Manaaki Whenua
Landcare Research

Assessing mustelid dispersal and the Predator Free Taranaki trapping programme using population genomics

Prepared for: Taranaki Mounga Project Ltd and Taranaki Regional Council

January 2023

Assessing mustelid dispersal and the Predator Free Taranaki trapping programme using population genomics

Contract Report: LC4241

Andrew J. Veale, Thomas R. Etherington

Manaaki Whenua – Landcare Research

Reviewed by:

Dr Amy Vaughan
Genomics postdoctoral fellow

Dr Chris Niebuhr
Vertebrate Ecologist
Manaaki Whenua – Landcare Research

Approved for release by:

Chris Jones
Portfolio Leader – Wildlife Management & Conservation Ecology
Manaaki Whenua – Landcare Research

Contents

Summary 1

1 Introduction 1

2 Objectives 2

3 Methods 2

3.1 Genomic analyses 2

3.2 Sex determination 2

3.3 Assembling the final data set to calculate relatedness 3

3.4 Spatial patterns for genetic similarity 3

3.5 Cost-surface analyses 3

3.6 Dispersal kernels 4

3.7 Description of the genomic results 4

3.8 Genetic data summary 5

4 Results 6

4.1 Species ratios and sex ratios 6

4.2 Relatedness calculations 6

4.3 Cost-surface analysis 13

4.4 Dispersal distances and kernels 13

5 Discussion 15

6 Conclusions 19

7 Recommendations 19

8 Acknowledgements 20

9 References 20

Appendix 23

Summary

Project and client

- Predator Free Taranaki is leading a project attempting to eradicate mustelids from the ring-plain around the mounga: the Taranaki Mustelid Control Programme.
- Manaaki Whenua – Landcare Research (MWLR) was contracted by Taranaki Mounga Project Ltd and the Taranaki Regional Council to investigate mustelid movements on the ring-plain and assess the efficacy of the trapping programme.

Objective

- To explore the potential for genomic analyses to determine the relatedness and dispersal of mustelids across the Taranaki ring-plain and mounga.

Methods

- We obtained 554 mustelid tissue samples from the Taranaki ring-plain and national park. These samples were taken from 357 stoats, 121 weasels, and 70 ferrets.
- All the stoat and weasel samples had DNA extracted.
- DNA was quantified and standardised.
- We performed genotyping by sequencing (GBS) to obtain genomic data for each sample.
- Pairwise relatedness was calculated between samples, identifying close kin pairs.
- Dispersal kernels (distributions) were created for each species from the distances between close kin.
- We compared genetic relatedness to geographical distance to understand how landscape affects connectivity.
- Finally, we compared the sex ratio of each species and the number of close kin detected in the sample to assess trapping efficiency.

Results

- GBS sequencing performed on 99 weasels and 289 stoats from across Taranaki resulted in over 400,000 variable genetic markers within each species, enabling fine-scale relatedness to be ascertained.
- We recorded 77 close kin (parent offspring or full sibling) pairs of stoats and two close-kin pairs of weasels.
- The sex ratio of stoats was c. 1:1, but for weasels it was considerably skewed, with 1 female to 2.4 males.
- No population structure or barriers to dispersal were detected for either species.
- No landscape features were observed that affect stoat or weasel dispersal: dispersal appears to be unaffected by landscape features.
- Long-distance dispersal (up to 40 km for stoats and up to 17.4 km for weasels) was recorded.

- Low sample numbers meant there was no statistically significant difference in dispersal distances between sexes, but for both species the maximum dispersal distances were recorded for males.
- The trapping network for stoats is likely to be sufficient to have a trap in all (or nearly all) home ranges, and so most stoats will encounter traps.
- Survival of stoats that avoid traps is still likely to be the primary source of recruitment, though there may also be some immigration due to the long potential dispersal distances.

Conclusions

- This is the first study of stoat and weasel dispersal in the North Island and in these habitats. The results indicate similar dispersal characteristics to the better-studied beech forest in the South Island.
- Stoats are capable of long-distance dispersal within the trapping area.
- Due to the dispersal capacity of stoats, combined with the habitat patch sizes in the Taranaki ring-plain, dispersal distance appears unaffected by landscape features: distance from potential sources of stoats is the primary factor influencing the likelihood of immigration.
- The trapping programme is dense enough to put the vast majority of stoats at risk (most stoats, including females, should encounter traps).
- The trapping programme is catching a much smaller proportion of weasels present and is particularly missing female weasels.

Recommendations

- Stoat dispersal up to, and possibly greater than, 40 km is occurring in the region, which means trapping must be maintained over large landscape scales to prevent reinvasion if an area is to be cleared of stoats (true eradication).
- Given that female stoat dispersal was only confirmed to c. 7 km, this should be regarded as a minimum trap network buffer size (though it is likely that small numbers of individuals will still go through such a buffer).
- From the genetic evidence, and based on what we know about stoat biology and from other stoat eradication programmes, it appears likely that most recruitment is from surviving residents rather than immigrants, so detecting and removing residents remains the primary concern for managers.
- The Taranaki trapping programme is dense enough to put the vast majority of stoats at risk and so should be maintained.
- Weasels are currently unmanaged with this trapping network. New traps specific to weasels and a pilot area attempting to control them are advisable.

1 Introduction

Mammal predator control programmes are currently being carried out in the Taranaki region within the overarching context of Predator Free 2050. The goal of Predator Free Taranaki is to remove all invasive mammalian predators from Egmont National Park, to work towards eradication of the surrounding ring-plain, and to prevent reinvasion into the region.

