Theor Appl Genet (2008) 116:1003–1013
DOI 10.1007/s00122-008-0731-5

ORIGINAL PAPER

Construction of a potato transcriptome map based on the cDNA–AFLP technique

E. Ritter · J. I. Ruiz de Galarreta · H. J. van Eck · I. Sánchez

Received: 7 March 2007 / Accepted: 9 February 2008 / Published online: 4 March 2008

Abstract The cDNA–AFLP technique can be used to monitor differential gene expression, but also for linkage mapping. Extending previous works, we have now constructed an integrated linkage map of the potato transcriptome based on the said technique that has a length of around 800 cM and contains nearly 700 transcriptome derived fragments (TDFs). At the same time, most of these markers are anchored to the bins of a highly saturated reference map in potato, combining in this way the information provided by different marker types. Moreover, we detected and confirmed an elevated degree of allelic fragments with this marker type, which was present in nearly half of all primer combinations and involved around 20% of all fragments. These properties were particularly useful to establish anchor points for integrating the individual parental linkage maps. Comparative expression profiling in different plant materials revealed that only a few additional TDFs were obtained which were specific for mature leaves or tubers compared to the TDFs present in whole in vitro plants. Since TDF markers are derived from coding regions, they generally also represent sequences with a biological function. In four case studies, co-migrating TDFs in different *Solanum* wild species always represented potential alleles based on elevated homologies among them. Two resistance gene homologs were identified by analysing TDFs, which were co-located with known QTLs.

Communicated by A. Schulman.

E. Ritter · J. I. Ruiz de Galarreta (✉) · I. Sánchez
NEIKER, The Basque Institute of Agricultural Research and Development, P.O. Box 46, E01080 Vitoria, Spain
e-mail: jiruiz@neiker.net

H. J. van Eck
Laboratory of Plant Breeding, Wageningen University, P.O. Box 386, 6700 AJ Wageningen, The Netherlands

Introduction

In potatoes, the first genetic linkage maps based on RFLP markers were published in 1988 by Bonierbale et al. for the interspecific cross *Solanum phureja* × (*Solanum tuberosum* × *Solanum chacoense*) and by Gebhardt et al. (1989) for a diploid *S. tuberosum* progeny. This latter map also included locations of known genes such as *PAL*, *Rubisco*, CoA *ligase* and *glutamine synthetase*. After further saturation it was possible to align potato and tomato maps using common probes descending from tomato (Tanksley et al. 1992; Gebhardt et al. 1992). Other reduced linkage maps were constructed in different genetic backgrounds and were aligned using codominant markers such as RFLPs, and particularly SSRs, provided together with a linkage map by Milbourne et al. (1998).

In order to improve the integration of available information and to provide a highly saturated reference map of potatoes, an ultra-high density (UHD) map of potato was recently established based on AFLP markers (Isidore et al. 2004; van Os et al. 2006). The UHD map permits the alignment of other maps in potatoes based on integrated SSR and RFLP reference markers or even by co-migrating AFLPs (van Eck et al. 1995). The linkage map contains over 10,000 markers derived from the analysis of 381 primer combinations in a population of 130 progeny genotypes (van Os et al. 2006).

With increasing marker density, the influence of scoring errors is also raised. These inflate map distances due to additional false recombinants. Therefore, in order to improve the quality of the map and to facilitate the control

of scoring values, a novel "bin" concept based on segregation patterns in the whole progeny was developed. This concept does not only consider the absolute numbers of recombination events between marker pairs, but also the particular genotypes in which recombination events occur (Isidore et al. 2004; van Os et al. 2006). All markers of the UHD map are organised in a total of 1,118 bins distributed over the 24 parental chromosomes of the mapping population.

However, linkage maps in potato are mainly based on neutral DNA markers and it is generally unknown if coding or non-coding regions of the genome are targeted by each marker. The cDNA–AFLP technique is analogous to the normal AFLP technique but uses, instead of genomic DNA, double-stranded cDNA derived from mRNA as a template. Therefore, the obtained cDNA–AFLP fragments or transcriptome derived fragments (TDFs; Bachem et al. 1996) target coding regions of the genome. The cDNA–AFLP technique allows the monitoring of differential gene expression, and to identify genes involved in, or controlling, various biological processes ranging from development to responses to environmental cues (Breyne and Zabeau 2001). The technique has been applied in different organisms, and a variety of examples exist for plants also (Van der Biezen et al. 2000; Durrant et al. 2000; Breyne and Zabeau 2001; Bachem et al. 1996, 2001; Trindade et al. 2003; Samuelian et al. 2004; Ranik et al. 2006).

However, in addition, the cDNA–AFLP technique can also be applied to obtain a genetic linkage map (Brugmans et al. 2002). The allelic variants of constitutively expressed genes can generate polymorphic cDNA–AFLP fragments that segregate in the progeny and can, therefore, be used like other markers for linkage mapping. However, the difference is that since these markers are derived from mRNA sequences of constitutively expressed genes, they generally also represent sequences with a biological function as previously shown by Brugmans et al. (2002).

Extending the original work of Brugmans et al. (2002), we report in this paper on the usefulness of cDNA–AFLP for linkage mapping and marker development. At the same time, allelic markers have been used to construct an integrated transcriptome map in the potato.

Materials and methods

Plant material

Parents and a subset of 90 progeny genotypes of the UHD mapping population described by Isidore et al. (2004) were used for transcriptome mapping with the cDNA–AFLP technique. The original mapping population was derived from the cross between two diploid heterozygous potato clones, SH83-92-488 × RH89-039-16 (hereafter referred to as SH × RH).

For transcriptome mapping, whole in vitro plants without roots were processed. In addition we performed comparative cDNA–AFLP analyses using in vitro plants, mature leaves of greenhouse plants and tubers from parents, and a subset of 12 progeny genotypes of the UHD mapping population. Materials were always of the same age, but possible differences at the physiological stage were not considered. Since the experimental conditions were homogenous for all materials under study, we assume that mRNA populations from genes that are continuously expressed (constitutive expression) were targeted and induced expression can be neglected.

