Genetica (2005) 123: 255–261

Quantitative assessment of heteroplasmy levels in *Senecio vulgaris* chloroplast DNA

Juerg E. Frey, Beatrice Frey & Didier Forcioli[1]
Agroscope FAW Waedenswil, Department of Plant Protection, Molecular Diagnostics Laboratory, Swiss Federal Research Station for Horticulture, CH8820 Wädenswil, Switzerland (Phone: +411-783-6332; Fax: +411-783-6434; E-mail: Juerg.frey@faw.admin.ch); [1]*Present address: UMR INRA-UNSA 1112 'ROSE', Université de Nice-Sophia Antipolis, Faculté des Sciences, Parc Valrose, F06108 Nice Cedex 2, France*

Received 7 June 2004 Accepted 17 September 2004

Key words: chloroplast heteroplasmy, common groundsel, *psbA*, triazine resistance, within-plant haplotype frequency variation

Abstract

Heteroplasmy in coding chloroplast DNA was only recently shown to occur and was so far not quantitatively assessed. We present a quantitative analysis of cpDNA heteroplasmy levels at a triazine-resistance determining site within and between individual *Senecio vulgaris* plants. Detectable levels of heteroplasmic haplotypes were observed in all tested plants. As expected, the levels of heteroplasmy vary greatly between plants. However, even within individual plants, the fraction of one haplotype may cover a range from below 1% to well over 90%. Our results suggest that heteroplasmy may be a common phenomenon in *S. vulgaris*. Possible consequences for molecular diagnostics of chloroplast encoded traits as well as evolutionary consequences of chloroplast heteroplasmy are discussed.

Introduction

Chloroplast heteroplasmy, the occurrence of more than one type of chloroplast within an individual plant, was once believed to be rare but is now established as quite common (Birky et al., 1989; Yu & Spreitzer, 1992; Fitter et al., 1996; Chat et al., 2002, and references therein). So far, the phenomenon was generally observed in non-coding cpDNA, mostly in the form of length polymorphisms. In an earlier study, we documented qualitative variation of heteroplasmy levels of a coding mutation within and between *Senecio vulgaris* L. (Asteraceae) plants (Frey, 1999; Frey et al., 1999). *S. vulgaris* is a self-compatible, strongly self-fertilizing annual described to be an important weed in many parts of the world (Müller-Schärer & Wyss, 1994; Mitich, 1995; Frey et al., 1999). The mutation involves a single base substitution in the *psbA* chloroplast gene which codes for a photosystem II membrane protein to which triazine herbicides bind (Hirschberg et al., 1984), thereby rendering its carrier highly resistant to this class of herbicides (e.g., Darmency, 1994, and references therein). Although the PCR–RFLP based technique used in this earlier study resulted in very interesting qualitative data, the technical characteristics of that method (as discussed in the discussion section) did not allow accurate quantification of the level of heteroplasmy nor of its variation within plants. At least two factors related to the PCR–RFLP technique may lead to a serious error in the estimate of the fraction of individual haplotypes in a mixed sample. The first factor may be restriction site preference of restriction enzymes. Site preference is a well known phenomenon in many restriction enzymes that may cover a broad range from only gradual effects to sometimes strong inhibition of restriction at adjacent restriction sites (Thomas & Davis, 1975; Armstrong & Bauer, 1982). Due to the limited resolution of gels, only pronounced preference may be noticed in most cases. However, if

