgc cag tta tc	gcc cca tac tga ccc att tg	56°	575
InvGF-4b	Supplementary Figure 2[b]	ggg ctt tgc cag tta tct tag tt	ggt tgt tca gat ggt aag tac tgg	61°	296
InvGF-4d	Supplementary Figure 2[b]	tgg gct ttg cca gtt atc tt	ggt tgt tca gat ggt aag tat ag	63°	248
InvGF-5	7984–8675	gtt aca gga atc aca cct gca c	ctt tcc acc agc acc aaa ac	56°	691
InvGE-6	3520–3952	gag caa ggg aga aat gtt tga	aaa cat ctt ggg cat aaa ggt c	59°	433
InvGE-7	3890–4362	caa gtt tga aca agg ccg aac	atc tac ata tcc agc aaa tga gg	59°	473
InvGE-8	3323–3529	cga aga att gta tgg ggt tgg	tcc ctt gct caa ctt ctt gtt	57°	206

[a] According to genomic sequence of GenBank accession AJ133765.
[b] Available at http://www.genetics.org/supplemental/.

flanked one or more of the five introns. Primer pair InvGE-8 was designed to amplify part of exon III of the *invGE* gene. This region is orthologous to the region of the tomato *Lin5* gene, which is responsible for the fruit-sugar-yield QTL *Brix9-2-5* (Fridman and Zamir 2003; Fridman *et al.* 2004). Amplicons generated with primers 1–7 were subjected to SSCP analysis. Scorable SSCP polymorphisms were identified with primers InvGF-4 (Figure 3) and InvGE-6 (not shown). The InvGF-4 and InvGE-6 PCR products of genotypes representing different SSCP fragment patterns in the populations were cloned and sequenced. SSCP analysis of individual invGF-4 and invGE-6 clones revealed five (Figure 3) and three (not shown) different sequence variants, respectively. These eight sequence variants were the most frequent and reliable ones. They did not represent the total number of *invGE/GF* alleles present in the populations. The aligned consensus sequences of *invGF-4* alleles *a, b, c, d,* and *e* (supplementary Figure 2 at http://www.genetics.org/supplemental/) and *invGE-6* alleles *f, g,* and *h* (supplementary Figure 3 at http://www.genetics.org/supplemental/) showed mainly insertion/deletion and single nucleotide polymorphisms in the introns and few amino acid changes in the flanking exon regions. Insertions and deletions within the introns resulted in variable intron length. Genetic mapping of alleles *invGF-4c* and *invGF-4e* confirmed the position of the *invGE/GF* locus on potato chromosome IX (not shown). The tetraploid genotypes were highly heterozygous, having up to four different invertase alleles (Figure 3). The apparent allele frequencies (presence/absence) are shown in Table 4. Alleles *invGF-4a, invGF-4c,* and *invGE-6h* had similar frequencies in all three breeding populations, with *invGF-4a* being the most abundant and *invGE-6h* being the least-frequent allele. The frequencies of the other alleles were more variable among the breeding populations. The alleles *invGF-4d* and *invGF-6f* were highly correlated. Only 6 of 177 genotypes scored for both alleles did not share the two alleles, indicating that *invGF-4d* and *invGF-6f* were in linkage disequilibrium. The InvGE-8 primers (Table 3, Figure 2) were used to generate amplicons from a subset of 20 breeding clones, 10 with and 10 without the alleles *invGF-4d* and *invGE-6f* (see below). The 20 amplicons were directly sequenced. The SNPs detected (not shown) translated into three polypeptide variants (Figure 4). Variant *invGE-8i* was identical to the deduced amino acid sequence of the characterized *invGE* gene (accession AJ133765). The variant *invGE-8f* had a unique histidine instead of proline at position 368. This variant was present in all 10 genotypes