Furthermore, we cloned and sequenced four co-migrating TDFs in several of the following *Solanum* wild species accessions: *S. venturii* (*vnt* 8239), *S. yungasense* (*yun* 2173), *S. weberbaueri* (*wbr* 2724), *S. multidissectum* (*mlt* 931), *S. pinnatisectum* (*pnt* 8175), *S. polyadenium* (*pld* 8182), *S. polytrichon* (*plt* 53650), *S. leptophyes* (*lph* 27215), *S. papita* (*pta* 15442). The three letters encode the species and the numbers represent the collection number (Ruiz de Galarreta et al. 1998).

Molecular techniques

An improved cDNA–AFLP technique described in detail by Breyne et al. (2003) was used and applied to parents and progeny genotypes of the mapping population. Total RNA was extracted from all materials using the method described by Bachem et al. (1998). Total RNA concentrations were estimated by a spectrophotometer and visualised on 1% agarose gels. Poly-A^+ RNA was obtained from 10 µg of total RNA using 5′ biotinylated oligo (dT) primer bound to paramagnetic beads coated with streptavidin (Dyanabeads M-280 Streptavidin, DYNAL, Oslo, Norway). First and second strand cDNAs were synthesized according to Sambrook et al. (1989). Double-stranded cDNA (ds-cDNA) was digested with *Ase*I and *Taq*I (NEB Biolabs Inc., New Brunswick, NE, USA) followed by ligation of *Ase*I and *Taq*I adapters with T4 DNA ligase (Invitrogen Inc., Barcelona, Spain). PCR reactions for cDNA–AFLP amplifications were carried out as described by Bachem et al. (1998). Primers were labelled for the specific amplifications with either of the fluorescent infrared dyes, IRD800 or IRD700 (LI-COR, Lincoln, Nebraska, USA). Amplification products were denatured and separated on 6% denaturing polyacrylamide (19:1) gels. They were visualised on a LI-COR 4200-S1 DNA Sequencer and Fragment Analysis System as described by the manufacturer (LI-COR, 1997).

Several cDNA-fragments were cloned and sequenced using standard methodology (Sambrook et al. 1989). The

TA Cloning Kit (Invitrogen Inc., Barcelona, Spain) was used for cloning in accordance with the instructions of the manufacturer. We sequenced three clones of each TDF in order to determine the consensus sequence.

Data analyses

Sequence homology searches were performed in the GenBank database (Benson et al. 2007) accessed through the NCBI homepage (http://www.ncbi.nlm.nih.gov/). We used the BLASTn and BLASTx search algorithms (Altschul et al. 1997). Multiple sequence alignments were carried out using the Clustal W algorithm (Chenna et al. 2003) through the WEB server of EBI (European Bioinformatics Institute; http://www.ebi.ac.uk/clustalw/).

Presence and absence of each segregating fragment was scored in parents and progeny genotypes of the mapping population. Linkage map construction was performed in several steps. Initially, parental linkage maps were established by arranging cDNA segregation patterns into the "bins" of the UHD map.

The methods to obtain a skeleton map of ordered bins are described in detail by van Os et al. (2006). Basically a "bin" represents a segregation pattern (presence/absence) of the markers contained by the bin, across the genotypes of the whole mapping progeny. The distance between two adjacent bins is one recombination in a specific genotype. If several recombination events occur between adjacent bins, so called "empty" bins are generated. We used the BINMAP Software (WAU, Laboratory for Plant Breeding, unpublished) to identify the most likely position of cDNA–AFLP markers relative to the skeleton bin map, by comparing the segregation pattern of the marker with the segregation patterns of the bins (van Os et al. 2006). This method also conveniently permitted the integration of markers with distorted segregations in the framework of existing bins.

In addition, we developed software *in house* (APODAT, unpublished) to generate graphical genotypes for the detection of singletons (unlikely double crossover events in adjacent bins for the same genotype; van Os et al. 2005) and to improve data quality. During iterations of tentative placement of markers, singleton checking in the lab, and renewed placement of markers, we finished with clean data, or discarded markers with ambiguous bin assignments.

On the other hand, all segregation data (presence/absence) were processed with the MAPRF Software (Ritter et al. 1990; Ritter and Salamini 1991) in order to produce an integrated transcriptome map of the population. For this purpose, initial marker orders were maintained within parental linkage groups as determined by BINMAP. The corresponding homologous linkage groups were integrated using the methods described by Ritter et al. (1990) and Ritter and Salamini (1991). In this case, anchor points were constituted by putative allelic cDNA–AFLP fragments shared by the two parents within the same primer combination (PC), and/or by common fragments (bridge markers) that showed recombination frequency (RF) values of zero with individual fragments of both parents. In this way, we took advantage of the more precise estimates of RF values for allelic configurations (Ritter and Salamini 1991). Moreover, some common markers were associated with the bins of individual markers if RF values of zero were observed between a common and an individual marker.

Initially, we have assumed allelism between fragments from the same PC when one of the following three conditions were met for our mapping population: (1) allelic fragments of the same parent were identified as cDNA fragments in repulsion with zero recombination, (2) putative alleles from both parents were detected involving one common allelic fragment in repulsion and with recombination values of zero, (3) when a maternal and a paternal marker mapped to a similar position on homologous chromosomes and similar distances, with a closely linked bridge marker, were observed. In addition, we sequenced several such fragments in order to verify their allelic nature.

Results

Polymorphisms

A total of 100 primer combinations (PCs) were processed for the analyses (Table 1). These PCs generated 749 segregating cDNA–AFLP fragments, which ranged between 1 and 20 fragments per PC. In total, 272 fragments descended from the SH parent, 251 from RH, and 226 bands were common to both parents of the UHD mapping population (Table 1). These proportions between maternal, paternal, and bridge markers resembled the proportions also observed with genomic AFLP markers in this material (van Os et al. 2006). Distortions were present in 107 fragments (14.2%), which showed significant deviations from the expected segregation ratios of 1:1 or 3:1, respectively.

