Photographic identification of individual Archey's frogs, *Leiopelma archeyi*, from natural markings

DOC SCIENCE INTERNAL SERIES 191

Kay Sara Bradfield

Published by
Department of Conservation
PO Box 10-420
Wellington, New Zealand

DOC Science Internal Series is a published record of scientific research carried out, or advice given, by Department of Conservation staff or external contractors funded by DOC. It comprises reports and short communications that are peer-reviewed.

Individual contributions to the series are first released on the departmental website in pdf form. Hardcopy is printed, bound, and distributed at regular intervals. Titles are also listed in the DOC Science Publishing catalogue on the website, refer http://www.doc.govt.nz under Publications, then Science and Research.

ISSN 1175–6519

ISBN 0–478–22632–2

In the interest of forest conservation, DOC Science Publishing supports paperless electronic publishing. When printing, recycled paper is used wherever possible.

This is a client report commissioned by Waikato Conservancy and funded from the Science Advice Fund. It was prepared for publication by DOC Science Publishing, Science & Research Unit; editing by Helen O'Leary and Geoff Gregory and layout by Geoff Gregory. Publication was approved by the Manager, Science & Research Unit, Science Technology and Information Services, Department of Conservation, Wellington.

CONTENTS

Photographic identification of individual Archey's frogs, *Leiopelma archeyi*, from natural markings

Kay Sara Bradfield

School of Tropical Biology, James Cook University, Townsville, Queensland 4811, Australia

ABSTRACT

The New Zealand endemic species, Archey's frog, *Leiopelma archeyi*, has recently undergone marked population declines, and is currently classified as Nationally Critical. This study assessed the potential of using natural markings to identify individual frogs for a capture-recapture monitoring programme. One to three photographs of each of 45 known individuals were used. Initially, a combination of characters that could be used to assign captures to one of 16 subgroups was identified. Two trials were then conducted to determine whether photo-matching could be used to identify individual *L. archeyi*, and whether individuals could consistently be allocated to subgroups (both within and among observers). Assigning captures to subgroups substantially reduced the number of photographs that an unidentified capture needed to be compared with, and the degree of consistency in assigning captures to the same subgroup(s), within and between observers, was generally high. Once captures were correctly assigned to subgroups, the success rate in photo-matching was very high (99.2% overall). Given that photograph quality was often poor in this study, enhanced image quality should improve these results. Recommendations are given for applying the technique to *L. archeyi*.

Keywords: Capture-recapture, individual identification, natural markings, photo-identification, *Leiopelma archeyi*, Archey's frog, threatened species, New Zealand.

Bradfield, K.S. 2004: Photographic identification of individual Archey's frogs, *Leiopelma archeyi*, from natural markings. *DOC Science Internal Series 191*. Department of Conservation, Wellington. 36 p.

1. Introduction

1.1 RECOGNITION OF INDIVIDUALS WITHIN A POPULATION

Accurate estimates of population size are essential for the effective management and conservation of a species, and capture-recapture methods are frequently employed in an attempt to obtain rigorous population estimates. The ability to recognise individuals within a population is fundamental to most capture-recapture methods. Individuals can potentially be recognised by artificial marks (e.g. tags in mammals and fishes, leg bands in birds) or, for species that exhibit sufficient phenotypic variation, by natural markings.

1.2 DISADVANTAGES OF ARTIFICIAL MARKS, WITH A FOCUS ON AMPHIBIANS

Artificially marking animals usually involves capturing and handling, which can stress individuals and/or lead to injury. Marking often creates a wound, which is a potential site of infection. In the case of amphibians, artificial marking usually involves tagging, toe-clipping, branding, tattooing, subcutaneous elastomer injections, or subcutaneous pit tags (see reviews in Ferner 1979 and Donnelly et al. 1994). Toe-clipping in particular is frequently used, as it is relatively easy and inexpensive (Donnelly et al. 1994).

Studies of the effects of these artificial marking techniques on behaviour and survival rates have reported conflicting results, indicating that it may be difficult to make broad generalisations about the effects of these techniques on amphibians. Reaser & Dexter (1996) found no evidence of significant effects of toe-clipping on spotted frogs, *Rana pretiosa*, and Lüddecke & Amézquita (1999) reported no significant effects of disc-clipping (i.e. removal of only the disc, or toe-pad) on the survival or behaviour of Andean frogs, *Hyla labialis*. Van Gelder & Strijbosch (1996) found that toe-clipping common toads, *Bufo bufo*, did not affect physical condition or result in immediate or longer-term inflammation, although they point out that *B. bufo* is not a highly vagile species, and that more vagile species may be adversely affected. A number of studies have documented adverse effects of toe-clipping on a variety of species, however. Clarke (1972) reported an inverse correlation between the number of toes removed from Fowler's toads, *B. woodhousei fowleri*, and the recapture rate, and concluded that this was likely to be the result of increased mortality (but see alternative hypotheses and comments in Reaser 1995 and van Gelder & Strijbosch 1996). Golay & Durrer (1994) found that 18% of recaptured natterjack toads, *B. calamita*, experienced infection or necrosis following toe-clipping. They reported that obvious inflammation did not occur immediately post-amputation; in 50% of cases, inflammation was not detected until at least one month after toe-clipping. Lemckert (1996) reported a very high rate of post-toe-clip swelling and necrosis (close to 100%) in smooth toadlets, *Uperoleia*

laevigata, but a very low rate of infection (approx. 1%) following toe-clipping in common eastern froglets, *Crinia signifera*. He noted that toe-clipping may still be problematic for *C. signifera* because approximately 25% of individuals emigrated from the breeding area within one to three days of being toe-clipped, returning within two to eight weeks (depending on weather conditions). Given that this species has a relatively high mortality rate, such delays in breeding may result in reduced reproduction and recruitment (Lemckert 1996). Davis & Ovaska (2001) found that toe-clipping western red-backed salamanders, *Plethodon vehiculum*, during spring affected their ability to take advantage of optimal foraging conditions and suggested that this could influence their ability to survive the adverse conditions of summer. They also reported that a number of toe-clipped individuals had swollen toe stumps for up to eight months post-marking, and that recapture rates were significantly higher for fluorescent-marked individuals than toe-clipped individuals, which suggests higher mortality for toe-clipped individuals.

These studies that indicate negative effects of artificial marking techniques on amphibians have significant implications for population monitoring via capture-recapture methods. Altered behaviour and increased mortality resulting from marking violate an assumption that underlies most capture-recapture methods, namely that the probability of recapture is not affected by marking (Caughley & Sinclair 1994). Mark-induced mortality or sub-lethal effects on fitness (e.g. reduced reproduction, decreased growth rate) are particularly unacceptable when the species concerned is endangered.

1.3 NATURAL MARKINGS, A VIABLE ALTERNATIVE

The use of natural features or markings to identify individuals within a population is non-invasive, and therefore does not pose the same risk as invasive artificial marking techniques. Individuals (or a particular region of their bodies) can be either drawn or photographed, and the resulting images compared with the images for all previous captures. This technique has been employed most frequently in studies on mammals, both marine (e.g. humpback whales, *Megaptera novaeangliae*, Glockner & Venus 1983; southern right whales, *Eubalaena australis*, Payne et al. 1983; Mediterranean monk seals, *Monachus monachus*, Forcada & Aguilar 2000) and terrestrial (e.g. zebras, *Equus burchelli*, Petersen 1972; lions, *Panthera leo*, Schaller 1972 in Kelly 2001; chimpanzees, *Pan troglodytes*, Goodall 1986; cheetahs, *Acinonyx jubatus*, Caro 1994; badgers, *Meles meles*, Dixon 2003). It has also been applied to birds (e.g. Bewick's swan, *Cygnus columbianus*, Scott 1978; ospreys, *Pandion haliaetus*, Bretagnolle et al. 1994; lesser white-fronted geese, *Anser erythropus*, Øien et al. 1996); reptiles (e.g. adders, *Vipera berus*, Sheldon & Bradley 1989; common garter snakes, *Thamnophis sirtalis sirtalis*, Hallmen 1999; five species of Central European lacertid lizards, Steinicke et al. 2000); and fishes (e.g. pipefish, *Corthoichthys intestinalis*, Gronell 1984; leafy seadragons, *Phycodurus eques*, Connolly et al. 2002).

With regards to amphibians, a number of studies have used natural markings/colour patterns to identify individual urodeles (e.g. the newts *Triturus cristatus*

and *T. vulgaris*, Hagström 1973; Blue Ridge dusky salamanders, *Desmognathus ochrophaeus*, Forester 1977, Tilley 1980; red-spotted newts, *Notophthalmus viridescens*, Gill 1978; western red-backed salamanders, *Plethodon vehiculum*, Davis & Ovaska 2001; *Eurycea bislineata wilderae*, Bailey 2004). However, few studies have used natural markings to identify individual anurans. Newman (1982) used markings along the upper lip to identify individual Hamilton's frogs, *Leiopelma hamiltoni*. Golay & Durrer (1994) used belly patterns to identify individual natterjack toads, *B. calamita*, while Denton & Beebee (1993) used throat-spot patterns to identify individuals of the same species. In the latter case, the authors validated the method by initially toe-clipping individuals as well as recording throat-spot patterns. In view of the use of natural markings to identify individuals of a wide variety of animals, it is clear that the technique has the potential to be useful for any anuran species that exhibits variability in markings on at least one region of the body.

1.3.1 Points to consider when assessing the technique

Points which should be kept in mind when assessing the use of natural markings to identify individuals of a particular species include the following:

As with any technique, identification of individuals must be efficient. Individuals are usually photographed or sketched in the field, and identifications are made at a later stage. When the catalogue of previous captures is relatively small, manual identification (i.e. identification entirely by eye) is rapid, but when the catalogue is large, it can take substantially longer. Computer-assisted matching can be used if photo-matching entirely by eye is too time-consuming (e.g. Whitehead 1990; Kelly 2001). Dividing individuals into subgroups can also facilitate rapid identification of unidentified captures when there are a large number of previous captures, because the observer only has to photo-match to a small subgroup rather than to all previous captures. For example, Gill (1978) was able to identify individual red-spotted newts, *Notophthalmus viridescens*, within 30 seconds, despite a catalogue of over 8500 individuals, because individuals could be assigned to subgroups based on the number of spots on each side of the dorsal surface. Another advantage of this approach is that it results in a higher degree of accuracy, as the larger the catalogue of photographs, the more likely it is that mismatching will occur.