In order to successfully manage these predators at this landscape scale, Predator Free Taranaki wish to understand how each species uses and disperses across the Taranaki ring-plain. To achieve this, fine-scale movement data are needed, particularly in relation to habitat use, dispersal, and reinvasion potential. Information on these parameters will enable future trapping regimes to target immigrants to the national park and will provide realistic scenarios of the reinvasion risks.

Tracking individual mustelids is a challenging task because they are generally at low population densities, they are neophobic, and they are small, which limits the devices they can wear without negative welfare impacts (García-Díaz & Niebuhr 2019; Niebuhr & Veale 2020). A further risk of relying solely on radio-tracked animals is that due to cost constraints only a small selection of animals will be able to be monitored. These animals may not represent average behaviour, nor are rarer long-distance dispersal movements that are important to long-term pest management likely to be observed. Tracking data will therefore be useful for monitoring individual fine-scale movement and habitat use, utilising updated technology (Niebuhr 2021), but population connectivity is likely to require other tools.

An alternative strategy to tracking individuals is to understand how mustelids use and move across the ring-plain using genetic measures to look at how connected populations are in this landscape. By correlating patterns of genetic relatedness to landscape features, it is possible to ascertain the movement of genes across that environment, and hence the animals these genes belong to (Etherington et al. 2014). Due to recent advances in genomic sequencing and analyses, we can now accurately describe the precise relationships between pairs of individuals (Flanagan & Jones 2019; Galla et al. 2020), which should increase the resolution of such analyses.

This technique has several advantages over more traditional tracking techniques. One is that the genetic landscape represents the true movement of genes across the whole landscape rather than the movements of specific individuals. Some age or sex classes may be missed through chance in a tracking study, and this is particularly likely for very young individuals during their dispersive phase. Also, the large number of samples enables a more general picture to be presented across a given landscape. Finally, genetic studies can be undertaken on dead individuals, and therefore does not require extra work to be undertaken in the field.

All of these factors mean that landscape genetic studies of pest species can provide different and complementary information on the movement of individuals across a landscape: tracking studies give fine detail on specific habitat use and movement patterns

of individuals, but genetic studies give broader generalisations of how individuals have, on average, moved across the landscape (Storfer et al. 2018).

2 Objectives

In this research we explore the potential for genomic analyses to determine the relatedness and dispersal of mustelids across the Taranaki ring-plain and in Egmont National Park, specifically to assess and inform eradication efforts.

3 Methods

3.1 Genomic analyses

A total of 554 mustelid tissue samples were obtained from both the ring-plain trapping (491 samples) conducted by Taranaki Regional Council, and trapping conducted by the Department of Conservation in Egmont National Park (63 samples). These samples came from 357 stoats (*Mustela erminea*), 121 weasels (*M. nivalis*), and 70 ferrets (*M. putorius*), along with a handful of unidentified mustelid samples. While most tissue samples were ears, preserved in 70% ethanol, some were tails, and a few of the more degraded samples were collections of miscellaneous unidentified tissue and fur.

DNA was extracted at the EcoGene laboratory at MWLR and sequenced at the Genomnz Animal Genomics Group (AgResearch, New Zealand). (See the Appendix for detailed laboratory and bioinformatic methods.) This sequencing produced a set of variable markers known as single nucleotide polymorphisms (SNPs) distributed across the genome.

3.2 Sex determination

To determine sex we used two different measures: the ratio of the average sequencing depth (the number of sequencing reads per location) on the X chromosome compared to that of the 21 autosomes, and the ratio of the average sequencing depth on the Y chromosome compared to that of the 21 autosomes. Males only have one X chromosome while females have two, so the sequencing depth will differ between them, and only males will have a Y chromosome (though any sequencing error or barcode jumping (where the barcode that identifies an individual swaps to another individual) could add a few reads that are inaccurately assigned to this chromosome). We found that, particularly for the Y chromosome for weasels (since the assembly was based on the stoat genome), there were issues with sexing, so we used a small section (248 SNPs) that appeared to best match the X-chromosome data.

3.3 Assembling the final data set to calculate relatedness

Once sex was determined for all samples we created filtered data sets, removing the two sex chromosomes along with any SNPs not assigned to the 21 autosomes, and conducted further filtering detailed in the appendix. This removal of the sex chromosomes is because if they were included they would lead to errors in measuring relatedness.

We used KGD, an R software package that calculates kinship (genetic relatedness) using GBS (genotyping-by-sequencing) with depth adjustment (Dodds et al. 2015) to calculate pairwise relatedness between all individuals. This software calculates relatedness, taking into account the level of support for each data point (the read-depth). This matrix of relatedness values was then compared to the pairwise geographical distances between samples. KGD provides a relatedness index that goes from 0 to 1, with 0 being unrelated for many generations and 1 being identical (either the same individual sampled twice or identical twins). Parent/offspring pairs and full sibling pairs will average 0.5 relatedness, with a recommended cut-off threshold for identifying close kin of 0.35 or greater (Dodds et al. 2015).

3.4 Spatial patterns for genetic similarity

For both stoats and weasels we defined kinship groups based on clustering in KGD (Dodds et al. 2015). For stoats, these consisted of family groups (which most likely originated from the same litter) having a kinship coefficient of greater than 0.35. For weasels, given the small number of individuals that had these levels of kinship, we lowered the threshold to display patterns of lower kinship (cousins, grandparents, etc.).

