Optimisation of a microsatellite panel for the individual identification of brushtail possums using low template DNA

Juan F. Dueñas[*], Robert Cruickshank and James Ross

Department of Ecology, Faculty of Agricultural and Life Sciences, Lincoln University, PO Box 84, Lincoln 7647, New Zealand
[*]Author for correspondence (Email: jduenas@yachaytech.edu.ec)

Abstract: The Australian brushtail possum *Trichosurus vulpecula* is a pervasive marsupial pest of New Zealand. Impacting on the native flora and fauna and the nation's livestock industry as a vector of bovine tuberculosis, *T. vulpecula* is a priority for control and eventual eradication. Possum control at present relies on conventional trapping and poisoning methods. Efficient allocation of control depends on accurate quantification of abundance, which could be achieved with the implementation of non-invasive sampling schemes. We evaluated the use of salivary DNA retrieved remotely as a source of DNA for microsatellite amplification. A panel of six loci were optimised using tissue samples from possums from three locations in the Canterbury Region, South Island, New Zealand. Optimised loci were then assembled into a multiplex PCR assay. Microsatellite diversity patterns revealed moderate to high polymorphism and heterozygosity, and a sufficiently low overall probability of identity adjusted for siblings ($PI_{sib} = 3.0 \times 10^{-3}$) to ensure a robust identification of individual possums based on their multi-locus genotype. While dilution of DNA extracted from tissue did not affect the results, the use of DNA from saliva significantly decreased the performance of the microsatellite amplification system. Altogether these results indicate that locus characteristics (i.e. amplicon size) and DNA quality are crucial factors affecting the sensitivity and reliability of this method.

Keywords: invasive species; population genetics; saliva; *Trichosurus vulpecula*; wildlife management

Introduction

The implementation of advanced monitoring methods for improved pest management of the Australian brushtail possum *Trichosurus vulpecula* is of particular importance for New Zealand. As a vector of bovine tuberculosis *T. vulpecula* constitutes a significant threat to the country's livestock industry (Montague & Warburton 2000). As an introduced species, a series of negative interactions of *T. vulpecula* with New Zealand indigenous biota have been documented (Cowan 1992; Cowan & Clout 2000). In response to these issues, extensive control operations are commonly undertaken in New Zealand rural areas (O'Reilly-Wapstra & Cowan 2010), primarily by aerial delivery of sodium fluoroacetate (1080) or ground-based deployment of toxic baits and traps. Once control operations take place, it is necessary to determine their effectiveness in order to guide subsequent control efforts and evaluate the work of private contractors (Warburton 2000). This requires a monitoring method that accurately estimates population size, or an index of abundance before and after control. In order to estimate abundance, the National Pest Control Agencies (NPCA) currently uses the residual trap catch-index (RTCI) (Warburton et al. 2004). The RTCI is a standardised index of abundance based on the number of possums captured in leg-hold traps, and is statistically robust. However, the RTCI suffers from several logistical and analytical limitations (Forsyth et al. 2005). Despite these limitations, equally robust and more cost effective methods for estimating possum abundance are currently unavailable for widespread adoption (but see Ruffell et al. 2015).

A possible alternative to the RTCI as a possum monitoring standard makes use of interference devices such as wax tags, and is under development by the NPCA (National Possum Control Agencies 2010). A wax tag is a small wax block attached to a plastic visual lure. When compared with leg-hold traps these devices are more effective at detecting possums while the indices of abundance derived from wax tags (i.e. tag station index and the bite-mark index) are consistent with RTCI (Thomas et al. 2003; Ogilvie et al. 2006). However, the accuracy of the indices derived from interference devices has also been questioned on several grounds: (1) wax tags appear to lose sensitivity in detecting population size changes when possum populations are either extremely large or small (Warburton et al. 2004; Morgan et al. 2007), and (2) the use of lures to attract possums has been shown to trigger a behavioural response known as contagion. Contagion occurs when an individual actively seeks and bites several different interference devices on a given sampling occasion, which if unnoticed will bias the bite-mark index estimate (Warburton 2000).

In an attempt to further develop the capacities of interference devices, Vargas et al. (2009) were able to amplify the barcoding region of the mitochondrial cytochrome oxidase sub-unit I (CO-I) gene from DNA found in traces of saliva collected using wax tags. A challenging limitation in using saliva from interference devices is that the DNA retrieved can often be of low quantity (< 20 ng/µL; sensu Morin et al. 2010) and/or quality (i.e. degraded DNA consisting of short fragments), yet this remains a promising option given that preliminary studies have established major logistical constraints on obtaining possum DNA from hair follicles and faeces (Gleeson et al. 2003; Morgan et al. 2007; but see Ramón-Laca & Gleeson 2014).

The ability to collect salivary DNA presents a new opportunity to test the potential of wax tags to improve allocation of possum control operations. When DNA of sufficient quantity and quality can be retrieved in order to conduct reliable genotyping assays, the census population size (N_C) can be directly determined. To achieve this goal, three major issues were addressed via a detailed pilot study. First, using a set of microsatellites developed in other studies, it was necessary to select the most informative markers to constitute a microsatellite panel with the capacity to identify individual possums. Parallel to this objective, microsatellites were multiplexed to reduce processing cost, while the reproducibility and reliability of the method was also assessed. Finally, a quality control system was implemented to monitor genotyping error.

Materials and methods

Study area and DNA collection

All tissue samples were collected in the Canterbury Region of the South Island of New Zealand between June and July 2011 (Fig. 1). Ear tissue was collected from 37 dead possums captured in three areas across Canterbury: Banks Peninsula ($n = 9$), Lewis Pass ($n = 9$) and Hororata ($n = 19$). Twenty of the possums were male, 14 female, and 3 pouch-young of undetermined sex. Two samples of approximately 10 mm of skin and cartilage per individual were preserved independently in 1.5-ml tubes containing 1 ml of 99% EtOH.

During April 2012, a total of 24 wax tags were presented to captive possums kept in individual pens at the Centre for Wildlife Management and Conservation at Lincoln University. The captive possums were captured in the Canterbury Region between March and April 2012. The tags were left in the pens

overnight and retrieved and taken to the laboratory the following morning. Tissue samples were not available for these possums because they were sacrificed as part of toxic-bait experiments.

DNA isolation and preparation

Genomic DNA was isolated from possum tissue and saliva at the Molecular Ecology Laboratory at Lincoln University. DNA from ear-tissue was extracted using a DNeasy Blood & Tissue Kit (QIAGEN), following the manufacturer's instructions. In order to standardise DNA concentration, 100-µl aliquots from a subset of the DNA extracts (27 out of 37) were obtained and measured using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Inc., Wilmington, Delaware). The minimum DNA concentration measured was 22.61 ng/µl, so extracts were diluted with volumes of Buffer AE (QIAGEN) into samples containing approximately 22.61 ng/µl of DNA each.

Extraction of DNA from bitten wax tags followed the QIAGEN protocol for blood and cultured cells implemented by Vargas et al. (2009), with slight modifications. Teeth markings on the wax blocks were excised using a surgical blade. The resulting pieces of wax, coated with saliva, were stored in sterile 15-ml screw-cap tubes. Phosphate-buffered saline (14 ml, pH 7.4) was added to each tube and centrifuged at 1500 g for 10 min. The supernatant and the wax were then discarded. The remaining 200 µl were transferred to a 1.5-µl micro centrifuge tube. The remainder of the extraction protocol followed the manufacturer's instructions except that samples were eluted twice in 50 µl of Buffer AE, ultimately yielding 100-µl solutions.