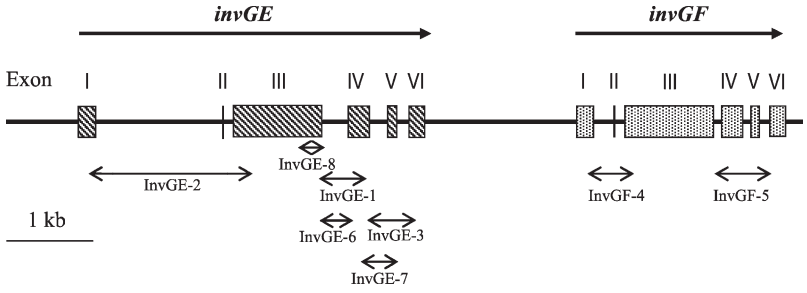

Figure 2.—Structure of the *invGE/GE* locus (adapted from Maddison *et al.* 1999). Size and position of PCR fragments InvGE-1, -2, -3, -6, -7, -8, and invGF-4 and -5 are indicated.

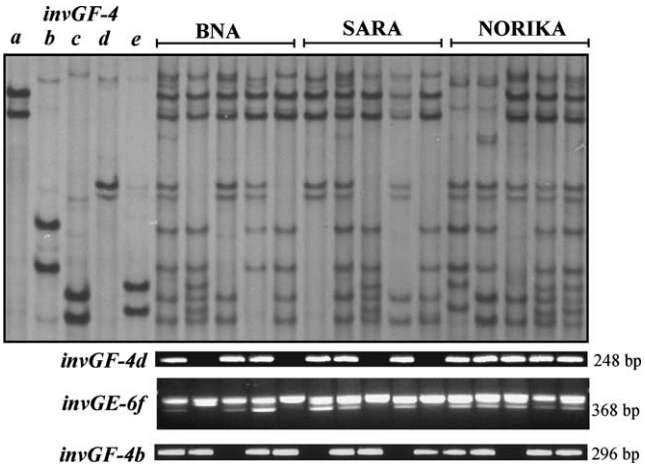

FIGURE 3.—SSCP patterns generated with InvGF-4 primers and allele-specific PCR products for *invGF-4d*, *invGF-4b*, and *invGE-6f*. Amplicons were generated from genomic DNA of 15 tetraploid breeding clones (BNA, SARA, NORIKA) and from plasmid DNA of cloned amplicons from *invGF* alleles *a, b, c, d,* and *e*. Amplicons were digested with *Mse*I. The fragments were separated on SSCP gel and silver stained. Allele-specific PCR products of *invGF-4d*, *invGE-6f*, and *invGF-4b*, which were generated from genomic DNA of the same 15 genotypes, are shown below the corresponding SSCP patterns. Alleles *invGF-4d* and *invGE-6f* are highly correlated with each other and both are associated with better chip quality. Allele *invGF-4b* is associated with lower tuber starch content. The allele-specific PCR products have the same distribution in the breeding clones as the corresponding SSCP fragments.

having alleles *invGF-4d* and *invGE-6f* and was absent in the other 10 genotypes lacking both alleles. *InvGE-6f* and *invGE-8f* were therefore different parts of the same invertase allele *invGE-f*.

Association test: The eight invertase SSCP alleles derived from primers InvGF-4 and InvGE-6 were tested for association with the tuber quality trait scores in individual sets of breeding clones and after combining the data for the same trait of all genotypes. The two highly correlated alleles, *invGF-4d* and *invGF-6f*, were significantly associated with chip quality in individual sets of breeding clones and in the combined populations, whether determined in autumn or in spring after cold storage (Table 5). Presence of these alleles was indicative of, on average, better chip quality. The same alleles

TABLE 4

Apparent frequency (presence/absence) of invertase alleles in the three sets of tetraploid breeding clones

	BNA		SARA		NORIKA	
SSCP alleles	Frequency	N^a	Frequency	*N*	Frequency	*N*
invGF-4a	0.90	61	0.84	51	0.96	67
invGF-4b	0.33	61	0.12	51	0.63	67
invGF-4c	0.51	61	0.61	51	0.51	67
invGF-4d	0.28	61	0.46	50	0.13	67
invGF-4e	0.26	58	0.46	48	0.34	64
invGE-6f	0.23	61	0.45	51	0.15	67
invGE-6g	0.44	61	0.25	51	0.40	67
invGE-6h	0.21	61	0.12	51	0.15	67

[a] *N*, number of genotypes scored.

were associated with decreased tuber starch content in the BNA population but not in any other data set. Allele *invGF-4b* was associated with decreased tuber starch content (TSC_B) and decreased relative starch yield (RSY_B) in the BNA population. In the combined populations, the negative association of *invGF-4b* with TSC was even more significant and explained 13% of the total variation, whereas the association with RSY was not detected any more (Table 5). Instead, a weak negative association was found with relative tuber yield (RTY).