Linkage mapping

Due to stringent quality control of the data, most of the segregating TDFs, including those with distorted segregations, could be integrated directly into the existing parental bins, since the corresponding segregation patterns were identical for the genotypes under study. The parental linkage groups contained 269 SH specific (98.8%) and 248 RH specific markers (98.8%). Moreover, 179 cDNAs common to both parents (79.2%) were associated with the UHD bins based on RF values of zero as described above. The cDNA fragments were located in 179 SH specific bins out of 569

Table 1 Observed segregating polymorphisms of cDNA–AFLP primer combinations

No.	PC	No SH	No RH	No CM	Total no.	Al-Loc	No.	PC	No SH	No RH	No CM	Total no.	Al-Loc
1	AC/AA	3	0	1	4	2	53	CC/CG	1	5	4	10	1
2	AC/AC	7	4	4	15	2	54	CC/CT	2	7	0	9	
3	AC/AG	7	10	2	19	1	55	CC/GA	3	2	4	9	2
4	AC/AT	5	4	3	12	1	56	CC/GC	4	8	4	16	
5	AC/CA	2	0	3	5		57	CC/GG	7	1	7	15	1
6	AC/CC	0	0	2	2		58	CC/GT	3	2	2	7	
7	AC/CG	1	2	0	3		59	CC/TA	5	6	6	17	3
8	AC/CT	2	2	0	4		60	CC/TC	9	3	6	18	1
9	AC/GA	11	5	1	17	1	61	CC/TG	2	3	0	5	1
10	AC/GC	5	2	1	8		62	CC/TT	6	8	6	20	4
11	AC/GG	3	1	2	6	1	63	CG/AA	2	4	0	6	1
12	AC/GT	2	4	2	8	1	64	CG/TAC	6	0	1	7	
13	AC/TC	3	4	9	16	1	65	CG/CG	1	0	0	1	
14	AC/TG	1	3	0	4		66	CG/CT	0	0	2	2	
15	AC/TT	2	1	4	7	1	67	CG/GA	1	2	1	4	
16	AG/AA	1	2	1	4		68	CG/GC	6	1	3	10	1
17	AG/AC	4	3	0	7	1	69	CG/GG	0	0	3	3	
18	AG/AG	3	3	1	7	1	70	CG/GT	2	1	0	3	
19	AG/AT	1	0	1	2		71	CG/TA	5	2	0	7	
20	AG/CA	1	1	1	3		72	CG/TC	1	5	1	7	1
21	AG/CC	2	4	1	7	1	73	CG/TG	2	1	2	5	
22	AG/CG	4	1	2	7	1	74	CT/AA	5	6	2	13	1
23	AG/CT	4	4	2	10		75	CT/AC	1	3	2	6	1
24	AG/GA	2	2	4	8		76	CT/AG	2	6	5	13	2
25	AG/GC	5	4	3	12	2	77	CT/AT	3	5	5	13	2
26	AG/GG	1	1	0	2		78	CT/CC	2	0	1	3	
27	AG/GT	4	1	2	7		79	CT/CG	4	7	2	13	2
28	AG/TA	1	0	0	1		80	CT/CT	0	0	2	2	
29	AG/TC	1	2	2	5	1	81	CT/GA	4	3	0	7	
30	AG/TG	2	2	1	5		82	CT/GC	7	1	0	8	
31	AG/TT	0	2	2	4	1	83	CT/GG	3	2	2	7	1
32	AT/AA	0	1	0	1		84	CT/GT	1	2	0	3	
33	AT/AC	0	3	0	3		85	CT/TG	0	0	3	3	1
34	AT/AG	3	2	0	5	1	86	GA/AA	6	1	3	10	
35	AT/AT	0	1	3	4		87	GA/AC	2	4	2	8	
36	AT/CA	2	1	4	7	1	88	GA/AG	1	1	0	2	
37	AT/CC	1	2	0	3	1	89	GA/AT	3	3	0	6	
38	AT/CG	1	1	2	4		90	GA/CA	1	2	2	5	1
39	AT/CT	2	1	0	3		91	GA/CC	3	2	0	5	
40	AT/GA	1	1	2	4		92	GA/CG	2	3	2	7	
41	AT/GC	7	2	1	10		93	GA/CT	1	4	7	12	1
42	AT/GG	1	3	6	10	2	94	GA/GA	1	2	3	6	1
43	AT/GT	5	4	1	10		95	GA/GC	0	1	1	2	
44	CA/AA	4	1	6	11		96	GA/GG	3	3	1	7	1
45	CA/AC	0	2	4	6		97	GA/GT	2	2	1	5	1
46	CA/AG	0	1	4	5		98	GA/TA	1	5	2	8	2
47	CA/AT	3	3	4	10	1	99	GA/TG	1	2	2	5	
48	CC/AA	6	5	5	16	1	100	GA/TT	4	4	5	13	2

Table 1 continued

No.	PC	No SH	No RH	No CM	Total no.	Al-Loc	No.	PC	No SH	No RH	No CM	Total no.	Al-Loc
49	CC/AC	1	2	5	8			Total	272	251	226	749	66
50	CC/AG	4	3	6	13	1		Mean	2.7	2.5	2.2	7.4	0.7
51	CC/AT	8	1	3	12	3		Min	0	0	0	1	1
52	CC/CA	1	0	2	3			Max	11	10	9	20	3

PC selective nucleotides of primer combination with Ase/Taq adaptors, *No SH*, *No RH*, *No CM* number of SH specific, RH specific, and common TDFs, respectively, *Al-Loc* no. of allelic loci revealed by the PC

(31.5%) and in 181 out of 549 RH specific bins (33.0%), respectively. In addition, 11 empty bins of the UHD map could be filled with 18 TDFs based on consistent recombination events with both neighboured bins. Seventeen markers located at 15 loci exceeded the distal bins of the 24 parental linkage groups by one or two recombinants each. On the other hand, three distal bins from SH and 21 from RH were not targeted by TDFs.

Based on the initial parental linkage groups and allelic loci or bridge markers as anchor points, an integrated linkage map was produced as described in "Materials and methods". The orders of fragments in the parental linkage groups were identical when analysed by MAPRF Software. However, recombination frequencies of the corresponding intervals showed smaller variations due to the reduced set of genotypes in our study. In general, reductions of distances were obtained due to missing recombinants in the additional genotypes of the UHD mapping population.

The characteristics of the integrated linkage map are summarised in Table 2. Projection of parental linkage groups was based on 89 anchor points, varying between five and 11 for pairs of homologous linkage groups. Total map length was 795.0 cM and the individual linkage groups varied between 55.1 and 81.3 cM in length (average length 66.3 cM). They contained between 37 and 85 markers each with an average of 58 markers per linkage group. Linkage groups contained between 12 and 45 SH specific markers, between 14 and 33 RH specific markers, and between 8 and 23 cDNA–AFLP fragments common to both parents. On average, a linkage group was composed of 14.9 common cDNA fragments and 22.4 and 20.7 SH and RH specific markers, respectively.