DNA from samples of bitten wax tags was not quantified because the samples were assumed to be contaminated with exogenous DNA (e.g. bacterial). However, these samples
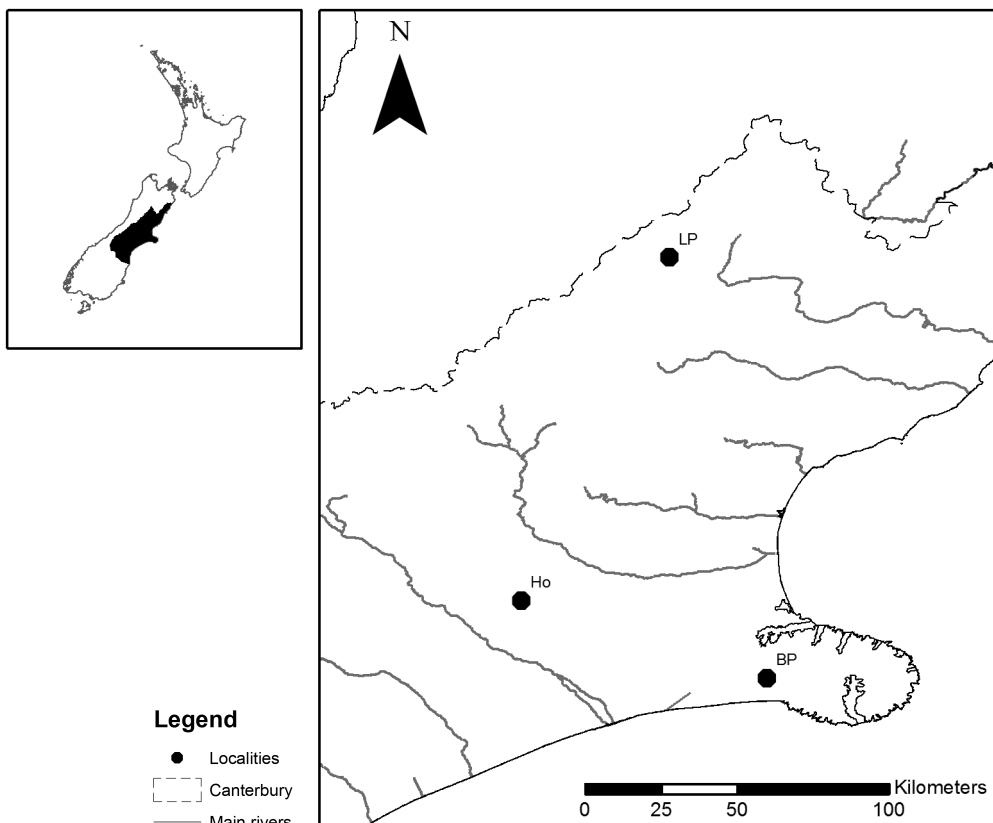
Figure 1. Geographic origin of samples collected in this study: BP is Banks Peninsula, Ho is Hororata, LP is Lewis Pass.

were screened for the presence of DNA by amplifying a fragment of undetermined length at the barcoding region of the mitochondrial cytochrome oxydase sub-unit I gene (CO-I). Amplicons were not further sequenced, since the aim was not to identify the taxon. All PCRs were performed in 10-µl reactions with a MultiGene TC9600-G Thermal Cycler (Labnet International, Inc., Edison, New Jersey) using an *i*-StarTaq kit (iNtRON). The thermocycler profile was the following: denaturation at 94°C for 2 min, then 33 cycles of denaturation at 94°C for 30 s, annealing at 45°C for 30 s and extension at 72°C for 1.30 min, followed by a final extension time at 72°C for 5 min. Reactions contained 1 µL of template, 2 mM of MgCl$_2$, 0.2 mM of each dNTP, 0.5 µM of primers MLepF1 (Hajibabaei et al. 2006) and HCO2198 (Folmer et al. 1994), and 0.4 U of polymerase.

Polymerase chain reaction (PCR) products were visualised by electrophoresis on 1.5% agarose gels. Samples with positive bands were then used as templates for microsatellite amplification.

Microsatellite panel optimisation

A panel of eight microsatellite loci were selected from the literature (Table 1). The 5′ end of the forward primers was labelled with a fluorescent dye (6-FAM, VIC, NED or PET; Applied Biosystems). Initial amplification was performed following protocols in Taylor & Cooper (1998) and Vargas et al. (2009), followed by empirical adjustment of MgCl$_2$ concentration, thermocycler profile and primer concentration.

Thirty-seven tissue extracts were amplified in 10-µl standard reactions (hereafter referred to as singleplex), using the *i*-StarTaq kit (iNtRON). Reactions contained ≈ 56.5 ng of DNA, 0.2–0.75 µM of forward and reverse primer solution, 2–3 mM of MgCl$_2$, 0.25 mM of each dNTP, and 0.4 U of polymerase. A minimum of two samples per surveyed region were randomly selected for blind replication of PCR reactions. The number of blind replicates represented approximately 20% (*n* = 8) of the total dataset.

Positive singleplex reactions were screened on 1.5% agarose gels. Capillary electrophoresis was carried out using an ABI PRISM® 3130xl sequencer (Applied Biosystems) with GeneScan™ 500 LIZ® Size Standard (Applied Biosystems). Fragments were visualised in a Peak Scanner™ v1.0 (Applied Biosystems) and results were then exported into the program Microsatelight (Palero et al. 2011), where binning and scoring was performed. Binning was undertaken using the AlleloBin (Prasanth et al. unpubl. ICRISAT report 2006) sub-routine. Finally, allele scoring was completed by visually inspecting all electropherograms and comparing these observations with results from the automated scoring routine.

Six primer-pair sets (Tv19, Tv27, Tv53, Tv58, Tv5.64 and TvM1) were then combined into two 100-µl solutions (i.e. one of forward primers and one of reverse primers). The same concentration in which a readable product was generated in singleplex reactions was used to build primer solutions. Multiplex Manager v.1.0 (Holleley & Geerts 2009) and empirical tests were performed to account for possible cross-reactivity of different primer pairs (data not shown). The combined primer solutions were used to amplify equalised DNA extracts in 10-µl reactions, using a Multiplex PCR Kit (QIAGEN). The reactions contained 56.5 ng of DNA, 1× of Multiplex PCR Kit master mix, 0.5 µl of RNase-free water, and 0.2–0.75 µM of each forward and reverse primer in the primer solution (see Table 1 for the final concentration of each primer). The thermocycler profile was the following: denaturation at 95°C for 15 min, then 25 cycles of denaturation at 94°C for 30 s, annealing temperature (57.8°C) for 90 s and extension at 72°C for 90 s, followed by a final extension time at 72°C for 10 min. PCR products were visualised with capillary electrophoresis and alleles were scored as described for the microsatellite panel optimisation.

To minimise contamination of samples, a series of protocol guidelines were implemented. Master-mix solutions were prepared under a UV-treated hood, stock solutions were stored as aliquots to prevent contamination of reagents, negative

Table 1 Characteristics of the eight loci selected to genotype possums.

