ORIGINAL ARTICLE

The utility of NBS profiling for plant systematics: a first study in tuber-bearing *Solanum* species

Miqia Wang · Ronald Van den Berg · Gerard Van der Linden · Ben Vosman

Abstract Systematic relationships are important criteria for researchers and breeders to select materials. We evaluated a novel molecular technique, nucleotide binding site (NBS) profiling, for its potential in phylogeny reconstruction. NBS profiling produces multiple markers in resistance genes and their analogs (RGAs). Potato (*Solanum tuberosum* L.) is a crop with a large secondary genepool, which contains many important traits that can be exploited in breeding programs. In this study we used a set of over 100 genebank accessions, representing 49 tuber-bearing wild and cultivated *Solanum* species. NBS profiling was compared to amplified fragment length polymorphism (AFLP). Cladistic and phenetic analyses showed that the two techniques had similar resolving power and delivered trees with a similar topology. However, the different statistical tests used to demonstrate congruency of the trees were inconclusive. Visual inspection of the trees showed that, especially at the lower level, many accessions grouped together in the same way in both trees; at the higher level, when looking at the more basal nodes, only a few groups were well supported. Again this was similar for both techniques. The observation that higher level groups were poorly supported might be due to the nature of the material and the way the species evolved. The similarity of the NBS and AFLP results indicate that the role of disease resistance in speciation is limited.

Keywords Potato · Phylogeny · Resistance gene analogs · *Solanum* · Speciation · NBS profiling

M. Wang · G. Van der Linden · B. Vosman
Plant Research International B.V. Wageningen University and Research Centre, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

M. Wang
Oil Crop Research Institute, Chinese Academy of Agricultural Sciences, 430062 Wuhan, Hubei, China

R. Van den Berg (✉)
Biosystematics Group, Wageningen University and Research Centre, Generaal Foulkesweg 37, 6703BL 700 ED Wageningen, The Netherlands
e-mail: ronald.vandenberg@wur.nl

Introduction

Recently, a novel molecular technique called Nucleotide Binding Site (NBS) profiling was developed (Van der Linden et al. 2004). This technique specifically targets resistance genes and their analogs. Resistance (R) genes containing an NBS are numerous in plants and are distributed over all chromosomes (Meyers et al. 2002, 2003; Monosi et al. 2004). The technique is based on amplification of DNA fragments starting from the conserved NBS domain towards an adaptor which is ligated to a restriction fragment. Primers based upon several conserved motifs (P loop, the kinase-2 motif, and the GLPL motif) within the NBS domain can be used as a starting point. As a consequence this technique produces gene targeted markers, while other marker techniques like AFLP produce markers randomly in the genome. NBS profiling generates a reproducible polymorphic multi-locus banding pattern and has already been successfully used to identify and map RGAs in potato, apple and lettuce (Van der Linden et al. 2004; Calenge et al. 2005; Syed et al. 2006).

Nucleotide Binding Site profiling was used recently to assess genetic diversity in accessions of durum wheat that are commonly used in breeding programs (Mantovani et al. 2007). Results were compared to AFLP and SSR analyses. In

general, the groups found with all three marker systems were similar. In this paper we describe the use of NBS profiling for phylogeny reconstruction and classification. As a model we used part of the section *Petota* of the genus *Solanum*, which also includes the cultivated potato. The potato (*Solanum tuberosum* L.) is a crop with a large secondary gene pool, which contains important traits that can be exploited in breeding programs. In the last few years the identification and cloning of late blight (*Phytophthora infestans*) resistance genes from wild relatives of the cultivated potato has been the subject of many studies (Ballvora et al. 2002; Song et al. 2003; Van der Vossen et al. 2003, 2005; Huang et al. 2005; Park et al. 2005). Systematic relationships within the group of tuber-bearing *Solanum* species are regarded as important criteria to select interesting materials. Based on morphological data, relationships among tuber-bearing *Solanum* species have been studied extensively resulting in a classification of 227 species in 19 series (Hawkes 1990). Spooner and Hijmans (2001) updated this by reducing the number of tuber-bearing species to 206. Recently, Spooner and Salas (2006) reduced the number to 188 wild and one cultivated species for section *Petota*, plus three species in section *Etuberosa*. Apparently, the boundaries between some of the species in this group are not very clear.
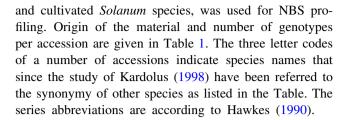
Relationships within the tuber-bearing *Solanum* species have been studied using different molecular markers resulting in new insights at different levels of potato taxonomy. These studies include Restriction Fragment Length Polymorphism (RFLP) markers of chloroplast DNA (cpDNA) (Hosaka et al. 1984; Sukhotu et al. 2004), RFLPs of the nuclear genome (Debener et al. 1990; Miller and Spooner 1999), Amplified Fragment Length Polymorphism (Kardolus et al. 1998), Simple Sequence Repeat (SSRs) (Raker and Spooner 2002), cpDNA SSRs (Bryan et al. 1999), sequence data of 5S ribosomal DNA (rDNA) (Volkov et al. 2001), and external transcribed spacer (ETS) (Volkov et al. 2003).

To explore the prospects of NBS profiling for systematic research we used previously obtained AFLP data (Kardolus 1998) as reference. As NBS profiling targets resistance genes and their analogs, we address the question whether this affects the outcome of the systematic analysis by comparing results from NBS profiling to results obtained from AFLP. The role of disease resistance in the evolution of species is discussed.

Materials and methods

Plant material

DNA material of one hundred and three accessions (Table 1) from Kardolus (1998), representing 49 wild

and cultivated *Solanum* species, was used for NBS profiling. Origin of the material and number of genotypes per accession are given in Table 1. The three letter codes of a number of accessions indicate species names that since the study of Kardolus (1998) have been referred to the synonymy of other species as listed in the Table. The series abbreviations are according to Hawkes (1990).