In KGD we also conducted a principal coordinate analysis (PCA) of kinship for each species, and mapped the first two principal coordinates and displayed these on maps of Taranaki using the package ggmaps (Kahle & Wickham 2013) in R. A principal coordinate analysis (PCA) is a statistical technique that takes multi-dimensional data (such as population genetic data, which have a high number of markers per individual) and reduces the number of dimensions while preserving the maximal amount of information. Each principal coordinate (PC) maximises the differences explained by the data (showing how much of this variation is explained by it as a percentage). We can then plot these PCs against each other in two of three dimensions to visualise these patterns. The distance between individuals in a PCA plot is a representation of how similar their multi-dimensional genetic data are to each other.

3.5 Cost-surface analyses

Cost-surfaces identify landscape features that are obstacles or conduits for dispersal (Etherington 2016) that can be developed using a variety of methods (Zeller et al. 2012). We developed a cost-surface using trapping data, on the basis that traps with more capture presences are located in landscape conduits for animal dispersal and traps with capture absences are located in landscape obstacles for animal dispersal.

We used data only from DOC200, DOC250, and podiTrap traps, because these require resetting and hence are checked regularly, so we can be confident absences are absences – unlike Goodnature A24 traps, which may have killed an animal but the carcass is often not found. To inform this model, we compared successful traps with unsuccessful traps nearby. We also only used kills within 25 m (the spatial resolution of the analysis) of a trap as presences to allow for location error. Absences were identified as traps without captures at least 25 m from a kill but within 500 m of a kill (representing a 78.5 ha home range, or core home range, of a stoat (King & Veale 2020). This helped to ensure that those traps used as absences were close to places with confirmed presence of the species, and hence for which absence of a kill would relate to movement preferences rather than complete absence of the species from that part of the landscape.

Landscape features were included in the cost-surface analysis at two different spatial scales. Site-scale variables were calculated at 25 m (the spatial resolution of the analysis), and they included elevation, terrain ruggedness, distance to cover, and distance to rivers. Home-range scale variables were calculated based on a 500 m radius around each point of interest to try to capture landscape characteristics at a scale comparable to a mustelid home-range. These home-range scale variables were calculated as the amount of cover and river habitat within the 500 m radius. Because all pairwise combinations of landscape variables were statistically independent for both species (Pearson correlations of <0.55), all 63 possible combinations of the landscape variables were assessed via a logistic regression model fitted to the landscape and trapping data. All logistic regression models were analysed with the Akaike Information Criterion (AIC), which was used to identify the best cost-surface model (Etherington et al. 2014), and the receiver operating characteristic area under the curve (AUC) was used to measure predictive ability (Fielding & Bell 1997).

3.6 Dispersal kernels

The pairwise genomic data were processed to identify independent dispersal events. Only pairwise combinations with a relatedness of at least 0.35 were used as they will represent close-kin individuals. We also only used pairwise combinations at least 1,000 m apart to ensure that the close-kin were sufficiently far apart to represent a dispersal event rather than a within home range movement. To prevent pseudo-replication, each animal was allowed to contribute to only one dispersal event, and where one animal was related to more than one other animal, the longest distance connection was selected. Dispersal events were categorised by sex as either male or female (where the related pair were of the same sex) or as unknown if they were different sexes.

Dispersal kernels quantify the likely dispersal distance of a species using a probability distribution. We used Mann-Whitney *U*-tests (Gregory 1978) to check for differences in dispersal values between sexes, with separate dispersal kernels fitted for each sex where a difference was detected.

3.7 Description of the genomic results

Genomic GBS sequences were successfully obtained for 99 weasels and 289 stoats after removing individuals with low-quality sequences (Figure 1). Due to the comparatively low

number of ferret samples, and their clustered distribution across the landscape effectively decreasing the data set, we were unable to analyse the landscape connectivity of this species.