Locus	Sequence 5'–3'	Size range[a]	Ta (ºC)	Pc	Source
TvM1	F:(6FAM)-GACCACAACCTGGGTCTAACCAACG R:CATGACACCTGGGCACTCAGGACT	195–223	55	0.5	Lam et al. (2000)
Tv27	F:(6FAM)-AGTGGAACCACATGTCAGGGC R:GGACTGAAATGACTGCACAAC	125–168	60–65	0.5	Taylor & Cooper (1998)
Tv53	F:(NED)-GGGAGTAGTTGTCTGAGTTCCC R:CCCTGGAGTTTGACAACCTG	233–263	60	0.2	Taylor & Cooper (1998)
Tv54	F:(NED)-GGGAGGCATAAAGTGCCAGA R:TGACCGACACTGACGACCCC	87–119	60	0.75	Taylor & Cooper (1998)
Tv5.64	F:(VIC)-TTTATCCCTACTAGAGGTAGGT R:ATTAGCGCTTACCAGAGTGC	122–168	55–60	0.5	Sarre et al. (2010)
Tv19	F:(PET)-CCTCCTCCCCATCCTTCCTG R:GTTCAATTGCAGGGCTATGG	214–254	55–60	0.5	Taylor & Cooper (1998)
Tv58	F:(PET)-GCACCCAAGGACCCCCAAGA R:CATATCACAGTGCTTGGCG	102–158	60	0.5	Taylor & Cooper (1998)
Tv16[b]	F:(VIC)-GAGGCTACCATTAGACGCAA R:AGATACTATCTGCATCCAGAG	83–115	–	–	Taylor & Cooper (1998)

[a] Allele size range is expressed in base pairs, Ta represents the optimal annealing temperature and Pc represents the final primer concentration expressed in moles per litre of solute (M).
[b] Locus Tv16 was excluded from the panel due to difficulties with PCR optimisation and its linkage to locus Tv27.

controls were included in each round of experiments to detect cross-contamination, and PCRs and extractions were performed on different days.

Evaluation of optimised protocol using low template DNA

Equalised DNA extracts obtained from possum tissue were subjected to a 1:6 dilution in Buffer AE. Using a NanoDrop ND-1000, diluted extracts were measured twice to provide an approximate range of their final concentration. These extracts, as well as genomic DNA obtained from bitten wax tags, were amplified using the Multiplex PCR protocol. The volume of reactions was increased to 12 µl, containing on average 10.4 ng of DNA, 1× Multiplex PCR Kit master mix, 0,5 µl of RNase-free water, and 0.2–0.75 µM of each forward and reverse primer in the primer solution. The number of PCR cycles in the thermocycler profile was increased to 40. PCRs were performed in duplicate.

Data analysis

Genetic diversity and the panel's ability to discriminate individuals

The number of alleles per locus, number of private alleles, observed (H_o) and expected heterozygosity (H_e) were estimated for each locus across sampling localities, using Genalex (Peakall & Smouse 2006). To detect departures from Hardy–Weinberg (HW) expectations, an HW exact probability test was performed (Guo & Thompson 1992; Weir 1996) in Genepop v.4.1.1 (Rousset 2008). Wright's F statistics were estimated by implementing the analysis of variance framework defined in Weir & Cockerham (1984) and Weir (1996), also in Genepop. To determine if F_{st} among localities significantly differed from zero, an exact test of population differentiation (100 000 Markov chain steps; Goudet et al. 1996) was performed using the program Arlequin (Excoffier & Lischer 2010).

As some of the *T. vulpecula* individuals collected for this study were likely to be related, the panel's discriminatory power was estimated using the probability of identity adjusted for siblings (PI_{sib}). PI_{sib} accounts for potential underestimation of PI and was estimated by the formulas given in Taberlet & Luikart (1999) and Waits et al. (2001).

Genotyping error characterisation and monitoring

Genotyping error parameters in the singleplex PCR dataset were estimated by comparing the multi-locus genotype of blindly replicated samples with their corresponding multi-locus genotype in the original dataset ($n = 8$). Multi-locus genotypes resulting from multiplex PCR were cross-referenced with their corresponding genotypes obtained with singleplex reactions ($n = 27$). As a result of this comparison, a consensus genotype was produced in which alleles were recorded only if they were observed in both the singleplex product and the multiplex product (Pompanon et al. 2005). The two-allele rule was adopted assuming that DNA obtained from tissue was of sufficient quality and quantity as to preclude the occurrence of allelic drop-out. Genotypes generated from the diluted extracts were compared with the consensus genotypes. A pair-wise comparison was performed between replicates of profiles obtained from wax tags ($n = 10$ pairwise comparisons between positive replicates). To avoid complications resulting from differences of automated binning routines, only raw allele-size estimates were used for comparison.

To assess which putative causes better explained the errors detected, discrepancies resulting from multi-locus genotype comparisons were categorised into two groups: (1) stochastic errors, which included allelic drop-out and false alleles; and (2) systematic errors, which covered scoring, contamination and sample confusion errors. Stochastic errors were accounted for according to the equations defined in Broquet & Petit (2004). Because of budgetary constraints, salivary samples could only be genotyped twice. Therefore, it was not possible to associate each error observed with a probable cause. The overall genotyping error was estimated according to the equations formulated by Pompanon et al. (2005).

To assess the compatibility of homozygote excess with the presence of null alleles, tests implemented by the program Microchecker (van Oosterhout et al. 2004) were applied to data from the localities that showed significant departures from HW equilibrium.

Reproducibility of results

The effect of implementing a multiplex PCR assay on the reproducibility of results was assessed. The mean and standard deviation of the difference between size estimates was estimated per locus. Allele size estimates obtained from singleplex and mutliplex approaches for the same individual were then plotted against allele size categories, assuming these represent the true size of each allele. The relationship between measurements was fitted with linear regression. Instances of sample confusion, contamination, allelic drop-out, and false alleles were excluded from the analysis.

Differences in size estimates between products of single versus multiplex amplifications for each locus were detected by analysis of covariance. A 95% confidence interval of linear regression parameters (i.e. slope and intercept) was estimated for each locus to assess whether the relationship between covariables had changed. The effect of replacing the amplification strategy (e.g. singleplex replaced by multiplex) on the mean error rate per locus was analysed by a generalised linear model with a binomial error distribution and logit link function. A binomial error distribution was used as the sample sizes were unbalanced and non-normally distributed.

Effects of reduced DNA quantity or quality on amplification success and frequency of error

Estimates of mean error rate per locus were classified into three different template DNA classes: (1) DNA of good quantity and quality (i.e. undiluted ear-tissue extracts); (2) DNA of lower quantity and good quality (i.e. diluted ear-tissue extracts); and (3) DNA of low quantity and quality (i.e. salivary DNA retrieved from wax tags). The effect of template DNA quantity or quality on the mean error rate per locus was analysed using a generalised linear model with a binomial error distribution and logit link function.

Sensitivity of the protocol to changes in the quantity and quality of template DNA was estimated by determining amplification success. The ratio of the number of positive and partially positive reactions to the total number of reactions attempted was estimated per locus. A positive reaction was defined as the observation of one allele per locus in at least four or more loci per sample. A partially positive reaction was defined if a minimum of one allele per locus was observed in fewer than four loci per sample.