NBS profiling procedure

Nucleotide binding site profiling was performed essentially as described by Van der Linden et al. (2004) with two minor modifications. Firstly, 200 instead of 400 ng of DNA was digested with a restriction enzyme. Secondly, in the study of Van der Linden et al. (2004), the restriction and adaptor ligation reactions were done separately, while in our study, the two were combined into one single reaction of 60 μl, consisting of 200 ng of DNA, 3 μl of adaptor (from a 75 μl reaction mix with 1.25 nmol of each adapter strand, heated to 90°C and allowed to cool down to room temperature), 1 mM ATP, 10U restriction enzyme, 5U T4 ligase and restriction/ligation buffer (10 mM Tris.HAc pH7.5, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 50 ng/μl BSA). Afterwards, the restriction/ligation product was two times diluted with MilliQ water and used as template for PCR reactions. The 25 μl PCR reaction mix consisted of 5 μl template DNA, 20 pmol of both adaptor primer and NBS primer, 200 μM dNTPs, 0.4 U HotStarTaq (Qiagen, Germany), and 2.5 μl HotStarTaq PCR buffer. The PCR program consisted of 30 cycles of 30 s at 95°C, 1 min 40 s at 55–60°C annealing, and 2 min at 72°C. Annealing temperature was 55°C for NBS5 and NBS9 primers and 60°C for NBS2 and NBS3 primers. After this, a second PCR reaction was performed using 5 μl of the 10 times diluted first PCR product as template, and the same NBS-specific primer but now ^{33}P radioactively labeled, and cycling conditions similar to the first PCR. Labeled products were separated on 6% polyacrylamide gels. X-ray films were exposed to the gels to visualize individual fragments. The presence or absence of polymorphic fragments was scored on the autoradiograms and transferred into a 1 (present) and 0 (absent) binary matrix for all accessions.

In total, four primers (NBS2, NBS3, NBS5 and NBS 9) and three enzymes (*MseI*, *RsaI* and *AluI*) were used (Table 2). NBS2, NBS5 and adaptor sequences were the same as described by Van der Linden et al. (2004). NBS3 and NBS9 primer sequences were as follows: NBS3 5′-GTWGTYTTICCYRAICCIGGCATICC-3′ and NBS9 5′-TGTGGAGGRTTACCTCTAGC-3′. Positions of the NBS primers in the NBS domain are shown in Fig. 1.

Table 1 Accessions used for phylogeny reconstruction

Code[a]	Series	Species	Origin	Source[b]	Chr no	# GT[c]
ach1	TUB	*S. achacachense* Cárdenas	Bolivia	B29617	24	10
acl7	ACA	*S. acaule* Bitter ssp. *acaule*	Bolivia	B28026	48	10
acl10	ACA	*S. acaule* ssp. *acaule*	Bolivia	B27206	48	10
acl12	ACA	*S. acaule* ssp. *acaule*	Bolivia	B27361	48	10
acl13	ACA	*S. acaule* ssp. *acaule*	Argentina	B16835	48	10
acl18	ACA	*S. acaule* ssp. *acaule*	Argentina	B17111	48	10
acl26	ACA	*S. acaule* ssp. *acaule*	Argentina	B47627	48	10
acl27	ACA	*S. acaule* ssp. *acaule*	Argentina	B17181	48	10
adg1	TUB	*S. tuberosum* L. ssp. *andigena* Hawkes	–	B7462	48	10
adg2	TUB	*S. tuberosum* ssp. *andigena*	–	B24677	48	9
aem1	ACA	*S. acaule* ssp. *aemulans* Hawkes et Hjert.	Argentina	B17129	48	10
aem5	ACA	*S. acaule* ssp. *aemulans*	Argentina	PI320280	48	10
ajh1	TUB	*S. ajanhuiri* Juz. et Bukasov	–	CIP702677	24	10
alb2	ACA	*S. albicans* Ochoa	Peru	CIP761438	72	1
alb3	ACA	*S. albicans*	Peru	PI365376	72	10
alb5	ACA	*S. albicans*	Peru	PI498194	72	10
bcp1	DEM	*S. brachycarpum* Corr.	Mexico	B8100	72	10
ber1	TUB	*S. berthaultii* Hawkes	Bolivia	B28009	24	9
ber2	TUB	*S. berthaultii*	Bolivia	B24578	24	10
ber3	TUB	*S. berthaultii*	Bolivia	B10063	24	10
blb	BUL	*S. bulbocastanum* Dunal	Mexico	B8009	24	3
blv3	MEG	*S. boliviense* Dunal	Bolivia	B27342	24	10
brc1	TUB	*S. brevicaule* Bitter	Bolivia	B18291	24	10
brc2	TUB	*S. brevicaule*	Bolivia	B28023	24	9
brd2	ETU	*S. brevidens* Phil.	Argentina	B17441	24	10
bst1	PIN	*S. brachistotrichium* (Bitter) Rydb.	Mexico	B7986	24	10
bst2	PIN	*S. brachistotrichium*	Mexico	B7987	24	10
buk1	TUB	*S. bukasovii* Juz.	Peru	B15424	24	10
buk2	TUB	*S. bukasovii*	Peru	B18294	24	10
can1	TUB	*S. canasense* Hawkes	Peru	B8105	24	10
can2	TUB	*S. canasense*	Peru	B7162	24	8
can3	TUB	*S. canasense*	Peru	B8012	24	9
cha1	TUB	*S. chaucha* Juz. et Bukasov	–	CIP 700145	36	1
cha4	TUB	*S. chaucha*	–	CIP701568	36	1
chc1	YNG	*S. chacoense* Bitter ssp. *chacoense*	Argentina	B17034	24	10
chc2	YNG	*S. chacoense* ssp. *chacoense*	Argentina	B17018	24	7
cop1	TUB	*S. coelestipetalum* Vargas	Peru	B7942	24	7
cop3	TUB	*S. coelestipetalum*	Peru	B7994	24	10
crc	CIR	*S. circaeifolium* Bitter ssp. *circaeifolium*	Bolivia	B27058	24	10
dms1	DEM	*S. demissum* Lindl.	Mexico	B10030	72	10
dms2	DEM	*S. demissum*	Mexico	B10022	72	10
dms3	DEM	*S. demissum*	Mexico	B9990	72	10
etb1	ETU	*S. etuberosum* Lindl.	Chili	B28476	24	9
etb2	ETU	*S. etuberosum*	Chili	B8082	24	10
fen	LON	*S. fendleri* A. Gray ssp. *fendleri*	Mexico	B7230	48	9
gnd	TUB	*S. gandarillasii* Cárdenas	Bolivia	B7174	24	9
grl2	TUB	*S. gourlayi* Hawkes ssp. *gourlayi*	Argentina	B17338	24	9
grl4	TUB	*S. gourlayi* ssp. *gourlayi*	Argentina	B16837	48	10