two restriction sites are only few base pairs apart from each other, steric hindrance may exaggerate low levels of site preference. A second important factor with serious effect on haplotype frequency estimates is a mere consequence of the PCR mechanism. When amplifying a template that consists of two different haplotypes, denaturation and re-hybridization will produce heteroduplexes in a binomial fraction. Most restriction enzymes will probably not accept heteroduplexes as target sites and hence, heteroduplexes will not be cut. Depending on the assay, this factor alone may be responsible for serious over- or underestimates up to a factor 2. Although these errors will not affect the interpretation of a qualitative analysis such as that presented by Frey (1999), they impede accurate quantification of individual haplotypes. To enable precise quantification of the level of heteroplasmy in *S. vulgaris,* we therefore developed a new assay based on nested-design quantitative real-time PCR. This assay is well suited to accurately assess the proportion of individual haplotypes or mutants in a DNA sample. The first-round PCR maintains the relative frequencies of different haplotypes (Chen et al., 2002) and is therefore widely used as the first step in single-nucleotide polymorphism detection in pooled DNA samples (reviewed in Sham et al., 2002) and in assessing allele frequencies together with a variety of follow-up detection methods, such as single base extension assays (Zhou et al., 2001; Yan et al., 2002), pyrosequencing (Gruber et al., 2002), mass spectrometry (Werner et al., 2002) or microarray technology (Lo et al., 2003). It was also shown that possible errors introduced by Taq polymerase (e.g., single base changes, insertions, deletions) do not negatively affect PCR-based molecular diagnostics (Gyllensten, 1989; Frey & Frey, 2004). Furthermore, the products of the first-round PCR are ideal templates for a follow-up nested allele-specific PCR as this reaction uses single strands that are produced during each denaturation step. It is thus not affected by possible heteroduplex formation that may cause serious distortion in quantitative analyses of PCR–RFLP products, as explained above. Our assay is designed to allow direct quantification of allele frequencies and to reliably work with reduced amplification efficiencies caused by mismatched primers as well as with low DNA amounts. Using this assay, we tested variation of cpDNA haplotype frequencies within and between leaves of individual *S. vulgaris* plants collected in an apple orchard.

The aim of this study was to assess whether heteroplasmic plants occur in the field and to quantify as accurately as possible the degree of heteroplasmy in individual plants using this new method. We found large variation in the degree of observed cpDNA heteroplasmy at all levels, i.e., within leaf, within plant and between plant. Some plants showed only minor variation or variation only at certain levels but not at others, while other plants showed highly significant variation in heteroplasmy at all levels.

Material and methods

Plant material

We collected nine *S. vulgaris* plants randomly from an apple orchard near Waedenswil (Canton Zuerich, Switzerland) that had a history of triazine treatments, i.e., the orchard was treated with triazine containing herbicides about every second year for ca. 15 years until four years ago, when herbicide treatment was discontinued. Two leaf disks of 2–3 mm^2 in area were punched out of selected leaves using a disposable Pasteur pipette. One to two samples per leaf of four leaves per plant were taken from each plant, with leaf 1 being one of the oldest leaves of the plant, i.e., closest to the ground, and leaf 4 being close to the top of the plant. The samples were placed in 0.5 ml Eppendorf tubes and kept in the freezer at −20 °C.

DNA extraction procedure

DNA from frozen leaf disks was homogenized in 1 µl lysis buffer (10 mM Tris HCl, 1 mM EDTA, 0.5% Tween 20, 50 µg/ml Proteinase K, pH 8.0) using a Mixer Mill (Retsch MM300, Qiagen; 25 Hz). The homogenate was then heat-treated for 10 min at 85 °C. One microliter of the homogenate corresponding to approximately 1–5 ng template DNA, was used for amplification.

Polymerase chain reaction (PCR)

PCR was performed with 1–5 ng template DNA and 0.5 µM of each of the two primes P1 (5′-ATGAGGGTTACAGATTTGGTC-3′) and P2 (5′-AGATTAGCACGGTTGATGATA-3′) described

by Cheung et al. (1993) in a total reaction volume of 20 µl using the Hot Star Taq Master Mix Kit (Qiagen). The DNA was amplified on a Applied Biosystems 9600 thermal cycler using the following protocol: a 15 min initial denaturation at 95 °C, then 35 cycles of 30 s at 95 °C, a 1 min annealing at 52 °C and a 1 min extension at 72 °C. The amplification was completed by a final 7 min extension step at 72 °C. About 5 µl of the amplification product was electrophoresed on a 1.4% agarose gel run at 9 V/cm for 60 min to control amplification success.