Digital photographs have a number of advantages over traditional slide or print film images, even when photo-matching is conducted entirely by eye. Markowitz et al. (2003) compared digital and slide film images of New Zealand dusky dolphins, and found a higher proportion of digital images were of suitable quality for use in photographic identification than slide film images taken by the same photographers. Digital images are also available for inspection immediately after they are taken (images can even be examined in the field), and they can be archived, accessed, and printed easily and rapidly.

Using natural markings to identify individuals often requires longer handling times than artificial marking techniques such as toe-clipping or pit-tagging (Reaser 1995). Capturing and handling amphibians can induce stress, and may cause acute changes in behaviour and physiology (Reaser 1995). Increased handling times could exacerbate such effects.

1.3.2 Potential sources of error

An assumption of capture-recapture methods is that all marked individuals are reported as such on recapture. There are two basic types of incorrect identification in studies using natural markings to identify individuals. False negative errors result from concluding that two captures of one individual actually represent two different individuals. When these errors occur, the number of marked individuals reported as recaptured is less than the number actually recaptured, resulting in inflated population estimates. False positive errors result from concluding that captures of different individuals actually represent a single individual (failed photo-match), and should be separated into false matches and mismatches (*sensu* Agler 1992). False matches are the result of identifying a new capture as a recapture of individual x, whereas mismatches are the result of identifying a recapture of individual x as a recapture of individual y. This distinction is important, as false matches will lead to population underestimates because the number of marked individuals reported as recaptured is greater than the number actually recaptured. In contrast, mismatches will not affect population estimates.

There are several potential sources of error associated with the use of natural markings to identify individuals. For example, image quality influences error rates, with poor images resulting in a higher number of incorrect identifications than high-quality images (e.g. Agler 1992; Forcada & Aguilar 2000; Gowans & Whitehead 2001; Stevick et al. 2001).

It is possible that two or more individuals in a population will have such similar natural markings that they cannot be distinguished from one another (Pennycuick 1978), resulting in false positive errors. The likelihood of this occurring increases with increasing population size, but decreases with increasing pattern complexity. The probability that a pattern will be repeated in a particular population can be estimated (e.g. Pennycuick 1978).

Variability in the degree of distinctiveness of individuals means that 'marked' individuals (i.e. those who have previously been captured and photographed/sketched) do not necessarily all have the same probability of being recognised, and this can potentially have serious effects when estimating abundance (Hammond 1986 in Friday et al. 2000). Only individuals distinctive enough to have equal probabilities of recognition should be considered as marked.

An assumption common to capture-recapture methods is that marks do not change over time. However, natural markings do have this potential, which would result in population overestimates. For example, Reaser (1995) found that spot patterns of adult California tiger salamanders, *Ambystoma californiense*, held in captivity changed over time. In contrast, Denton & Beebee (1993) found that throat-spot patterns did not significantly change over time in adult natterjack toads, *B. calamita*, and Stephenson & Stephenson (1957) noted that colour patterns in leiopelmid frogs did not change during two years in captivity.

1.4 A CANDIDATE FOR THE USE OF NATURAL MARKINGS TO IDENTIFY INDIVIDUALS

The family Leiopelmatidae is endemic to New Zealand, and contains only four extant species, all of which belong to the genus *Leiopelma* (Frost 2002). These four species are among the most morphologically primitive living anurans (Cannatella 1995), and are therefore of considerable evolutionary significance.

Leiopelma archeyi, or Archey's frog, is a small nocturnal frog (maximum SVL 37 mm). It is terrestrial, and inhabits damp native forest, subalpine vegetation, and open ridge tops (Stephenson & Stephenson 1957; Bell 1978). The species is currently known to occur in only two areas: the Coromandel Peninsula (Moehau and Colville Ranges south to ranges near Paeroa) and Whareorino Forest (Bell et al. 1998). The populations inhabiting Mt. Moehau (Chappell unpublished 2001a, b) and the central Coromandel Ranges (Bell 2001) appear to have undergone substantial declines in recent years, while the population in Whareorino Forest appears to be relatively stable (T. Thurley, DOC, Te Kuiti, pers. comm.). The species is classified as Nationally Critical (i.e. highest threat category) by the Department of Conservation (DOC) (Hitchmough 2002).

Effective conservation and management of *L. archeyi* requires ongoing monitoring of populations throughout its range. However, an effective monitoring technique has yet to be developed for this species. Simple nocturnal counts of emerged frogs are so variable, even between consecutive nights, that they are of limited value for detecting population trends (T. Thurley, DOC, Te Kuiti, pers. comm.). Capture-recapture methods have the greatest potential to provide rigorous population estimates. The conservation status of this species means that toe-clipping and other invasive marking techniques are not acceptable, but it may be possible to use natural markings to identify individuals within a population.

1.5 OBJECTIVES

The objectives of this study were to determine whether photographic identification of naturally marked animals can be used to identify individual *L. archeyi*, and whether individuals of *L. archeyi* can be divided into subgroups on the basis of their markings to increase the efficiency of identification.

2. Methods

2.1 DEVELOPING A METHODOLOGY

A number of *L. archeyi* from Whareorino Forest are currently in captivity at the University of Canterbury, Christchurch. Between one and three sets of digital photographs, all taken between 23 May and 4 June 2003, were supplied for each

of 45 individuals. Each set consisted of six photographs per frog: dorsal, ventral, anterior (facial), posterior, right lateral, and left lateral views. Sets 1 and 3 were taken using a flash, and set 2 was taken under ambient light conditions. Inspection of the photographs showed that the images in set 3 were generally of higher quality than those in sets 1 and 2; photographs in set 3 (which only contained images of 43 of the 45 frogs) were therefore used to develop the method. However, it should be noted that, even in set 3, image quality was often poor (i.e. out of focus, underexposed, or with a considerable amount of glare reflected off the frog), particularly for the dorsal and lateral photographs. For the purposes of this report, all photographs of one individual from one set are referred to as a 'capture'. Thus, the dorsal, ventral, anterior, posterior, right lateral, and left lateral photographs of an individual taken using flash on 28 May represent one capture, photographs of the same individual taken with ambient light on 28 May represent another capture, and those of the same individual taken using flash on 4 June represent a third capture, as this is what they would represent if taken during a population monitoring programme. As three of the 45 frogs were only photographed on two of the three occasions, there was a total of 132 captures. Each capture was assigned a unique identification number.

Photographs of twenty of the 43 individuals in set 3 were selected at random, and designated as the 'initial captures'. These photographs were used to identify characters suitable for inclusion in a key that would allow each frog to be allocated to one of 16 subgroups (A–P). The use of 16 subgroups means that, even if a field data set contains hundreds of previous captures, each subgroup should contain a relatively small number of individuals (provided that individuals are divided among subgroups fairly evenly), while still only using a small number of characters in the key. Subgroups containing relatively small numbers of individuals should result in rapid and accurate matching, and the fewer characters that are used to assign individuals to subgroups, the quicker it is to do this, and the lower the likelihood that observers will make errors or that different observers will interpret characters differently. Following an initial inspection of these photographs, it was decided to use only the dorsal, anterior (facial), and right and left lateral views, as individuals appeared to have distinct markings on these body regions which were clear in these photographs. Ventral patterns were not considered at this stage, as obtaining ventral images requires additional manipulation of individuals. A number of characters that could potentially be used to assign frogs to subgroups were identified (Table 1; see also Appendix 5, Fig. A5–1) and assessed for each of the 20 initial captures. Many of the characters were categorised as either continuous or discontinuous; in order to avoid subjective decisions by observers, a character was classed as continuous if the markings of interest were joined by a line *of any thickness* and *of the same colour (i.e. black)*.

Of the characters that were found to be useful, the combination that divided the 20 frogs most evenly among the 16 subgroups was chosen.

Assessment of the 15 characters identified as potentially of use for assigning individuals to subgroups (Table 1) revealed that 12 were unsuitable for inclusion in these trials (Appendix 1). Reasons for unsuitability were: inability to define the character without involving subjective decisions by observers for

some/all captures; all or the majority of individuals assessed were either continuous or discontinuous for that character; and poor quality of photographs. It should be noted that characters deemed to be unsuitable for assigning captures to subgroups are still likely to be useful when photo-matching within subgroups. The characters suitable for inclusion in the trials were: continuous/discontinuous right and left eye and upper lip markings (E-ULM); continuous/discontinuous right and left nare and upper lip markings (N-ULM); and continuous/discontinuous right and left dorso-lateral stripes between the posterior edge of the eye and the posterior edge of the base of the forelimb (D-L stripe (E-FL)). The combination of these characters that divided the 20 frogs most evenly among 16 subgroups was: right & left N-ULM + right & left D-L stripe (E-FL), so these characters were used to create a key. The specific combinations of continuous and discontinuous characters used to define each subgroup are detailed in Appendix 2.

TABLE 1. POTENTIAL CHARACTERS FOR ASSIGNING FROGS TO SUBGROUPS.