Table 1 continued

Code[a]	Series	Species	Origin	Source[b]	Chr no	# GT[c]
hje	LON	*S. hjertingii* Hawkes	Mexico	B8088	48	9
ifd1	CUN	*S. infundibuliforme* Phil.	Argentina	B17212	24	4
juz3	TUB	*S. juzepczukii* Bukasov	–	CIP701895	36	1
ktz1	TUB	*S. kurtzianum* Bitter et Wittm.	Argentina	B17585	24	10
ktz2	TUB	*S. kurtzianum*	Argentina	B16861	24	10
ktz3	TUB	*S. kurtzianum*	Argentina	B17580	24	10
les	POL	*S. lesteri* Hawkes et Hjert.	Mexico	B55219	24	10
lgl	LIG	*S. lignicaule* Vargas	Peru	B8106	24	12
lph1	TUB	*S. leptophyes* Bitter	Argentina	B7184	24	8
lph2	TUB	*S. leptophyes*	Bolivia	B27176	24	10
lph3	TUB	*S. leptophyes*	Bolivia	B27211	24	10
mcd1	TUB	*S. microdontum* Bitter	Bolivia	B31189	24	10
mcd2	TUB	*S. microdontum*	Argentina	B24649	24	10
mcq1	TUB	*S. mochiquense* Ochoa	Peru	B32672	24	8
mcq2	TUB	*S. mochiquense*	Peru	B8142	24	10
mga3	MEG	*S. megistacrolobum* Bitter ssp. *megistacrolobum*	Argentina	B17642	24	10
mgl	MGL	*S. maglia* Schlechtd.	Chili	B23571	24	10
mlt1	TUB	*S. multidissectum* Hawkes	Peru	B8145	24	10
opl1	TUB	*S. oplocense* Hawkes	Argentina	B16868	72	10
opl2	TUB	*S. oplocense*	Argentina	B24650	72	10
opl3	TUB	*S. oplocense*	Argentina	B16879	72	10
oxc	CON	*S. oxycarpum* Schiede	Mexico	B53011	48	10
pcs1	PIU	*S. paucissectum* Ochoa	Peru	B8162	24	6
pcs2	PIU	*S. paucissectum*	Peru	B55216	24	10
phu1	TUB	*S. phureja* Juz. et Bukasov	–	B15482	24	10
phu2	TUB	*S. phureja*	–	B50199	24	10
pne1	ACA	*S. acaule* ssp. *punae* Hawkes et Hjert.	Peru	PI365312	48	10
pne2	ACA	*S. acaule* ssp. *punae*	Peru	B7958	48	10
pne4	ACA	*S. acaule* ssp. *punae*	Peru	PI473442	48	10
pnt1	PIN	*S. pinnatisectum* Dunal	Mexico	B8168	24	10
qum	CIR	*S. circaeifolium* ssp. *quimense* Hawkes et Hjert.	Bolivia	B27034	24	10
rap1	MEG	*S. raphanifolium* Cárdenas et Hawkes	Peru	B15445	24	10
rap3	MEG	*S. raphanifolium*	Peru	B7207	24	10
sct1	MEG	*S. sanctae-rosae* Hawkes	Argentina	B15454	24	10
sct2	MEG	*S. sanctae-rosae*	Argentina	B17568	24	7
sct3	MEG	*S. sanctae-rosae*	Argentina	B17051	24	10
spg2	TUB	*S. spegazzinii* Bitter	Argentina	B24694	24	8
spg3	TUB	*S. spegazzinii*	Argentina	B16905	24	10
spl1	TUB	*S. sparsipilum* (Bitter) Juz. et Bukasov	Bolivia	B8209	24	10
spl2	TUB	*S. sparsipilum*	Bolivia	B8150	24	10
spl3	TUB	*S. sparsipilum*	Bolivia	B15455	24	10
stn1	TUB	*S. stenotomum* Juz. et Bukasov ssp. *stenotomum*	Bolivia	B27165	24	10
stn2	TUB	*S. stenotomum* ssp. *goniocalyx* Hawkes	Peru	B7478	24	10
sto	LON	*S. stoloniferum* Schlechtd. et Bche.	Mexico	B7229	48	10
tar1	YNG	*S. tarijense* Hawkes	Argentina	B17423	24	10
tar2	YNG	*S. tarijense*	Argentina	B8229	24	6
tbr1	TUB	*S. tuberosum* L. ssp. *tuberosum*	–	'Certa' × 'Gloria'	48	1
ver1	TUB	*S. verrucosum* Schlechtd.	Mexico	B8255	24	10

