THE ROYAL SOCIETY

An extreme cytoplasmic bottleneck in the modern European cultivated potato (*Solanum tuberosum*) is not reflected in decreased levels of nuclear diversity

J. Provan, W. Powell*, H. Dewar, G. Bryan, G. C. Machray and R. Waugh

Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, UK

We have used the polymorphic chloroplast (cp) and nuclear simple sequence repeats (SSRs) to analyse levels of cytoplasmic and nuclear diversity in the gene pool of the European cultivated potato (*Solanum tuberosum* ssp. *tuberosum*). Primers designed from the complete chloroplast sequence of tobacco (*Nicotiana tabacum*) were used to amplify polymorphic products in a range of potato cultivars. Combining the data from seven polymorphic cpSSR loci gave 26 haplotypes, one of which (haplotype A) accounted for 151 out of the 178 individuals studied and corresponded to the T-type cytoplasm previously identified in cultivated potatoes using chloroplast restriction fragment length polymorphism analysis. Phylogenetic and diversity analyses of the relationships between cpSSR haplotypes confirmed much higher levels of cytoplasmic diversity outwith the T-type group. Diversity levels at eight nuclear SSR loci, however, were not significantly different between cytoplasmic groups, suggesting a severe maternal bottleneck in the evolution of the modern cultivated potato. These results highlight the importance in quantifying levels of cytoplasmic as well as nuclear diversity and confirm the need for a change in breeding practices to increase levels of non-T-type cytoplasm in the cultivated gene pool, thus helping reduce problems associated with pollen sterility. This may be facilitated by germplasm analysis using cpSSRs, which will allow efficient selection of diverse cytoplasm donors.

Keywords: bottleneck; chloroplast; microsatellite; potato; simple sequence repeat

1. INTRODUCTION

The European cultivated potato (*Solanum tuberosum* ssp. *tuberosum*) is known to have arisen from a limited number of introductions (Simmonds 1979; Glendinning 1983), resulting in a low level of genetic diversity compared to the potato gene pool of the potato. After its introduction, identification of genotypes which produced tubers under long day-length conditions, combined with selection for superior agronomic characters, further narrowed the effective gene pool. In these early days, lack of disease pressure resulted in the virtual eradication of any effective resistance, a situation which culminated in the catastrophic late blight epidemic of the 1840s (Simmonds 1979; Grun 1990). The late 19th century was characterized by an intensified breeding effort from a much reduced genetic base in which deliberate selection for non-berrying varieties resulted in a substantial decline in sexual fertility.

Chloroplast restriction fragment length polymorphism (cpRFLP) analysis of various wild species in relation to the modern cultivated potato led Hosaka & Hanneman (1988*a*,*b*) to identify five major cytoplasm types. One of these, the T-type cytoplasm, was found in the majority of cultivated ssp. *tuberosum* accessions, suggesting the existence of a narrow cytoplasmic base in the European cultivated gene pool. Subsequent studies on chloroplast diversity using cpRFLPs (Waugh *et al.* 1990; Powell *et al.* 1993) have confirmed these low levels of cytoplasmic diversity, although RFLP analysis of chloroplast DNA has historically revealed little intraspecific diversity due to extremely low levels of nucleotide substitution in the chloroplast genome (Wolfe *et al.* 1987; Soltis *et al.* 1992). Low levels of fertility and the observation of reciprocal differences in crosses between *S. tuberosum* ssp. *tuberosum* and *S. tuberosum* ssp. *andigena* (Waugh *et al.* 1990; Maris 1989), suggest that significant performance improvements may be realized by widening the cytoplasmic genetic base. A high-resolution cytoplasmic marker system to distinguish between different cytoplasm types would be a valuable asset in such endeavours.

Recently, we have reported the occurrence of polymorphic mononucleotide repeats in the chloroplast genomes of soyabeans (Powell *et al.* 1995*a*,*b*), rice (Provan *et al.* 1996*a*, 1997) and pines (Powell *et al.* 1995*b*) which are analogous to nuclear microsatellites or simple sequence repeats (SSRs; Powell *et al.* 1996*a*; Morgante & Olivieri 1993). Analysis of length polymorphism at nuclear SSR loci has been used to study levels of genetic diversity in many major crop species (Powell *et al.* 1996*a*,*b*), and PCR-mediated amplification of regions of the organellar genome which contain mononucleotide repeats can

*Author for correspondence (wpowel@scri.sari.ac.uk).

Received 26 October 1998 *Accepted* 26 November 1998

provide new insights into cytoplasmic variation at the intraspecific level. In addition, the high levels of conservation of the chloroplast genome may make it possible to design primers based on one of the published sequences which even work in fairly distantly related species (Provan *et al.* 1999).

In this study we have used chloroplast simple sequence repeat (cpSSR) primers derived from the complete sequence of the tobacco (*N. tabacum*) chloroplast genome (Shinozaki *et al.* 1986) to study the range and magnitude of cytoplasmic diversity present in the majority of potato cultivars grown in the UK over the past two centuries. In addition, we have used eight polymorphic nuclear SSRs to directly compare levels of nuclear and cytoplasmic diversity.