REGION OF BODY	CHARACTER	ABBREVIATION
Anterior (facial)	Black marking below eye and black upper lip markings continuous or discontinuous (right and left sides); Appendix 5, Fig. A5-1A	E-ULM
Anterior (facial)	Black stripe that passes through the nare and black upper lip markings continuous or discontinuous (right and left sides); Appendix 5, Fig. A5-1B	N-ULM
Anterior (facial)	Frontal stripe (runs from the posterior edge of the right eye, through the eye, around the snout, through the left eye, ending at the posterior edge of the left eye) continuous or discontinuous; Appendix 5, Fig. A5-1C	FS
Anterior (facial)	Number of breaks in the frontal stripe	FS breaks
Anterior (facial)	Number of black spots between the frontal stripe and the upper lip markings; Appendix 5, Fig. A5-1D	UL spots
Anterior (facial)	Number of black spots in the 'frontal triangle' (area between the eyes and the snout); Appendix 5, Fig. A5-1E	FT spots
Anterior (facial)	Central upper lip marking and adjacent upper lip marking continuous or discontinuous along edge of lip (right and left sides); Appendix 5, Fig. A5-1F	CULM & adjacent ULM
Anterior (facial)	Upper lip markings continuous or discontinuous along edge of lip	ULM
Anterior (facial)	Black edge of lower lip continuous or discontinuous	LLM
Dorsal	Base of dorsal triangle continuous or discontinuous; Appendix 5, Fig. A5-1G	DT base
Dorsal	Corners on the base of the dorsal triangle complete or incomplete (right and left sides); Appendix 5, Fig. A5-1H	DT corners
Dorsal	1st and 2nd paired dorsal markings continuous or discontinuous (right and left sides); Appendix 5, Fig. A5-1I	1st & 2nd PDM
Dorsal	Number of black stripes on the forelimb (right and left sides); Appendix 5, Fig. A5-1J	FL stripes
Lateral	Dorso-lateral stripe (the black stripe that lies immediately below the dorso-lateral ridge) continuous or discontinuous	D-L stripe
Lateral	Dorso-lateral stripe continuous or discontinuous between the posterior edge of the eye and the posterior edge of the attachment point of the forelimb (right and left sides); Appendix 5, Fig. A5-1K	D-L stripe (E-FL)

2.2 TRIALLING THE METHODOLOGY

Sixteen folders were labelled (A–P), and photographs of the 20 initial captures were allocated to the appropriate folder to serve as an initial set of 'previous captures' for comparison.

Two trials were conducted. In the first, the primary observer (KSB) identified each capture in the entire set of test captures (i.e. photographs of all 45 frogs from all three sets of photographs, minus the 20 initial captures used to develop the method; a total of 112 test captures) as either a new capture or a recapture. In the second trial, four observers (KSB and three DOC Waikato Conservancy staff members) identified each of 24 test captures (a subset from the entire set of 112 test captures used in Trial 1) as either a new capture or a recapture. Trial 1 assessed one observer's ability to consistently assign different photographs of a particular individual to the same subgroup(s) and correctly photo-match within subgroups when a reasonably large number of captures is involved. Trial 2 assessed intra- and inter-observer differences in both assigning individuals to subgroups and photo-matching within subgroups.

The 112 test captures were put in random order. Observers in both trials were thus unaware of which of the three sets of photographs a given capture belonged to, and therefore could not have any expectations about whether it was a new capture or a recapture.

Observers used worksheets (Appendix 3) to assign each test capture to a subgroup(s), and then photo-matched the test capture to the previous captures in that subgroup by eye. In some instances, classification of one or more characters was questionable due to either observer uncertainty or photo quality. When this occurred, the capture in question was assigned to a primary subgroup (the one it was considered most likely to belong to) as well as to all other subgroups it could potentially belong to (referred to as alternative subgroups). In this situation, the photographs of the capture were placed into the folder for the primary subgroup, and the individual's number and primary subgroup were noted on the outside of all alternative subgroup folders. This allowed observers to quickly identify other previous captures that the test capture in question should be compared to.

In these trials, false negative errors could result from either assigning two or more captures of one individual to different subgroups (observer inconsistency), or from assigning different captures of one individual to the same subgroup, but failing to recognise that they are the same individual (failed photo-match within a subgroup).

2.2.1 Trial 1. Intra-observer consistency (primary observer)

The primary observer (KSB) viewed all 112 test captures and identified each as either a new capture or a recapture. The dorsal, anterior, and lateral views for each capture were printed in colour, and these images were used to identify individuals. Because the 112 test captures consisted of multiple captures of each of 45 individual frogs, each test capture could be identified as a new capture, a recapture of an initial capture (i.e. one of the set of 20 frogs that were used to develop the method, which represent 'previous captures' for the purposes of this trial), or a recapture within the set of 112 test captures. Upon

completion of the trial, the percentage of times a capture was assigned to two or more subgroups was calculated in order to determine whether assigning captures to subgroups actually increases efficiency of identification (note that all 132 captures were included, i.e. the 20 initial captures used to develop the method and the 112 test captures, in order to maximise sample size). The accuracy of identifications was also determined and the reason(s) for errors were identified. There were five possible types of error:

- the position of the frog obscured the region of interest for one of the characters used to assign captures to subgroups (e.g. frog twisted to one side),
- intra-observer inconsistency (i.e. the observer assigned two different captures of one individual to different subgroups),
- the observer failed to make a photo-match within a subgroup,
- the observer made an incorrect photo-match, and
- other observer error (e.g. observer recorded the subgroup letter incorrectly).

2.2.2 Trial 2. Intra- and inter-observer consistency

A random subset of 24 captures from the set of 112 test captures assessed in Trial 1 was used in this trial. Four observers (KSB and three DOC Waikato Conservancy staff members) identified each of the 24 test captures as either a new capture, a recapture of an initial capture, or a recapture within the set of 24 test captures.

As mentioned previously, photographs were frequently of poor quality, and the three inexperienced observers were only provided with written definitions of the characters used to assign frogs to subgroups. They were not provided with photographs or diagrams to clarify the definitions, nor did they receive any advice or training on how to classify captures for each character. Thus, the results of these observers should represent the worst-case scenario.

For each observer, the percentage of times they assigned a capture to two or more subgroups was calculated, as were the number and percentage of incorrect identifications. Test captures correctly identified as new captures by these observers were only considered to be correct identifications when they were assigned to the same primary and/or alternative subgroups as by the primary observer. This is not implying that the primary observer is correct and the inexperienced observers are incorrect in such instances. Rather, it is used to assess the frequency with which two different observers assigned the same capture to different subgroups. This is an important factor to consider because a long-term monitoring programme may involve different observers, and inconsistencies in classification will result in false negative errors and inflated population estimates. Incorrect identifications that were the result of an inexperienced observer assigning a new capture to a different subgroup from that assigned by the primary observer are referred to as 'potential false negatives'.

The cause(s) of errors were noted. There were six types of error: the five noted in Trial 1 (see above), and inter-observer inconsistency (i.e. primary observer and observer in question assigned one capture, or two different captures of the one individual, to different subgroups).

2.2.3 Post-trial assessment of characters

In the initial assessment of characters for assigning captures to subgroups, only one set of photographs of each of 20 individuals was used (out of the three sets of photographs per frog and 45 frogs). Once the trials were completed, it was possible to use all 132 captures to assess potentially useful characters. This increased the sample size from 20 individuals to 45, thus increasing the likelihood of identifying the combination of characters that would be of most use in future studies. It also increased the chances of classifying characters as definitely continuous or definitely discontinuous for all individuals, as the complete set of 132 captures usually included at least one capture for each individual where the photographs were of sufficient quality that all characters of interest could be assessed.

The initial assessment of characters indicated that, in addition to the N-ULM and the D-L stripe (E-FL) characters, the eye-upper lip markings (E-ULM) character was also suitable, and that the 1st and 2nd paired dorsal markings (1st & 2nd PDM) character was likely to be suitable with improved photograph quality (Appendix 1). Also, during the course of the trials, it was noted that the E-ULM and N-ULM characters could be combined into a single character, referred to as the frontal stripe-upper lip markings (FS-ULM). This character was classed as continuous if the frontal stripe (beginning from the posterior edge of each eye and running forward through the nare on each side to the mid-line of the snout) and the upper lip markings were connected by a line of any thickness, and discontinuous if the frontal stripe and upper lip markings were not connected. Left and right sides were considered separately. Examination of the results for the 20 frogs included in the initial assessment indicated that this new character would have divided those frogs more evenly among the 16 subgroups than either the E-ULM or the N-ULM. The FS-ULM character was therefore included in a second assessment using all 132 captures, along with the four above-mentioned characters from the initial assessment (the E-ULM and N-ULM characters were included even though the FS-ULM character appeared to be more suitable in case problems are encountered using the FS-ULM character). All relevant pairwise combinations were assessed to determine the combination of characters most useful for dividing individuals into subgroups.

3. Results

3.1 TRIAL 1. INTRA-OBSERVER CONSISTENCY

Of the 132 total captures, 63.5% were assigned to only one subgroup (i.e. a primary subgroup only), 33% were assigned to two subgroups (i.e. a primary subgroup and one alternate subgroup, indicating that classification of one of the four characters was uncertain), and 3.5% were assigned to four subgroups (i.e. a primary subgroup and three alternate subgroups, indicating that classification of two of the four characters was uncertain). When considering the left and right sides of each character separately (i.e. right N-ULM is one character, left N-ULM is another character), classification of the N-ULM was uncertain 7% of the

time, and classification of the D-L stripe (E-FL) was uncertain 13% of the time. Poor photograph quality was likely to be responsible for this uncertainty 37% of the time for the N-ULM, and 88% of the time for the D-L stripe (E-FL).

Within subgroups, photo-matching was successful 100% of the time. The single incorrect identification (a false negative) was the result of assigning two test captures to different subgroups, despite the fact they were two captures of one individual. This occurred because the left D-L stripe (E-FL) appeared to be continuous for one test capture (Appendix 5, Fig. A5-2A), but was obviously discontinuous for the other test capture (Appendix 5, Fig. A5-2B). The left lateral photograph of the former test capture was taken with the camera on an angle, rather than held parallel to the side of the frog, and the frog was also twisted to the left. As this error was the result of the position of the frog in one of the photographs, and not the result of the observer interpreting the character differently for each capture, it is not counted as an error due to intra-observer inconsistency in assigning captures to subgroups. Therefore, the primary observer consistently assigned all captures of a single individual to the same subgroup.