Table 1 continued

Code[a]	Series	Species	Origin	Source[b]	Chr no	# GT[c]
ver2	TUB	*S. verrucosum*	Mexico	B8246	24	7
ver3	TUB	*S. verrucosum*	Mexico	B8254	24	4
vid1	TUB	*S. gourlayi* ssp. *vidaurrei* Hawkes et Hjert.	Argentina	B16831	24	10
vid2	TUB	*S. gourlayi* ssp. *vidaurrei*	Argentina	B18528	24	9
vrn1	TUB	*S. vernei* Bitter et Wittm.	Argentina	B15451	24	8
vrn2	TUB	*S. vernei* ssp. *ballsii* Hawkes et Hjert.	Argentina	B17536	24	10
vrn3	TUB	*S. vernei* ssp. *vernei* Bitter et Wittm.	Argentina	B17542	24	10

[a] Series and taxon code abbreviations according to Hawkes (1990)

[b] Accessions with a prefix of B were obtained from the Braunschweig Genetic Resources Collection; Accessions with prefix CIP were obtained from the International Potato Centre, Peru; Accessions with prefix PI showed the Plant Introduction Number after Bamberg et al. (1996)

[c] Number of genotypes used to represent the accession in the DNA sample
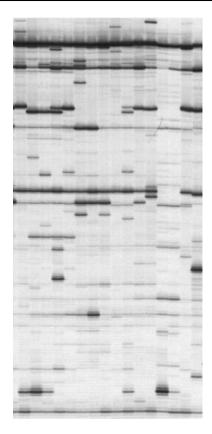
Table 2 Number of polymorphic NBS markers for different primer/enzyme combinations

Enzyme/Primer	NBS5	NBS9	NBS2	NBS3
MseI	32	36	66	37
RsaI	31	23	29	ND
AluI	37	40	ND	31
RGA content (%)	94	57	63	61

Bottom row shows RGA content for the four primers

ND no data

Fig. 1 Schematic representation of the nucleotide-binding site (NBS) of disease resistance genes and positions of primers used in this study. P-loop, kinase-2 and GLPL refer to highly conserved motifs in the NBS. Primer positions are indicated by *horizontal arrows*

Annotation of NBS profiling sequences

To determine the RGA content of NBS profiling markers, 384 bands were randomly chosen and excised from the gel, re-amplified with PCR conditions identical to the first PCR of the NBS profiling protocol and purified with Qiaquick PCR purification spin columns (Qiagen). After inspection on agarose gels fragments were directly sequenced using the adaptor primer as a sequencing primer with the BigDye Terminatior kit and an ABI 3700 automated sequencer from Applied Biosystems (USA). Sequences of poor quality were excluded from further analysis, and the remaining sequences were compared to archived sequences in the NCBI nucleotide databases (11 Jan 2006) using

BLASTN and BLASTX (Altschul et al. 1997). Sequences that were significantly similar to known R genes and RGAs (either E value lower than 1e-05 for BLASTN, or E value lower than 1e-03 for BLASTX) were regarded as RGAs.

NBS profiling and AFLP data analysis

AFLP data from three primer combinations were available from Kardolus (1998). Two AFLP datasets (from the primer combinations E32M49 and E35M48) and one NBS profiling dataset were used in this study. Neighbor joining (NJ) and parsimony analyses were performed with PAUP* version 4.0 b10 (Swofford 2001). For the parsimony analysis, we used a two-step search strategy. (1) The first heuristic search was conducted with 10,000 random additions, holding ten trees, and saving five trees per search replicate. (2) Resulting trees from the first search were used as starting trees to swap to completion with MULTREES and TBR. A jackknife analysis (10,000 replicates) was performed with the same settings as the heuristic search. In the parsimony analyses, three accessions from the series *Etuberosa* Juz. were used as outgroups.

Congruence between AFLP and NBS profiling was assessed in three ways: (1) visual qualitative comparisons of the trees obtained, (2) distance matrix-based comparisons, and (3) character-based comparisons.

For the distance matrix-based comparisons, the program SIMQUAL was used to compute similarity matrices using the DICE option, which ignores shared absent bands, and is an appropriate algorithm for dominant markers like AFLP or NBS profiling markers. A pairwise comparison for these matrices with the Mantel test in NTSYS-pc version 2.10j (Rohlf 1992) was made. For the character-based comparisons, the partition homogeneity test (Farris et al. 1995) was performed in PAUP* version 4.0 b10 (Swofford 2001). This test is also called the incongruence length difference

Fig. 2 Part of an NBS profiling gel obtained with primer NBS9 and restriction enzyme *MseI*

(ILD) test. ILD tests were performed with 100 replicates, the heuristic search option, and TBR and MULPARS in effect.

Results

NBS profiling on *Solanum* species

Nucleotide binding site profiling produced well scorable banding patterns on a gel; part of such a gel is shown in Fig. 2. For the whole set of 103 accessions, ten primer/

enzyme combinations were tested. Each combination produced 23–66 scorable polymorphic bands (Table 2). A dataset of 362 characters was produced.

To obtain information on the RGA content of the bands that make up the NBS profiles, 384 bands were excised, reamplified and sequenced. Of these, 232 produced a readable sequences and 155 bands (67%) could be annotated as RGA, indicating that NBS profiles indeed largely consisted of markers derived from R-genes and their analogs. The six bands with the highest similarity to known resistance genes are listed in Table 3. Most of the remaining fragments (33%) could not be annotated. Only in a few cases significant homologies to other known genes, like retrotransposons, were found. Table 2 also shows the percentage of RGA-derived bands for each primer. The RGA content of the bands obtained with the NBS5 primer was 94%. For NBS9, NBS2 and NBS3, these figures were 57, 63 and 61%, respectively.