2. MATERIALS AND METHODS

(a) *Genetic material*

The 178 accessions in this study represent 95% of the cultivars on the UK National List, as well as several older varieties, and are summarized in table 1. An additional accession of *S. tuberosum* ssp. *andigena* was also included. DNA was isolated from fresh leaf material using the CTAB method of Saghai-Maroof *et al.* (1984).

(b) *Primer design and polymerase chain reaction*

The complete sequence of the tobacco chloroplast genome (EMBL accession number CHNTXX) was searched for mononucleotide repeats of ten or more bases using the FINDPATTERNS program (Genetics Computer Group). Primers were designed using PRIMER v. 0.5 and are shown along with annealing temperatures [T_m] in table 2. In addition, eight pairs of primers from previous studies (Provan *et al.* 1996*b*; Milbourne *et al.* 1998) were used to analyse levels of nuclear SSR diversity (table 2).

For cpSSRs, polymerase chain reaction (PCR) was carried out in a total volume of 10 µl containing 1 × PCR buffer (10 mM Tris-HCl, 1.5 mM $MgCl_2$, 50 mM KCl, pH 8.3), 200 µM dNTPs, 10 pmol [^{32}P]-ATP end-labelled forward primer, 10 pmol reverse primer, 0.1 U *Taq* polymerase (Promega) and 50 ng genomic DNA. Reactions were carried out on a MJ Research PTC 200 DNA engine thermal cycler using the following parameters: (i) initial denaturation at 94 °C for 3 min; (ii) 30 cycles of denaturation at 94 °C for 15 s, annealing at [T_m] for 15 s and extension at 72 °C for 30 s; (iii) final extension at 72 °C for 5 min. After the addition of 10 µl loading buffer (95% formamide), products were resolved on 6% denaturing polyacrylamide gels containing 1 × TBE buffer and 8 M urea at 80 W constant power for 2 h. Gels were transferred onto 3 MM blotting paper (Whatman), dried and exposed to X-ray film overnight without intensification screens.

For nuclear SSRs, one of the primers was end-labelled with TET, FAM or HEX fluorophores (Applied Biosystems). PCR was carried out in a total volume of 10 µl containing 1 × PCR buffer (10 mM Tris-HCl, 1.5 mM $MgCl_2$, 50 mM KCl, pH 8.3), 200 µM dNTPs, 10 pmol 5′-fluor end-labelled forward primer, 10 pmol reverse primer, 0.05 U *Taq* polymerase (Boehringer Mannheim) and 10 ng genomic DNA. Reactions were carried out on a Perkin-Elmer PE9600 thermal cycler using the following parameters: (i) initial denaturation at 94 °C for 150 s; (ii) 30 cycles of denaturation at 94 °C for 30 s, annealing at [T_m] for 30 s and extension at 72 °C for 30 s; (iii) final extension at 72 °C for 5 min. PCR products were resolved on a Prism 373 automated sequencer (Applied Biosystems) and fragment sizes were calculated using the Genescan software.

Table 1. *Potato accessions used in this study*

Accent	P55/7	Cleopatra
Andorra	Penta	Colleen
Ailsa	Pentland Crown	Colmo
Argos	Pentland Dell	Concurrent
Arran Banner	Pentland Hawk	Cosmos
Arran Comet	Pentland Ivory	Costella
Arran Consul	Pentland Javelin	Cultra
Arran Pilot	Pentland Lustre	Diamant
Arran Victory	Pentland Squire	Diana
Avalanche	Pink Fir Apple	Ditta
Avondale	Rathlin	Edzina
Baillie	Record	Estima
Ballydoon	Red Cara	Europa
Bimonda	Red Craigs Royal	Famosa
Bintje	Red King Edward	Felsina
British Queen	Red Pontiac	Fresco
Brodick	Redskin	Granola
Brodie	Rocket	Jaerla
Buchan	Roscor	Karlena
Carlingford	Royal Kidney	Lady Rosetta
Carrick	Rubicon	Larissa
Catriona	Russet Burbank	Lido
Craigs Alliance	Sangre	Liseta
Cramond	Sharpe's Express	Lola
Crispin	Shula	Lutitia
Desiree	Sierra	Manna
Di Vernon	Skirza	Marfona
Doon Star	Sovereign	Mariana
Duke of York	Stamina	Minerva
Dunbar Rover	Stemster	Mirakel
Dunbar Standard	Stirling	Mistral
Dunrod	Stroma	Mona Lisa
Dundrum	Swift	Mondial
Epicure	Teena	Nikita
Fambo	Tristar	Obelix
Fianna	Ulster Chieftain	Optima
Figaro	Ulster Prince	Panda
Foremost	Ulster Sceptre	Pepo
Glamis	Up To Date	Picasso
Golden Wonder	Valor	Premier
International Kid	Vanessa	Proloog
Jamila	Winston	Remarka
Junior	Yukon Gold	Robinta
Karama	Agria	Romano
Kennebec	Aida	Rooster
Kestrel	Aiko	Rubis
King Edward	Ajiba	Samba
Kingston	Ambo	Sandra
Lumpers	Anna	Slaney
Majestic	Asterix	Spunta
Maris Bard	Atlanta	Timate
Maris Peer	Ausonia	Tomensa
Maris Piper	Baraka	Tulla
Maxine	Barna	Ukama
Morene	Batina	Van Gogh
Myatt's Ashleaf	Berber	Wilja
Nadine	Bright	Yesmin
Navan	Caesar	
Nienta	Carlita	ssp. *andigena*