3.2 TRIAL 2. INTRA- AND INTER-OBSERVER CONSISTENCY

In trial 2, between 21% and 71% of captures were assigned to a single subgroup (mean = 51%), between 25% and 41.5% were assigned to a primary subgroup and one alternate subgroup (mean = 33%), between 4% and 37.5% were assigned to a primary subgroup and three alternate subgroups (mean = 14%), and between 0% and 4% were assigned to a primary subgroup and seven alternate subgroups (mean = 2%) (Table 2). Of the 64 occasions when the classification of a character was uncertain, only seven were for the N-ULM, while 57 were for the D-L stripe (E-FL). Poor photograph quality was frequently responsible for this uncertainty (86% of the time for the N-ULM and 54% of the time for the D-L stripe (E-FL); Table 2).

TABLE 2. TRIAL 2: RESULTS FOR ASSIGNMENT OF CAPTURES TO SUBGROUPS FOR EACH OBSERVER.

OBSERVER	PERCENTAGE OF CAPTURES ASSIGNED TO *x* SUBGROUPS, WHERE *x* EQUALS				NO. OF TIMES A CAPTURE WAS ASSIGNED TO > 1 SUBGROUP DUE TO UNCERTAIN CLASSIFICATION*		NO. OF TIMES PHOTO QUALITY PROBABLY RESPONSIBLE FOR UNCERTAIN CLASSIFICATION	
	1	2	4	8	N-ULM	D-L stripe (E-FL)	N-ULM	D-L stripe (E-FL)
1 (primary observer)	71	25	4	0	1	7	0 / 1	5 / 7
2	50	38	8	4	3	13	3 / 3	10 / 13
3	63	29	4	4	3	9	3 / 3	7 / 9
4	21	41.5	37.5	0	0	28*	0 / 0	9 / 28
Means/Totals	51	33	14	2	7	57	6 / 7 (86%)	31 / 57(54%)

* For some captures, both the N-ULM and the D-L stripe (E-FL) were questioned on one or both sides.

Observer 1 incorrectly identified one (4%) of the 24 test captures, Observers 2 and 4 made two (8%) incorrect identifications, and Observer 3 incorrectly identified nine captures (37.5%) (Table 3). Overall, test captures were incorrectly identified on 14 of the total of 96 (24 test captures × 4 observers) occasions, or 14.5% of the time.

The incorrect identifications were either false negatives (11/14) or potential false negatives (3/14) (Table 3). One of the false negatives was the result of failed photo-matching within a subgroup (Table 3). Overall, photo-matching within subgroups was successful 98% of the time. The remaining ten false

TABLE 3. TRIAL 2: INCORRECT IDENTIFICATIONS BY EACH OBSERVER.

OBSERVER	INCORRECT IDENTIFICATIONS		NO. OF FALSE POSITIVES	NO. OF POTENTIAL FALSE NEGATIVES[a]	NO. OF FALSE NEGATIVES	NO. OF FALSE NEGATIVES DUE TO		
	NO.	%				POSITION OF FROG	INTRA-OBSERVER INCONSISTENCY	OTHER FACTORS[b]
1 (primary observer)	1	4	0	-	1	1		
2	2	8	0	0	2	1	1	
3	9	37.5	0	3	6	2	1	3
4	2	8	0	0	2	1	1	
TOTALS	14/96	14.5	0/14	3/14	11/14	5/11	3/11	3/11

[a] All potential false negatives are the result of inter-observer inconsistency.

[b] One each for false negatives due to inter-observer inconsistency, failed photo-match, and error.

negatives were all the result of the initial capture and recapture being assigned to different subgroups. The position of the frog in one of the photographs was responsible for the discrepancy on five of the ten occasions, a mistake on the part of the observer was responsible once, intra-observer inconsistency in the classification of one or more characters used to assign captures to subgroups was responsible on three occasions, and inter-observer inconsistency in the classification of one or more characters used to assign captures to subgroups was responsible once (Table 3; further details provided in Appendix 4). Potential false negatives are, by definition, the result of a second observer assigning a capture to a different subgroup(s) from that done by the primary observer. Inter-observer inconsistency was therefore responsible for four errors compared with three from intra-observer inconsistency. It is important to note, however, that there were more than twice as many opportunities for errors to occur as the result of inter-observer inconsistency than for intra-observer inconsistency. Inter-observer inconsistency only occurred on 8% of possible occasions, while intra-observer inconsistency occurred on 12.5% of possible occasions.

3.3 POST-TRIAL ASSESSMENT OF CHARACTERS

Assessment of all relevant pairwise combinations of the five characters revealed that the combination used in Trials 1 and 2 was not the optimal combination (Table 4). There were three combinations that best divided the individuals

TABLE 4. RESULTS OF THE ASSESSMENT OF EIGHT PAIRWISE COMBINATIONS OF CHARACTERS FOR ASSIGNING CAPTURES TO SUBGROUPS. BEST COMBINATIONS ARE IN BOLD.

COMBINATION OF CHARACTERS	NO. OF SUBGROUPS USED (out of 16)	PERCENTAGE OF INDIVIDUALS IN *x* LARGEST SUBGROUPS		
		x = 1	*x* = 2	*x* = 5
E-ULM + N-ULM	11	56	64	84
E-ULM + D-L stripe (E-FL)	13	27	47	76
E-ULM + 1st & 2nd PDM	13	27	44	76
N-ULM + D-L stripe (E-FL)	14	29	53	80
N-ULM + 1st & 2nd PDM	11	31	53	82
FS-ULM + D-L stripe (E-FL)	**15**	24	40	62
FS-ULM + 1st & 2nd PDM	**15**	22	38	62
D-L stripe (E-FL) + 1st & 2nd PDM	14	**18**	**33**	**60**

among subgroups: D-L stripe (E-FL) + 1st & 2nd PDM, FS-ULM + D-L stripe (E-FL) and FS-ULM + 1st & 2nd PDM. The combination of D-L stripe (E-FL) + 1st & 2nd PDM divided the 45 individuals among 14 of the 16 subgroups, with no more than 18% belonging to any one subgroup. The five largest subgroups together accounted for only 60% of all individuals (Table 4). The FS-ULM + D-L stripe (E-FL) and FS-ULM + 1st & 2nd PDM combinations both divided the frogs into 15 of the 16 subgroups, with just under a quarter of individuals belonging to the largest subgroup, and less than half belonging to the two largest subgroups. For both combinations, the five largest subgroups accounted for 62% of individuals (Table 4).

4. Discussion

4.1 TRIALS

These trials demonstrated that natural markings can be used to assign individual *L. archeyi* to subgroups according to a simple key, and that once a capture has been assigned to a subgroup, its markings can be used to correctly match it to previous captures of the same individual present in that subgroup. Thus, the technique shows great promise for application to capture-recapture studies of this species.

In most instances, observers were able to assign captures to a very small number of subgroups (between one and four). There were no instances of a capture being assigned to all 16 subgroups in either trial. Thus, assigning captures to subgroups substantially reduces the number of previous captures that an unidentified capture needs to be compared with. The degree of consistency in assigning captures to the same subgroup(s), within and between observers, was generally high (see below for discussion of errors). Together, these results indicate that it should be possible to efficiently process photographs of unidentified captures in a full-scale monitoring programme by

using a key to identify subgroups to search in order to determine the identity of any given capture. Assigning individuals to subgroups can, however, increase the chances of making a false negative error, because there is the potential for different captures of one individual to be assigned to different subgroups. Observers must therefore exercise due diligence when assigning captures to subgroups, ensuring that any characters that cannot be classified with absolute certainty are recorded as such, and that each capture is assigned to all potentially relevant subgroups. This will greatly increase the chances that each capture will be compared to all potential matches. A simple question mark can be used to indicate that character classification is uncertain (refer to Appendix 3), or a more detailed measure of character fidelity could be incorporated into the assessment of characters (e.g. high/1 = character classification is certain, medium/2 = character classification is fairly certain, low/3 = character classification is uncertain).

It should be kept in mind that these trials were conducted under worst-case conditions (poor photograph quality; minimal instruction for inexperienced observers). Poor photograph quality was frequently responsible for the uncertainty regarding one or more of the characters used to assign captures to subgroups (see Table 2). Therefore, improved photograph quality should greatly reduce the number of captures assigned to more than one subgroup. In Trial 2, the primary observer assigned 71% of the 24 captures to a single subgroup, which was a higher percentage than for any of the inexperienced observers (Table 2). This suggests that appropriate observer training should also reduce the number of captures assigned to more than one subgroup by inexperienced observers. Finally, due to an error on the part of the author, Observer 4 was not provided with the written definitions of the characters while completing the trial, and therefore found the dorso-lateral stripe (eye-forelimb) character very difficult to use. This explains why Observer 4 assigned a relatively high percentage of individuals to four subgroups (resulting from uncertain classification of both left and right D-L stripe (E-FL)), compared with the other three observers (Table 2).

In Trial 1, the primary observer exhibited a very low identification error rate (< 1%), indicating that an experienced observer can identify individuals with a high degree of accuracy. It is worth noting that the single incorrect identification was actually correctly identified during the trial despite the fact the initial capture and recapture were assigned to different subgroups; the primary observer recognised the individual on sight as a recapture, and, when a match could not be found within the subgroup the test capture was assigned to, all other subgroups were checked until the match was found. On finding the match, the source of the problem was identified (i.e. frog position hindering correct character classification). The primary observer subsequently assigned all other similar captures to both relevant subgroups (character in question scored as either continuous or discontinuous) to ensure that matches were not missed. This explains why, in Trial 2, the primary observer correctly identified test capture 137 as a recapture of an initial capture, while the other three observers identified it as a new capture (Appendix 4). It is also worth noting that, when a photograph identified as a recapture was of better quality than the photograph of the initial capture for that individual, the primary observer replaced the initial capture photograph with the higher-quality photograph in

the subgroup folder. This should have increased the efficiency and accuracy of all subsequent photo-matching comparisons to that individual.