Cladistic results

The NBS profiling dataset was composed of 362 characters, 351 of which were parsimony informative. Sixteen most parsimonious 3,553-step trees were produced, with a consistency index of 0.102 and retention index of 0.464. The strict consensus tree obtained is shown in Fig. 3. Numbers above the branches indicated jackknife values >50%.
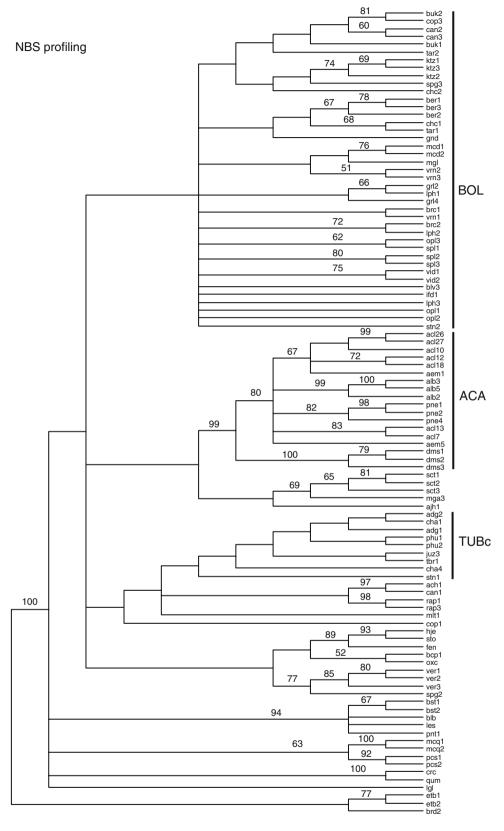
The three non-tuber-bearing *Solanum* species (outgroup species: brd2, etb1 and etb2) were basal to the tree. Next came a polytomy consisting of (1) the diploid Mexican species (bst1, bst2, blb, les, pnt1), (2) diploid South American species (mcq1, mcq2, pcs1, pcs2), (3) the representatives of series *Circaeifolia* (crc and qum), (4) the species *S. lignicaule*, and (5) all other species. Within the latter group four clades could be distinguished but none of them had statistical support: a group of Mexican polyploids including the diploid *S. verrucosum*, a group consisting of taxa from series *Acaulia*, including *S. demissum* from series *Demissa*, with a

Table 3 NBS profiling bands with high similarity to known resistant (R) genes and R gene cluster members

Band Code	Primer	Homologue	Identity (number of nucleotides)	E value for BLASTN
65	nbs2	*Solanum tuberosum* potato resistance-like protein I2GA-SH23-3	90% (324)	5.00E-82
1	nbs2	*Lycopersicon esculentum* BAC clone Clemson_Id 127E11	92% (455)	2.00E-162
g0739 mq21	nbs5	*Solanum bulbocastanum* Rpi-blb2 gene	97% (281)	8.00E-137
g0732 mq14	nbs5	*Solanum demissum* chromosome 11 clone PGEC591C22 map MAP_LOC	96% (457)	0
G14	nbs9	*Lycopersicon esculentum* Tm-2 ToMV resistant protein (Tm-2nv) gene	92% (276)	7.00E-97
18c-7	nbs9	*Solanum acaule* Rx2.ac15 gene	97% (223)	1.00E-93

Fig. 3 NBS profiling strict
consensus tree. *Numbers above
the branches* are jackknife
supports based on 10,000
replicates (%). Jackknife
support values lower than 50%
are not shown. ACA and TUB
are abbreviations for series
Acaulia and *Tuberosa* as
outlined by Hawkes (1990).
Clade ACA includes a group of
taxa from series *Acaulia*
together with *S. demissum* from
series *Demissa*. TUBc indicates
the cultivated species from
series *Tuberosa*. BOL includes
accessions from series *Tuberosa*
from Bolivia, Argentina and
Chile

number of species of series *Megistacroloba* as a sister
clade, and two groups with mainly species from series
Tuberosa, divided in a Bolivian and Peruvian group, the
latter also containing a number of cultivated species.
Only the *Acaulia* group had a high jackknife support
of 99%.

The AFLP datasets resulting from the two chosen primer combinations were found to be congruent with each other both for the Mantel test and ILD test (data not shown). For this reason we combined the two AFLP datasets for the comparison with NBS profiling.

The AFLP dataset consisted of 591 characters, among which 539 were informative. Parsimony analysis of AFLP data yielded eight equally parsimonious 1,548-step trees with a consistency index of 0.169 and retention index of 0.542. The strict consensus tree is shown in Fig. 4.

The outgroup species are again at the base of the tree, followed by the Mexican diploids, the taxa of series *Circaeifolia*, and relatively primitive South American species. The remainder of the tree was subdivided into three clades: a group of Peruvian species of series *Tuberosa* with the

accessions of the cultivated species, a group of taxa from series *Acaulia*, including *S. demissum*, forming a polytomy with a number of species from series *Megistacroloba*, and a large polytomy including mainly the Bolivian representatives of series *Tuberosa*, with the Mexican polyploids nested within.

In both the NBS profiling and AFLP based strict consensus trees, the more basal branches either had poor jackknife support (lower than 50%) or form a polytomy. To investigate the influence of polyploid taxa on the poor resolution of the trees, we excluded all polyploids from the analysis. Two new datasets, containing the marker data of 72 accessions from the diploid species were produced. Parsimony and jackknife analysis were performed in the same way as with the complete dataset. Excluding the

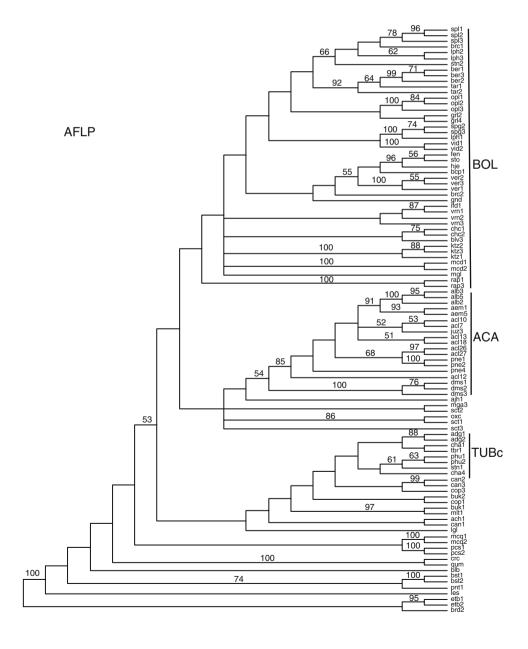**Fig. 4** AFLP strict consensus tree. For details see Fig. 3