The amplification success ratios thus obtained were then ranked by the length of the locus amplicon. Two categories were considered: loci producing amplicons < 200 base pairs and loci producing amplicons ≥ 200 base pairs. The interaction

of amplicon length and template DNA class and its effect on locus amplification success were analysed with a generalised linear model using a binomial error distribution and logit link function.

When generalised linear models indicated significant differences, pair-wise comparisons of the mean values were undertaken using Fisher's restricted LSD test at $\alpha = 0.05$.

All the analyses were conducted in R (R Development Core Team 2011) as implemented in the program RKWard (Rödiger et al. 2012).

Results

Protocol optimisation

Tissue samples yielded variable, but relatively high, DNA concentrations (median = 57.22 ng/µl, range: 22.61–168.00 ng/µl). Singleplex PCR conditions were successfully optimised for seven out of eight microsatellite primer pairs. Locus Tv16 generated electropherograms that were difficult to interpret and could not be optimised. Because of this difficulty, and given that locus Tv16 was previously reported to be linked to locus Tv27 (Taylor et al. 2004), it was excluded from the panel. The automated binning routine performed well for all markers, as revealed in the low magnitudes of normalised variance estimates (Idury & Cardon 1997; data not shown).

Overall, the remaining loci in the panel showed moderate levels of allelic diversity (mean ± SE: 5.57 ± 0.5) and H_o (0.64 ± 0.06). The only exception was locus Tv54, which showed low diversity across localities (Table 2). Loci in two of the sampled localities showed significant departures from HW equilibrium (Hororata $\chi^2 = 33.98$; d.f. = 14; $P = 0.0021$, and Lewis Pass $\chi^2 = 23.11$; d.f. = 12; $P = 0.026$). When these results were partitioned in locus/locality combinations, 6 out of 21 combinations showed significant departures from HW expectations (Table 2). However, after the sequential Bonferroni correction (Holm 1979), only locus Tv58 in Hororata continued to exhibit a significant departure ($P = 0.019$).

There was little partitioning of genetic diversity among localities (multi-locus $F_{st} = 0.03$). The exact test of population differentiation showed that F_{st} did not differ significantly from zero ($P = 0.179$). However, a detailed locus-by-locus analysis of F_{st} parameters revealed moderate levels of genetic structure for locus Tv19 (Table 2).

The PI_{sib} estimated for all localities was 3.0×10^{-3}. As an expected consequence of its low allelic diversity, exclusion of locus Tv54 from the panel did not affect the magnitude of PI_{sib} significantly (3.9×10^{-3}). This indicates that a panel of six loci is able to provide sufficient information to identify individual possums on the basis of their multi-locus genotype. The PI_{sib} value estimated here is below the cut-off that could cause the shadow effect (≥ 0.01; Mills et al. 2000).

Genotyping error

Variations in protocol methodology did not increase the incidence of genotyping error. Neither the implementation of a multiplex assay ($\chi^2 = 0.65$, d.f. = 1, $P = 0.42$) nor the reduction of template concentration ($\chi^2 = 2.66$, d.f. = 1, $P = 0.1$) had any significant effect on the mean error rate per locus. However, the use of DNA retrieved from saliva as the template for microsatellite amplification did produce a significant increase in mean error rate per locus ($\chi^2 = 16.26$, d.f. = 2, $P < 0.001$) (Fig. 2).

The degree to which each error type affected the generation of genotypes varied according to the protocol implemented. Table 3 reports the incidence of stochastic and overall genotyping error found using four different protocols.

In singleplex PCR reactions, the average mean error rate per locus (e_l) was 4.76% yielding an overall high observed error rate per multi-locus genotype (e_{obs}) (33.33%), while in multiplex PCR reactions (using the undiluted templates) the average e_l was 7% and the e_{obs} dropped to 25%. When diluted DNA was used as the template for multiplex reactions, e_l increased to 10%. Notably, as the observed discrepancies were concentrated in just 15 of these samples, the e_{obs} only increased to 30%.

Genotyping error in singleplex amplifications was associated in equal measure with both stochastic and systematic causes. In contrast, when multiplex PCR was implemented, the incidence of allelic drop-out remained unaltered but the rate of false alleles increased by 5%. Visual inspection of the profiles generated with multiplex PCR confirmed that the majority of these discrepancies were due to stuttering peaks that were mistakenly scored as true alleles. Diluted DNA samples were more affected by allelic drop-out than false alleles. Allele mobility shifts were detected, but only affected locus Tv19. Finally, seven instances of contamination were detected.

No loci from Hororata or Lewis Pass showed evidence of the presence of null alleles according to Microchecker tests. Banks Peninsula data were not tested with Microchecker.

Table 2 Genetic diversity of the six loci successfully amplified with singleplex PCR for the genotyping of Trichosurus vulpecula samples obtained in Canterbury Region.

Locus[b]	Banks Peninsula[a]								Hororata[a]								Lewis Pass[a]								
	n	A	PA	H_o	H_e	F_{is}	HW P	PI_{sib}	n	A	PA	H_o	H_e	F_{is}	HW P	PI_{sib}	n	A	PA	H_o	H_e	F_{is}	HW P	PI_{sib}	F_{st}
Tv19	9	4	0	0.78	0.64	-0.80	0.01	0.48	19	5	0	0.79	0.67	-0.04	0.01	0.46	9	6	1	0.78	0.80	0.34	0.01	0.37	0.11
Tv27	9	6	0	0.67	0.77	-0.21	0.31	0.39	19	10	2	0.79	0.80	0.12	0.28	0.37	9	9	1	0.89	0.85	0.22	0.16	0.34	0.04
Tv53	9	7	0	0.89	0.80	0.17	0.54	0.37	19	8	0	0.79	0.75	-0.14	0.39	0.40	9	6	0	0.56	0.80	0.23	0.01	0.37	0.03
Tv54	9	2	0	0.11	0.10	-0.14	1.00	0.90	19	3	1	0.58	0.43	-0.26	0.72	0.64	9	1	0	0.00	0.00	0.00	--	1.00	0.07
Tv58	9	4	0	0.22	0.38	0.42	0.03	0.66	19	8	2	0.42	0.66	0.45	0.00	0.46	9	8	3	0.67	0.66	0.02	0.65	0.45	-0.04
Tv5.64	9	4	1	0.44	0.60	-0.33	0.78	0.51	19	4	0	0.79	0.67	0.16	0.47	0.46	9	4	0	0.78	0.70	-0.20	1.00	0.44	0.00
TvM1	9	5	0	0.89	0.77	-0.12	1.00	0.39	18	7	0	0.83	0.78	0.06	0.83	0.38	9	6	1	0.78	0.80	-0.03	0.79	0.37	0.00