Quantitative allele-specific real-time PCR

The amplification products were cleaned using spin columns (QIAquick PCR Purification Kit, Qiagen) and the DNA concentration of the cleaned amplification fragment was quantified on a photometer (BioPhotometer, Eppendorf). The allele-specific forward primer for amplification of the R-haplotype, TRIAS/Rf4 (GATTGATCTTCCAATAGGCGG), was designed to end at the diagnostic base and included a mismatch T to G at position 2 and another mismatch T to G at position 5 from the 3′ end. The reverse primer TRIAS/Rr1 (TGCTGATACCTAAAGCAGTGA) was a full match to the template region 103 bases upstream of the P2 primer. The concentration of the first-round amplification product was adjusted to 1 ng/µl and 0.5 µl corresponding to 0.5 ng template DNA was used together with 0.16 µM of each of the two primers TRIAS/Rf4 and TRIAS/Rr1 in a final volume of 25 µl for the allele-specific nested real-time PCR based on the CybrGreen PCR Master Mix (Applied Biosystems). Four replicates were prepared from each individual DNA sample as follows: 100 µl of the amplification mix containing everything except for the DNA was prepared and distributed in two vials. The DNA was pipetted into each of these two vials to enable assessment of variation caused by pipetting errors. Each of the two vials was mixed well and its contents further distributed in two replicate 0.2 µl amplification tubes (to test variation of the system itself) resulting in a total of four replicates per DNA sample. The quantitative PCR was performed on a Applied Biosystems ABI5700 sequence detection system with an initial 2 min at 50 °C and 10 min 95 °C, followed by 50 cycles of 15 S at 95 °C and 1 min at 60 °C. Standards of 100, 25, 5, and 1 % R-haplotype were prepared as follows: clones of R- and S-plants were sequenced to confirm the haplotype. The clones were amplified using the same procedure as above, i.e., using primers P1 and P2 to obtain the fragment encompassing the resistance mutation, cleaning and quantifying the amplification products and diluting them to 1 ng/µl. This DNA was then mixed at the appropriate amounts to generate the above concentrations for the controls. Four replicates per concentration, prepared as described above for the plant samples, were included in each amplification for accurate quantification. A fifth standard with 0% R was also included, again with four replicates. The selectivity of Taq polymerase with respect to single base mismatches is dependent on its strong bias against mismatched primers (Rhodes et al., 1997). This bias is not absolute and therefore, the allele-specific PCR always produced a (false-positive) product after ca. 35 cycles, which represented the system's limit of sensitivity to detect small amounts of R-haplotype in a S-haplotype background. Consequently, the concentration of 0% R was not used for establishing the standard curve. Data below 1% R are thus extrapolated values; if they were equal or smaller than the 0% R control value (calculated as the median quantity of the four replicates of each standard concentration) then they were set to zero. Extrapolation below 1% R was linear down to at least 0.2% R as the deviation of the measured ct-value of 0.2% R (median ct = 29.700) from the ct-value expected by extrapolation (ct = 29.932) was less than 1%.

Data analysis

The data were analyzed with the in-built software of the ABI5700 sequence detection system according to the manufacturers recommendations. The statistical analysis was performed with the JMP V.5 software using nested univariate analysis of variance (ANOVA) based on the square-root arcsin transformed proportions (Sokal & Rohlf, 1981).

Results

Origin of assay variation

The variation observed in the detection limit was the result of small errors of the analysis itself and/or of the reaction preparation steps. Our four

replicates per reaction were composed of two 'pipetting' replicates and two 'chemical' replicates per reaction. The difference between the 'pipetting' replicates was that the reaction mix, after careful mixing, was distributed into two different vials before the DNA was pipetted into each of them separately. Thus, variation observed between these two replicate groups may be caused by pipetting errors, a well-known error source. In contrast, the 'chemical' replicates differed only in that each of the two reaction mixes containing the DNA was further distributed, again after careful mixing, into two different reaction tubes, resulting in a total of four reaction tubes with the same components. Possible error sources between the 'chemical' replicates were thus restricted to difference in location of the tubes in the thermocycler and/or stochastic variation in enzyme kinetics. Therefore, variation between the 'chemical' replicates should be very small. As expected, the variation between the 'chemical' replicates (average percent difference in the quantity calculated for the two replicate pairs ± STD; 14.1 ± 7.5) was lower than that of the 'pipetting' replicates (percent difference between the averages of the two replicate pairs ± STD; 24.5 ± 13.3; data not shown). A paired *t*-test comparing these differences between the replicates of each DNA sample indicates that pipetting contributes significantly to the variation due to the summed effects of thermocycler position and stochastic chemical effects ($df = 41$, $t = -4.321$, $p < 0.0001$).