Table 4 Common groups recognized both from AFLP and NBS profiling tree

Group	Common groups	NBS profiling jacknife (%)	AFLP jackknife (%)
1	acl 26 acl27	99	97
2	alb2 3 5	99	100
3	alb3 alb5	100	95
4	ber1 ber3	78	71
5	crc qum	100	100
6	dms1 dms2	79	76
7	dms1 2 3	100	100
8	etb1 etb2	77	95
9	hje sto fen	89	96
10	ktz1 2 3	74	100
11	mcd1 mcd2	76	100
12	mcq1 mcq2	100	100
13	pcs1 pcs2	92	100
14	pne1 pne2	98	100
15	rap1 rap3	98	100
16	ver1 2 3	85	100
17	vid1 vid2	75	100

polyploid species did not improve the trees, the polytomies and basal nodes with jackknife support lower than 50% were still present.

At the lower (accession) level, common well supported subclades can be found in both the NBS profiling and AFLP strict consensus trees. In total, 17 subclades (containing two or three accessions) with a jackknife support higher than 70% were observed (Table 4).

Phenetic results

NBS profiling and AFLP phenograms had a very high cophenetic correlation coefficient of 0.96 and 0.93, indicating excellent fits of the similarity matrices to the resulting phenograms. The NJ trees of both NBS profiling and AFLP data basically showed the same groups as the cladograms. Except for the branch leading to the outgroups many of the interconnecting branches were very short and jackknife support was low. Average similarities for AFLP and NBS profiling were 0.29 and 0.52, respectively.

Congruence between AFLP and NBS profiling derived data

As described above, a visual inspection of the trees produced from NBS profiling and AFLP data showed that they were similar in many aspects, although many details were different. The correlation coefficient between NBS profiling and AFLP similarity matrices was 0.737, and between

the cophenetic value matrices derived from the NBS profiling and AFLP trees 0.762. Mantel test showed that both were significantly correlated at $\alpha = 0.05$ ($P = 0.001$), indicating both NBS profiling and AFLP derived similarity matrices and tree topologies are congruent.

In contrast to the Mantel test, the ILD test produced different results. P value of 0.01 was observed, showing that NBS profiling and AFLP were incongruent with each other ($\alpha = 0.05$).

Discussion

NBS profiling in *Solanum*

When introducing a new tool for systematic analysis one-first needs to compare the results from the new tool with results from established techniques, like AFLP (Vos et al. 1995; Kardolus et al. 1998). Our study aimed at such comparison. We first investigated whether NBS profiling produces markers derived from resistance genes and their analogs with a high frequency. Table 2 clearly shows that the majority of the bands were derived from resistance genes and their analogs. The primers giving the lower RGA content (NBS2, NBS3 and NBS9) generate fragments outside of the conserved NBS domain. Annotation of the fragments as putative RGAs relies on sequence similarity with other RGAs. Within the NBS domain, sequence conservation is high (hence the high RGA content within the set of fragments generated with NBS5 from within the NBS domain), whereas outside of the NBS domain, sequence conservation between different RGAs is much lower. Therefore, the observed RGA content is most likely an underestimation. This was supported by the observation that some bands could not be annotated during our first analysis in August 2004, while the second analysis, with an updated version of the database (in January 2006), yielded a positive identification for several previously unidentifiable bands. Also, the annotation became more accurate. An example was the band g0739 mq21 obtained with the primer NBS5 in the *S. bulbocastanum* BGRC 8009 (Table 3). The first analysis showed that this band had the highest similarity with the Mi-gene of tomato. During the second annotation round, this fragment had highest similarity (97%) with the late blight resistance gene rpi-blb2 (van der Vossen et al. 2005).

Systematic relationships inferred from NBS profiling and AFLP

The topology of the cladogram from our AFLP analysis is comparable to Kardolus (1998), which is not surprising since we studied a subset of his material (because not all of

the DNA was still available). Kardolus (1998) distinguished more groups in his trees but also in his results the Mexican diploid species and primitive South American species like *S. circaeifolium* and *S. mochiquense* are placed in a basal group. Furthermore, a group of representatives of series *Acaulia* can be recognized, closely related to the Mexican hexaploid species *S. demissum*, and the species belonging to the series *Tuberosa* are separated in a Peruvian (including the cultivated material) and a Bolivian group. He only showed bootstrap support values for the groups in his NJ tree, and generally the values for deeper branches are lower than 50%. The pattern from our analyses of the NBS profiling dataset is similar to that of Kardolus (1998).

This general pattern in our AFLP and NBS profiling results supplements the results of the most extensive cpDNA RFLP study of section *Petota* (Spooner and Castillo 1997). They recovered four clades, with representatives of the Mexican diploids in clade 1 and 2, species like *S. mochiquense* and *S. paucissectum* in clade 3, and a large polytomy of the other investigated species in clade 4. Spooner et al. (2005a) performed phylogenetic analyses of AFLP data of 362 individual wild (261) and landrace (98) members of section *Petota*. The strict consensus tree also has bootstrap values lower than 50% for the deeper branches (except the designated outgroup). In rDNA ETS results (Volkov et al. 2003) only three structural variants were found, with variant A present in the non-tuber-bearing species of series *Etuberosa* and in the representatives of the Mexican diploid series, variant B in series *Circaeifolia*, and variant C in all other investigated species. The dendrograms presented show many polytomies, indicating that resolution within the groups is mostly lacking.