[a]Banks Peninsula $PI_{sib} = 7.9 \times 10^{-3}$, Hororata $PI_{sib} = 3.4 \times 10^{-3}$ and Lewis Pass $PI_{sib} = 3.3 \times 10^{-3}$
[b]Sample size (n), number of alleles (A), number of private alleles (PA), observed (H_o) and expected (H_e) heterozygosity, inbreeding coefficient (F_{is}), the HW exact probability test P-values (HW P), and the probability of identity adjusted for siblings (PI_{sib}) for every locus. The population structure coefficient (F_{st}) is also included.
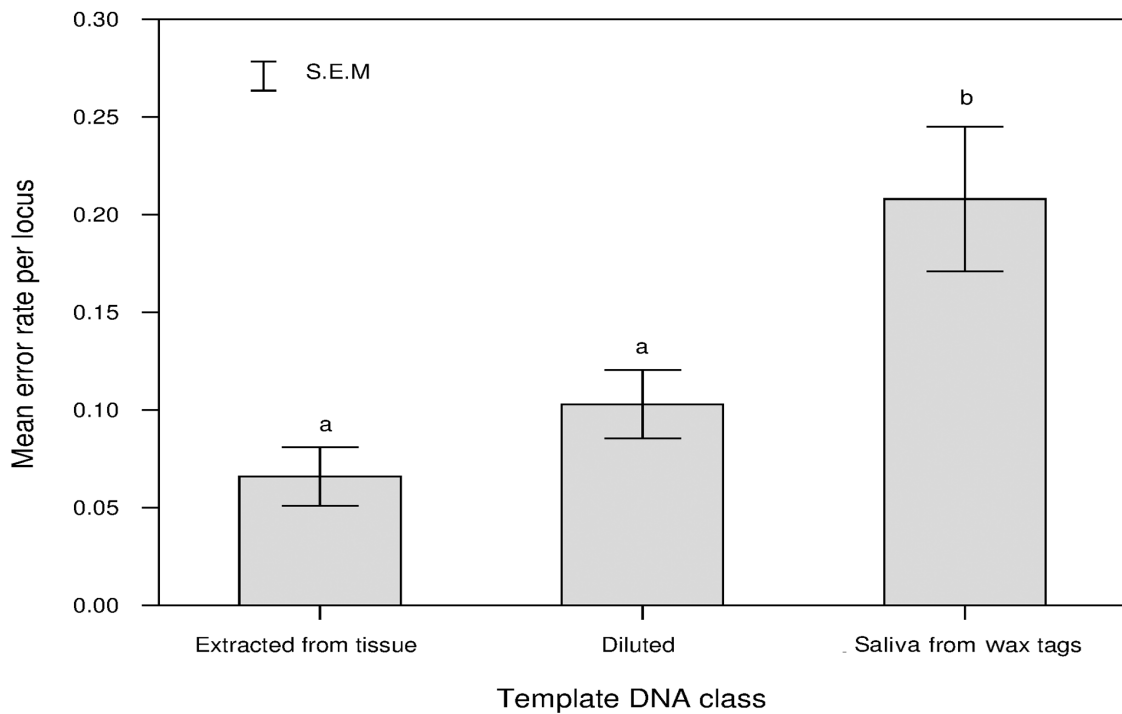
Figure 2. Increase in mean error rate per locus (e_l) according to the quantity and quality of template DNA used. Letter 'b' indicates significant differences between means according to Fisher's LSD test ($\alpha = 0.05$).

Reproducibility of results

Capillary electrophoresis produced different allele-size estimates for fragments amplified by singleplex and multiplex PCR. Locus Tv53 showed the largest mean difference between estimates (mean, range: 0.95, −2.05 to 1.45) followed by loci Tv58 and TvM1 (0.68, −1.42 to 0.75; 0.11, −0.44 to 0.30, respectively). The largest difference observed between size estimates for a particular allele was 5.75 base pairs at locus Tv19, although the mean difference was small (0.06, −5.75 to 5.58). We attribute the large difference observed to a mobility shift, possibly originated by the submission of that particular sample to a different sequencer. Cross-platform differences in the chemical conditions of capillary electrophoresis could be the cause of the observed large difference, as has been suggested in the literature (e.g. Moran et al. 2006). Loci Tv27 and Tv5.64 size estimates were more consistent regardless of the amplification method (0.07, −0.37 to 0.35; 0.02, −0.98 to 0.96, respectively).

There were significant covariate effects across all loci and a significant treatment effect at three loci (Tv53, Tv58 and TvM1). While a significant covariate effect simply suggests a single regression line regardless of the factor level, the significant treatment effect (i.e. singleplex vs multiplex) indicated that the y-intercepts of the linear regressions for the above-mentioned three loci are different (Fig. 3). Intercept differences confirm that size estimates obtained from different amplification regimes were indeed different. However, examination of the y-intercept confidence intervals of regression lines at loci Tv53, Tv58 and TvM1 overlap, suggesting the relationship with the covariate is still approximately 1:1 and the amplification method is not important.

### Evaluation of protocol using low template DNA

Diluted DNA concentration was low and differed slightly on both measuring occasions (median, range: 2.61 ng/µl, 0.36–5.2 ng/µl; 2.62 ng/µl, 0.28–5.7 ng/µl, respectively); however, the difference was not significant ($t = -0.69$, $P = 0.49$). Use of low-template-DNA concentration and quality affected the sensitivity of the multiplex protocol. Of the 27 diluted extracts, 26 (96, 29%) consistently showed positive peaks for four or more loci. The only sample failing to show positive peaks for all loci had a DNA concentration of ≈ 1.93 ng/µl. We attribute reaction failure to low target DNA concentration during PCR set-up, possibly due to inadequate sample homogenisation. Amplification success varied among loci, with loci Tv53 and Tv19 the most affected by reaction failure (Table 3). In contrast, of the 24 salivary samples, only 18 (74.9%) were positive for the CO-I DNA mini-barcoding region. When these 18 samples were used as template DNA for multiplex reactions, only 10 (55.5%) generated positive profiles. The remaining 8 (44.4%) generated partially positive genotypes.

Amplification success of DNA extracted from salivary samples was statistically lower than that of diluted DNA ($\chi^2 = 5.68$, d.f. = 1, $P = 0.01$). Similarly, amplification success of alleles >200 base pairs was significantly lower than alleles <200 base pairs. The post hoc comparison of means indicated that the interaction between amplicon length and template DNA class was highly significant for samples retrieved from wax tags ($\chi^2 = 31.9$, d.f. = 3, $P < 0.001$).