Figure 1 shows a graphical representation of the integrated linkage map displaying the location of parent-specific and common cDNA markers, anchor points, and allelic fragments. The figure also shows bin assignments of the markers in brackets.

In total, 307 loci were detected, containing on average 2.3 markers each. A total of 146 single marker loci (47.6% of all loci) were observed, but also 18 loci with six or more markers distributed over all parts of the linkage groups. The largest cluster containing 14 markers was detected on linkage group (LG) XI. Seven to ten TDFs were found for two loci of LG IV and VI and for three loci on LG XII. The average distance between markers was 1.14 cM. A total of 403 markers (57.9%) were linked with a recombination

Table 2 Characteristics of the integrated transcriptome map of the SH × RH population

LG	SH-M	RH-M	CM	TM	AP	cM	CL	Fr NA	SH bins	RH bins	Empty bins
I	45	17	23	85	11	81.3	5	11	24	19	2
II	22	26	11	59	8	64.9	4	8	17	18	1
III	15	16	16	47	6	78.4	3	9	13	12	0
IV	24	33	19	76	10	72.7	11	26	17	21	0
V	18	17	15	50	7	63.9	4	9	16	13	1
VI	26	16	19	61	8	62.9	5	12	13	15	2
VII	21	14	13	48	8	65.0	6	13	16	11	1
VIII	15	20	9	44	5	61.9	5	11	12	13	0
IX	23	25	17	65	6	62.9	5	10	15	20	0
X	12	17	8	37	5	63.9	3	7	11	15	0
XI	23	24	20	67	10	62.1	9	23	13	14	3
XII	25	23	9	57	5	55.1	6	12	12	10	1
Total	269	248	179	696	89	795.0	66	151	179	181	11
Mean	22.4	20.7	14.9	58.0	7.4	66.3	5.5	12.6	14.9	15.1	0.9
Min	12	14	8	37	5	55.1	3	7	11	10	0
Max	45	33	23	85	11	81.3	11	26	24	21	3

LG linkage group; *SH-M*, *RH-M*, *CF*, *TM* SH specific, RH specific, common, and total markers, respectively; *AP* number of anchor points; *cM* LG lengths in cM (Kosambi units); *CL* codominant loci based on allelic fragments; *Fr NA* total number of fragments composing the CL; *SH bins*, *RH bins* number of SH and RH specific bins which contain at least one TDF; *Empty bins* number of new, former empty bins filled by TDFs

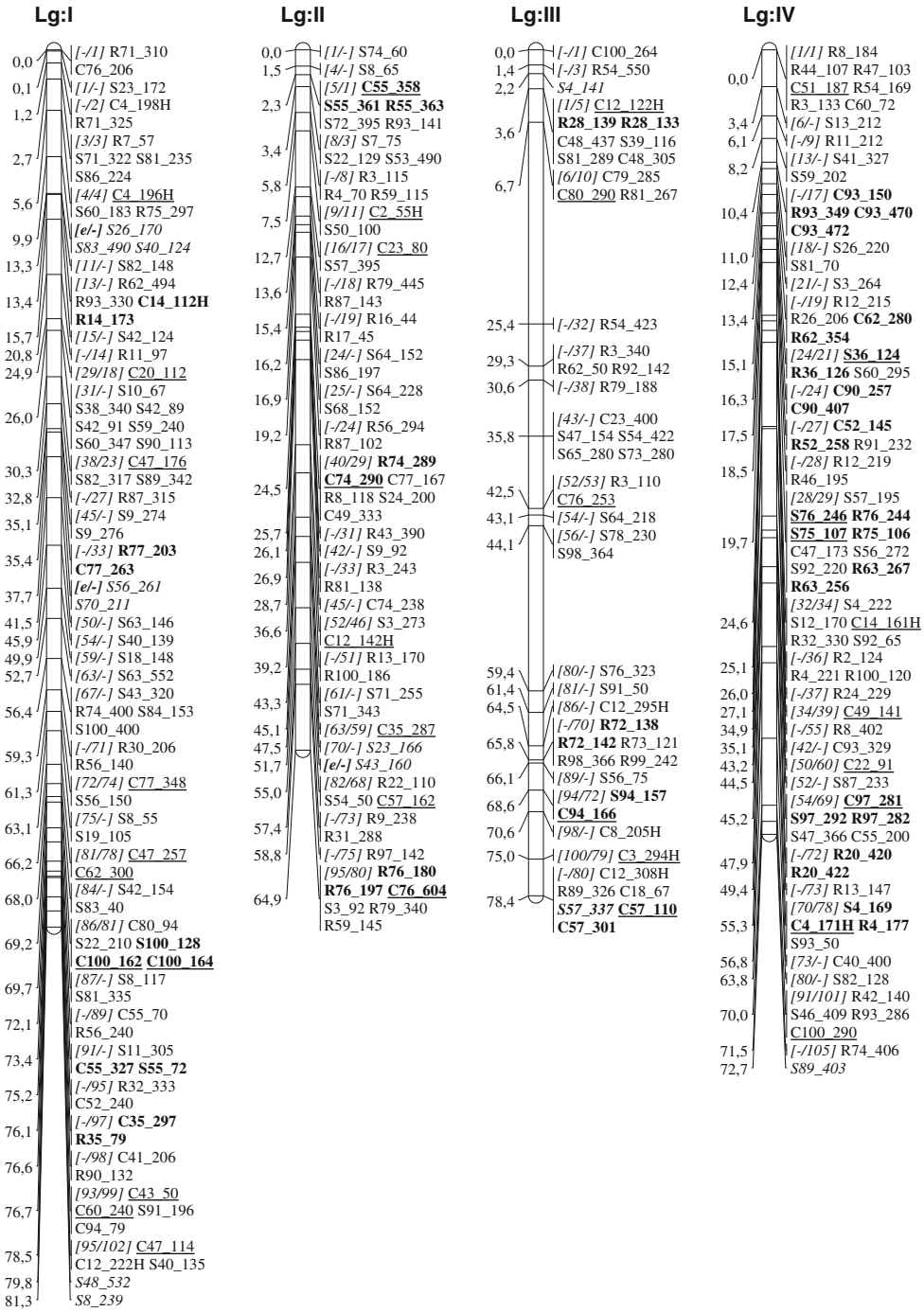
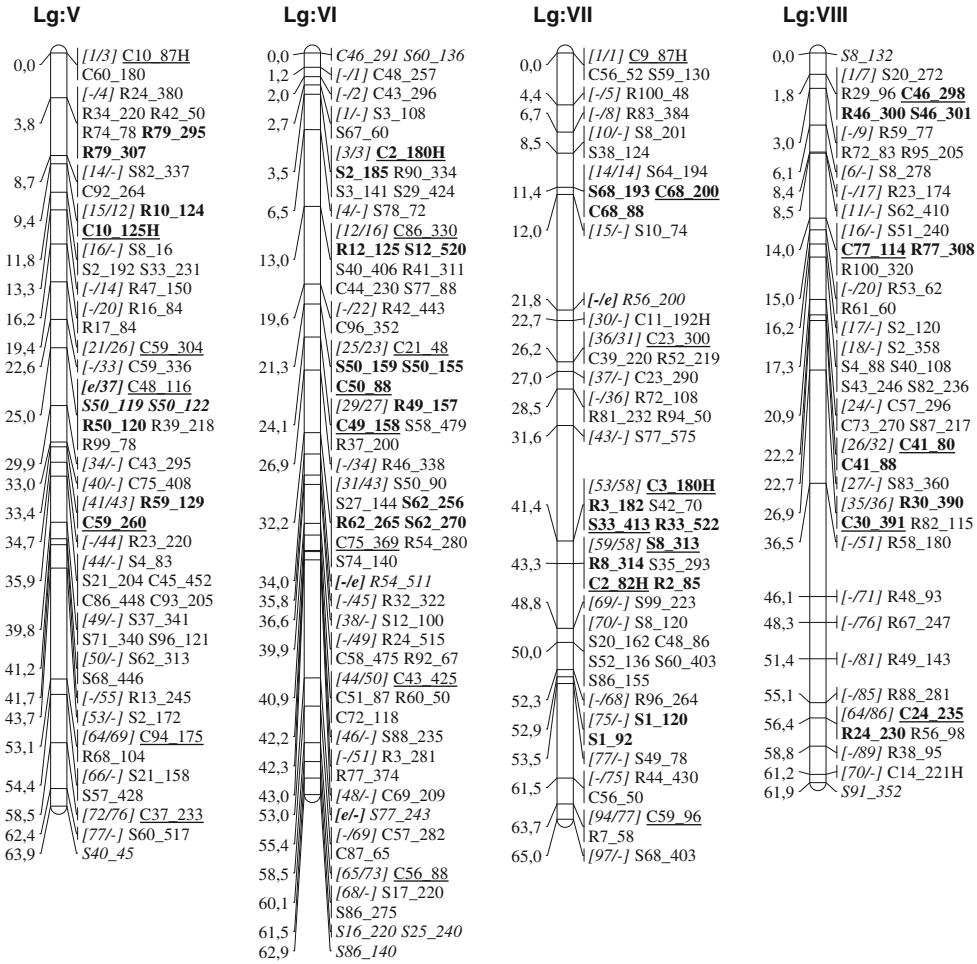