Diagnostic PCR-based marker assays: Allele-specific primers were designed for alleles *invGF-4d* and *invGF-4b* on the basis of sequence polymorphisms in intron I (supplementary Figure 2 at http://www.genetics.org/supplemental/). Allele *invGE-6f* was assayed directly by PCR due to the shorter length of intron III (Figure 3 and supplementary Figure 3 at http://www.genetics.org/supplemental/). The distribution of the allele-specific markers in the breeding clones was identical when compared to the original SSCP fragments (Figure 3).

DISCUSSION

In contrast to inbred lines of maize (REMINGTON *et al.* 2001), tetraploid potato cultivars and breeding clones are noninbred and originate from intercrossing heterozygous parental clones. Recombinant genotypes with su-

```
                                        348                 368
                                         ▼                   ▼
S. pennellii LA716    (AJ272305) VLPDDDIKKGWAGIQGIPRQVWLDLSGKQLVQWPIEELETLRKQKV QLNNKKL
S. lycopersicum       (AJ272304) VLPDDEIKKGWAGIQGIPRQVWLNLSGKQLLQWPIEELETLRKQKV.QLNNKKL
S. tuberosum invGE-8i(AJ133765) VLPDDEIKKGWAGIQAIPRKVWLDPSGKQLIQWPIEELETLRKQKII QLNNKKL
S. tuberosum invGE-8k           VLPDDEIKKGWAGIQAIPRKVWLDPSGKQLIQWPIEELETLRKQKI.QLNNKKL
S. tuberosum invGE-8f           VLPDDEIKKGWAGIQAIPRKVWLDHSGKQLIQWPIEELETLRKQKI.QLNNKKL
```

FIGURE 4.—Amino acid sequence alignment of potato alleles *invGE-8* with orthologous tomato *Lin5* alleles of *Solanum pennellii* LA716 and *S. lycopersicum* (FRIDMAN *et al.* 2004). The amplicon includes that part of exon III of the *invGE* gene that corresponds to exon III of the *Lin5* gene containing the QTN 348. Polymorphic amino acids are shaded. The unique aspartic acid at position 348 of tomato *Lin5* and a unique histidine at position 368 of potato *invGE-f* are shown against a solid background. The invertase allele *invGE-f* was associated with better chip quality.

TABLE 5

***P*-values for the association of invertase alleles *invGF-4b*, *invGF-4d*, and *invGE-6f* with chip quality in autumn (CQA), in spring after cold storage (CQS), with tuber starch content (TSC), and relative tuber and starch yield (RTY, RSY) in populations BNA, NORIKA, and SARA**

Population	Trait	N^a	*invGF-4b* (direction of effect)	*invGF-4d* (direction of effect)	*invGE-6f* (direction of effect)
BNA	CQA_B	61	NS[b]	0.000 (+)	0.000 (+)
	CQS_B	61	NS	0.000 (+)	0.000 (+)
	TSC_B	61	0.010 (−)[c]	0.004 (−)	0.002 (−)
	RTY_B	68	NS	NS	NS
	RSY_B	68	0.010 (−)	NS	NS
NORIKA	CQA_N	69	NS	0.012 (+)	0.012 (+)
	CQS_N	69	NS	0.011 (+)	0.012 (+)
	TSC_N	69	NS	NS	NS
	RTY_N	69	NS	NS	NS
	RSY_N	69	NS	NS	NS
SARA	CQA_S	28	NA[b]	NS	NA
	CQS_S	40	NS	NS	0.009 (+)
	TSC_S	55	NA	NS	NS
	RTY_S	55	NA	NS	NS
	RSY_S	55	NA	0.036 (−)	NS
Combined	CQA	155	NS	0.000 (+)	0.000 (+)
	CQS	167	NS	0.000 (+)	0.000 (+)
	TSC	178	0.000 (−)	NS	NS
	RTY	176	0.027 (−)	NS	NS
	RSY	176	NS	NS	NS

