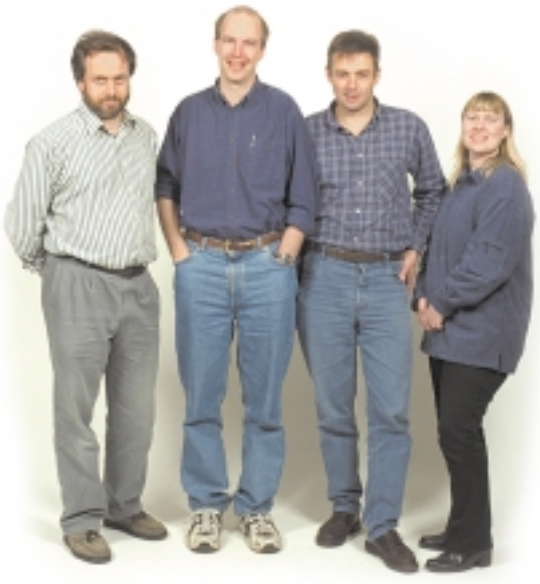
Potato genomics: a general strategy for the molecular genetic characterisation of *Solanum* germplasm

G. Bryan, W. De Jong, J. Provan, D. Milbourne, J. McNicoll, J. Davidson, G. Ramsay, R. Waugh

In recent years, a major goal of potato genetic research at SCRI has been the development of PCR-based molecular genetic markers. The primary use of these markers has been for the analysis of segregating populations of potato, which has allowed the genetic locations of markers to be determined. In combination with phenotypic data obtained from such populations, it is also possible to identify molecular markers linked to traits of biological and agronomic interest. A further objective is to use these powerful genetic markers for further analysis of populations and the genetic characterisation of potato germplasm. This germplasm may take the form of tetraploid potato cultivars and breeding lines, as well as accessions of wild and primitive cultivated potato species of various ploidy levels, such as material from the Commonwealth Potato Collection (CPC).

Simple Sequence Repeats (SSRs)

At the present time, approximately 150 microsatellite or simple sequence repeat (SSR) loci have been identified and characterised in potato at SCRI. Progress in mapping and characterising these SSRs has been reported in previous Annual Reports (e.g. Ann. Rep. 1997/8, 86-88). Microsatellites offer the advantage of showing a codominant mode of inheritance, meaning that several alleles can be identified for each locus, subject to the degree of polymorphism shown by that locus. This property is particularly useful in tetraploids, and allows the possibility of assigning microsatellite genotypes (the combination of alleles present at a set of SSR loci) for any given accession. SSR markers have been used very effectively to anchor genetic maps containing large numbers of high-volume, 'single-dose', markers such as Amplified Fragment Length Polymorphisms (AFLP). AFLP technology has become the molecular workhorse for the generation of high volume markers in potato genetic mapping. AFLP markers typically exhibit a 'dominant' (i.e. presence/absence) pattern of inheritance that makes the establishment of genetic maps based on AFLPs alone somewhat problematical, particularly in tetraploids. Incorporation of co-dominant SSR marker locus data greatly reduces the problem of map construction by, for example, allowing the rapid assignation of linkage groups to chromosomes. Microsatellite markers are eminently suitable for the analysis of potato germplasm, from cultivars and breeding material to wild species.



Chloroplast SSRs

In recent years, SCRI has led the way in the development of microsatellite markers from chloroplast genomes of several plant species, including maize and pine (Ann. Rep. 96/97, 93-4). Chloroplast genome-based microsatellite markers are attractive in that they are easy to deploy (owing to the high copy number of the cpDNA molecule per cell) and, due to the lack of recombination in the chloroplast genome, it is very straightforward to assign a cpSSR 'haplotype' to any given accession. We have identified a set of 36 primer pairs, which amplify polymorphic PCR products from the chloroplast genomes of most Solanaceous plants, including potato and tobacco. The origin of these primers is the completely sequenced tobacco chloroplast genome, but the chloroplast genome sequence is sufficiently conserved among Solanaceous plants to permit cross-species PCR amplification. Figure 1 shows an example of a cpSSR locus PCR assay on a set of potato germplasm from the

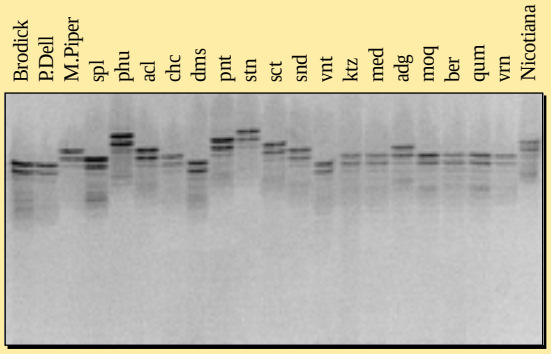


Figure 1 Chloroplast microsatellite NTCP12 applied to a diverse set of potato/tobacco germplasm.