Fig. 1 Integrated transcriptome map from the cross SH × RH. Marker names: SH specific, RH specific and common TDFs are indicated with prefixes "S", "R" and "C", respectively. The number following these corresponds to the PC number in Table 1 and the observed migration values [bp] are indicated in the marker name extension. Bin assignments *[x/y]*: The values in brackets indicate the assignments of markers to SH bin *x* and RH bin *y*. Values in *bold* and showing an *"e"* represent assignments to empty bins of the UHD map. Allellic fragments are indicated in *bold* and have the same PC number. Markers exceeding the distal bins of the UHD map are indicated in *italics* and have a proper cM position without bin assignment. Markers representing an anchor point are *underlined*

frequency of zero to one or several markers. Distances between adjacent TDFs were within 5.3 cM for 95% of all markers. The two largest distances of 18.7 and 15.3 cM were detected on LG III, but also, on LG VII and VIII, two distances between 9 and 10 cM were observed. According to χ^2 tests, the distributions of marker distances and marker number per loci revealed highly significant deviations from uniform distributions.

An elevated number of putative allelic fragments were observed based on the three criteria mentioned above. We identified 66 loci involving 151 allelic fragments (21.7% of all mapped fragments). In 49% of all PCs, at least one locus composed of allelic fragments was detected. Several PCs showed up to three and even four loci (Table 1). We cloned and sequenced several allelic fragments from the three mentioned configuration types and determined their homologies. We performed homology searches in public databases to determine their possible biological function. Table 3 shows the results. In all 11 cases studied, the allelic nature of the segregating TDFs could be verified. The

Fig. 1 continued



average sequence identities from Clustal W alignments of the analysed fragments ranged from 100 to 84.1% in ten cases. In one case, the identity was lower (49.6%), but the homology search with known genes revealed an identical association, indicating that overlapping partial cDNAs of the same gene had been targeted. In nine cases, large homologies with known genes were observed. These include ribosomal genes, transporter proteins, and enzymes (Table 3).