[a] Number of breeding clones included in the association test.
[b] NS, $P > 0.05$; NA, not analyzed either due to low allele frequency or small population size.
[c] Direction of effect. "Plus" and "minus" indicate whether the allele was associated with increased or decreased trait value.

perior agronomic performance are selected in F_1 progeny and the genotypes are fixed by vegetative propagation. Between one and four alleles are present at each locus. Allele dosage varies from one (simplex) to four (quadruplex). The number of meiotic generations separating different cultivars is limited (Gebhardt *et al.* 2004; Simko 2004). Neighbor-joining trees based on genetic distances, principal component analysis, and tests for population substructure with the method of Pritchard *et al.* (2000) using random AFLP markers did not reveal population substructure in the genetic materials used in our association study. Most individuals were clearly separated from each other by similar genetic distances, except few highly similar genotype pairs. Upon closer scrutiny, those genotype pairs turned out to be, in fact, identical, the few AFLP differences being due to scoring errors. The uniform population structure as revealed in the sample of 179 breeding clones from three selection programs probably results from the high heterozygosity of the clones and should be rather typical for cultivated potato, at least in the material bred in temperate zones of Europe and North America (Simko *et al.* 2004). The risk of false-positive associations due to population substructure (Pritchard *et al.* 2000) is therefore considered low in the sample of cultivated potato studied here.

The DNA polymorphisms associated with chip quality in the invertase genes *invGE* and *invGF* were separated by ~3 kbp (Maddison *et al.* 1999). This short physical distance explains the coinheritance of the alleles *invGE-f* and *invGF-d* in 97% of the breeding clones. The physical distance on both sides of the *invGE/GF* locus, over which linkage disequilibrium decays, and the flanking genomic sequences are currently unknown. We therefore cannot exclude the possibility that a gene physically closely linked to *invGE/GF* is causal for the observed effect on chip quality. The association test did not clearly resolve which, if any, of the two invertase genes may be causal for the effect on chip quality. The association between allele *invGE-f* and chip quality was slightly more significant and more consistent than the association with allele *invGF-d* when tested in individual breeding populations. This difference is based on only six genotypes and needs confirmation in larger populations. However, evidence based on gene expression patterns and orthology with the tomato *Lin5/Lin7* locus suggests that *invGE* rather than *invGF* is causal for cold-sweeten-

ing QTL *Sug9*. Genes controlling tuber sugar content are expected to be expressed in tuber tissue, and transcript, protein, and/or enzyme activity levels are expected to be affected by low temperature. Histochemical expression analysis in transgenic potato plants transformed with chimeric constructs of the β-glucuronidase (GUS) reporter gene fused to either the *invGF* or the *invGE* promoter showed that *invGF* is expressed only in anthers, whereas *invGE* is expressed in addition to anthers in several other tissues, particularly in the basal vascular tissue of axillary buds (Maddison *et al.* 1999). In the tuber, which is a modified stem growing from an underground axillary bud, the GUS-stained tissue corresponding to this region was underneath the "eyes" (leaf scars with subtended lateral buds) (Maddison *et al.* 1999). In developing tubers, GUS staining was restricted to the apical bud region (Viola *et al.* 2001). Responsiveness of *invGE* expression or invGE enzyme activity to low temperature has not been examined. However, alternative splicing upon cold treatment has been demonstrated for the related potato invertase gene *invCD111* (Bournay *et al.* 1996). *InvCD111* is the invertase gene on potato chromosome X that corresponds to *invGE* on chromosome IX as part of the segmental duplication between chromosomes IX and X, which includes tomato loci *Lin5/Lin7* and *Lin8/Lin6*, respectively (Fridman and Zamir 2003). Moreover, potato *invGE* is most closely related (92% amino acid similarity) and orthologous to tomato invertase gene *Lin5*, which is causal for fruit-sugar-yield QTL *Brix9-2-5* (Fridman *et al.* 2000). Unlike *Lin5*, the *invGE-f* allele was not polymorphic at the codon identified as the quantitative trait nucleotide (QTN) for *Brix9-2-5* (Fridman *et al.* 2004). Interestingly, the invGE-f-deduced protein instead has 20 amino acids downstream from a specific replacement of proline by histidine. Whether or not this replacement is relevant for the functionality of invGE-f needs further investigation.