Precision and sensitivity of the assay

The precision of the assay was in average 13.3 ± 9.4 % (ranging from 0.01 to 46.20%), as determined by calculating the percent deviation from the average quantity measured for each DNA sample, averaged over all DNA samples (data not shown). To optimize the performance of our assay, we included a full set of standards on each 96-well microplate used in our real-time PCR experiments (as opposed to applying standard curves across plates). The detection limits, i.e., the sensitivity of the assay (defined as the median of the 0% R values), of the three 96-well real-time PCR experiments were 0.10 ± 0.02% R (plate 1, plants 1–3), 0.36 ± 0.05% R (plate 2, plants 4–7), and 0.09 ± 0.01% R (plate 3, plants 8 and 9), respectively, averaging 0.18 ± 0.15% R. The lower detection limit, i.e., the sensitivity of our assay was thus variable and ranged from 0.10 to 0.41% R (average ± standard deviation).

Variation in heteroplasmy levels between plants

The overall variation in the level of heteroplasmy, i.e., the percentage of R-haplotypes among the entire population of haplotypes, between the tested plants was very strong ($SS = 24.566$, $F = 8.67$, $p < 0.0001$; $df = 8$). The main reason for this is that the plants form two distinct groups with respect to this characteristic, i.e., one with high and a second with very low heteroplasmy levels (Figure 1(a and b)).

Variation in heteroplasmy levels within plants

Significant variation in the degree of heteroplasmy was also observed within plants. Two-thirds of all tested plants, i.e., plants F_01 and F_05 to F_09, show highly significant variation in the % R between individual leaves of each plant (Table 1 and Figure 1). Only low levels of variation were detected in the remaining plants, i.e., plants F_02 to F_04 (Table 1). The lack of within-plant variation may be partly due to the fact that these plants belong to the low-level heteroplasmy group, however, significant variation was also observed in plant F_05 that actually showed only low heteroplasmy levels. These results suggest that variable levels of heteroplasmy may be common.

Variation in heteroplasmy levels within leaves

Within-leaf variation in heteroplasmy levels was tested in only two plants and was highly significant in both cases (F_08: $SS = 2.418$, $F = 33.33$, $p < 0.0001$; F_09: $SS = 2.442$, $F = 34.87$, $p < 0.0001$; $df = 4$ for both plants).

Discussion

Although it is now well established that heteroplasmy occurs in plants, it has not so far been quantitatively assessed. In an earlier study, PCR–RFLP was used to study cpDNA heteroplasmy in *S. vulgaris* (Frey, 1999). However, for several reasons outlined in the Introduction section, this