Discussion

The development of methodologies that allow accurate monitoring of shifts in population size provides additional tools to manage and eradicate pests. The combination of DNA genotyping technology with the ability to retrieve DNA non-invasively promises to overcome the limitations of traditional monitoring methods, particularly when targeted species are elusive. Our results indicate that it is possible to implement an
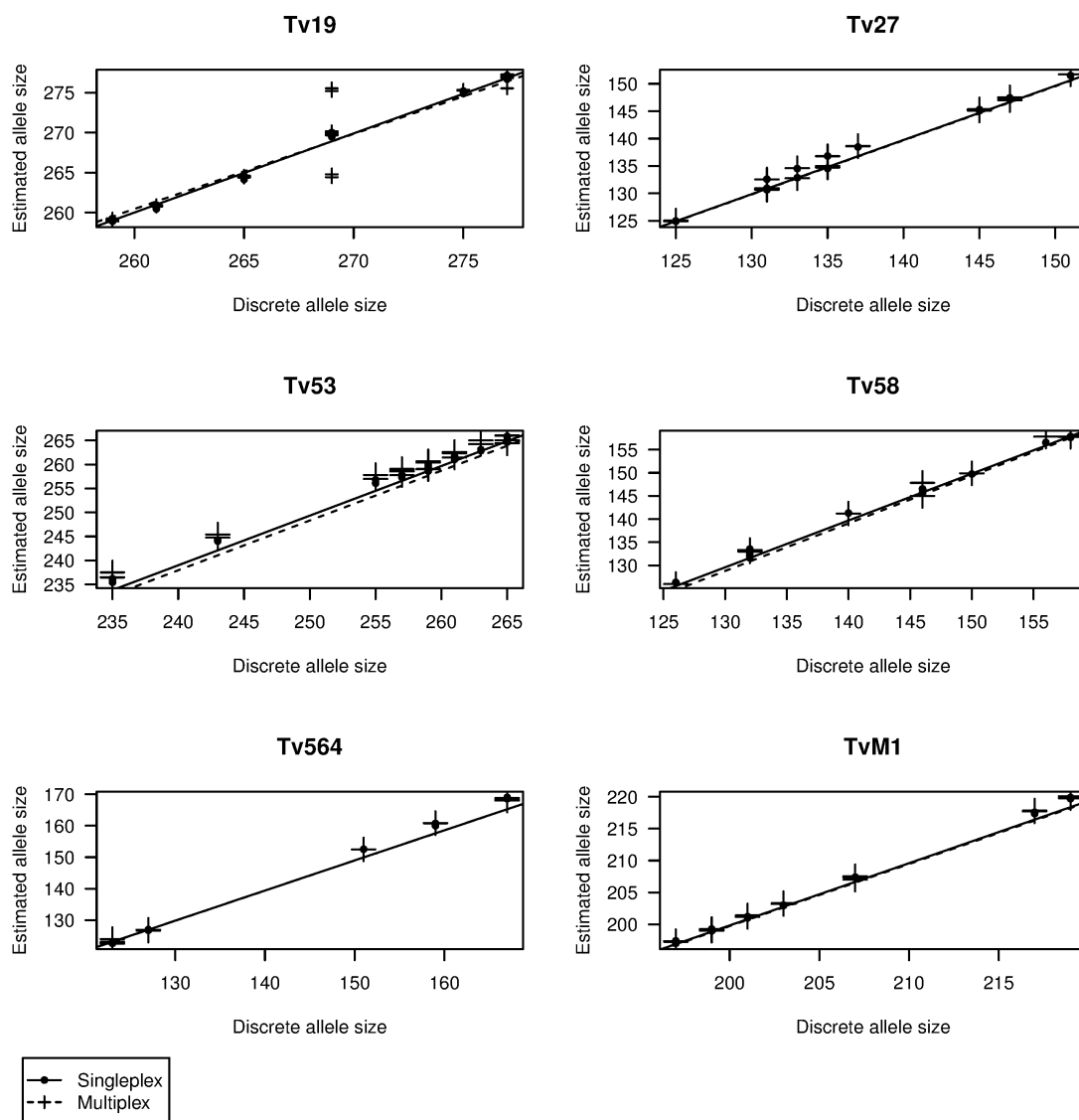
Figure 3. Linear regression models of allele-size estimates of products generated by singleplex and multiplex PCR approaches versus discrete allele categories.

efficient, reproducible, and reliable microsatellite amplification system capable of identifying individual possums. However, the sensitivity and reliability of the methodology are compromised when template DNA is not of sufficient quantity and quality. The reduced amplification success and increased incidence of genotyping errors observed here could be related to factors such as the presence of exogenous DNA (Herráez & Stoneking 2008), PCR inhibitors (Alaeddini 2012), or the inability of interference devices to collect enough DNA to override the deleterious effects of these other factors.

The patterns of genetic diversity observed are in agreement with Taylor et al. (2004) who concluded that the introduction of possums to the South Island occurred repeatedly and was supplemented by subsequent introductions of mixed stocks. Gene flow between otherwise isolated populations could have also contributed to reduce the genetic structure of the studied populations. Possums in their native range, and in captivity, exhibit territorial and social behaviour that is compatible with assortative mating (Stow et al. 2006; Clinchy et al. 2004). Assuming this is the default behaviour, possum populations would eventually show signs of structuring and inbreeding. However, previous studies on the mating system of possums in New Zealand suggests that reproduction occurs randomly

(Taylor et al. 2000; Sarre et al. 2000), and that home range and mobility expansion can occur following population control (Ji et al. 2001; Blackie et al. 2011).

Moderate levels of genetic diversity confirm that a panel of six microsatellite loci has sufficient statistical power to allow identification of individual possums. The PI_{sib} estimated in this study guarantees that at least 250 possums in a given population can be identified by their multi-locus genotype, even if some of these are full siblings. If one considers the greater genetic diversity known in populations in the North Island (Taylor et al. 2004) then this protocol allows monitoring of possum populations at least across New Zealand's main islands. Despite demonstrating the statistical reliability of this panel, it is anticipated that the inclusion of additional loci will be required in order to allow gender identification (e.g. Ramón-Laca & Gleeson 2014) and extend the monitoring capacity to areas with very high possum densities.

The incidence of stochastic and systematic genotyping error remained relatively constant despite changes during protocol development. Significant increases in e_1 and e_{obs} were observed only when the concentration and quality of the template DNA was reduced. These observations were expected and suggest that the incidence of genotyping error

Table 3 Summary of genotyping error metrics estimated during protocol optimisation.

	Tv19	Tv27	Tv53	Tv54	Tv58	Tv5.64	TvM1	Average
Singleplex								
Amplication success	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
per locus allele drop-out rate	0.00	0.17	0.00	0.00	0.20	0.00	0.00	0.05
per locus false allele rate	0.00	0.00	0.06	0.06	0.00	0.00	0.00	0.02
mean error rate per locus	0.00	0.08	0.08	0.08	0.08	0.00	0.00	0.05
Multiplex								
Amplication success	1.00	1.00	1.00	–	1.00	1.00	1.00	1.00
per locus allele drop-out rate	0.00	0.09	0.11	–	0.13	0.00	0.00	0.05
per locus false allele rate	0.00	0.04	0.06	–	0.02	0.04	0.04	0.07
mean error rate per locus	0.08	0.08	0.10	–	0.04	0.04	0.04	0.07
Multiplex (*diluted*)[a]								
Amplication success	0.86	0.93	0.83	–	0.97	0.93	0.93	0.91
per locus allele drop-out rate	0.09	0.13	0.16	–	0.38	0.00	0.11	0.10
per locus false allele rate	0.00	0.00	0.03	–	0.00	0.00	0.00	0.01
mean error rate per locus	0.12	0.10	0.20	–	0.08	0.04	0.08	0.10
Multiplex PCR (*wax tag*)[b]								
Amplication success	0.78	0.89	0.44	–	0.94	0.83	0.50	0.73
mean error rate per locus	0.25	0.25	0.25	–	0.05	0.20	0.25	0.21

[a]Refers to multiplex PCR performed with diluted tissue extracts.
[b]Refers to multiplex PCR performed with DNA retrieved from saliva.

cannot be attributed to improvements in the amplification protocol (e.g. multiplexing). Although the rates of genotyping error observed for tissue samples were high, these rates fall into the range reported by similar studies (Bonin et al. 2004; Hoffman & Amos 2005; Soulsbury et al. 2007). While an increase in the occurrence of false alleles with the multiplex protocol was observed, we suspect that this was due to the high concentration of template DNA used. The low incidence of false alleles when the template DNA concentration was reduced supports this interpretation. Nonetheless, since the panel developed here was constituted mainly of microsatellites with dinucleotide repeats, it is possible that the occurrence of false alleles is related to inherent characteristics of these loci (Broquet et al. 2007). Altogether these observations suggest an optimal concentration of template DNA is required to mitigate the occurrence of stochastic genotyping error. We observed that high-template-DNA concentrations can increase the occurrence of false alleles, while a decrease in DNA quantity can increase the incidence of allelic drop-out. The existence of an optimal DNA concentration threshold for microsatellite amplification systems has been confirmed in studies of both wildlife and humans (Kline et al. 2005; MacDonald et al. 2011). The optimal concentration for this system is yet to be determined.