On the basis of a comparative sequence analysis with other plant invertase genes, both *invGE* and *invGF* encode cell-wall-bound invertase, which is located extracellularly in the apoplastic space. There is no experimental evidence on the subcellular localization of invertase proteins in tuber tissue. In developing tubers, the activity of cell-wall-bound invertase is restricted to the apical bud region and does not seem to play a major role for the import of sucrose into the sink tuber, where it is converted into hexose sugars for starch biosynthesis (Viola *et al.* 2001; Roitsch and González 2004). Cold sweetening so far has been attributed to soluble, vacuolar invertase rather than to cell-wall-bound invertase (Pressey and Shaw 1966). The association of alleles of the invertase gene *invGE* with sugar content may indicate a novel regulatory role for this gene or a subcellular localization other than that inferred from DNA sequence similarity.

Our results further support a causal role of the *invGE/GF* locus for the potato cold-sweetening QTL *Sug9a* and demonstrate the value of the candidate gene approach as an alternative to map-based cloning of QTL, which is difficult to perform in noninbred, polyploid species. We show that natural DNA variation of candidate genes can be used to develop diagnostic, PCR-based markers for complex agronomic characters in a polyploid crop species. The basis for this is the finding of associations between molecular diversity at a candidate locus and phenotypic evaluation data as generated routinely within breeding programs. In contrast to QTL mapping, no time-consuming generation and phenotypic evaluation of specific experimental populations was required. However, the candidate gene approach is limited to traits for which good functional models are available on the basis of physiological, biochemical, and genetic characterization, such as sugar and starch content of potato tubers.

The *invGE/GF* locus explains only 1 of ~20 cold-sweetening QTL mapped in diploid, experimental populations (Menendez *et al.* 2002). Selecting superior invertase alleles at this singular locus does not yet guarantee better chip quality, as additive effects and interactions with other QTL for tuber quality traits are likely to be relevant for overall chip quality. Similar studies such as those for *invGE/GF* have therefore been initiated for further candidate loci. With each additional association identified, the combinatorial possibilities multiply and, correspondingly, the possibility of selecting and combining superior QTL alleles at several loci. This will eventually lead to haplotype-based design and selection of better-adapted potato cultivars.

The authors appreciate the critical reading of the manuscript by Maarten Koornneef. Part of this work was supported by grant 0313038 Genomanalyse in Biologischen System Pflanze-CHIPS from Bundesministerium für Bildung und Forschung (German Federal Ministry for Education and Research) carried out in the department of plant breeding research and yield physiology headed by Francesco Salamini. L.L. gratefully acknowledges support by a fellowship of the Max-Planck Society.

LITERATURE CITED

Bachmann, K., 1997 Nuclear DNA markers in plant biosystematic research. Opera Botanica **132:** 137–148.

Bonierbale, M. W., R. L. Plaisted and S. D. Tanksley, 1988 RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics **120:** 1095–1104.

Bournay, A.-S., P. E. Hedley, A. Maddison, R. Waugh and G. C. Machray, 1996 Exon skipping induced by cold stress in a potato invertase gene transcript. Nucleic Acids Res. **24:** 2347–2351.

Chen, X., F. Salamini and C. Gebhardt, 2001 A potato molecular function map for carbohydrate metabolism and transport. Theor. Appl. Genet. **102:** 284–295.

Coffin, R. H., Y. Yada, K. L. Parkin, B. Grodzinski and D. W. Stanley, 1987 Effect of low temperature storage on sugar concentrations and chip color of certain potato processing potato cultivars and selections. J. Food Sci. **52:** 639–645.