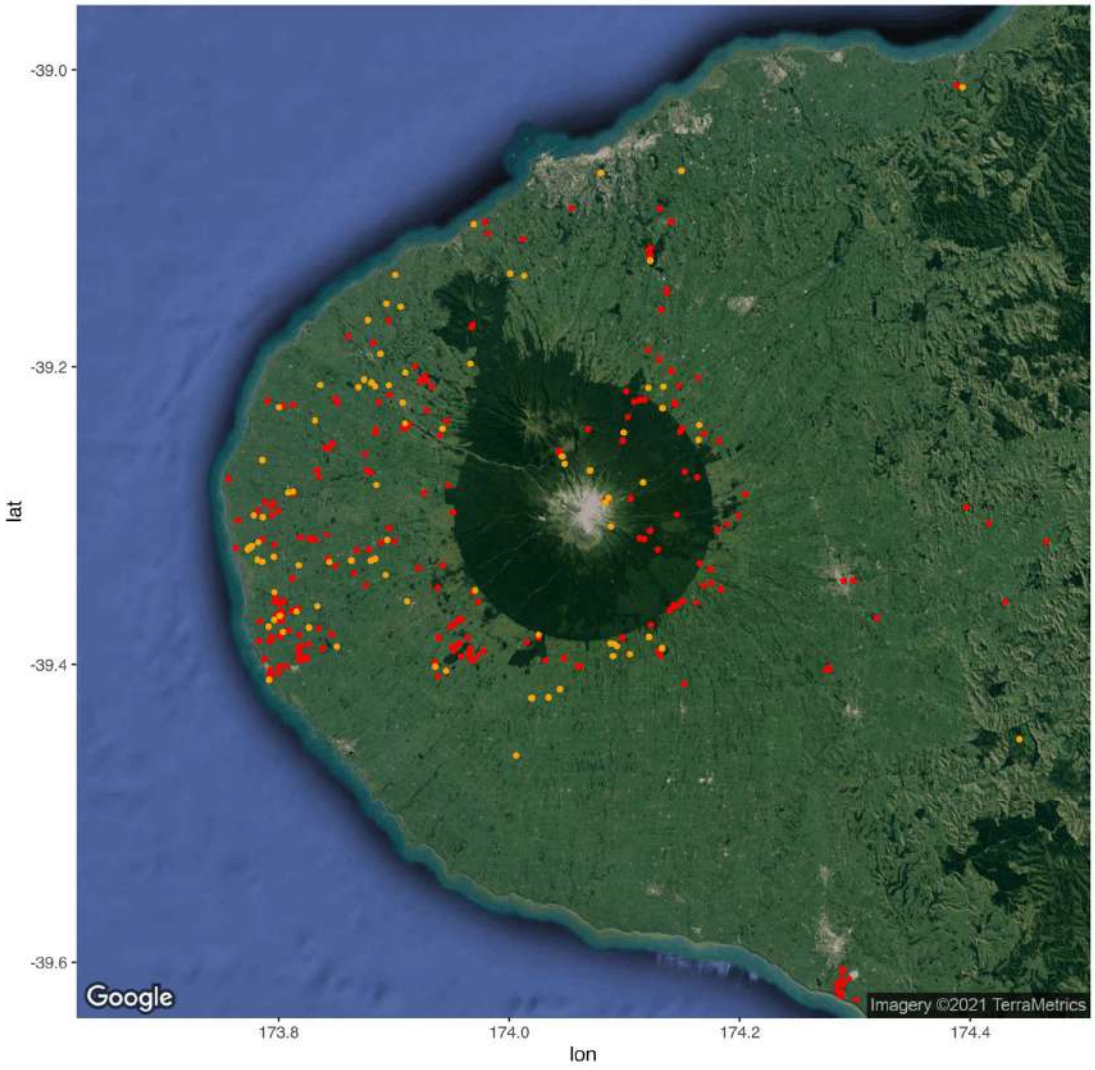

Figure 1. Map of the Taranaki region showing source locations of all mustelid samples that were successfully sequenced. Stoats = red, weasels = orange.

3.8 Genetic data summary

Once all individuals were identified to the correct species, the two data sets were separated and reassembled and analysed. For the stoat data set, 55,880 SNPs were retained after filtering, while for the weasel data set 141,825 SNPs were retained. Given the different sizes of the two data sets and the filtering parameters that are affected by sample size, these numbers may not reflect the relative genetic diversity of these two species accurately. However, it appears that weasels have considerably more genetic variation than stoats in Taranaki. All quality metrics for each SNP were high, with the mean

sample depth (8.4) and mean co-call rate (0.98) both indicating high reliability of the data, and comparable to the livestock genetic studies performed at AgResearch. Mean sample depth is the average number of sequence reads for each genetic locus in each individual, and co-call rate is the proportion or percentage of samples in which a confident genotype call could be made.

4 Results

4.1 Species ratios and sex ratios

The number of stoats successfully sequenced was 289 (out of 357) and the number of weasels successfully sequenced was 99 (out of 121). The ratio of weasels to stoats caught across the programme is therefore 1 weasel to 2.95 stoats. The success rate for sequencing stoats and weasels was identical. For stoats the sex ratio recorded was 1 female to 1.05 males, which is statistically indistinguishable from 1:1 (Chi squared = 0.085 with 1 degree of freedom, $P = 0.7705$). For weasels, however, the sex ratio was 1 female to 2.4 males, which is a marked, significant departure from 1:1 (Chi squared equals 11.879 with 1 degree of freedom, $P = 0.0006$). There is no chance that this was caused by random variation or small sample sizes – more male weasels are being caught than females. Given the sex ratio in nature is 1:1, this means that male weasels are 2.4 times more likely than females to be caught in this trapping programme.

4.2 Relatedness calculations

The number of close-kin pairs detected for stoats was 77, while for weasels it was only two. The number of close-kin pairs detected in a population is related to the sample size (as sample size increases, the chances of detecting close-kin pairs goes up). Therefore, to compare this pattern accurately between species we randomly subsampled the genotyped stoats (with 10 replicates) to give 10 randomised stoat data sets identical in size to the weasel data set to make further comparisons. From these stoat subsamples we detected, on average, 49 close-kin pairs (range 43–55).

Forty-nine close-kin pairs for stoats compared to two close-kin pairs for weasels is a stark difference that requires explanation. This is not due to differences in the ability to detect close-kin pairs in each species in terms of genetic resolution: it is an attribute of the trapping programme. Having a low proportion of close-kin pairs detected in a sample (assuming similar biology; e.g. litter size, natural mortality) indicates that a smaller proportion of the population is being trapped.

4.2.1 Stoats

We were able to accurately assess the pairwise relationships of all stoats sequenced. This is key to understanding dispersal based on genomic data, because close kin (parent offspring or full siblings) were originally in the same location before they dispersed. We can see this dispersal by identifying pairs of close kin and examining the distances

between them. From these data we found 77 pairs of close kin among the stoats, and we were able to determine more distant (e.g. half sibling, cousin) relationships.

A matrix of the relatedness of all samples (comparing all samples with each other) is shown in Figure 2. In this plot, individuals that are highly related to each other are positioned next to each other, and the relatedness value will be redder, while unrelated pairs will be yellow. On the diagonal each individual is compared to itself. In this plot, if there was population subdivision you would see large square patterns in these data, indicating that these separate groups of individuals are highly related to each other, with low relatedness to other groups. Only one such feature exists which is the group C1, highlighted. Other smaller kinship groups, such as C2 and C3, are likely to be single litters of highly related individuals.