Wax tags showed a limited capacity for providing DNA of sufficient quantity and quality to allow reliable genotyping. Studies that have obtained DNA by buccal swabbing of free-ranging animals have obtained concentrations of DNA comparable with those of tissue samples (Yannic et al. 2011; Prunier et al. 2012). It is clear from this evidence that salivary DNA can yield a large enough quantity of DNA when collected appropriately. The low amplification success of fragments ≥200 base pairs that was observed when salivary DNA was used as the template suggests DNA degradation was occurring. Suboptimal preservation and extraction methods, as well as individual variation, could have exacerbated the degradation of DNA present in our salivary samples (Piggott 2004). The presence of PCR inhibitors, and/or large amounts of exogenous DNA (e.g. bacterial DNA), may have also been a factor inhibiting amplification.

Low DNA quantity might be exacerbating the problems of poor DNA quality. From a separate batch of samples deployed in the field (Hororata and Banks Peninsula) we estimated that wax tags yield very low quantities of total DNA (median, range: 0.93 ng/µl, 0.1–9.22 ng/µl, $n = 30$). We excluded these samples from the analysis since we suspected wax tags were bitten by more than one individual and showed clear signs of interaction with non-target species (i.e. rats). Nonetheless, we present these data to illustrate the point that wax tags have a DNA yield comparable with hair follicles (Broquet et al. 2007). Accordingly, the implementation of a more sensitive screening method (e.g. quantitative PCR; Morin et al. 2010) is advised. This new screening method will enable us to determine the actual concentration of DNA that possum saliva yields under a range of different conditions.

The robustness, efficiency, and reproducibility of the protocol developed here is encouraging, yet this pilot study is still incomplete. We have not yet addressed different collection and preservation strategies, the number of replicate PCRs required to obtain a reliable multi-locus genotype, or the impacts of the estimated error rates on demographic estimation. Since implementation of the multiplex protocol was estimated to reduce costs by 92.3% (data not shown), replicating PCRs would not impose a substantial increase in genotyping costs. However, analytical approaches, such as the maximum likelihood method for error estimation (Johnson & Haydon 2007), can avoid the need for extensive replication and should be considered.

Although these issues demand attention, we feel that it is first necessary to improve the current salivary DNA collection methods before this can become a suitable alternative or supplement to field-based monitoring methodologies. We have shown that despite substantial calibration of laboratory protocols, interference devices such as wax tags perform poorly as DNA collection devices. Consequently, we recommend the design of an improved saliva collection device that allows interaction with only a single individual and minimises DNA degradation while collecting enough DNA to conduct reliable genotyping assays.

Acknowledgements

References

Alaeddini R 2012. Forensic implications of PCR inhibition—A review. Forensic Science International: Genetics 6: 297–305.

Blackie HM, Russell JC, Clout MN 2011. Maternal influence on philopatry and space use by juvenile brushtail possums (*Trichosurus vulpecula*). Journal of Animal Ecology 80: 477–483.

Bonin A, Bellemain E, Eidesen PB, Pompanon F, Brochmann C, Taberlet P 2004. How to track and assess genotyping errors in population genetics studies. Molecular Ecology 13: 3261–3273.

Broquet T, Petit E 2004. Quantifying genotyping errors in noninvasive population genetics. Molecular Ecology 13: 3601–3608.

Broquet T, Ménard N, Petit E 2007. Noninvasive population genetics: a review of sample source, diet, fragment length and microsatellite motif effects on amplification success and genotyping error rates. Conservation Genetics 8: 249–260.

Clinchy M, Taylor AC, Zanette LY, Krebs CJ, Jarman PJ 2004. Body size, age and paternity in common brushtail possums (*Trichosurus vulpecula*). Molecular Ecology 13: 195–202.

Cowan PE 1992. The eradication of introduced Australian brushtail possums, *Trichosurus vulpecula*, from Kapiti Island, a New Zealand nature reserve. Biological Conservation 61: 217–226.

Cowan PE, Clout MN 2000. Possums on the move: activity patterns, home ranges and dispersal, In: Montague TL ed. The brushtail possum: biology, impact and management of an introduced marsupial. Lincoln, Manaaki Whenua Press. Pp. 24–34.

Excoffier L, Lischer HEL 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10: 564–567.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3: 294–299.

Forsyth DM, Link WA, Webster R, Nugent G, Warburton B 2005. Nonlinearity and seasonal bias in an index of brushtail possum abundance. Journal of Wildlife Management 69: 976–984.

Gleeson D, Byrom A, Howitt R, Nugent G 2003. Monitoring vertebrate pests by using their DNA. Kararehe Kino Vertebrate Pest Research 2: 13–15.

Goudet J, Raymond M, de Meeüs T, Rousset F 1996. Testing differentiation in diploid populations. Genetics 144: 1933–1940.

Guo SW, Thompson EA 1992. Performing the exact test of Hardy–Weinberg proportion for multiples alleles. Biometrics 48: 361–372.

Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PDN 2006. DNA barcodes distinguish species of tropical Lepidoptera. Proceedings of the National Academy of Sciences (USA) 103: 968–971.

Herráez DL, Stoneking M 2008. High fractions of exogenous DNA in human buccal samples reduce the quality of large-scale genotyping. Analytical Biochemistry 383: 329–331.

Hoffman JI, Amos W 2005. Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. Molecular Ecology 14: 599–612.

Holleley CE, Geerts PG 2009. Multiplex Manager 1.0: a crossplatform computer program that plans and optimizes multiplex PCR. Biotechniques 46: 511–517.

Holm S 1979. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics 6: 65–70.

Idury RM, Cardon LR 1997. A simple method for automated allele binning in microsatellite markers. Genome Research 7: 1104–1109.

Ji W, Sarre SD, Aitken N, Hankin RKS, Clout MN 2001. Sex-biased dispersal and a density-independent mating system in the Australian brushtail possum, as revealed by minisatellite DNA profiling. Molecular Ecology 10: 1527–1537.

Johnson PCD, Haydon DT 2007. Maximum-likelihood estimation of allelic dropout and false allele error rates from microsatellite genotypes in the absence of reference data. Genetics 175: 827–842.

Kline MC, Duewer DL, Redman JW, Butler JM 2005. Results from the NIST 2004 DNA Quantitation Study. Journal of Forensic Sciences 50: 571–578.

Lam MKP, Hickson RE, Cowan PE, Cooper DW 2000. A major histocompatibility (MHC) microsatellite locus in brushtail possums (*Trichosurus vulpecula*). Online Journal of Veterinary Research 4: 139–141.