When comparing the cladograms from AFLP and NBS profiling data (Figs. 3, 4) visually, there were congruencies at the lower level with accessions grouping together in the same way in both trees. These groups were supported with jackknife values >70% (Table 4). When looking at the more basal nodes, we found only a few groups (P and ACA) which were well supported by relatively high jackknife values. A number of other groups were present, but they lacked statistical support. Again this was similar for both trees. Two different statistical tests were used to evaluate the congruency between the AFLP and NBS profiling based trees. The outcome of the test was different. Whereas congruency was indicated by the Mantel test, the ILD test indicated incongruency. Similar observations were made by Spooner et al. (2005b), who compared AFLPs with other markers for phylogenetic inference in wild tomatoes. In their Mantel test, the comparison of cpDNA/GBSSI (granule-bound starch synthase gene) gives a high matrix correlation coefficient (0.831), but fails to pass the ILD test. The suitability of the ILD test is also questioned

by other researchers (Graham et al. 1998; Yoder et al. 2001; Barker and Lutzoni 2002; Darlu and Lecointre 2002). In view of problems with the ILD test and the significant correlation found with the Mantel test we conclude that there is congruency between the AFLP and NBS profiling derived tree topologies.

The interesting observation that the overall pairwise similarity based on NBS profiling markers was 25% higher than the overall similarity based on AFLP markers, suggests that NBS profiling markers are more conserved than AFLP markers. Targeting more conserved regions renders the marker system more appropriate than AFLP when materials are more diverse, as it will reduce the chance for homoplasy.

The fact that the basal branches either had poor jackknife support (lower than 50%) or formed a polytomy in both the AFLP and NBS profiling tree, probably results from the nature of the studied material. Apparently, *Solanum* species are more similar to each other than expected on the basis of morphological characters, even when the polyploid species were excluded from the analysis. Extensive hybridization and introgression among species might be the reason for the poor resolution at the more basal nodes. An alternative explanation could be that many *Solanum* species have evolved in a relatively short period of time after rapid radiation over South America, which would result in species with distinct characters but with no clear sequential branching order (apart from originating from the common ancestor).

The role of disease resistance in speciation

Many NBS profiling markers were shown to be RGA-related (Table 2). NBS profiling preferentially generates markers in resistance genes. These genes are likely to be under selection, which might influence the outcome of the phylogenetic analysis. Disease resistance might play a role in the speciation process and it may thus be challenged whether NBS profiling can be used for phylogeny reconstruction. Our results demonstrated that systematic relationships from NBS profiling data do not essentially differ from that from AFLP data. This congruence between AFLP and NBS profiling may not be as unexpected as it appears to be. Plants have to deal with many different pathogens during their lifetime and are thus exposed to selective pressures in different directions. This is also evidenced by the large number of NBS-LRR resistance genes present in plants. However, it is possible that a single resistance gene was essential for species survival, or a speciation event. The R-gene(s) that may have been under selection after a period of disease pressure by a pathogen may have spread relatively fast in neighboring species through hybridization. The specific effect of selective pressure on R-genes will therefore be only detectable on a very short evolutionary time scale, and is diluted when many markers

are analyzed phylogenetically. In addition, the selective advantage of retaining a specific R-gene will most likely be reflected by the absence or presence of one or a few markers in NBS profiling. The outcome of the phylogenetic analysis is based on all NBS profiling markers, the majority of which was not affected by selective pressure. A single or a few markers very likely will not influence the outcome.

In conclusion, NBS profiling is at least as good as AFLP for phylogeny reconstruction and might even be superior when more diverse material is used, as it will reduce the chance for homoplasy. The observation that higher level groups were poorly supported might be due to the nature of the material and the way the species evolved.

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

Ballvora A, Ercolano MR, Weiss J, Meksem K, Bormann CA, Oberhagemann P, Salamini F, Gebhardt C (2002) The R1 gene for potato resistance to late blight (*Phytophthora infestans*) belongs to the leucine zipper/NBS/LRR class of plant resistance genes. Pl J 30:361–371

Bamberg JB, Martin MW, Schartner JJ, Spooner DM (1996) Inventory of tuber-bearing *Solanum* species: catalog of potato germplasm. Potato Introduction Station, NRSP-6, Sturgeon Bay, Wisconsin

Barker FK, Lutzoni FM (2002) The utility of the incongruence length difference test. Syst Biol 51:625–637

Bryan GJ, McNicoll J, Ramsay G, Meyer RC, De Jong WS (1999) Polymorphic simple sequence repeat markers in chloroplast genomes of Solanaceous plants. Theor Appl Genet 99:859–867

Calenge F, Van der Linden CG, Van de Weg E, Schouten HJ, Van Arkel G, Denance C, Durel CE (2005) Resistance gene analogues identified through the NBS-profiling method map close to major genes and QTL for disease resistance in apple. Theor Appl Genet 110:660–668

Darlu P, Lecointre G (2002) When does the incongruence length difference test fail? Molec Biol Evol 19:432–437

Debener T, Salamini F, Gebhardt C (1990) Phylogeny of wild and cultivated *Solanum* species based on nuclear restriction fragment length polymorphisms (RFLPs). Theor Appl Genet 79:360–368

Farris JS, Källersjö M, Kluge AG, Bult C (1995) Testing significance of incongruence. Cladistics 10:315–319

Graham SW, Kohn JR, Morton BR, Eckenwalder JE, Barrett SCH (1998) Phylogenetic congruence and discordance among one morphological and three molecular data sets from Pontederiaceae. Syst Biol 47:545–567

Hawkes JG (1990) The potato-evolution, biodiversity and genetic resources. Belhaven Press, London

Hosaka K, Ogihara Y, Matsubayashi M, Tsunewaki K (1984) Phylogenetic relationship between the tuberous *Solanum* species as revealed by restriction endonuclease analysis of chloroplast DNA. Jpn J Genet 59:349–369