CPC. Figure 2 provides information on levels of polymorphism obtained with some of the more polymorphic cpSSR primer pairs. These primers almost certainly will prove useful in all Solanaceous plant species. Owing to the ease of deployment of these markers, it is possible to 'multiplex' several cpSSR primer pairs in a single PCR reaction, meaning that a cpSSR haplotype can be obtained using just a few PCR reactions. These markers will have several applications, such as examining the levels of chloroplast genome variation in the cultivated potato gene pool (see below), as well as for phylogenetic and diversity analysis of potato germplasm.

Deployment of these markers in the analysis of potato germplasm:

1. SURVEY OF POTATO CULTIVARS AND COMMONWEALTH POTATO COLLECTION In order to assess their utility for germplasm analysis, the 36 cpSSR primer pairs described above were applied to the analysis of a representative set of 25 wild species from the CPC and a set of 30 potato cultivars. The cultivars were selected to represent a wide range of chloroplast diversity; these had been identified in a previous study, based on chloroplast derived RFLPs. Our preliminary survey, using these 36 primer pairs, showed that 26 (72%) detected some level of polymorphism among the 55 accessions. This study showed that the cpSSRs are very suitable for the analysis of potato and tobacco germplasm at several taxonomic levels[1]. A cluster analysis of material from the CPC shows good agreement with previously published phylogenies of potato. Secondly, it appears that amongst the diverse set of potato cultivars, there is a predominant 'cytotype', which is thought to result from the widespread use of the cultivar Rough Purple Chili as a female parent in the latter half of the 19th century (Fig. 3). A third observation is that some cultivars and breeding lines

Locus	Av. Heterozygosity	No. of alleles
NTCP3	0.59	4
NTCP23	0.66	5
NTCP4	0.62	4
NTCP6	0.59	7
NTCP7	0.58	3
NTCP8	0.66	5
NTCP9	0.70	8
NTCP30	0.54	4
NTCP12	0.58	5
NTCP14	0.69	5
NTCP18	0.66	3
NTCP39	0.62	4

Figure 2 Levels of polymorphism among 12 of the most polymorphic cpSSR loci. Average heterozygosity ($\hat{H}$) has been calculated as follows : $\hat{H} = n/n\text{-}1(1\text{-}\Sigma p_i^2)$
where p_i = frequency of i^{th} allele
n = number of samples

have a cpSSR haplotype identical to that of particular wild species of potato, which were used in the past to introgress genes of agronomic importance into cultivated *S. tuberosum*. This latter observation suggests that a large part of the chloroplast diversity still present in cultivated potato is derived from wild or primitive cultivated species.

2. UK POTATO NATIONAL LIST SURVEY A combination of nuclear SSRs and cpSSRs has been applied to the analysis of 178 potato cultivars on the UK

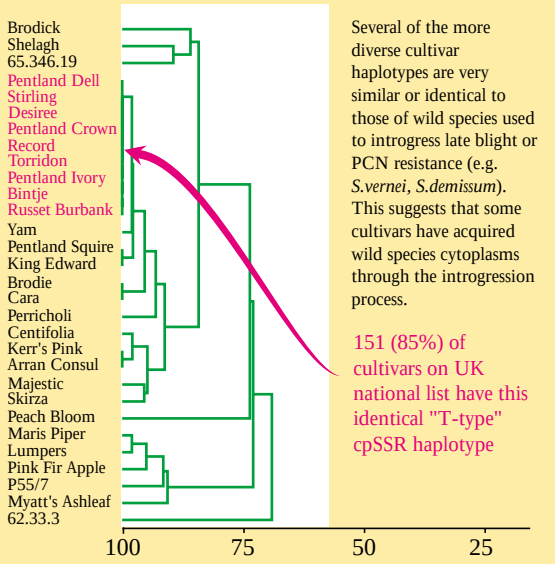


Figure 3 Cluster analysis of chloroplast haplotypes of 30 potato cultivars. Haplotypes have been clustered using a UPGMA method based on similarity estimates obtained using a 'city-block' method.

National List. This study has highlighted a paucity of chloroplast genetic variability in the cultivated potato in Europe, which is not seen at the nuclear level[2]. A very large proportion of these cultivars (151 i.e. 85%) has exactly the same type of 'T-type' cytoplasm described above. The high levels of nuclear SSR variability do not differ markedly among different cpSSR haplotypes and, moreover, the large group with identical cytoplasms shows no less nuclear polymorphism than the remaining cultivars. These observations suggest that, unless diverse chloroplast types are actively chosen as parents in potato breeding schemes, there is a strong possibility that the cultivated potato in the UK, and possibly Europe, will ultimately be represented by a single chloroplast haplotype.