MacDonald AJ, Sarre SD, FitzSimmons NN, Aitken N 2011. Determining microsatellite genotyping reliability and mutation detection ability: an approach using small-pool PCR from sperm DNA. Molecular Genetics and Genomics 285: 1–18.

Mills LS, Citta JJ, Lair KP, Schwartz MK, Tallmon DA 2000. Estimating animal abundance using noninvasive DNA sampling: promises and pitfalls. Ecological Applications 10: 283–294.

Montague TL, Warburton B 2000. Non-toxic techniques for possum control. In Montague TL ed. The brushtail possum: biology, impact and management of an introduced marsupial. Lincoln. Manaaki Whenua Press. Pp. 164–174.

Moran P, Teel DJ, LaHood ES, Drake J, Kalinowski S 2006. Standardising multi-laboratory microsatellite data in Pacific salmon: an historical view of the future. Ecology of Freshwater Fish 15: 597–605.

Morgan DR, Nugent G, Gleeson D, Howitt RLJ 2007. Are some possums untrappable, unpoisonable, and therefore unmonitorable? Landcare Research Contract Report LC0607/143 for the Animal Health Board (R-10623),

Wellington, New Zealand. Available online at http://www.tbfree.org.nz/.

Morin PA, Martien KK, Archer FI, Cipriano F, Steel D, Jackson J, Taylor BL 2010. Applied conservation genetics and the need for quality control and reporting of genetic data used in fisheries and wildlife management. Journal of Heredity 101: 1–10.

National Possum Control Agencies 2010. Possum population monitoring using the WaxTag® method. Wellington, National Possum Control Agencies. 28 p.

Ogilvie SC, Paterson AM, Ross JG, Thomas MD 2006. Improving techniques for the WaxTag® possum (*Trichosurus vulpecula*) monitoring index. New Zealand Plant Protection 59: 28–33.

O'Reilly-Wapstra JM, Cowan P 2010. Native plant/herbivore interactions as determinants of the ecological and evolutionary effects of invasive mammalian herbivores: the case of the common brushtail possum. Biological Invasions 12: 373–387.

Palero F, González-Candelas F, Pascual M 2011. MICROSATELIGHT—pipeline to expedite microsatellite analysis. Journal of Heredity 102: 247–249.

Peakall R, Smouse PE 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6: 288–295.

Piggott MP 2004. Effect of sample age and season of collection on the reliability of microsatellite genotyping of faecal DNA. Wildlife Research 31: 485–493.

Pompanon F, Bonin A, Bellemain E, Taberlet P 2005. Genotyping errors: causes, consequences and solutions. Nature Reviews Genetics 6: 847–859.

Prunier J, Kaufmann B, Grolet O, Picard D, Pompanon F, Joly P 2012. Skin swabbing as a new efficient DNA sampling technique in amphibians, and 14 new microsatellite markers in the alpine newt (*Ichthyosaura alpestris*). Molecular Ecology Resources 12: 524–531.

Ramón-Laca A, Gleeson D 2014. Use of forensic genetics to detect a potential incursion of the brushtail possum onto Great Barrier Island. New Zealand Journal of Ecology 38: 152–155.

R Development Core Team 2011. R: a language and environment for statistical computing. Vienna, Austria, R Foundation for Statistical Computing. http:www.r-project.org/

Rödiger S, Friedrichsmeier T, Kapat P, Michalke M 2012. RKWard: A comprehensive graphical user interface and integrated development environment for statistical analysis with R. Journal of Statistical Software 49: 1–34.

Rousset F 2008. GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources 8: 103–106.

Ruffell J, Innes J, Didham RK 2015. Efficacy of chew-track-card indices of rat and possum abundance across widely varying pest densities New Zealand Journal of Ecology 39(1): in press.

Sarre SD, Aitken NC, Clout MN, Ji W, Robins J, Lambert DM 2000. Molecular ecology and biological control: the mating system of a marsupial pest. Molecular Ecology 9: 723–733.

Sarre SD, Aitken NC, Quinn AE, MacDonald AJ, Cowan P 2010. *Trichosurus vulpecula* microsatellite Tv5.64 sequence, direct submission to GenBank. URL http://www.ncbi.nlm.nih.gov/nuccore/300488325. Genetic analyses for limiting the spread of bovine tuberculosis by brushtail possums in New Zealand.

Soulsbury CD, Iossa G, Edwards KJ, Baker PJ, Harris S 2007. Allelic drop-out from a high-quality DNA source. Conservation Genetics 8: 733–738.

Stow AJ, Minarovic N, Eymann J, Cooper DW, Webley LS 2006. Genetic structure infers generally high philopatry and male-biased dispersal of brushtail possums (*Trichosurus vulpecula*) in urban Australia. Wildlife Research 33: 409–415.

Taberlet P, Luikart G 1999. Non-invasive genetic sampling and individual identification. Biological Journal of the Linnean Society 68: 41–55.

Taylor AC, Cooper DW 1998. Microsatellite markers for the Phalangerid marsupial, the common brushtail possum (*Trichosurus vulpecula*). Molecular Ecology 7: 1780–1782.

Taylor AC, Cowan PE, Fricke BL, Cooper DW 2000. Genetic analysis of the mating system of the common brushtail possum (*Trichosurus vulpecula*) in New Zealand farmland. Molecular Ecology 9: 869–879.

Taylor AC, Cowan PE, Fricke BL, Geddes S, Hansen BD, Lam M, Cooper DW 2004. High microsatellite diversity and differential structuring among populations of the introduced common brushtail possum, *Trichosurus vulpecula*, in New Zealand. Genetical Research 83: 101–111.

Thomas MD, Brown JA, Maddigan FW, Sessions LA 2003. Comparison of trap-catch and bait interference methods for estimating possum densities. New Zealand Plant Protection 56: 81–85.

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4: 535–538.

Vargas ML, Cruickshank RH, Ross JG, Holyoake AJ, Ogilvie SC, Paterson AM 2009. Noninvasive recovery and detection of possum *Trichosurus vulpecula* DNA from bitten bait interference devices (WaxTags). Molecular Ecology Resources 9: 505–515.

Waits LP, Luikart G, Taberlet P 2001. Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. Molecular Ecology 10: 249–256.

Warburton B 2000. Monitoring possum populations. In: Montague TL ed. The brushtail possum: biology, impact and management of an introduced marsupial. Lincoln, Manaaki Whenua Press. Pp. 132–142.

Warburton B, Barker R, Coleman M 2004. Evaluation of two relative-abundance indices to monitor brushtail possums in New Zealand. Wildlife Research 31: 397–401.

Weir BS 1996. Genetic data analysis II: methods for discrete population genetic data. Sunderland, MA, Sinauer. 445 p.

Weir BS, Cockerham CC 1984. Estimating *F*-statistics for the analysis of population structure. Evolution 38: 1358–1370.

Yannic G, Sermier R, Aebischer A, Gavrilo MV, Gilg O, Miljeteig C, Sabard B, Strøm H, Pouivé E, Broquet T 2011. Description of microsatellite markers and genotyping performances using feathers and buccal swabs for the Ivory gull (*Pagophila eburnea*). Molecular Ecology Resources 11: 877–889.