In Trial 2, Observer 3 incorrectly identified a substantially higher number of test captures than did either of the other inexperienced observers (Table 3). Post-trial discussion revealed that this observer had mistakenly classified the right and left *eye*-upper lip markings character rather than the right and left *nare*-upper lip markings character, which accounts for three of the nine errors (test captures 51, 89, and 97). All three of these incorrect identifications were initially classed as errors due to inter-observer inconsistency in the classification of the N-ULM character (Appendix 4). As this was not the case, there was actually only one error due to inter-observer inconsistency, and inter-observer inconsistency therefore only occurred on 2% of possible occasions.

All errors made during the course of both trials were false negative errors. False-negative errors inflate population estimates, which is particularly undesirable when dealing with declining or endangered species. Most of the sources of error identified in this study can be reduced or eliminated in future studies. Errors due to the position of the frog are avoidable, once the photographer and observers are aware of the potential for this problem to occur. Errors due to mistakes on the part of the observer are also avoidable, if due care is taken and identifications are independently verified by a second observer. Errors due to inter-observer inconsistency in the classification of a character used to assign captures to subgroups can be reduced, or eliminated, by observer training. Also, a reference set of captures can be used to determine whether new observers are interpreting the characters used to assign captures to subgroups in the same way as experienced observers (i.e. following training, all new observers would have to complete an exercise similar to Trial 2 in the present study). This exercise would also indicate how frequently each observer fails to photo-match two captures of one individual, or assigns different captures of the one individual to different subgroups. If necessary, error rates could then be calculated for each observer and incorporated into population estimates (see Agler 1992 and Stevick et al. 2001 for examples of corrected estimators). The use of two independent observers should substantially reduce the chances of false-negative errors due to failed photo-matches or intra-observer inconsistencies, as it is highly unlikely that two different observers would both miss the same photo-match, or both assign the same two captures of one individual to different subgroups. Improved photograph quality should also reduce the number of errors due to both failed photo-matches (see below for further details) and intra-observer inconsistencies.

Once captures were correctly assigned to subgroups, the error rate in photo-matching captures was extremely low (0.8% overall). As noted in the introduction, previous studies using photographic identification have found that the accuracy of photo-matching increases as photograph quality increases (e.g. Agler 1992; Forcada & Aguilar 2000; Gowans & Whitehead 2001; Stevick et al. 2001). Thus, it is anticipated that the use of higher-quality photographs of *L. archeyi* in future studies will decrease the error rate for photo-matching even further. It is also worth noting that all four observers frequently used only the anterior (facial) view to compare captures, as images of this view tended to be of higher quality. The high success rate of photo-matching thus indicates that individuals can be recognised on the basis of markings on this region of their

body alone. Close inspection of dorsal and lateral photographs of sufficient quality showed individuals can also be distinguished on the basis of each of these regions. If observers initially compare markings on one region of the body to determine whether two captures are a match, and then compare markings on the other regions to confirm their conclusion, the chance of making false-positive errors will be remote, as it is highly unlikely that two different individuals in a population will have similar or identical markings on all four regions, given the high degree of pattern complexity in this species. Such a procedure should also reduce the chance of making false-negative errors as a result of failed photo-matches, because it should be considerably less likely that an observer will fail to match four regions than a single region.

4.2 POST-TRIAL ASSESSMENT

Analysis of all photographs of all 45 frogs identified three combinations (out of a possible eight) which resulted in optimal allocation of individuals to subgroups (Table 4). The decision as to which of these combinations to use in future studies will depend on the clarity of the frontal stripe–upper lip markings, dorso-lateral stripes, and paired dorsal markings in photographs taken in the field. It is recommended that the combination of frontal stripe–upper lip markings and 1st & 2nd paired dorsal markings is considered initially, because, as noted previously, the position of the frog can influence the appearance of the dorso-lateral stripe immediately above the attachment point of the forelimb. Although it should be possible to align each frog so that it is parallel to the camera and its upper body is not twisted to one side, and to adjust the position of its limbs, this will necessitate additional manipulation of the frog, thereby increasing stress and the chance that the individual will move out of the photograph frame.

4.3 POTENTIAL PROBLEMS WITH USING NATURAL MARKINGS TO IDENTIFY INDIVIDUALS

If two or more individuals in a population have such similar markings that they cannot be distinguished from one another, false positive errors will result. *Leiopelma archeyi* has quite complex patterns made up of numerous characters, which reduces the likelihood of this occurring.

Unequal probabilities of recognition due to variability in individual distinctiveness can bias estimates of population size in capture-recapture studies (Hammond 1986 in Friday et al. 2000). Thus, the proportion of individuals in a population that lack distinct markings is an important factor to consider when conducting capture-recapture studies using natural markings. In the present study, observers frequently compared only a single body region (anterior/facial) during photo-matching, but still had a very high success rate (99.2% across both trials). Also, the single capture that was not recognised as a recapture due to failed photo-matching by one observer was correctly identified by the other three observers, indicating that this individual did have recognisable markings. Despite the relatively small sample size (n = 45), the

results indicate that the proportion of individuals in the population that lack distinctive markings on the face is probably very low. Given that inspection of the dorsal, right lateral, and left lateral photographs of the 45 individuals showed that these regions also possess distinct markings, the chance that successive captures of an individual *L. archeyi* would not be accurately identified is very low, at least within the source population for the individuals included in this study. Thus, it appears unlikely that heterogeneity in individual distinctiveness is a significant issue for photographic identification using natural markings in *L. archeyi*. If, however, individuals lacking recognisable markings are encountered, population estimates can be adjusted to account for this 'unmarked' component of the population (e.g. Whitehead et al. 1997). It is therefore advisable to include an assessment of capture distinctiveness in the database of identifications (e.g. high/1 = this capture has distinct, immediately recognisable markings, therefore it is highly unlikely that the capture has been incorrectly identified as a result of an error during photo-matching; low/2 = this capture does not have distinct, easily recognisable markings, therefore it may be relatively easy to miss a match or make an incorrect photo-match). As with assigning captures to subgroups, it would be necessary to evaluate how consistent different observers are in their assessments of capture distinctiveness.

The potential for an individual's markings to change over time warrants consideration, particularly given that a number of *L. archeyi* included in this study had short, very thin lines connecting markings of interest. If such markings do change over time, small sections of thin line could disappear between sampling occasions, and successive captures of the one individual would be assigned to different subgroups. Stephenson & Stephenson (1957) noted that *Leiopelma* colour patterns did not change during two years in captivity, but it is not clear whether they were referring to detailed outlines of markings or coloured areas, or to general observations such as 'dorsal surface mostly bright green'. Captive colonies of *L. archeyi* can be used to determine whether or not markings change with age or state of health. In the interim, it would be advisable to assign all individuals that have thin sections of line connecting markings of interest to both relevant subgroups (i.e. the subgroup where *x* side of *y* character is continuous, and the alternate subgroup where *x* side of *y* character is discontinuous).

Reaser (1995) noted that handling times were frequently longer when using natural markings to identify individuals than when using artificial marks such as toe-clips, and suggested that stress resulting from capture and handling could induce acute behavioural and physiological changes. *Leiopelma archeyi* is a small, cryptically coloured species, and therefore marked individuals will need to be caught and handled on each sampling trip in order to determine their identity regardless of the marking technique used. While the interval between capture and release may potentially be longer when using natural markings (due to the time it will take to position the frog and photograph it), the actual amount of handling should not be significantly greater than that involved with reading toe-clips, elastomer marks, etc. (which requires manipulation of the animal within the hand), and should be less than that involved with the initial application of artificial marks.

5. Conclusions

Two major conclusions regarding the use of natural markings to identify individual *L. archeyi* can be drawn from this study. First, assigning unidentified captures to subgroups is feasible and will substantially reduce the number of previous captures an unidentified capture must be compared with, thus reducing the time taken to identify it. Second, photo-matching within subgroups is highly successful (99.2% of the time overall in this study).

6. Recommendations

In the field, frog handling times, and the total time between capture and release, should be kept to an absolute minimum. If possible, images should be taken using a digital camera.

Increasing the efficiency and accuracy of identification

Only high-quality images should be used.

Captures should be assigned to subgroups prior to photo-matching, with observers ensuring that any characters that cannot be classified with absolute certainty are recorded as such, and that each capture is assigned to all potentially relevant subgroups.

To ensure consistency, all observers should be trained in the classification of the characters used to assign captures to subgroups. Training should conclude with a formal evaluation (using a reference set of captures) of the newly trained observer's ability to (i) correctly classify individuals and (ii) assess capture distinctiveness.

In cases where the markings of interest are connected only by a thin line, the capture should be assigned to both relevant subgroups (i.e. *x* side of *y* character continuous and *x* side of *y* character discontinuous).

When captures are assigned to more than one subgroup, photographs should be included in each subgroup folder (as opposed to placing a photograph in the folder for the primary subgroup only, and writing the details on the outside of the folder for each alternate subgroup, as was the case in this study). This should increase efficiency and decrease the risk of observer error. A divider could be included in each folder, with photographs of previous captures that were assigned to that subgroup as the primary subgroup located in front of the divider, and photographs of previous captures that were assigned to that subgroup as an alternative subgroup located behind the divider.

When a photograph that is identified as a recapture is of better quality than the photograph of that individual included in the subgroup folder, the poorer quality image should be replaced with the higher quality image.

When photo-matching, observers should initially compare markings on one region of the body to determine whether two captures are a match, and then compare markings on the other regions to confirm their conclusion.

A scale should be included in photographs, as this will allow SVL to be used as a secondary feature to verify identification.

Due to uncertainty regarding the stability of colour patterns over time, these should (at this stage) only be used as a secondary feature to verify identification when the time interval between successive captures is relatively short.