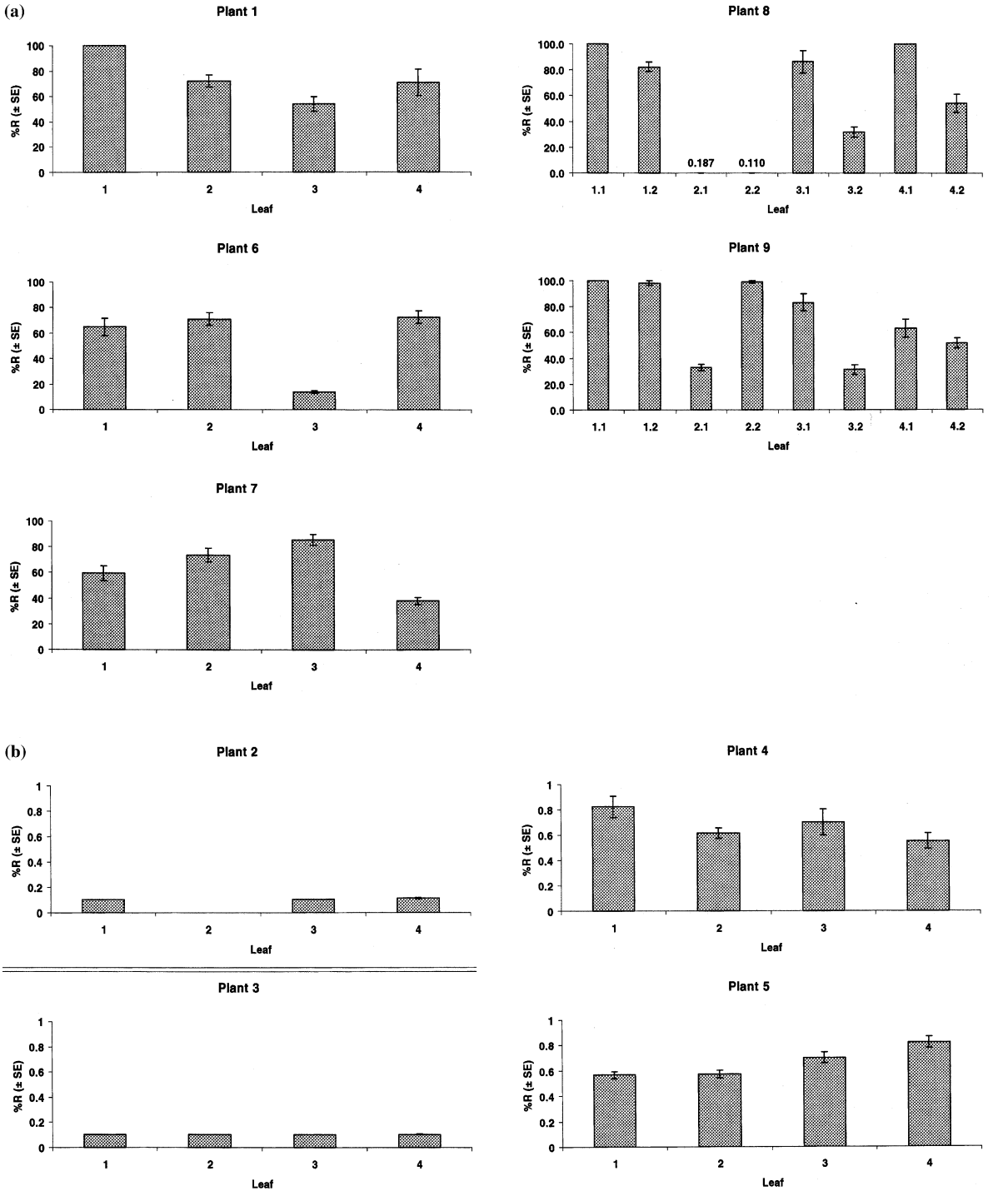

Figure 1. Percent R-haplotype of individual tissue samples from different *S. vulgaris* plants demonstrating between-plant as well as within-plant variation in heteroplasmy levels. (a) plants with high levels of the R-haplotype; (b) plants with low levels of the R-haplotype.

Table 1. ANOVA analysis of within-plant variation of the level of cpDNA heteroplasmy in *S. vulgaris*

Plant	df	SS	*F*	*P*
F_01	3	1.20819620	11.2329	0.0008
F_02	3	0.00000678	2.9271	0.1049
F_03	3	0.00000092	0.9586	0.4437
F_04	3	0.00058971	2.3176	0.1273
F_05	3	0.00064760	10.4763	0.0011
F_06	3	1.13405210	29.3859	< 0.0001
F_07	3	0.62196462	16.1480	0.0002
F_08	3	6.33091710	116.3517	< 0.0001
F_09	3	2.01119320	38.2858	< 0.0001

df = degrees of freedom; SS = sum of squares.

method is not adequate for true quantitative analysis of heteroplasmy levels within plants.

In the present study, we show that within-plant heteroplasmy is widespread in the sampled population of *S. vulgaris* and that the levels of heteroplasmy encountered both within and between individual plants covers a broad range. Variation in the fraction of resistant haplotypes within a single plant can be below 1% yet may reach levels of up to 100% (see for example plant 8 in Figure 1(a)). The quantitative analysis of the field plants revealed patterns of heteroplasmy that very closely match those reported previously on laboratory-reared specimens using PCR–RFLP as a semi-quantitative technique (Frey, 1999; Frey et al., 1999).