(c) *Data analysis*

Diversity values based on phenotype frequencies were calculated for each nuclear SSR locus using Nei's unbiased statistic (Nei 1987):

Table 2. *Nuclear and chloroplast primers used in this study*

locus	repeat	location	primers (5′–3′)	T_m (°C)	expected size
chloroplast					
NTCP6	$(A)_{11}$	*rps*16/*trn*Q intergenic region	GGTTCGAATCCTTCCGTC GATTCTTTCGCATCTCGATTC	60	176
NTCP7	$(A)_{11}$	ORF98/*trn*S intergenic region	TGATCCCGGACGTAATCC CGAATCCCTCTCTTTCCG	60	175
NTCP8	$(T)_{11}$	*trn*G intron	ATATTGTTTTAGCTCGGTGG TCATTCGGCTCCTTTATG	55	251
NTCP9	$(T)_{10}$	*trn*G/*trn*R intergenic region	CTTCCAAGCTAACGATGC CTGTCCTATCCATTAGACAATG	55	237
NTCP10	$(T)_{13}$	*atp*F intron	TGCTGAATCGACGACCTA AATATTCGGAGGACTCTTCTG	55	120
NTCP12	$(T)_{10}$ $(A)_{13}$	*rps*2/RF862 intergenic region	TGGTTTGGGTCGTGTATC CCATTTTAGGATTCCATTTC	60	236
NTCP14	$(T)_{11}$	ORF154/*trn*D intergenic region	AATCCGTAGCCAGAAAAATAAA CCGATGCATGTAATGGAATC	60	152
NTCP18	$(T)_{11}$	*psb*C/*trn*S intergenic region	CTGTTCTTTCCATGACCCCTC CCACCTAGCCAAGCCAGA	60	186
NTCP25	$(A)_{13}$	atp*B-E/rpc*L intergenic region	TTAGTCAGGTATTTCCATTTC CTTTTCATAGGAATCTTTCACA	50	190
NTCP28	$(T)_{14}$	*rpl*20/ORF73 intergenic region	TCCAATGGCTTTGGCTA AGAAACGAAGGAACCCAC	55	170
NTCP37	$(A)_{13}$	5SrDNA/*trn*R intergenic region	TTCCGAGGTGTGAAGTGG CAGGATGATAAAAAGCTTAACAC	55	143
NTCP40	$(A)_{14}$	*rpl*23/*trn*H intergenic region	TAATTTGATTCTTCGTCGC GATGTAGCCAAGTGGATCA	55	163
nuclear					
ST13ST	$(AT)_{11}$	*def*4 deficiens analogue intron V	TATTCCCCCTTCCTACTCAA TCTTCCACATTCCTAACCTG	58	202
STGBSS	$(TCT)_9$	granule-bound starch synthase intron I	AATCGGTGATAAATGTGAATGC ATGCTTGCCATGTGATGTGT	58	138
STMYB	$(TTG)_6$	*myb*St1 transcriptional activator	ATACAGGACCTTAATTTCCCCAA TCAAAACCCAATTCAATCAAATC	60	143
STPATP1	$(AT)_{22}$	patatin pseudogene intron II	TGCAATGTGTCGAACAATCA TAATTGGATAGGTCGGCCTG	56	203
STPOACUTR	$(AT)_{13}$	*PoAc*85 actin 3′-UTR	TGTGTTTGTTTTTCTGTAT AATTCTATCCTCATCTCTA	48	276
STMSAC47	$(AC)_{12}$ $(AT)_5$ $(AG)_8$	unknown—chromosome VII	CATAAATGGTTATACACGCTTTGC TAATGGAGGTTCCTGAAAAGAAAGG	60	145
STMSAC65	$(TA)_3$ $(TG)_{12}$	unknown—chromosome II	AACTCTAGCAGTATTTGCTTCA TTATTTAGCGTCAAATGCATA	54	108
LERBCO	$(TG)_4$ $(TA)_6$	tomato RUBISCO intron I	CATTACCTCCATTGCTA GTCTCGTACTTCTTCAT	48	172

$$\hat{H} = \frac{n\left(1 - \sum \left[p_i^2\right]\right)}{n-1}, \tag{1}$$

where n = number of individuals analysed and p_i is the frequency of the ith allele. The same measure was used to calculate cytoplasmic diversity based on cpSSR haplotype frequencies. The sampling variance was calculated as

$$V(\hat{H}) = 2\left[2(n-2)\left[\sum p_i^3 - \left(\sum p_i^2\right)^2\right] + \sum p_i^2 - \left(\sum p_i^2\right)^2\right]. \tag{2}$$