All identifications should be independently verified by a second observer to reduce the chances of incorrect photo-matches resulting in false positive errors, and simple mistakes, failed photo-matches or intra-observer inconsistencies resulting in false negative errors.

An assessment of distinctiveness for each capture should be included in the database of identifications.

Identifying the combination of characters to use in a key

The combination of the frontal stripe-upper lip markings (FS-ULM) with the first and second paired dorsal markings (1st & 2nd PDM) should be assessed initially. If this combination turns out to be unsuitable, the combination of FS-ULM with the dorso-lateral stripe (eye-forelimb) (D-L stripe (E-FL)) or of D-L stripe (E-FL) with 1st & 2nd PDM should be examined.

If all three of these combinations are unsuitable, the continuous/discontinuous upper lip markings (ULM), continuous/discontinuous lower lip markings (LLM), and continuous/discontinuous central upper lip marking & adjacent upper lip marking (CULM & adjacent ULM) characters could be assessed. However, it should be kept in mind that, as these characters all require a clear view of the line of the mouth, individuals may need to be manipulated into the necessary position (i.e. head parallel to the substrate or angled slightly upwards, rather than angled downwards as has been observed in a number of the photographs in this study), thus increasing handling times.

7. Acknowledgements

I wish to thank A. Holzapfel, who gave me the opportunity to complete this study and provided valuable assistance along the way, J. Webster and B. Waldman, who supplied the photographs used in this study, and M. Crossland and L. Marshall, who also provided valuable assistance on several occasions.

8. References

Agler, B.A. 1992: Testing the reliability of photographic identification of individual fin whales (*Balaenoptera physalus*). *Report of the International Whaling Commission 42*: 731–737.

Bailey, L.L. 2004: Evaluating elastomer marking and photo identification methods for terrestrial salamanders: marking effects and observer bias. *Herpetological Review 35*: 38–41.

Bell, B.D. 1978: Observations on the ecology and reproduction of the New Zealand Leiopelmid frogs. *Herpetologica 34*: 340-354.

Bell, B. D. 2001. Recent population declines of Archey's frog *Leiopelma archeyi* in the central Coromandel Ranges. April 2001 report to Native Frog Recovery Group, Department of Conservation, Wellington, New Zealand (unpublished), 16 p.

Bell, B.D.; Daugherty, C.H.; Hitchmough, R.A. 1998: The taxonomic identity of a population of terrestrial *Leiopelma* (Anura: Leiopelmatidae) recently discovered in the northern King Country, New Zealand. *New Zealand Journal of Zoology 25*: 139–146.

Bretagnolle, V.; Thibault, J-C.; Dominici, J-M. 1994: Field identification of individual ospreys using head marking pattern. *Journal of Wildlife Management 58*: 175–178.

Cannatella, D. 1995: Leiopelma. The Tree of Life Web Project, http://tolweb.org/tree?group=Leiopelma&contgroup=Salientia.

Caro, T.M. 1994: Cheetahs of the Serengeti Plains: group living in an asocial species. University of Chicago Press, Chicago. 478 p.

Caughley, G.; Sinclair, A.R.E. 1994: Wildlife ecology and management. Blackwell Science, Massachusetts. 334 p.

Chappell, R. 2001a: Survey of Archey's frogs *Leiopelma archeyii*, Mt Moehau 15 January 2001. Report to Department of Conservation, Hauraki Area Office (unpublished), 7 p.

Chappell, R. 2001b: Survey of Archey's frogs *Leiopelma archeyii*, Mt Moehau 26 July 2001. Report to Department of Conservation, Hauraki Area Office (unpublished), 4 p.

Clarke, R.D. 1972: The effect of toe clipping on survival in Fowler's toad (*Bufo woodhousei fowleri*). *Copeia 1972*: 182–185.

Connelly, R.M.; Melville, A.J.; Keesing, J.K. 2002: Abundance, movement and individual identification of leafy seadragons, *Phycodurus eques* (Pisces: Syngnathidae). *Marine and Freshwater Research 53*: 777–780.

Davis, T.M.; Ovaska, K. 2001: Individual recognition of amphibians: effects of toe clipping and fluorescent tagging on the salamander *Plethodon vehiculum*. *Journal of Herpetology 35*: 217–225.

Denton, J.S.; Beebee, T.J.C. 1993: Reproductive strategies in a female-biased population of natterjack toads, *Bufo calamita*. *Animal Behaviour 46*: 1169–1175.

Dixon, D.R. 2003: A non-invasive technique for identifying individual badgers *Meles meles*. *Mammal Review 33*: 92–94.

Donnelly, M.A.; Guyer, C.; Juterbock, J.E.; Alford, R.A. 1994: Techniques for marking amphibians. Appendix 2, pp. 277–284 in Heyer, W.R.; Donnelly, M.A.; McDiarmid, R.W.; Hayek, L-A.C.; Foster, M.S. (Eds): Measuring and monitoring biological diversity: standard methods for amphibians. Smithsonian Institution Press, Washington DC. 364 p.

Ferner, J.W. 1979: A review of marking techniques for amphibians and reptiles. *Herpetological Circular 9*. Society for the Study of Amphibians and Reptiles Publications, Salt Lake City. 41 p.

Forcada, J.; Aguilar, A. 2000: Use of photographic identification in capture-recapture studies of Mediterranean monk seals. *Marine Mammal Science 16*: 767–793.

Forester, D.C. 1977. Comments on the female reproductive cycle and philopatry by *Desmognathus ochrophaeus* (Amphibia, Urodela, Plethodontidae). *Journal of Herpetology 11*: 311–316.

Friday, N.; Smith, T.D.; Stevick, P.T.; Allen, J. 2000: Measurement of photographic quality and individual distinctiveness for the photographic identification of humpback whales, *Megaptera novaeangliae*. *Marine Mammal Science 16*: 355–374.

Frost, D.R. 2002: Amphibian species of the world: an online reference. V2.21 (15 July 2002). Electronic database available at http://research.amnh.org/herpetology/amphibia/index.html.

Gill, D.E. 1978: The metapopulation ecology of the red-spotted newt, *Notophthalmus viridescens* (Rafinesque). *Ecological Monographs 48*: 145–166.

Glockner, D.A.; Venus, S.C. 1983: Identification, growth rate, and behaviour of humpback whale (*Megaptera novaeangliae*) cows and calves in the waters off Maui, Hawaii, 1977–79. Pp. 223–258 in Payne, R. (Ed.): Communication and behavior of whales. Westview Press, Colorado. 643 p.

Golay, N.; Durrer, H. 1994: Inflammation due to toe-clipping in natterjack toads (*Bufo calamita*). *Amphibia–Reptilia 15*: 81–83.

Goodall, J. 1986: The Chimpanzees of Gombe: patterns of behaviour. The Belknap Press of Harvard University Press, Cambridge, Massachusetts. 673 p.

Gowans, S.; Whitehead, H. 2001: Photographic identification of northern bottlenose whales (*Hyperoodon ampullatus*): sources of heterogeneity from natural marks. *Marine Mammal Science 17*: 76–93.

Gronell, A.M. 1984: Courtship, spawning and social organization of the pipefish, *Corthoichthys intestinalis* (Pisces: Syngnathidae) with notes on two congeneric species. *Zeitschrift für Tierpsychologie 65*: 1–24.

Hagström, T. 1973: Identification of newt specimens (Urodela, *Triturus*) by recording the belly pattern and a description of photographic equipment for such registrations. *British Journal of Herpetology 4*: 321–326.

Hallmen, M. 1999: Individualerkennung melanistischer Tiere bei der gewöhnlichen Strumpfbandnatter *Thamnophis sirtalis sirtalis* (Serpentes: Colubridae). *Salamandra 35*: 113–122.

Hitchmough, R. (Comp.) 2002: New Zealand Threat Classification System lists 2002. *Threatened Species Occasional Publication 23*. Department of Conservation, Wellington, New Zealand. 210 p.

Kelly, M.J. 2001: Computer-aided photograph matching in studies using individual identification: an example from Serengeti cheetahs. *Journal of Mammalogy 82*: 440–449.

Lemckert, F. 1996: Effects of toe-clipping on the survival and behaviour of the Australian frog *Crinia signifera*. *Amphibia-Reptilia 17*: 287–290.

Lüddecke, H.; Amézquita, A. 1999: Assessment of disc clipping on the survival and behaviour of the Andean frog *Hyla labialis*. *Copeia 1999*: 824–830.

Markowitz, T.M.; Harlin, A.D.; Würsig, B. 2003: Digital photography improves the efficiency of individual dolphin identification. *Marine Mammal Science 19*: 217–223.

Newman, D.G. 1982: New Zealand herpetological research, the work of the NZ Wildlife Service. *Herpetofauna 14*: 1–10.

Øien, I.J.; Aarvak, T.; Lorentsen, S-H.; Bangjord, G. 1996: Use of individual differences of lesser white-fronted geese *Anser erythropus* in belly patches in population monitoring at a staging ground. *Fauna Norvegica, Series C 19*: 69–76.

Payne, R.; Brazier, O.; Dorsey, E.M.; Perkins, J.S.; Rowntree, V.J.; Titus, A. 1983: External features in southern right whales (*Eubalaena australis*) and their use in identifying individuals. Pp. 371–445 in Payne, R. (Ed.): Communication and behavior of whales. Westview Press, Colorado. 643 p.

Pennycuick, C.J. 1978: Identification using natural markings. Pp. 147–159 in Stonehouse, B. (Ed.): Animal marking: recognition marking of animals in research. MacMillan Press, London. 243 p.

Petersen, J.C.B. 1972: An identification system for zebra (*Equus burchelli*, Gray). *East African Wildlife Journal 10*: 59–63.

Reaser, J. 1995: Marking amphibians by toe-clipping: a response to Halliday. *Froglog 12*: 1–2.

Reaser, J.K.; Dexter, R.E. 1996: *Rana pretiosa* (spotted frog). Toe clipping effects. *Herpetological Review 27*: 195–196.