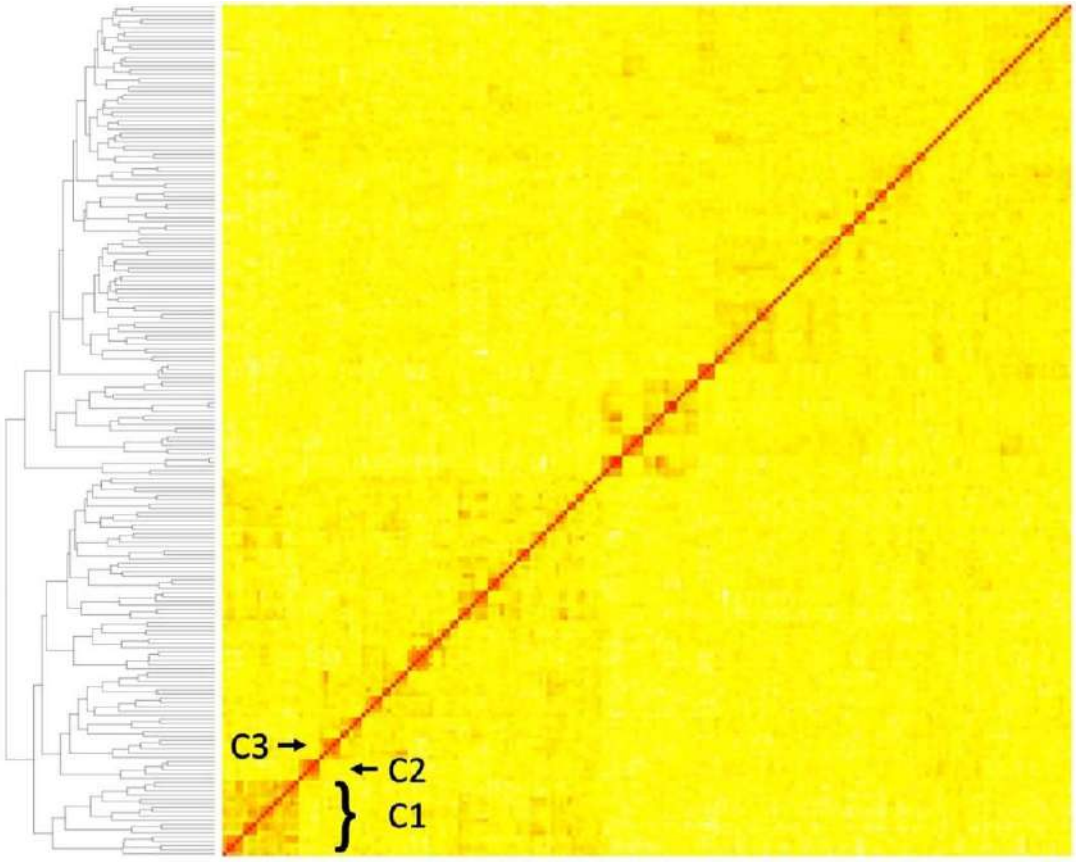

Figure 2. Relatedness heat map for stoats in Taranaki. This shows the relatedness of all individuals vs all individuals, with 'self-relatedness' on the diagonal. Yellow is low relatedness, red is high relatedness. The neighbour-joining tree to the left shows pairwise similarities between individuals. Some groups of individuals displayed in Figure 3 are highlighted (C1–C3). C1 is a group of related individuals showing potential population structure (segments of the population that are differentiated from each other), while C2 and C3 are examples of litters of related stoats.

From mapping these kinship groups we can see that while individuals are often caught close to individuals they are related to (relatives are not randomly distributed across the trapping programme), related individuals can also be caught at a considerable distance from each other (40+ km). Genetic population structure is when two or more geographic

regions of populations become somewhat genetically differentiated because they comparatively rarely exchange genes. The only hint of population structure (showing possible decreased connectivity between regions) was the separation of group C1; however, stoats from other clusters were caught nearby, with both C3 and C7 having overlapping ranges with C1. This result indicates that this cluster of related individuals is not a landscape population structure, but instead all of the individuals share some recent common ancestors, such as having a father/grandfather that sired multiple litters in the area. Sampling was particularly dense in this area, meaning that it was more likely that a larger number of somewhat related stoats (e.g. cousins) would be sampled, increasing the likelihood of observing this pattern.