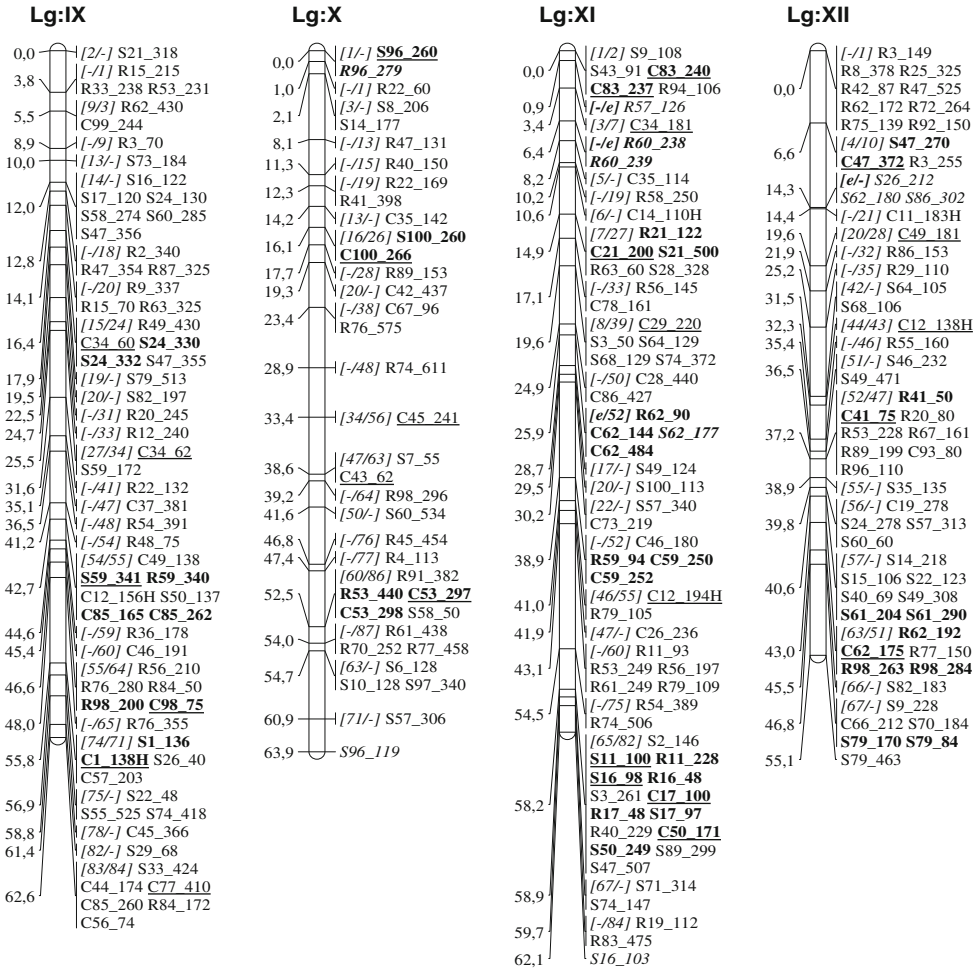
Moreover, in four case studies, we cloned and sequenced co-migrating cDNA–AFLP markers in different accessions of *Solanum* wild species (SWS) as specified in "Materials and methods" and performed homology searches and comparisons. The results are also indicated in Table 3. In all four case studies, we could confirm that the co-migrating TDFs indeed represented alleles of the same gene. The average sequence identities identified by Clustal W alignments ranged from 98.7 to 90.3%. We had chosen TDFs for this analysis, which could be potentially co-located with known QTLs for resistance (see "Discussion") and found in two cases (no. 1 and 2 in Table 4) large homologies with known resistance genes.

Finally, we conducted comparative expression analyses of constitutively expressed genes in different plant materials. Table 5 summarises the results of expression profiling for in vitro plants without roots, mature leaves and dormant tubers. A total of 14 primer combinations were evaluated in these materials using parents and 12 progeny genotypes in each case. Over 88.7% (102) of the 115 segregating TDFs observed in all materials were also present in the cDNAs derived from in vitro plants. Fewer fragments were found in mature leaves and tuber materials (88 and 61 bands, respectively). In 50% of all PCs, one or two segregating TDFs, specific for mature leaves and/or tubers, were detected, but only 13 of such extra bands were observed in these materials (Table 5).

Discussion

The adopted methodology described above led to the establishment of an integrated transcriptome map of the SH × RH mapping population which displays the relative position of specific markers from different parents within

Fig. 1 continued



the genome. The integrated map has a considerable density of nearly 700 markers and an average number of 58 markers per linkage group. At the same time, we could assign nearly all of the individual cDNA fragments and most of the common fragments to the bins of the UHD reference map, linking in this way the information provided by both mapping systems.

Compared with the distribution of genomic AFLP markers in the UHD map, we obtained good genome coverage, although with less markers and more gaps. Additional TDF markers located in empty bins or situated in positions exceeding distal bins compensated for missing TDFs in distal bins of the UHD map. Some of the observed gaps were also found in the parental linkage groups of the UHD map (van Os et al. 2006). Clustering of TDFs was observed and confirmed by the significant deviations from uniform distributions of marker distances and marker number per loci. One explanation for these findings could be the presence of hot and cold spots of recombination that have been described in humans (Jeffreys et al. 2005) and plants (Schmidt et al. 1995). We can also discard visually an association with putative centromere regions of the map of van Os et al. (2006). Lack of centromeric clustering of cDNA–AFLP markers was also observed by Brugmans et al. (2002).

Compared with genomic AFLP markers, we found within identical PCs, in many cases, TDFs that could represent allelic fragments based on the criteria described above. Sequence analyses confirmed their codominant properties. These findings are not surprising if we consider that TDFs are analogous to expressed sequence tag (EST) markers. EST markers have codominant properties and are conserved between species to a certain degree (Temesgen et al. 2001; Fulton et al. 2002). While map positions of TDFs are known but not their sequences, the opposite is the case with EST markers.

However, we cannot discard the possibility that instead of true alleles, a closely linked paralog has been targeted by chance. Nevertheless, considering the resolution of the map, this would not alter the results and the codominant properties of TDFs make them useful for aligning parental linkage maps as in our study.

In all four case studies, large homologies between TDFs in different SWS accessions were observed, indicating their allelic nature. A detailed co-migration analysis of cDNA–AFLP fragments is useful in order to determine the degree

Table 3 Homologies with known genes and sequence identities of cDNA–AFLP fragments from the same PC representing putative alleles in different *Solanum* accessions

Case no.	Origin/allele 1	Origin/allele 2	Clustal W (%)	Homologous gene	Blast X accession no.	e-value
Homologies between putative allelic fragments						
1. Different alleles of the same parent						
1	SH CC/AT-119	SH CC/AT-122	100	Unknown		
2	SH GA/GT-292	CF GA/GT-281	93.7	SCD1 (stomatal cytokinesis-defective)	NP_850959	1E-35
3	RH CA/AT-300	CF CA/AT-298	94.8	50S ribosomal protein L13. (CL13)	P12629	9E-27
4	CF CC/GG-301	CF CC/GG-110	49.5	Alanine aminotransferase	AAR05449	9E-46
5	CF GA/TT-162	CF GA/TT-164	97.5	Unknown		