Heteroplasmy in *S. vulgaris* has now been confirmed by two different methods and in greenhouse-grown plants originally collected in Washington and Oregon, USA (Frey et al., 1999) as well as in field plants from an apple orchard near Waedenswil, Switzerland, suggesting that it may be a common phenomenon. The observed pattern of within-plant variation may have two consequences of potentially high significance. First, with respect to molecular diagnostics of chloroplast-DNA encoded traits. Our study shows that if only small plant parts are used for genotyping, the results for such traits may, in essence, be completely contradictory depending on the specific situation at the location of the sample. For example, sampling leaf 1 of plant 8 would suggest a highly resistant plant individual, whereas if sampling leaf 2 of the same plant individual would indicate the exact opposite, i.e., a highly susceptible plant. Thus, although PCR technology allows using very small tissue samples, it would be advisable to routinely sample several leaves of a single plant to enable a rough estimate of its overall chloroplast genotypic status. Second, the observed pattern of within-plant variation in heteroplasmy levels may have evolutionary consequences. The co-occurrence within an individual plant of two or more chloroplast haplotypes coding for different variants of an enzyme may represent a system to allow fine-tuned adaptation to specific environmental conditions within the lifetime of a plant individual, as was proposed by Frey (1999). As the mutation conferring resistance to triazine imposes fitness-costs (Gressel & Segel, 1990; Jasieniuk et al., 1996) it should be expected to go extinct unless selection to promote it occurs. If selection does occur, the frequency of resistant haplotypes is expected to increase. A further expectation is that this increased proportion of resistant haplotypes is transmitted to a large fraction of the next generation, i.e., that epigenetic inheritance acts on chloroplast (and possibly mitochondrial) haplotype frequencies. In a preliminary study to investigate the genetics of heteroplasmy in *S. vulgaris,* we found that the progeny of heteroplasmic mothers is also heteroplasmic (data not shown). However, we found no effect of selection on the frequency of the resistant haplotype in triazine-treated plants. These preliminary assays were performed with plants that possessed only low levels ($< 1\%$) of resistant haplotypes which may have been below the frequency necessary for survival of a triazine dose that would kill susceptible individuals, i.e., we may have used too soft selection to see an effect. We intend to repeat this experiment with plants containing intermediate levels of both haplotypes and to extend it to test if the progeny of selected plants show increased levels of resistant haplotypes.

Acknowledgements

We thank Rolf Blapp for providing us with the field samples and Monika Pfunder, Paul Wolf and an anonymous referee for valuable comments.

References

Armstrong, K. & W.R. Bauer, 1982. Preferential site-dependent cleavage by restriction endonuclease Pstl. Nucleic Acids Res. 1982: 993–1007.

Birky, C.W. Jr., P. Fuerst & T. Maruyama, 1989. Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. Genetics 121: 613–627.

Chat, J., S. Decroocq, V. Decroocq & R.J. Petit, 2002. A case of chloroplast heteroplasmy in Kiwifruit (*Actinidia deliciosa*) that is not transmitted during sexual reproduction. J. Hered. 93: 293–300.

Chen, G.K., N.J. Lacayo, G.E. Durán, Y. Wang, C.D. Bangs, S. Rea, M. Kovacs, A.M. Cherry, J.M. Brown & B.I. Sikic, 2002. Preferential expression of a mutant allele of the amplified *MDR*I (*ABC*I) gene in drug-resistant variants of a human sarcoma. Gene Chromosome Canc. 34: 372–383.

Cheung, W.Y., J-C. Côté, D.L. Benoit & B.S. Landry, 1993. A rapid assay for chloroplast-encoded triazine resistance in higher plants. Plant Mol. Biol. Rep. 11: 142–155.

Darmency, H., 1994. Genetics of herbicide resistance in weeds and crops, pp. 262–297 in Herbicide Resistance in Plants, edited by S.B. Powels & J.A.M. Holtum. CRC Press, Boca Raton Fla.

Fitter, J.T., M.R. Thomas, R.J. Rose & N. Steele-Scott, 1996. Heteroplasmy of the chloroplast genome of *Medicago sativa* L. cv 'Regen S' confirmed by sequence analysis. Theor. Appl. Genet. 93: 685–690.