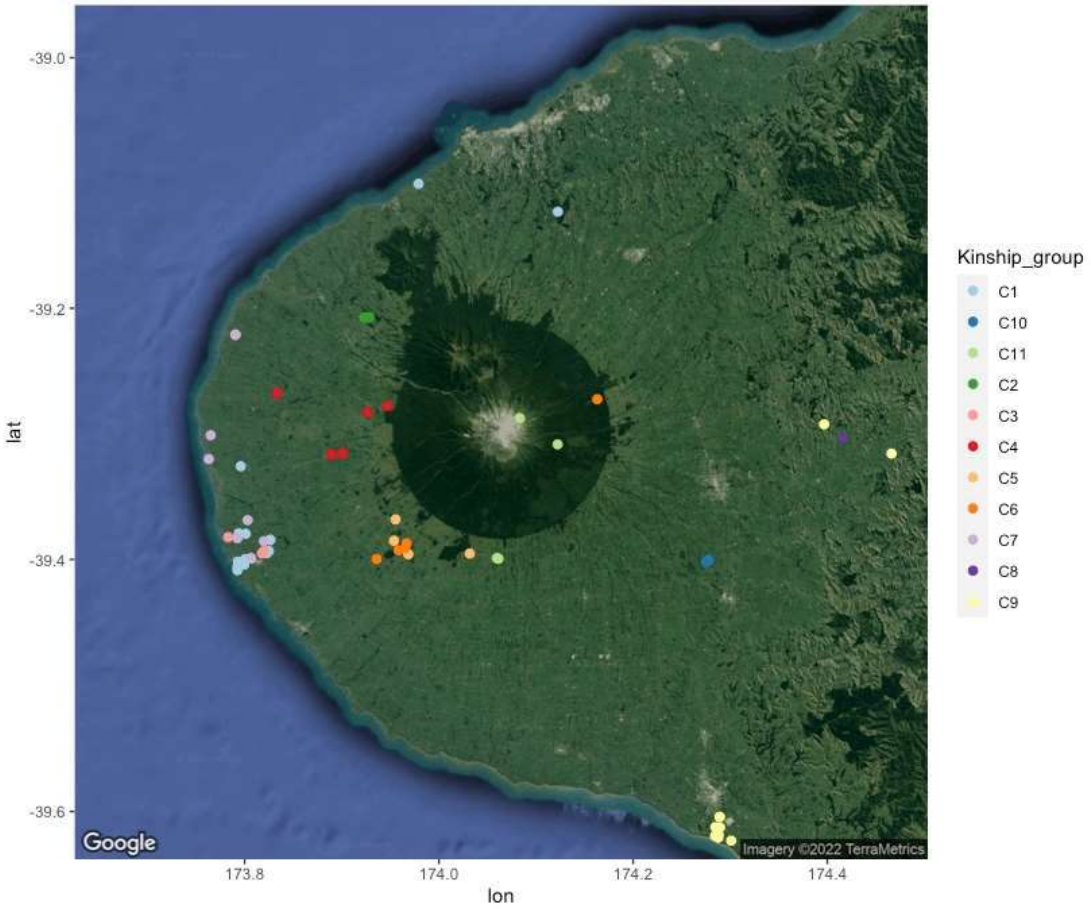

Figure 3. Kinship groups of stoats (with at least four members) in Taranaki, represented by a different colour for each group. These kinship groups are not necessarily strict family groups, but are groups of related individuals that cluster together (extended family groups). The primary large-scale population structure observed in Figure 2 was the grouping of C1, whereby the stoats are largely distributed close to each other to the southwest of the mountain near the coast, but two males from this cluster are found to the north.

As noted in 3.4, above, the distance between individuals in a principal components analysis plot is a representation of how similar their multi-dimensional genetic data are to each other. In practice, PCAs will generally reveal if there are genetically distinct clusters of

individuals: the higher the PCs, the smaller the distinct groups (if these exist). The first PC (PC1) explains the highest proportion of the data, followed by PC2, etc.