Distances between individuals were calculated from cpSSR haplotypes using the computer program MICROSAT v. 1.5 (Eric Minch, Stanford University, USA) to calculate the absolute distance (D_{AD}) metric which is based on the sum of the number of repeat differences between cpSSR haplotypes and is similar to the $\delta\mu^2$ metric of Goldstein *et al.* (1995)

$$\hat{d}_{ij} = \frac{\sum_{k=1}^{L} |a_{ik} - a_{jk}|}{L}, \tag{3}$$

where a_{ik} and a_{jk} are the repeat sizes of the ith and jth individuals at the kth locus, respectively, and L is the number of loci analysed. Amplification of locus NTCP9 gave products differing in size by units of 10 bp and subsequent sequencing of these alleles revealed that these were due to expansion and contraction of a 10 bp motif adjacent to the mononucleotide repeat (G. Bryan and W. de Jong, unpublished data). For D_{AD} analyses, these 10 bp shifts were treated as single repeat steps. The genetic

Table 3. *Chloroplast SSR haplotypes based on seven polymorphic loci*

			locus						
haplotype	cultivar(s)	RFLP[c]	NTCP6	NTCP7	NTCP8	NTCP9	NTCP12	NTCP14	NTCP18
A	151 individuals[a]	B	174	171	253	280	236	152	189
B	Brodick	A	175	172	253	280	236	153	189
C	Carlingford	—	175	172	254	300	236	152	189
D	Cramond	—	175	172	253	280	236	152	189
E	Kennebec	—	174	171	253	280	237	152	189
F	Maris Peer/Sierra	—	175	172	254	280	236	154	189
G	Maris Piper	A	175	172	252	280	236	154	189
H	Morene[b]	—	174/175	171/172	253/254	280	236	152	189
I	Pentland Lustre	—	175	171	253	280	236	152	189
J	Pink Fir Apple	A	174	172	252	290	237	154	187
K	Roscor	—	174	172	254	290	236	153	187
L	Stamina	—	174	173	253	280	236	152	189
M	Aida	—	176	172	253	280	236	152	189
N	Asterix	—	175	172	253	300	236	154	189
O	Jaerla/Lady Rosetta	—	175	172	254	300	236	154	189
P	Obelix	—	176	172	254	280	236	153	189
Q	Optima	—	175	171	253	280	236	154	189
R	Rubis	—	174	171	253	300	236	152	189
S	Sandra	—	175	171	254	300	236	152	189
T	Timate	—	176	172	254	280	236	154	189
U	Myatt's Ashleaf	C	175	173	249	290	238	156	187
V	Skirza	C	174	172	253	280	236	156	188
W	Shelagh	A	175	171	250	310	236	154	189
X	Lumpers	A	175	172	250	290	238	156	187
Y	P55/7	A	175	171	249	290	238	152	187
Z	ssp. *andigena*	—	174	172	250	290	237	154	188

[a] Haplotype A was found in all accessions studied with the exception of those listed here.
[b] Cultivar Morene exhibited heteroplasmy at loci NTCP6, NTCP7 and NTCP8.
[c] RFLP groupings after Powell *et al.* (1993). Group B was assigned to individuals with T-type cytoplasm. —, the cultivar was not studied.

diversity within cytoplasm types was calculated by averaging the pairwise D_{AD} values and correcting for small sample size

$$D_{AD} = \frac{n}{n-1}\left(\frac{\sum_{i=1}^{k}\sum_{j=i+1}^{k}\sum_{k=1}^{L}|a_{ik}-a_{jk}|}{L}\right), \quad (4)$$

where n is the number of individuals analysed, a_{ik} and a_{jk} are the repeat sizes of the ith and jth individuals at the kth locus, respectively, and L is the number of loci analysed. An analysis of molecular variance (AMOVA) (Michalakis & Excoffier 1996) was carried out to estimate the extent of genetic differentiation (R_{ST}) (Slatkin 1995) between cytoplasm types, using the ARLEQUIN software package (v. 1.1: Stefan Schneider, University of Geneva, Switzerland). A UPGMA (unweighted pair-group method of analysis) tree of relationships between cpSSR haplotypes based on D_{AD} genetic distances was constructed using the NEIGHBOR and DRAWTREE options in the PHYLIP package (v. 3.57c; Joe Felsenstein, University of Washington, USA).