Dominguez, I., E. Graziano, C. Gebhardt, A. Barakat, S. Berry *et al.*, 2003 Plant genome archeology: evidence for conserved ancestral chromosome segments in dicotyledonous plant species. Plant Biotechnol. J. **1:** 91–99.

Douches, D., and R. Freyre, 1994 Identification of genetic factors influencing chip color in diploid potato *Solanum spp.* Theor. Appl. Genet. **89:** 474–480.

Felsenstein, J., 1985 Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:** 783–791.

Fridman, E., and D. Zamir, 2003 Functional divergence of a syntenic invertase gene family in tomato, potato, and Arabidopsis. Plant Physiol. **131:** 603–609.

Fridman, E., T. Pleban and D. Zamir, 2000 A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. Proc. Natl. Acad. Sci. USA **97:** 4718–4723.

Fridman, E., F. Carrari, Y.-S. Liu, A. Fernie and D. Zamir, 2004 Zooming-in on a quantitative trait nucleotide (QTN) for tomato yield using wild species introgression lines. Science **305:** 1786–1789.

Gebhardt, C., E. Ritter, A. Barone, T. Debener, B. Walkemeier *et al.*, 1991 RFLP maps of potato and their alignment with the homeologous tomato genome. Theor. Appl. Genet. **83:** 49–57.

Gebhardt, C., B. Walkemeier, H. Henselewski, A. Barakat, M. Delseny *et al.*, 2003 Comparative mapping between potato (*Solanum tuberosum*) and *Arabidopsis thaliana* reveals structurally conserved domains and ancient duplications in the potato genome. Plant J. **34:** 529–541.

Gebhardt, C., A. Ballvora, B. Walkemeier, P. Oberhagemann and K. Schüler, 2004 Assessing genetic potential in germplasm collections of crop plants by marker-trait association: a case study for potatoes with quantitative variation of resistance to late blight and maturity type. Mol. Breed. **13:** 93–102.

Greiner, S., T. Rausch, U. Sonnewald and K. Herbers, 1999 Ectopic expression of a tobacco invertase inhibitor homolog prevents cold-induced sweetening of potato tubers. Nat. Biotechnol. **17:** 708–711.

Hedley, P. E., G. C. Machray, H. V. Davies, L. Burch and R. Waugh, 1993 cDNA cloning and expression of a potato (*Solanum tuberosum*) invertase. Plant. Mol. Biol. **22:** 917–922.

Hedley, P. E., G. C. Machray, H. V. Davies, L. Burch and R. Waugh, 1994 Potato (*Solanum tuberosum*) invertase-encoding cDNAs and their differential expression. Gene **145:** 211–214.

Heineke, D., U. Sonnewald, D. Bussis, G. Gunter, K. Leidreiter *et al.*, 1992 Apoplastic expression of yeast-derived invertase in potato. Plant Physiol. **100:** 301–308.

Jaccard, P., 1908 Nouvelle recherches sur la distribution florale. Bull. Soc. Vaud. Sci. Nat. **44:** 223–270.

Lynch, M., and B. G. Milligan, 1994 Analysis of population genetic structure with RAPD markers. Mol. Ecol. **3:** 91–99.

Maddison, A. L., P. E. Hedley, C. R. Meyer, N. Aziz, D. Davidson *et al.*, 1999 Expression of tandem invertase genes associated with sexual and vegetative growth cycles in potato. Plant Mol. Biol. **41:** 741–751.

Menendez, C. M., E. Ritter, R. Schäfer-Pregl, B. Walkemeier, A. Kalde *et al.*, 2002 Cold sweetening in diploid potato: mapping quantitative trait loci and candidate genes. Genetics **162:** 1423–1434.

Orita, M., Y. Suzuki, T. Sekiya and K. Hayashi, 1989 Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics **5:** 874–879.

Pflieger, S., V. Lefebvre and M. Causse, 2001 The candidate gene approach in plant genetics: a review. Mol. Breed. **7:** 275–291.

Pressey, R., and R. Shaw, 1966 Effect of temperature on invertase, invertase inhibitor, and sugars in potato tubers. Plant Physiol. **41:** 1657–1661.