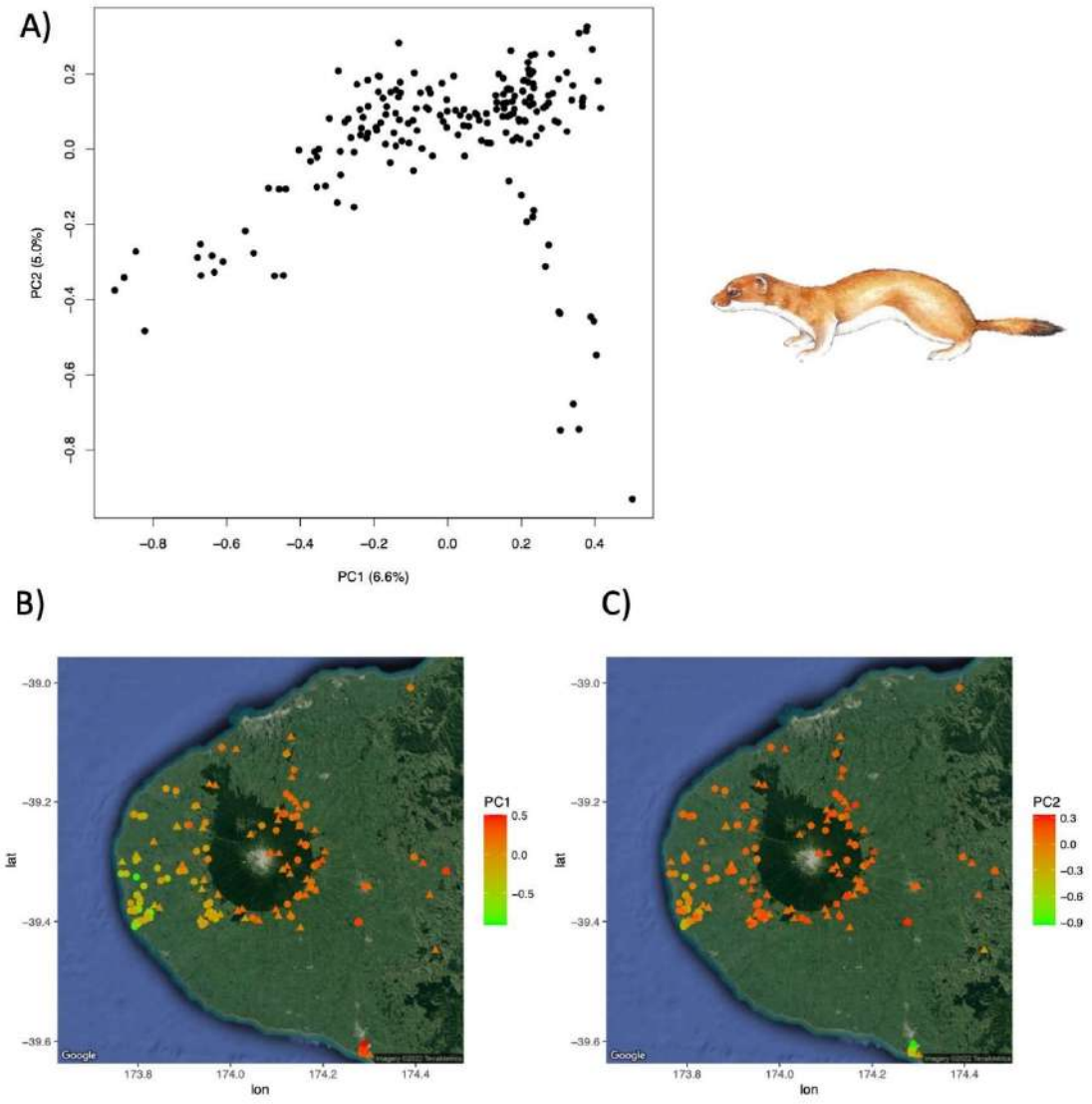

Figure 4. Principal coordinate analyses (PCAs) of genetic distances between stoats. A) PC1 vs PC2; B) PC1 displayed as a gradient over a map of Taranaki; C) PC2 displayed as a gradient over a map of Taranaki.

When we look for population structure (indicating differentiation of groups) for stoats in the PCA (Figure 4A), we see that there is some (albeit low) differentiation across the landscape, with a gradient of similarity across PC1 and a differentiated group spreading across PC2. If there were barriers to dispersal with significant landscape effects we would see clusters separated by gaps in the PCA, and abrupt changes in the colour of individuals displayed on the map around features that inhibit dispersal. If we compare the positions geographically for each stoat with their position on the PCs, we see that PC1 is highly correlated with longitude, with low values in the west and higher values in the east, and a smooth transition across the landscape (Figure 4B). PC2 is somewhat correlated with

latitude, with the separate group consisting of the individuals caught around Hawera in the south being the divergent group (Figure 4C). While not displayed, further PCs did not show any clustering or spatial pattern.

All of these patterns together show that there are no strong barriers to dispersal across this landscape for stoats: individuals that are caught near each other are generally slightly more related than random, while individuals caught further away from each other are generally less related, a pattern known as isolation by distance (IBD). The only genetic clustering we see in the PCA is the cluster caught to the south near Hawera; these stoats were clustered together geographically, with no stoats caught between this cluster and the mountain. Therefore, the physical distance of this cluster from other samples explains this pattern.

4.2.2 Weasels

We were able to accurately assign relationships between the weasels sequenced. However, only two pairs of close-kin (parent–offspring or full siblings) individuals were recorded (Figure 5), along with two samples that were clearly mislabelled replicates.

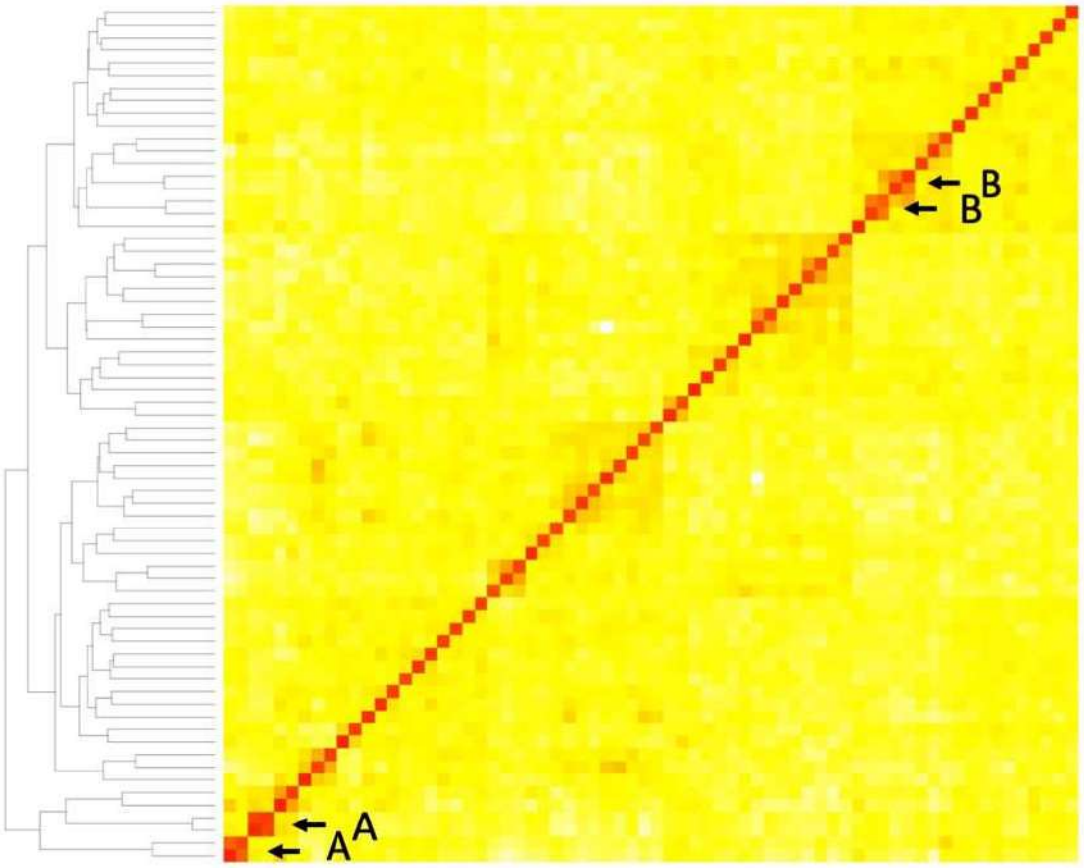

Figure 5. Heat map of relatedness for weasels in Taranaki, with a neighbour-joining similarity tree shown to the left. Two duplicated individuals where one was supplied mislabelled are indicated by the two As, and the two close-kin pairs are indicated by the Bs.