3. RESULTS

(a) *Chloroplast and nuclear SSR polymorphism detected in cultivated potato*

Twelve pairs of primers designed from the tobacco chloroplast sequence were tested. Of these, one failed to generate the expected amplification product, four gave monomorphic products in the 178 accessions studied and seven were polymorphic, with between two and five variants detected. Heteroplasmy was detected in a single cultivar, Morene, with two variants observed at loci NTCP6, NTCP7 and NTCP8. A comparison of patterns of variation observed in this cultivar with the other accessions studied confirms that they could not have arisen from the contamination of DNA samples, since the alternative allele combination is not found in any other individual. As no recombination occurs in the chloroplast molecule, data from the seven polymorphic loci were combined to give 26 haplotypes (table 3) and it can be seen that haplotype A predominated, being present in 151 out of the 178 individuals examined. The American cultivar, Russet Burbank, possessed cpSSR haplotype A and comparisons with previous cpRFLP studies which included this cultivar (Hosaka & Hanneman 1988*b*; Waugh *et al.* 1990; Powell *et al.* 1993) confirm that the predominant A haplotype corresponds to the T-type cytoplasm.

Between four and nine alleles were detected at the eight nuclear SSR loci examined. The tetraploid nature of the cultivated potato precluded accurate determination of the allele composition of the majority of the accessions studied and thus all subsequent nuclear diversity analyses were based on phenotype frequencies elucidated from band patterns obtained with individual primer pairs.

(b) *Relationships between cpSSR haplotypes*

The UPGMA phylogenetic tree based on D_{AD} genetic distances is shown in figure 1. The absolute distance

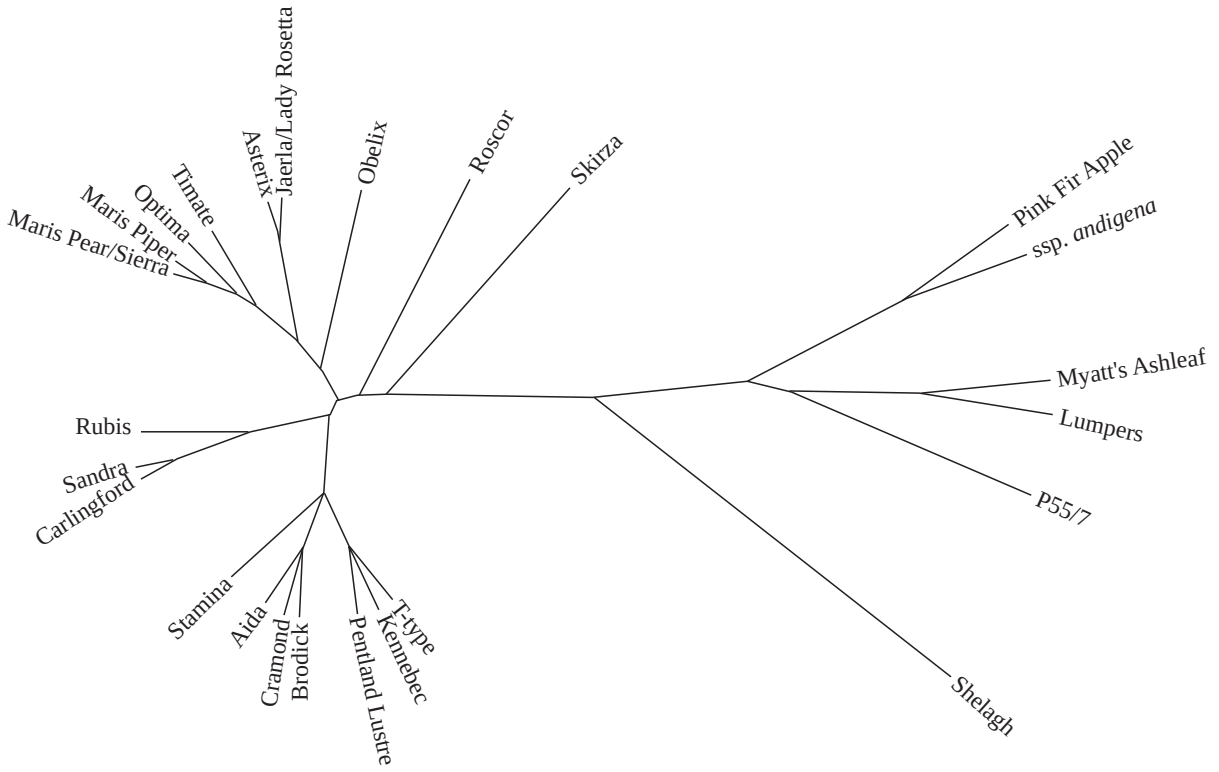

Figure 1. Phylogenetic tree showing relationships between potato accessions.

metric was used since it is most appropriate for closely related individuals, in this case below the subspecies level, which have diverged fairly recently (Goldstein *et al.* 1995). The majority of the cultivars (151 out of 178), exhibited haplotype A, which corresponds to the T-type cytoplasm. Kennebec appears adjacent to the T-type individuals and was classified as belonging to this group by cpRFLP analysis (Hosaka & Hanneman 1988*b*). This is likely to be the case, since it possesses haplotype E which differs by a single step mutation from haplotype A, and which probably represents a recent mutation. The same is true for haplotype I (Pentland Lustre) which also differs from haplotype A by a single mutation. Haplotypes J and U through Z exhibited much higher levels of genetic differentiation, with haplotype W (cultivar Shelagh) being the most divergent. Cultivar Morene was heteroplasmic and was thus excluded from this analysis. An analysis of molecular variance confirmed the high levels of differentiation between haplotypes with 83% of the total genetic variation being attributed to differences between the T-type cytoplasm and the remaining chloroplast haplotypes (R_{ST}=0.829).