Scott, D.K. 1978: Identification of individual Bewick's swans by bill patterns. Pp. 160–168 in Stonehouse, B. (Ed.): Animal marking: recognition marking of animals in research. MacMillan Press, London. 243 p.

Sheldon, S.; Bradley, C. 1989: Identification of individual adders (*Vipera berus*) by their head markings. *Herpetological Journal 1*: 392–396.

Steinicke, H.; Ulbrich, K.; Henle, K.; Grosse, W-R. 2000: Eine neue Methode zur fotografischen Individualidentifikation mitteleuropäischer Halsbandeidechsen (Lacertidae). *Salamandra 36*: 81–88.

Stephenson, E.M.; Stephenson, N.G. 1957: Field observations on the New Zealand frog, *Leiopelma*. *Transactions of the Royal Society of New Zealand 84*: 867–882.

Stevick, P.T.; Palsbøll, P.J.; Smith, T.D.; Bravington, M.V.; Hammond, P.S. 2001: Errors in identification using natural markings: rates, sources, and effects on capture-recapture estimates of abundance. *Canadian Journal of Fisheries and Aquatic Science 58*: 1861–1870.

Tilley, S.G. 1980: Life histories and comparative demography of two salamander populations. *Copeia 1980*: 806–821.

van Gelder, J.J.; Strijbosch, H. 1996: Marking amphibians: effects of toe clipping on *Bufo bufo* (Anura: Bufonidae). *Amphibia–Reptilia 17*: 169–174.

Whitehead, H. 1990: Computer assisted individual identification of sperm whale flukes. *Reports of the International Whaling Commission*, *special issue 12*: 71–77.

Whitehead, H.; Gowans, S.; Faucher, A.; McCarrey, S. 1997: Population analysis of northern bottlenose whales in the Gully, Nova Scotia. *Marine Mammal Science 13*: 173–185.

Appendix 1

Assessment of the suitability of various characters for inclusion in a key to divide frogs into subgroups

CHARACTER*	SUITABLE	IF UNSUITABLE, REASON
E-ULM	Yes	
N-ULM	Yes	
FS	No	Discontinuous for all 20
FS breaks	No	Cannot always determine path of frontal stripe through a broken area with certainty, so different observers could follow different paths and end up with different numbers of breaks (Appendix 5, Figure A5-3)
UL spots	No	Cannot clearly and easily define boundaries of area, because cannot always determine path of frontal stripe with certainty (as above); angle of head and lighting also influence visibility and appearance of spots
FT spots	No	As for UL spots
CULM & adjacent ULM	Not for this trial, but may be suitable with better quality photographs	Can be clearly defined, but it was not possible to use this character in this study because the anterior (facial) photographs were often taken from a high angle and/or the frog's head was angled downwards, so that the edge of the upper lip was not visible; also, when visible, these markings usually appeared to be continuous
ULM	As for CULM & adjacent ULM	As for CULM & adjacent ULM
LLM	As for CULM & adjacent ULM	Can be clearly defined, but it was not possible to use this character in this study because the edge of the lower lip was not visible in a number of photographs; also, when visible, these markings were usually discontinuous
DT base	Not for this trial, and unlikely to be useful at all	Can be clearly defined, but was continuous in only four individuals, so is not particularly good for dividing individuals into subgroups (80% of individuals are contained within 50% of subgroups)
DT corners	Not for this trial, and unlikely to be useful at all	Can be clearly defined, but only one individual had complete corners on both sides, and three others had complete corners on the left side only (two of which were questionable due to poor photo quality), so this character was not useful for dividing individuals into subgroups
1st & 2nd PDM	As for CULM & adjacent ULM	Can be clearly defined, but poor quality of dorsal photographs made it difficult to use this character in these trials
FL stripes	No	Cannot clearly and easily define what constitutes a single stripe (How far around the limb does the marking need to extend to be considered a stripe? How far around the limb do two adjacent stripes need to be joined before they are considered a single stripe?) (Appendix 5, Figure A5-4); even with an unambiguous definition of the character, the position of the arm can change the appearance of stripes
D-L stripe	No	Individuals often held their hind limbs close to their sides, rendering it impossible to see the posterior section of the stripe
D-L stripe (E-FL)	Yes	

* See Table 1 for explanation of abbreviations used

Appendix 2

Combination of characters used to define each of the 16 subgroups

SUBGROUP	CHARACTER	CONTINUOUS OR DISCONTINUOUS
A	right N-ULM	continuous
	left N-ULM	continuous
	right D-L stripe (E-FL)	continuous
	left D-L stripe (E-FL)	continuous
B	right N-ULM	continuous
	left N-ULM	continuous
	right D-L stripe (E-FL)	continuous
	left D-L stripe (E-FL)	discontinuous
C	right N-ULM	continuous
	left N-ULM	continuous
	right D-L stripe (E-FL)	discontinuous
	left D-L stripe (E-FL)	continuous
D	right N-ULM	continuous
	left N-ULM	continuous
	right D-L stripe (E-FL)	discontinuous
	left D-L stripe (E-FL)	discontinuous
E	right N-ULM	continuous
	left N-ULM	discontinuous
	right D-L stripe (E-FL)	continuous
	left D-L stripe (E-FL)	continuous
F	right N-ULM	continuous
	left N-ULM	discontinuous
	right D-L stripe (E-FL)	continuous
	left D-L stripe (E-FL)	discontinuous
G	right N-ULM	continuous
	left N-ULM	discontinuous
	right D-L stripe (E-FL)	discontinuous
	left D-L stripe (E-FL)	continuous
H	right N-ULM	continuous
	left N-ULM	discontinuous
	right D-L stripe (E-FL)	discontinuous
	left D-L stripe (E-FL)	discontinuous
I	right N-ULM	discontinuous
	left N-ULM	continuous
	right D-L stripe (E-FL)	continuous
	left D-L stripe (E-FL)	continuous

* See Table 1 for explanation of abbreviations used

SUBGROUP	CHARACTER	CONTINUOUS OR DISCONTINUOUS
J	right N-ULM	discontinuous
	left N-ULM	continuous
	right D-L stripe (E-FL)	continuous
	left D-L stripe (E-FL)	discontinuous
K	right N-ULM	discontinuous
	left N-ULM	continuous
	right D-L stripe (E-FL)	discontinuous
	left D-L stripe (E-FL)	continuous
L	right N-ULM	discontinuous
	left N-ULM	continuous
	right D-L stripe (E-FL)	discontinuous
	left D-L stripe (E-FL)	discontinuous
M	right N-ULM	discontinuous
	left N-ULM	discontinuous
	right D-L stripe (E-FL)	continuous
	left D-L stripe (E-FL)	continuous
N	right N-ULM	discontinuous
	left N-ULM	discontinuous
	right D-L stripe (E-FL)	continuous
	left D-L stripe (E-FL)	discontinuous
O	right N-ULM	discontinuous
	left N-ULM	discontinuous
	right D-L stripe (E-FL)	discontinuous
	left D-L stripe (E-FL)	continuous
P	right N-ULM	discontinuous
	left N-ULM	discontinuous
	right D-L stripe (E-FL)	discontinuous
	left D-L stripe (E-FL)	discontinuous

Appendix 3

Worksheet to aid identification of individual frogs

Leiopelma archeyi **PHOTO ID**

WORKSHEET FOR DETERMINING WHICH SUBGROUP A FROG BELONGS TO

FROG FIELD NO. ________________
R nare – mouth markings **con / discon / ?**
L nare – mouth markings **con / discon / ?**
R dorso-lateral stripe **con / discon / ?**
L dorso-lateral stripe **con / discon / ?**
SUBGROUP ____________________

FROG FIELD NO. ________________
R nare – mouth markings **con / discon / ?**
L nare – mouth markings **con / discon / ?**
R dorso-lateral stripe **con / discon / ?**
L dorso-lateral stripe **con / discon / ?**
SUBGROUP ____________________

FROG FIELD NO. ________________
R nare – mouth markings **con / discon / ?**
L nare – mouth markings **con / discon / ?**
R dorso-lateral stripe **con / discon / ?**
L dorso-lateral stripe **con / discon / ?**
SUBGROUP ____________________

FROG FIELD NO. ________________
R nare – mouth markings **con / discon / ?**
L nare – mouth markings **con / discon / ?**
R dorso-lateral stripe **con / discon / ?**
L dorso-lateral stripe **con / discon / ?**
SUBGROUP ____________________

FROG FIELD NO. ________________
R nare – mouth markings **con / discon / ?**
L nare – mouth markings **con / discon / ?**
R dorso-lateral stripe **con / discon / ?**
L dorso-lateral stripe **con / discon / ?**
SUBGROUP ____________________

FROG FIELD NO. ________________
R nare – mouth markings **con / discon / ?**
L nare – mouth markings **con / discon / ?**
R dorso-lateral stripe **con / discon / ?**
L dorso-lateral stripe **con / discon / ?**
SUBGROUP ____________________

FROG FIELD NO. ________________
R nare – mouth markings **con / discon / ?**
L nare – mouth markings **con / discon / ?**
R dorso-lateral stripe **con / discon / ?**
L dorso-lateral stripe **con / discon / ?**
SUBGROUP ____________________

FROG FIELD NO. ________________
R nare – mouth markings **con / discon / ?**
L nare – mouth markings **con / discon / ?**
R dorso-lateral stripe **con / discon / ?**
L dorso-lateral stripe **con / discon / ?**
SUBGROUP ____________________