3. MULTITRAIT RECURRENT SELECTION SCHEME An analysis of genetic variation among 23 of the parental lines used in a multitrait recurrent breeding scheme has been initiated. To provide a comparison, representative samples of 24 old cultivars (pre-1900) and 20 popular cultivars currently grown in the UK have been included in this study. These accessions have been analysed with nuclear and chloroplast SSRs as well as AFLP markers. AFLP and nuclear SSR analysis show that the parents of the multitrait scheme and, to a lesser extent, modern cultivars, show significantly greater numbers of 'rare alleles' than the old cultivars (Fig. 4). It is also clear that only a small number of rare alleles, present in the older material, have been lost from the cultivated potato genepool. A very encouraging observation from this study, provided by cpSSR analysis, is that the parents of the multitrait breeding scheme are represented by the presence of 8 distinct cpSSR haplotypes. This compares with two and three haplotypes in modern and old sets of cultivars respectively. The potato breeding industry as a whole should be encouraged to adopt a similar strategy in the choice of parents for potato breeding schemes.

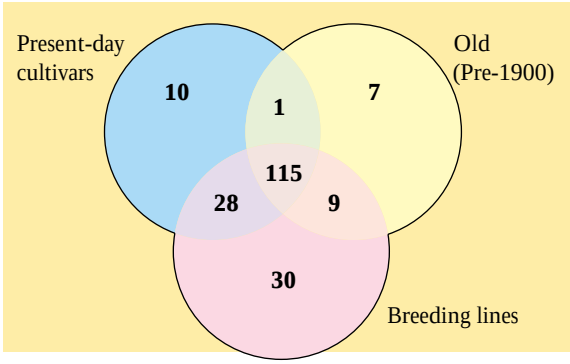


Figure 4 Distribution of polymorphic AFLP markers among three sets of adapted potato germplasm.

4. CPC SURVEY In the last year, we have started to use microsatellite markers to examine additional material from the CPC. These studies were designed to test further the utility of SSRs for the analysis of potato germplasm across relatively wide genetic distances. For example, a set of nuclear and chloroplast SSRs was used to study 33 accessions from series Longipedicellata, which contains several species. This study has indicated that there appear to be no distinct boundaries between groups of accessions representing different species from this complex, suggesting that the classification into separate species should be re-examined. Recently, we have initiated a collaboration with Dr David Spooner (University of Wisconsin, USA), who is currently examining accessions from series Longipedicellata, to carry out an extended molecular and morphological analysis of material from this complex. This work has implications for potato taxonomists and for the maintenance of larger germplasm collections, such as the CPC, where the accurate identification of accessions is highly important, particularly when the costs of preserving such collections are considered. A second study, recently initiated, is to examine potential progenitor species of cultivated potato, in an

attempt to elucidate further the likely origins of the cultivated potato.

Potato germplasm evaluation - a general strategy

Microsatellites remain the recommended type of marker for the examination of potato germplasm, owing to their ability to detect allelic variants and high levels of genetic polymorphism. The need for examining chloroplast DNA variation independently of the nucleus is highlighted by the results outlined above. Differences in nuclear genotypes may not preclude identical cytoplasms, although accessions with different cytoplasms are very unlikely to have the same nuclear constitution. This suggests that a sensible strategy may be to deploy cpSSRs as a primary screen for polymorphism. If higher numbers of markers are required, say for fingerprinting more closely related accessions, the use of AFLPs may also be an option. For most purposes, however, the combination of the 10 most polymorphic cpSSRs in combination with 10-20 nuclear SSRs will be sufficient.

The future

We will continue to use molecular markers for the analysis of potato germplasm at SCRI. Further studies will involve adapted material, such as potato cultivars and breeding lines, segregating populations, and material from the CPC. Ultimately, we would like to genetically 'fingerprint' the entire CPC. This would provide a much stronger basis for the use of the CPC as a source of novel genes (e.g. disease resistance). It would also provide a stronger means for a rational framework on which to maintain the collection. This would be of even greater benefit if other large potato collections (e.g. CIP, Peru) could be examined with the same genetic markers. This would permit quantifiable assessments of the relative levels of genetic diversity within and between different germplasm collections.

References

[1]Bryan, G.J., McNicoll, J., Meyer, R.C., Ramsay, G. & De Jong, W.S. (1999). *Theoretical and Applied Genetics* 99, 859-867.

[2] Provan, J., Powell, W., Dewar, H., Bryan, G., Machray, G.C. & Waugh, R. (1999). *Proceedings of the Royal Society Series B* 266, 633-639.