Huang S, Van der Vossen EA, Kuang H, Vleeshouwers VG, Zhang N, Borm TJ, Van Eck HJ, Baker B, Jacobsen E, Visser RG (2005) Comparative genomics enabled the isolation of the R3a late blight resistance gene in potato. Pl J 42:251–261

Kardolus JP (1998) A biosystematic analysis of *Solanum* acaule. Ph.D. thesis. Wageningen Agricultural University, Wageningen, The Netherlands

Kardolus JP, Van Eck HJ, Van den Berg RG (1998) The potential of AFLPs in biosystematics, a first application in *Solanum* taxonomy (Solanaceae). Pl Syst Evol 210:87–103

Mantovani P, Van der Linden CG, Maccaferri M, Sanguineti MC, Tuberosa R (2007) Nucleotide-binding site (NBS) profiling of genetic diversity in durum wheat. Genome 49:1473–1480

Meyers BC, Morgante M, Michelmore RW (2002) TIR-X and TIR-NBS proteins: two new families related to disease resistance TIR-NBS-LRR proteins encoded in *Arabidopsis* and other plant genomes. Pl J 32:77–92

Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW (2003) Genome-wide analysis of NBS-LRR encoding genes in Arabidopsis. Pl Cell 15:809–834

Miller JT, Spooner DM (1999) Collapse of species boundaries in the wild potato *Solanum brevicaule* complex (Solanaceae, *S.* sect. *Petota*): molecular data. Pl Syst Evol 214:103–130

Monosi B, Wisser RJ, Pennill L, Hulbert SH (2004) Full genome analysis of resistance gene homologues in rice. Theor Appl Genet 109:1434–1447

Park TH, Gros J, Sikkema A, Vleeshouwers VG, Muskens M, Allefs S, Jacobsen E, Visser RG, Van der Vossen EA (2005) The late blight resistance locus Rpi-bib3 from *Solanum bulbocastanum* belongs to a major late blight R gene cluster on chromosome 4 of potato. Mol Plant Microbe Interact 18:722–729

Raker CM, Spooner DM (2002) Chilean tetraploid cultivated potato, *Solanum tuberosum*, is distinct from the Andean populations: microsatellite data. Crop Sci 42:1451–1458

Rohlf FJ (1992) NTSYS-pc, numerical taxonomy and multivariate system. Exeter Publishing, Ltd, New York

Song J, Bradeen JM, Naess SK, Raasch JA, Wielgus SM, Haberlach GT, Liu J, Kuang H, Austin-Phillips S, Buell CR, Helgeson JP, Jiang J (2003) Gene RB cloned from *Solanum bulbocastanum* confers broad spectrum resistance to potato late blight. Proc Natl Acad Sci 100:9128–9133

Spooner DM, Castillo RT (1997) Reexamination of series relationships of South American wild potatoes (Solanaceae, *Solanum* S. sect. *Petota*), evidence from chloroplast DNA restriction site variation. Amer J Bot 84:671–685

Spooner DM, Hijmans RJ (2001) Potato systematics and germplasm collecting, 1989–2000. Amer J Potato Res 78:237–268

Spooner DM, Mclean K, Ramsay G, Waugh R, Bryan GJ (2005a) A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. Proc Natl Acad Sci 102:14694–14699

Spooner DM, Peralta IE, Knapp S (2005b) Comparison of AFLPs with other markers for pylogenetic inference in wild tomatoes [*Solanum* L. section *Lycopersicon* (Mill.)] Wettst.]. Taxon 54:43–61

Spooner DM, Salas A (2006) Structure, biosytematics and genetic resources. In: Gopal J, Khurana SMP (eds) Handbook of potato production, improvement, and postharvest management. The Haworth Press, Binghamton, p 17

Sukhotu T, Kamijima O, Hosaka K (2004) Nuclear and chloroplast DNA differentiation in Andean potatoes. Genome 47:46–56

Swofford DL (2001) PAUP* Phylogenetic analysis using parsimony (*and other methods) version 4b10. Sinauer, Sunderland

Syed NH, Sorensen AP, Antonise R, Van de Wiel C, Van der Linden CG, Van't Westende W, Hooftman DA, Nijs HCMD, Flavell AJ (2006) A detailed linkage map of lettuce based on SSAP, AFLP and NBS markers. Theor Appl Genet 112:517–527

Van der Linden CG, Wouters DCAE, Mihalka V, Kochieva EZ, Smulders MJM, Vosman B (2004) Efficient targeting of plant disease resistance loci using NBS profiling. Theor Appl Genet 109:384–393

Van der Vossen EA, Sikkema A, Hekkert BL, Gros J, Stevens P, Muskens M, Wouters D, Pereira A, Stiekema W, Allefs S (2003) An ancient R gene from the wild potato species *Solanum bulbocastanum* confers broad-spectrum resistance to *Phytophthora infestans* in cultivated potato and tomato. Pl J 36:867–882

Van der Vossen EA, Gros J, Sikkema A, Muskens M, Wouters D, Wolters P, Pereira A, Allefs S (2005) The Rpi-blb2 gene from *Solanum bulbocastanum* is an *Mi–1* gene homolog conferring broad-spectrum late blight resistance in potato. Pl J 44:208–222

Volkov RA, Zanke C, Panchuk II, Hemleben V (2001) Molecular evolution of 5S rDNA of *Solanum* species (sect. *Petota*), application for molecular phylogeny and breeding. Theor Appl Genet 103:1273–1282

Volkov RA, Komarova NY, Panchuk II, Hemleben V (2003) Molecular evolution of rDNA external transcribed spacer and phylogeny of sect. *Petota* (genus *Solanum*). Molec Phylogenet Evol 29:187–202

Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP, A new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Yoder AD, Irwin JA, Payseur BA (2001) Failure of the ILD to determine data combinability for slow loris phylogeny. Syst Biol 50:408–424