Frey, J.E., 1999. Genetic flexibility of plant chloroplasts. Nature 398: 115–116.

Frey, J.E. & B. Frey, 2004. Origin of intra-individual variation in PCR-amplified mitochondrial cytochrome oxidase I of *Thrips tabaci* (Thysanoptera: Thripidae): mitochondrial heteroplasmy or nuclear integration? Hereditas 140: 92–98.

Frey, J.E., H. Müller-Schärer, B. Frey & D. Frei, 1999. Complex relation between triazine-susceptible phenotype and genotype in the weed *Senecio vulgaris* may be caused by chloroplast DNA polymorphism. Theor. Appl. Genet. 99: 578–586.

Gressel, J. & L.A. Segel, 1990. Modelling the effectiveness of herbicide rotations and mixtures as strategies to delay or preclude resistance. Weed Technol. 4: 186–198.

Gruber, J.D., P.B. Colligan & J.K. Wolford, 2002. Estimation of single nucleotide polymorphism allel frequency in DNA pools by using pyrosequencing. Hum. Genet. 110: 395–401.

Gyllensten, U. 1989. Direct sequencing of *in vitro* amplified DNA, pp. 45–60 in PCR Technology – Principles and Applications for DNA Amplification, edited by H.A. Ehrlich. Stockton Press, London.

Hirschberg, H., A. Bleecker, D.J. Kyle, L. McIntosh & C.J. Arntzen, 1984. The molecular basis of triazine resistance in higher plant chloroplasts. Z. Naturforsch. 45c: 412–419.

Jasieniuk, M., A.L. Brûlé-Babel & J.N. Morrison, 1996. The evolution and genetics of herbicide resistance in weeds. Weed Sci. 44: 176–193.

Lo, H.S., Z. Wang, Y. Hu, H.H. Yang, S. Gere, K.H. Buetow & M.P. Lee, 2003. Allelic variation in gene expression is common in the human genome. Genome Res. 13: 1855–1862.

Mitich, L.W. 1995. Common groundsel (*Senecio vulgaris*). Weed Technol. 9: 209–211.

Müller-Schärer, H. & G.S. Wyss, 1994. Das Gemeine Kreuzkraut (*Senecio vulgaris* L.): Problemunkraut und Möglichkeiten der biologischen Bekämpfung. Z. Pflanzenkrank. Pflanzenschutz 14: 201–209.

Rhodes, C.H., C. Honsinger, D.M. Porter & G.D. Sorenson, 1997. Analysis of the allele-specific PCR method for the detection of neoplastic disease. Diagn. Mol. Pathol. 6: 49–57.

Sham, P., J.S. Bader, I. Craig, M. O'Donovan & M. Owen, 2002. DNA pooling: a tool for large-scale association studies. Nat. Rev. Genet. 3: 862–871.

Sokal, R. & F. Rohlf, 1981. Biometry, 2nd edn. Freeman, San Francisco, CA.

Thomas, M. & R.W. Davis, 1975. Studies on the cleavage of bacteriophage DNA with EcoRI restriction endonuclease. J. Mol. Biol. 91: 315.

Werner, M., M. Sych, N. Herbon, T. Illig, I.R. König & M. Wjst, 2002. Large-scale determination of SNP allele frequencies in DNA pools using MALDI-TOF mass spectrometry. Hum. Mutat. 20: 57–64.

Yan, H., W. Yuan, V.E. Velculescu, B. Vogelstein & K.W. Kinzler, 2002. Allelic variation in human gene expression. Science 297: 1143.

Yu, W. & R.J. Spreitzer, 1992. Chloroplast heteroplasmy is stabilized by an amber-supressor tryptophan tRNACUA. Proc. Natl. Acad. Sci. USA 89: 3904–3907.

Zhou, G.-H., M. Kamahori, K. Okano, G. Chuan, K. Harada & H. Kambara, 2001. Quantitative detection of single nucleotide polymorphisms for a pooled sample by a bioluminometric assay coupled with modified primer extension reactions (BAMPER). Nucleic Acids Res. 29, 19 *e93*.