(c) *Comparison of levels of chloroplast and nuclear diversity*

Two different measures were used to quantify the levels of cytoplasmic diversity revealed by cpSSRs. Nei's unbiased measure (Nei 1987) showed that the diversity within the T-type accessions (0.026 ± 0.018) was much lower than the diversity within the other accessions (0.987 ± 0.023; table 4). These lower levels of diversity within the T-type cytoplasm were also evident with the D_{AD} measure, which takes into account differences in allele sizes at SSR loci and thus provides a measure of genetic diversity. Using this metric, the levels of diversity

Table 4. *Frequency of size variants and diversity statistics for seven polymorphic cpSSRs*

		frequency	
locus	size variants (bp)	T-type ($n=153$)	others ($n=24$)[a]
NTCP6	174	0.993	0.263
	175	0.007	0.579
	176	—	0.158
NTCP7	171	1.000	0.208
	172	—	0.709
	173	—	0.083
NTCP8	249	—	0.083
	250	—	0.125
	252	—	0.083
	253	1.000	0.333
	254	—	0.375
NTCP9	280	1.000	0.458
	290	—	0.250
	300	—	0.250
	310	—	0.042
NTCP12	236	0.993	0.792
	237	0.007	0.083
	238	—	0.125
NTCP14	152	1.000	0.292
	153	—	0.125
	154	—	0.458
	156	—	0.125
NTCP18	187	—	0.208
	188	—	0.083
	189	1.000	0.709
diversity		0.026 ± 0.018	0.987 ± 0.023
D_{AD}		0.006	0.864

[a] Cultivar Morene excluded due to heteroplasmy.

Table 5. *Diversity statistics for nuclear SSRs*

	diversity	
locus	T-type ($n=153$)	others ($n=24$) [a]
ST13ST	0.806±0.017	0.785±0.086
STGBSS	0.883±0.015	0.848±0.052
STMYB	0.707±0.024	0.758±0.079
STPATP1	0.704±0.031	0.745±0.101
STPOACUTR	0.583±0.034	0.765±0.081
STMSAC47	0.881±0.009	0.837±0.057
STMSAC65	0.827±0.030	0.774±0.083
LERBCO	0.757±0.017	0.778±0.061

[a] Cultivar Morene excluded due to cpSSR heteroplasmy.

in the T-type accessions were over 100-fold less (0.006 compared with 0.864) than those calculated for the other accessions (table 4). An examination of the size variants present in the T-type cytoplasm shows that they would appear to have arisen from a very limited subset of the variants in the total gene pool.

A comparison of the levels of diversity calculated from nuclear SSR polymorphism data for the same cytoplasmic groups does not reflect the apparent genetic depletion observed with chloroplast markers (table 5). A significantly lower level of nuclear diversity within the T-type individuals was only observed at one locus (STPOACUTR) and an AMOVA based on phenotype frequencies for each locus did not show any significant divergence between the phenotypes present in individuals with T-type cytoplasm and those in the other cultivars.

4. DISCUSSION

To our knowledge, this study represents by far the largest analysis of cytoplasmic diversity within the modern cultivated potato. We have demonstrated the potential for high-resolution cytoplasmic analysis of potatoes using primers designed to amplify cpSSRs from the complete sequence of another member of the same family (*Nicotiana tabacum*). The high resolving power of cpSSRs compared to RFLPs allowed a more detailed analysis of patterns of cytoplasmic variation than has previously been possible using cpRFLPs. Even at this level of detail, however, the predominance of a single chloroplast cytotype was still observed. This feature was also found in previous cpRFLP studies (Hosaka & Hanneman 1988*b*; Waugh *et al.* 1990; Powell *et al.* 1993) and the dominant haplotype corresponds to the T-type cytoplasm found in ssp. *tuberosum* and a limited number of ssp. *andigena* accessions, but not in any diploid or tetraploid wild species. The absence of any alleles in the T-type accessions which are not found in the other cultivars suggests that this cytoplasm may have evolved monophyletically from the total gene pool, rather than representing a separate introduction.