Appendix 4

Trial 2: details for captures that were identified incorrectly

OBSERVER	TEST NO.	TYPE OF ERROR	REASON FOR ERROR	FURTHER DETAILS
1	79	false negative	Initial capture and recapture assigned to different subgroups	Due to **position of frog** in photograph of initial capture (TC64) (the frog was twisted to the left, so that the left D-L stripe (E-FL) appeared continuous, and TC64 was therefore assigned to a different subgroup to TC79)
2	104	false negative	Initial capture and recapture assigned to different subgroups	Due to **intra-observer inconsistency** (the observer made different decisions about the right N-ULM, and thus assigned the initial capture and the recapture to different subgroups)
	137	false negative	Initial capture and recapture assigned to different subgroups	Due to **position of frog** in photograph (the frog was twisted to the right, so that the right D-L stripe (E-FL) appeared continuous)
3	51	potential false negative	Assigned to different subgroups by the observer in question and the primary observer	Due to **inter-observer inconsistency** (Observer 3 reached a different conclusion regarding the left and right N-ULM to the primary observer, so the primary observer assigned this capture to one subgroup, while Observer 3 assigned it to a different subgroup)
	79	false negative	Initial capture and recapture assigned to different subgroups	Due to **position of frog** in photograph of initial capture (TC64) (the frog was twisted to the left, so that the left D-L stripe (E-FL) appeared continuous, and TC64 was therefore assigned to a different subgroup from TC79)
	84	false negative	Initial capture and recapture assigned to different subgroups	Due to **intra-observer inconsistency** (the observer made different decisions about the left D-L stripe (E-FL), and thus assigned the initial capture and the recapture to different subgroups)
	89	false negative	Initial capture and recapture assigned to different subgroups	Due to **inter-observer inconsistency** (Observer 3 reached a different conclusion regarding the left N-ULM from the primary observer, so the primary observer assigned the initial capture to one subgroup, and Observer 3 assigned the recapture to a different subgroup)
	97	potential false negative	Assigned to different subgroups by the observer in question and the primary observer	Due to **inter-observer inconsistency** (Observer 3 reached a different conclusion regarding the right N-ULM from the primary observer, so the primary observer assigned this capture to one subgroup, while Observer 3 assigned it to a different subgroup)
	114	false negative	Failed to photo-match this capture to the initial capture of the same individual (T9)	
	115	potential false negative	Assigned to different subgroups by the observer in question and the primary observer	Due to **inter-observer inconsistency** (Observer 3 reached a different conclusion regarding the right N-ULM and the right D-L stripe (E-FL) from the primary observer, so the primary observer assigned the initial capture to one subgroup, and Observer 3 assigned the recapture to a different subgroup)

OBSERVER	TEST NO.	TYPE OF ERROR	REASON FOR ERROR	FURTHER DETAILS
3 (continued)	133	false negative	Initial capture and recapture assigned to different subgroups	Due to **error** (recorded the subgroup incorrectly for the initial capture—wrote down E instead of H)
	137	false negative	Initial capture and recapture assigned to different subgroups	Due to **position of frog** in photograph (the frog was twisted to the right, so that the right D-L stripe (E-FL) appeared continuous)
4	133		Initial capture and recapture assigned to different subgroups	Due to **intra-observer inconsistency** (Observer 4 made different decisions about the left N-ULM, and thus assigned the initial capture and the recapture to different subgroups)
	137		Initial capture and recapture assigned to different subgroups	Due to **position of frog** in photograph (the frog was twisted to the right, so that the right D-L stripe (E-FL) appeared continuous)

Appendix 5

Photographs showing characters used in identifying individual frogs

Figure A5-1 (opposite). Photographs illustrating potential characters for assigning captures to sub-groups. Superimposed lines identify areas/markings for inspection. See Table 1 for explanation of abbreviations used.

A. E-ULM (continuous),

B. N-ULM (continuous),

C. FS (discontinuous),

D. UL spots,

E. FT spots,

F. CULM & adjacent ULM (right side continuous, left side discontinuous),

G. DT base (continuous),

H. DT corners (right and left both complete)

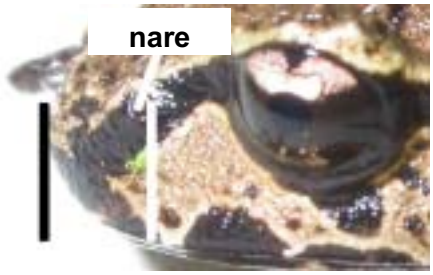
I. 1st & 2nd PDM (right side continuous, left side discontinuous),

J. FL stripes,

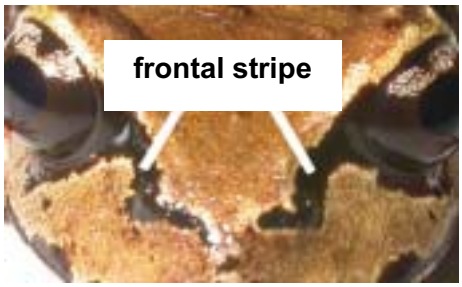
K. D-L stripe (E-FL) (discontinuous).

A

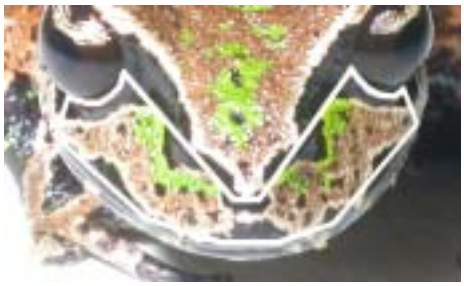
B



C

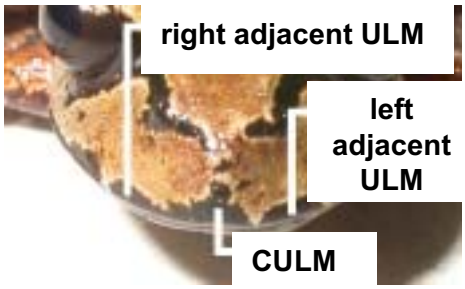


D



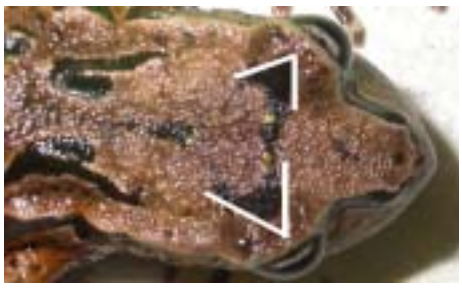
E

F

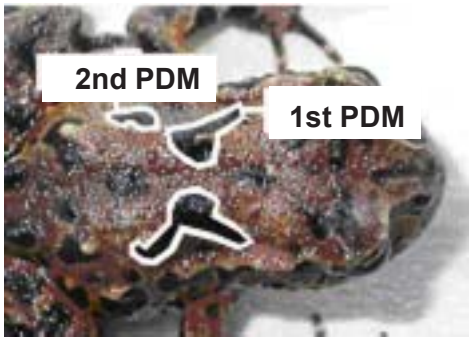


G

H



I



J

K

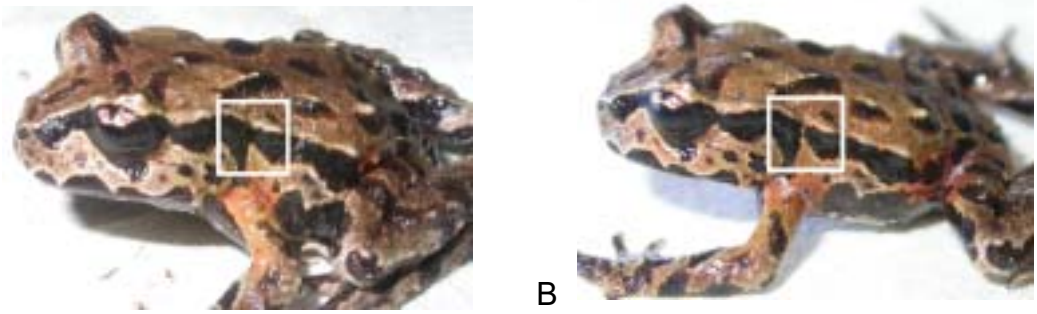


Figure A5-2. Two different photographs (test captures) of a single individual; the white box superimposed on each photograph indicates the region of interest.
A. The left D-L stripe (E-FL) appears to be continuous.
B. The left D-L stripe (E-FL) is clearly discontinuous.

Figure A5-3. Photograph illustrating the problem with the FS breaks character. The lines superimposed on the photograph illustrate two potential paths of the frontal stripe through the break between the nares: if the frontal stripe followed the path of the white line, there is only a single break in the stripe, if the frontal stripe followed the path of the grey line (incorporating the two black spots directly above the central upper lip marking), there are three breaks in the stripe.

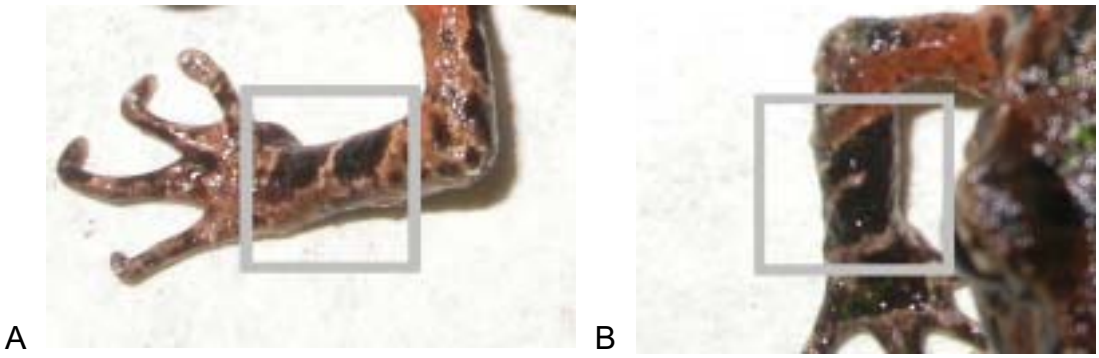


Figure A5-4. Photographs illustrating two problems with the FL stripes character; the grey box superimposed on each photograph indicates the region of interest.
A. How far do markings have to extend around the forelimb to be counted as stripes?
B. To what extent do two adjacent stripes need to be connected to one another before they are considered to be a single stripe?