Case no.	SH allele 1	RH allele 2	CF allele 3	Clustal W (%)	Homologous gene	Blast X accession no.	e-value
2. Different alleles in both parents							
6	CC/GA-361	CC/GA-363	CC/GA-358	90.3	Unknown	NP_195785	6E-42
7	GA/GT-292	GA/GT-282	GA/GT-281	92.4	Os01g0833500 (*Oryza sativa*)	NP_001044713	1E-69
8	CA/AT-301	CA/AT-300	CA/AT-298	80.6	Copper transporter protein	CAA90018	3E-21
9	CC/GA-361	CC/GA-363		95.1	Proteosome/cyclosome regulative subunit	ABE89010	2E-35
10	GA/GT-292	GA/GT-282		92.7	Putative nitrate transporter	AAB95302	5E-63
11	CA/AT-301	CA/AT-300		84.1	EMB2394; structural constituent of ribosome	NP_172011	6E-38

CF fragment common to both parents, *Clustal W* average sequence identity obtained from Clustal W alignments

Table 4 Homologies with known genes and sequence identities of cDNA–AFLP fragments from the same PC representing potential comigrating alleles in different *Solanum* accessions

No.	Comigrating fragment	In genotypes				Clustal W (%)	Homologous gene	Blast X accession no.	e-value
Homologies of co-migrating cDNA–AFLP fragments									
1	AT/CA-135	SH	*wbr* 2724	*vnt* 8239	*pnt* 8175	93.0	Class III chitinase	ASD27874	2E-05
2	CC/TA-240	SH	*yun* 2173	*mlt* 931	*pld* 8182	98.7	LRR protein	NP_191196	1E-36
3	CC/TA-115	RH	*pta* 15442	*plt* 53650		90.3	SNF2 transcriptional factor	NP192575	2E-05
4	AT/CA-293	SH	*pta* 15442	*plt* 53650	*lph* 27215	95.4	*S*-adenosyl-L-methionine sinthetase	ABB02634	1E-61

CF fragment common to both parents, *Clustal W* average sequence identity obtained from Clustal W alignments

of transferability to a broader genetic background. Rouppe van der Voort et al. (1997) observed locus specificity at least for co-migrating genomic AFLP markers in a varying genetic background. However, compared with genomic AFLP markers, we observed a 4-fold reduction in the number of markers per primer combination. Since coding sequences are under purifying selection, they are much more conserved as compared to non-coding sequences, which are targeted by genomic AFLP.

The relationships between TDFs and ESTs should also allow them to predict each other by considering parameters such as sequence information of the EST, restriction site sequences, and size of expected or observed TDFs. The software GeneEST (Qin et al. 2005) has been developed based on these hypotheses.

Transcriptome mapping was conducted using in vitro plants cultured under homogenous conditions in order to avoid artefacts due to variation in gene expression levels. Comparative expression analyses in different tissues revealed only small amounts of additional polymorphisms in mature leaves and tubers. Considering the increase in workload for preparing additional plant materials, the same efficiency can be achieved by increasing the number of analysed PCs using whole in vitro plants. Nevertheless, for detecting tissue specific genes, differential cDNA–AFLP analyses would be the approach of choice.

The origin of absence/presence polymorphisms in transcripts could be due to differences in gene expression or sequence polymorphisms in the coding region. Brugmans et al. (2002) investigated this aspect and found that the vast

Table 5 Segregating cDNA–AFLP fragments in different plant materials (in vitro plants, mature leaves and tubers)

PC	"In vitro" plants Tot Fr	Mature leaves		Tubers		Total bands
		Sh Fr	Ext Fr	Sh Fr	Ext Fr	
AC/AA	4	3	1	4	1	6
AC/AC	15	13	1	8	1	17
AC/AG	19	17	0	9	0	19
AC/AT	12	11	0	5	0	12
AC/CA	5	4	0	4	2	7
AC/CC	2	2	1	2	1	3
AC/CG	3	4	1	5	2	5
AC/CT	4	4	1	3	2	6
AG/AA	4	4	0	2	0	4
AG/AC	7	4	0	4	0	7
AG/CA	3	2	0	2	0	3
AG/CC	7	7	0	3	0	7
AG/CG	7	5	1	4	1	9
AG/CT	10	8	0	6	0	10
Total	102	88	6	61	10	115
Mean	7.3	6.3		4.4		8.2

PC selective nucleotides of primer combination with Ase/Taq adaptors, *Tot Fr* total number of segregating fragments, *Sh Fr* segregating fragments shared with in vitro plants, *Ext Fr* additional specific fragments in the particular material

majority of these polymorphisms resulted from genomic sequence polymorphisms. Comparing the available sequence data of putative alleles and co-migrating allelic fragments in different backgrounds, we always detected smaller sequence divergences and even SNPs supporting this hypothesis. Also, the low amount of additional polymorphisms detected in different tissues indicates that the TDFs represent mainly highly expressed house-keeping genes. It would be interesting to conduct the same mapping experiment after, for example, pathogen inoculation of the whole progeny and to compare the variation in transcript abundance and nature of the TDFs.

Compared with neutral markers, TDFs generally have a biological meaning representing a particular gene as shown in Table 3 and as previously also observed by Brugmans et al. (2002). TDF markers that map to bins where QTLs are located may represent candidate genes controlling the particular trait in the mapping population. It is also possible to project QTL locations from other experiments to certain bin regions in a reference map based on common marker intervals in both maps. We have chosen for homology analyses (Table 4) TDFs that are co-located with known QTLs for resistance. Although the probability that these TDFs explain co-localized QTLs is very low on a case-by-case basis, we found in two cases homologies with resistance genes. Considering that families of resistance genes are frequently organised in clusters (Gebhardt and Valkonen 2001), the chance of finding a target gene of interest is higher in this case. If the TDFs do not represent such candidate genes, then they constitute at least useful markers for marker-assisted selection.

Acknowledgments Part of this work was financed within the frame of EU projects QLK5-CT02-01849 (Apophys) and FOOD-CT-2005-513959 (Bioexploit).