The predominance of the T-type cytoplasm in the modern European cultivated gene pool can be traced back to the introduction of the American cultivar, Rough Purple Chili, after the late blight epidemic of the 1840s and its derivatives, particularly cultivars Early Rose and Garnet Chili. The rapid spread of Rough Purple Chili or T-type cytoplasm through the modern gene pool may have resulted from a number of factors. Grun (1979) has suggested that this is probably due to the pollen sterility of derivatives of Rough Purple Chili, which has led to the extensive use of these genotypes as female parents in crosses. Other studies on reciprocal crosses between subspecies *tuberosum* and *andigena* have shown that ssp. *tuberosum* cytoplasm derivatives are higher yielding, but that this increased yield is coupled with an increase in levels of male sterility (Maris 1989). This may have led to the deliberate selection for non-berrying varieties, as well as the possibility that any berries used from open pollinated crosses in the field may have resulted from selfing. Grun also highlighted the role of nuclear–cytoplasmic interactions between ssp. *tuberosum* cytoplasmic factors which confer sterility in the presence of dominant nuclear genes from ssp. *andigena* (Grun 1990).

A previous study on cytoplasmic diversity in European potato cultivars by Powell *et al.* (1993) identified two other groups of cultivars which did not possess the typical or common cytoplasm. The cpSSR technique has allowed further resolution of these groups into clades containing accessions which possess cytoplasm types with various degrees of divergence. This is highlighted by examination of the cultivars Lumpers, Pink Fir Apple, Maris Piper, Brodick and Shelagh, and the breeding line P55/7, which were identical in the earlier study based on cpRFLPs but displayed different, unique cpSSR haplotypes. The apparent difference between the group of cultivars mentioned above and another group comprising the cultivars Myatt's Ashleaf and Skirza, based on cpRFLPs, was not observed in this study, since they appear in different clades outwith the T-type group. The cultivars Pink Fir Apple (1850), Myatt's Ashleaf (1847) and Lumpers (1806) all have haplotypes (J, U and X, respectively) which are fairly divergent from the T-type haplotype. This is presumably representative of the wider '*andigena*'-type cytoplasm which existed prior to the blight epidemics of the 1840s. An earlier discrepancy between cpRFLP studies concerning the cytoplasm type of cultivar Estima would also appear to be resolved using cpSSRs. Waugh *et al.* (1990) found a unique cytoplasm type in Estima, whereas Powell *et al.* (1993) assigned it, along with its sport Famosa, the T-type cytotype, which is consistent with our findings where both cultivars share haplotype A. The apparent heteroplasmy detected in cultivars Skirza and Myatt's Ashleaf using cpRFLPs was not observed with cpSSRs but was detected, however, in cultivar Morene which appeared to possess a combination of the T-type cpSSR haplotype and another not detected elsewhere in this study.

Genetic bottlenecks are a frequent consequence of the domestication process in modern crop species. Tanksley & McCouch (1997) have highlighted the narrow genetic base of many crops as well as the concept of supplementing cultivated gene pools with 'exotic' alleles from wild ancestors of cultivated crops. The cytoplasmic bottleneck in the modern cultivated potato revealed by cpSSRs, however, was not reflected in decreased levels of nuclear SSR diversity. This suggests the use of a very limited number of maternal lineages in breeding programmes but high levels of paternal diversity, and may reflect the

somewhat random nature of early, open-pollinated crosses. In addition, the fact that potato is a highly heterozygous tetraploid means that the nuclear genome should be better able than the haploid chloroplast genome to counter the effects of any inbreeding. Our results highlight the importance of the analysis of cytoplasmic diversity and Buckner and Hyde (1985) have stressed the need to increase levels of chloroplast variability found in the modern potato gene pool. However, breeding constraints imposed by the pollen sterility of many of the T-type derivatives may suggest the need for a change in breeding practices involving replacement of the dominant cytoplasm with more diverse material, either from within the cultivated gene pool or from wild species and/or Neo-*tuberosum*. Germplasm analysis using highly polymorphic cpSSR markers will allow the efficient selection of cytoplasm donors to replenish depleted stocks and to generate male- and female-fertile breeding lines.

This work was supported by the Scottish Office Agriculture, Environment and Fisheries Department. We thank John Bradshaw for his comments on the manuscript and Walter de Jong, Jane McNicol and Rhonda Meyer for sharing unpublished data.

REFERENCES

Buckner, B. & Hyde, B. B. 1985 Chloroplast DNA variation between the common cultivated potato (*Solanum tuberosum* ssp. *tuberosum*) and several South American relatives. *Theor. Appl. Genet.* **71**, 527–531.

Glendinning, D. R. 1983 Potato introductions and breeding up to the early 20th century. *New Phytol.* **94**, 479–505.

Goldstein, D. B., Ruíz-Linares, A., Feldman, M. & Cavalli-Sforza, L. L. 1995 Genetic absolute dating based on microsatellites and the origin of modern humans. *Proc. Natl Acad. Sci. USA* **92**, 6720–6727.

Grun, P. 1979 The potato. In *Biology and taxonomy of the Solanaceae* (ed. J. G. Hawkes, R. N. Lester & A. D. Skelding), pp. 655–665. London: Academic Press.

Grun, P. 1990 The evolution of cultivated potatoes. *Econ. Bot.* **44** (Suppl.), 39–55.