Pritchard, J. K., M. Stephens and P. Donnelly, 2000 Inference of population structure using multilocus genotype data. Genetics **155:** 945–959.

Remington, D. L., J. M. Thornsberry, Y. Matsuoka, L. M. Wilson, S. R. Whitt *et al.*, 2001 Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proc. Natl. Acad. Sci. USA **98:** 11479–11484.

Roitsch, T., and M.-C. Gonzáles, 2004 Function and regulation of plant invertases: sweet sensations. Trends Plant Sci. **9:** 606–613.

Saitou, N., and M. Nei, 1987 The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:** 406–425.

Sanguinetti, C. J., E. Dias Neto and A. J. G. Simpson, 1994 Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. Biotechniques **17:** 914–921.

Schäfer-Pregl, R., E. Ritter, L. Concilio, J. Hesselbach, L. Lovatti *et al.*, 1998 Analysis of quantitative trait loci (QTL) and quantitative trait alleles (QTA) for potato tuber yield and starch content. Theor. Appl. Genet. **97:** 834–846.

Scheffler, J. A., H. Hesselbach, H. Hemme and F. Salamini, 1992 Sampling potato genotypes that maintain chip quality under low temperature storage. J. Genet. Breed. **46:** 103–110.

Simko, I., 2004 One potato, two potato: haplotype association mapping in autotetraploids. Trends Plant Sci. **9:** 441–448.

Simko, I., S. Costanzo, K. G. Haynes, B. J. Christ and R. W. Jones, 2004 Linkage disequilibrium mapping of a *Verticillium dahliae* resistance quantitative trait locus in tetraploid potato (*Solanum tuberosum*) through a candidate gene approach. Theor. Appl. Genet. **108:** 217–224.

Sonnewald, U., M. R. Haijrezaei, J. Kossmann, A. Heyer, R. N. Trethewey *et al.*, 1995 Increased potato tuber size resulting from apoplastic expression of a yeast invertase. Nat. Biotechnol. **15:** 794–797.

Studier, J. A., and K. J. Keppler, 1988 A note on the neighbor-joining algorithm of Saitou and Nei. Mol. Biol. Evol. **5:** 729–731.

Tymowska-Lalanne, Z., and M. Kreis, 1998 The plant invertases: physiology, biochemistry and molecular biology. Adv. Bot. Res. **28:** 71–117.

Van De Peer, Y., and R. De Wachter, 1997 Construction of evolutionary distance trees with TREECON for Windows: accounting for variation in nucleotide substitution rate among sites. Comput. Appl. Biosci. **13:** 227–230.

Viola, R., A. G. Roberts, S. Haupt, S. Gazzani, R. D. Hancock *et al.*, 2001 Tuberization in potato involves a switch from apoplastic to symplastic phloem unloading. Plant Cell **13:** 385–398.

Von Scheele, C., G. Svensson and J. Rasmusson, 1937 Die Bestimmung des Stärkegehalts und der Trockensubstanz der Kartoffel mit Hilfe des spezifischen Gewichts. Landw. Vers. Station **127:** 67–96.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. Van De Lee *et al.*, 1995 AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. **23:** 4407–4414.

Weir, B. S., 1996 *Genetic Data Analysis II.* Sinauer Associates, Sunderland, MA.

Zhou, D., A. Mattoo, N. Li, H. Imaseki and T. Solomos, 1994 Complete nucleotide sequence of potato tuber acid invertase cDNA. Plant Physiol. **106:** 397–398.

Zimnoch-Guzowska, E., W. Marczewski, R. Lebecka, B. Flis, R. Schäfer-Pregl *et al.*, 2000 QTL analysis of new sources of resistance to Erwinia carotovora ssp. atroseptica in potato done by AFLP, RFLP and resistance-gene-like markers. Crop Sci. **40:** 1156–1167.

Zrenner, R., K. Schüler and U. Sonnewald, 1996 Soluble acid invertase determines the hexose-to-sucrose ratio in cold-stored potato tubers. Planta **198:** 246–252.

Communicating editor: P. J. Oefner