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Bachem C, Oomen R, Visser R (1998) Transcript imaging with cDNA–AFLP: a step by step protocol. Plant Mol Biol Rep 16:157–173

Bachem C, van der Hoeven RS, de Bruijn SM, Vreugdenhil D, Zabeau M, Visser RG (1996) Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. Plant J 9:745–753

Bachem C, Horvath B, Trindade L, Claassens M, Davelaar E, Jordi W, Visser R (2001) A potato tuber-expressed mRNA with homology to steroid dehydrogenases, gibberellin levels and plant development. Plant J 25:595–604

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2007) GenBank. Nucleic Acids Res 35(Database issue):D21–D25

Bonierbale M, Plaisted RL, Tanksley SD (1988) RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics 120:1095–1103

Breyne P, Zabeau M (2001) Genome-wide expression of plant cell cycle modulated genes. Curr Opin Plant Biol 4:136–142

Breyne P, Dreesen R, Cannoo B, Rombaut D, Vandepoele K, Rombauts S, Vandeerhaeghen R, Inze D, Zabeau M (2003) Quantitative cDNA–AFLP analysis for genome-wide expression studies. Mol Genet Genomics 269:173–179

Brugmans B, Fernandez del Carmen A, Bachem CW, van Os H, van Eck HJ, Visser RG (2002) A novel method for the construction of genome wide transcriptome maps. Plant J 31:211–222

Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD (2003) Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res 31:3497–3500

Durrant W, Rowland O, Piedras P, Hammond-Kosack K, Jones J (2000) cDNA-AFLP reveals a striking overlap in race specific resistance and wound response gene expression profiles. Plant Cell 12:963–977

Fulton TM, Van der Hoeven R, Eannetta NT, Tanksley SD (2002) Identification, analysis, and utilization of conserved ortholog set markers for comparative genomics in higher plants. Plant Cell 14:1457–1467

Gebhardt C, Valkonen JP (2001) Organization of genes controlling disease resistance in the potato genome. Annu Rev Phytopathol 39:79–102

Gebhardt C, Ritter E, Debener T, Schachtschabel U, Walkemeier B, Uhrig H, Salamini F (1989) RFLP analysis and linkage mapping in *Solanum tuberosum*. Theor Appl Genet 78:65–75

Gebhardt C, Ritter E, Barone A, Debener T, Walkemeier B, Schachtschabel U, Kaufmann H, Thompson RD, Bonierbale MW, Ganal MW, Tanksley SD, Salamini F (1992) RFLP maps of potato and

their alignment with the homeologous tomato genome. Theor Appl Genet 83:49–57

Isidore E, van Os H, Andrzejewski S, Bakker J, Barrena I, Bryan GJ, Caromel B, van Eck H, Ghareeb B, de Jong W, van Koert P, Lefebvre V, Milbourne D, Ritter E, van der Voort JR, Rousselle-Bourgeois F, van Vliet J, Waugh R (2004) Toward a marker-dense meiotic map of the potato genome: lessons from linkage group I. Genetics 165:2107–2116

Jeffreys AJ, Neumann R, Panayi M, Myers S, Donnelly P (2005) Human recombination hot spots hidden in regions of strong marker association. Nat Genet 37:601–606

Milbourne D, Meyer RC, Collins AJ, Ramsay LD, Gebhardt C, Waugh R (1998) Isolation, characterisation and mapping of simple sequence repeat loci in potato. Mol Gen Genet 259:233–245

Qin L, Prins P, Helder J (2005) Linking cDNA–AFLP-based gene expression patterns and ESTs. Methods Mol Biol 317:123–138

Ranik M, Creux NM, Myburg AA (2006) Within-tree transcriptome profiling in wood-forming tissues of a fast-growing Eucalyptus tree. Tree Physiol 26:365–375

Ritter E, Salamini F (1991) The calculation of recombination frequencies in crosses of allogamous plant species with application to linkage mapping. Genet Res 67:55–65

Ritter E, Gebhardt C, Salamini F (1990) Estimation of recombination frequencies and construction of RFLP linkage maps in plants from crosses betwen heterozygous parents. Genetics 224:645–654

Rouppe van der Voort JN, van Zandvoort P, van Eck HJ, Folkertsma RT, Hutten RC, Draaistra J, Gommers FJ, Jacobsen E, Helder J, Bakker J (1997) Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. Mol Gen Genet 255:438–447

Ruíz de Galarreta JI, Carrasco A, Salazar A, Barrena I, Iturritxa E, Marquinez E, Legorburu FJ, Ritter E (1998) Wild *Solanum* species as resistance sources against different pathogens of potato. Potato Res 41:57–68

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press, New York

Samuelian S, Kleine M, Ruyter-Spira CP, Klein-Lankhorst RM, Jung C (2004) Cloning and functional analyses of a gene from sugar beet up-regulated upon cyst nematode infection. Plant Mol Biol 54:147–156

Schmidt R, West J, Love K, Lenehan Z, Lister C, Thompson H, Bouchez D, Dean C (1995) Physical map and organization of *Arabidopsis thaliana* chromosome 4. Science 270:480–483

Tanksley SD, Ganal MW, Prince JP, de Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141–1160

Temesgen B, Brown GR, Harry DE, Kinlaw CS, Sewell MM, Neale DB (2001) Genetic mapping of expressed sequence tag polymorphism (ESTP) markers in loblolly pine (*Pinus taeda*). Theor Appl Genet 102:664–675

Trindade LM, Horvath B, Bachem C, Jacobsen E, Visser RG (2003) Isolation and functional characterization of a stolon specific promoter from potato (*Solanum tuberosum* L.). Gene 303:77–87

van der Biezen EA, Juwana H, Parker JE, Jones JDG (2000) cDNA–AFLP display for the isolation of *Pernospora parasitica* genes expressed during infection in *Arabidopsis thaliana*. Mol Plant Microbe Interact 13:895–898

van Eck HJ, Rouppe van der Voort J, Draaistra J, van Zandvoort P, van Enckevort E, Segers B, Peleman J, Jacobsen E, Helder J, Bakker J (1995) The inheritance and chromosomal location of AFLP markers in a non-inbred potato offspring. Mol Breed 1:397–410

van Os H, Stam P, Visser RG, van Eck HJ (2005) SMOOTH: a statistical method for successful removal of genotyping errors from high-density genetic linkage data. Theor Appl Genet 112:187–194

van Os H, Andrzejewski S, Bakker E, Barrena I, Bryan GJ, Caromel B, Ghareeb B, Isidore E, de Jong W, van Koert P, Lefebvre V, Milbourne D, Ritter E, Rouppe van der Voort JNAM, Rousselle-Bourgeois F, van Vliet J, Waugh R, Visser RGF, Bakker J, van Eck HJ (2006) Construction of a 10,000-marker ultradense genetic recombination map of potato: providing a framework for accelerated gene isolation and a genomewide physical map. Genetics 173:1075–1087