Hosaka, K. & Hanneman, R. E. 1988*a* Origin of chloroplast DNA diversity in the Andean potatoes. *Theor. Appl. Genet.* **76**, 333–340.

Hosaka, K. & Hanneman, R. E. 1988*b* The origin of the cultivated tetraploid potato based on chloroplast DNA. *Theor. Appl. Genet.* **76**, 172–176.

Maris, B. 1989 Analysis of an incomplete diallel cross among three ssp. *tuberosum* varieties and seven long-day adapted ssp. *andigena* clones of the potato (*Solanum-tuberosum*-L.). *Euphytica* **41**, 163–182.

Michalakis, Y. & Excoffier, L. 1996 A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* **142**, 1061–1064.

Milbourne, D., Meyer, R. C., Collins, A. J., Ramsay, L. D., Gebhardt, C. & Waugh, R. 1998 Isolation, characterisation and mapping of simple sequence repeat based markers in potato. *Mol. Gen. Genet.* **259**, 233–245.

Morgante, M. & Olivieri, A. M. 1993 PCR amplified microsatellites in plant genetics. *Plant J.* **3**, 175–182.

Nei, M. 1987 *Molecular evolutionary genetics*. NY: Columbia University Press.

Powell, W., Baird, E., Duncan, N. & Waugh, R. 1993 Chloroplast DNA variability in old and recently introduced potato cultivars. *Ann. Appl. Biol.* **123**, 403–410.

Powell, W., Morgante, M., Andre, C., McNicol, J. W., Machray, G. C., Doyle, J. J., Tingey, S. V. & Rafalski, J. A. 1995*a* Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. *Curr. Biol.* **5**, 1023–1029.

Powell, W., Morgante, M., McDevitt, R., Vendramin, G. G. & Rafalski, J. A. 1995*b* Polymorphic simple sequence repeat regions in chloroplast genomes: applications to the population genetics of pine. *Proc. Natl Acad. Sci. USA* **92**, 7759–7763.

Powell, W., Machray, G. & Provan, J. 1996*a* Polymorphism revealed by simple sequence repeats. *Trends Plant Sci.* **1**, 215–222.

Powell, W., Morgante, M., Doyle, J. J., McNicol, J. W., Tingey, S. V. & Rafalski, J. A. 1996*b* Genepool variation in genus *Glycine* subgenus *soja* revealed by polymorphic nuclear and chloroplast microsatellites. *Genetics* **144**, 793–803.

Provan, J., Corbett, G., Waugh, R., McNicol, J. W., Morgante, M. & Powell, W. 1996*a* DNA fingerprints of rice (*Oryza sativa*) obtained from hypervariable chloroplast simple sequence repeats. *Proc. R. Soc. Lond.* B **263**, 1275–1281.

Provan, J., Powell, W. & Waugh, R. 1996*b* Microsatellite analysis of relationships within cultivated potato (*Solanum tuberosum*). *Theor. Appl. Genet.* **92**, 1078–1084.

Provan, J., Corbett, G., McNicol, J. W. & Powell, W. 1997 Chloroplast variability in wild and cultivated rice (*Oryza* spp.) revealed by polymorphic chloroplast simple sequence repeats. *Genome* **40**, 104–110.

Provan, J., Soranzo, N., Wilson, N. J., McNicol, J. W., Morgante, M. & Powell, W. 1999 The use of uniparentally inherited simple sequence repeat markers in plant population studies and systematics. In *Advances in plant molecular systematics* (ed. P. Hollingsworth, R. Gornall & R. Bateman). London: Chapman & Hall. (In the press.)

Saghai-Maroof, M. A., Soliman, K. M., Jorgensen, R. A. & Allard, R. W. 1984 Ribosomal DNA spacer length polymorphism in barley: Mendelian inheritance, chromosomal location and population dynamics. *Proc. Natl Acad. Sci. USA* **81**, 8014–8018.

Shinozaki, K. (and 22 others) 1986 The complete sequence of the tobacco chloroplast genome—its gene organization and expression. *EMBO J.* **5**, 2043–2049.

Simmonds, N. W. 1979 Potatoes. In *Evolution of crop plants* (ed. N. W. Simmonds), pp. 279–283. New York: Longman.

Slatkin, M. 1995 A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 457–462.

Soltis, D. E., Soltis, P. S. & Milligan, B. G. 1992 Intraspecific chloroplast DNA variation: systematic and phylogenetic implications. In *Molecular systematics of plants* (ed. D. E. Soltis, P. S. Soltis & J. J. Doyle), pp. 117–150. New York: Chapman & Hall.

Tanksley, S. D. & McCouch, S. R. 1997 Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* **277**, 1063–1066.

Waugh, R., Glendinning, D. R., Duncan, N. & Powell, W. 1990 Chloroplast DNA variation in European potato cultivars. *Potato Res.* **33**, 505–513.

Wolfe, K. H., Li, W. H. & Sharp, P. M. 1987 Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast and nuclear DNA. *Proc. Natl Acad. Sci. USA* **84**, 9054–9058.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.