On average, the weasels were less related to each other than stoats were to each other in this region, as well as there being very few highly related pairs of weasel individuals. This can be seen by the low levels of orange on Figure 5 compared to Figure 2: most weasels sampled had no related individuals sampled, even out to more distant relationship categories, while most stoats had these relatives sampled.

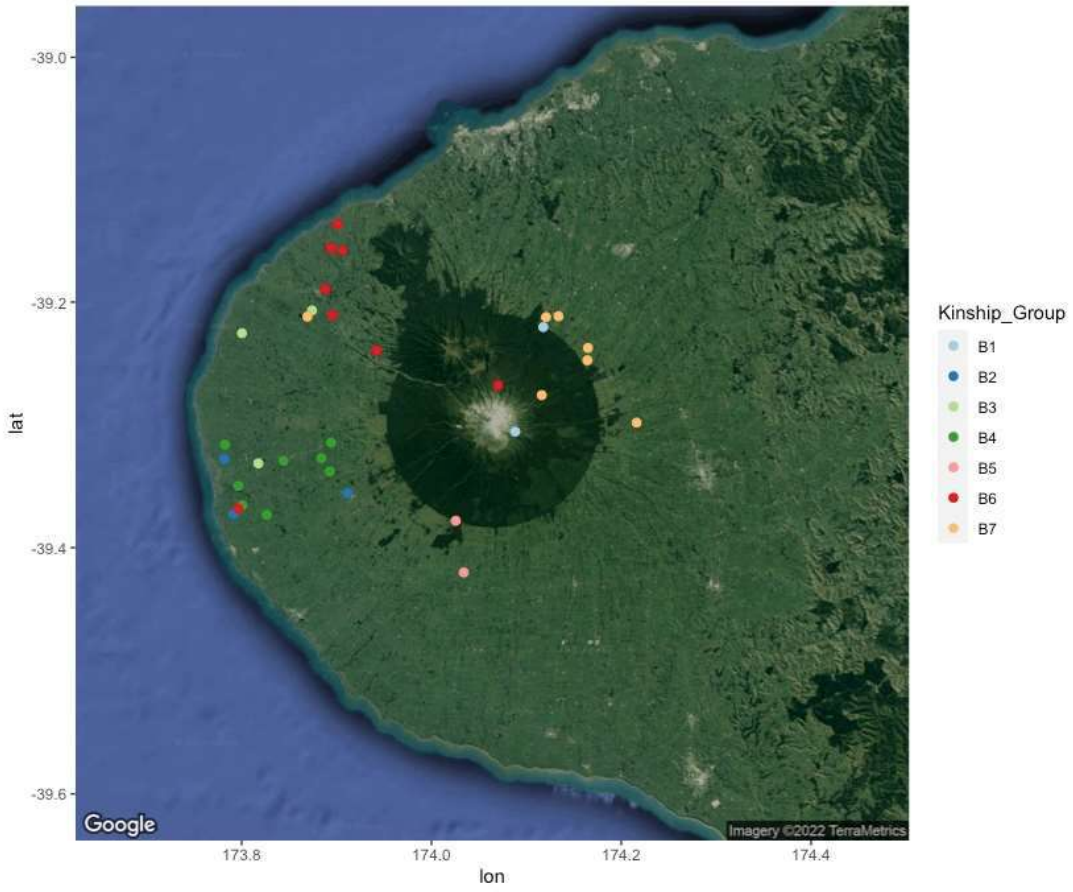

Figure 6. Extended kinship groups for weasels in Taranaki. These kinship groups are not necessarily strict family groups, but instead are groups of related individuals that cluster together (extended family groups).

From the map of kinship groups for weasels (Figure 6) we see a similar pattern as in stoats, with related individuals generally clustered together geographically but with overlap between kinship groups, showing that these represent family groups rather than true spatial structure of the population. This provides no evidence for any boundaries to dispersal or changes in dispersal ability across these landscapes. The criterion for similarity within a kinship group for display was slightly lowered for weasels compared to stoats, as pairwise relatedness values were generally lower, and this lowering of similarity enabled a better pattern to be shown visually.

When we look for population structure (indicating differentiation of groups) for weasels in the PCA (Figure 7A), we see that there is minimal population structure, with no clear clusters. Again, if we compare the positions geographically for each weasel with their

position on the PCs, as we did for the stoats, we see that both PC1 and PC2 have reasonably consistent gradients, with PC1 generally higher in the west and low in the east (Figure 7B), and PC2 lower in the southwest and higher in the northeast (Figure 7C). Again, no strong clustering is evident, and no dramatic changes of colour (representing each PC) related to landscape features are obvious. Visually, it may appear that those north of the Stoney River are redder and those south of the Stoney River greener for the plot of PC2 (Figure 7C), but we tested for this effect and found no significant difference: the pattern of genetic differentiation could be entirely explained by distance, with no effect of the river.

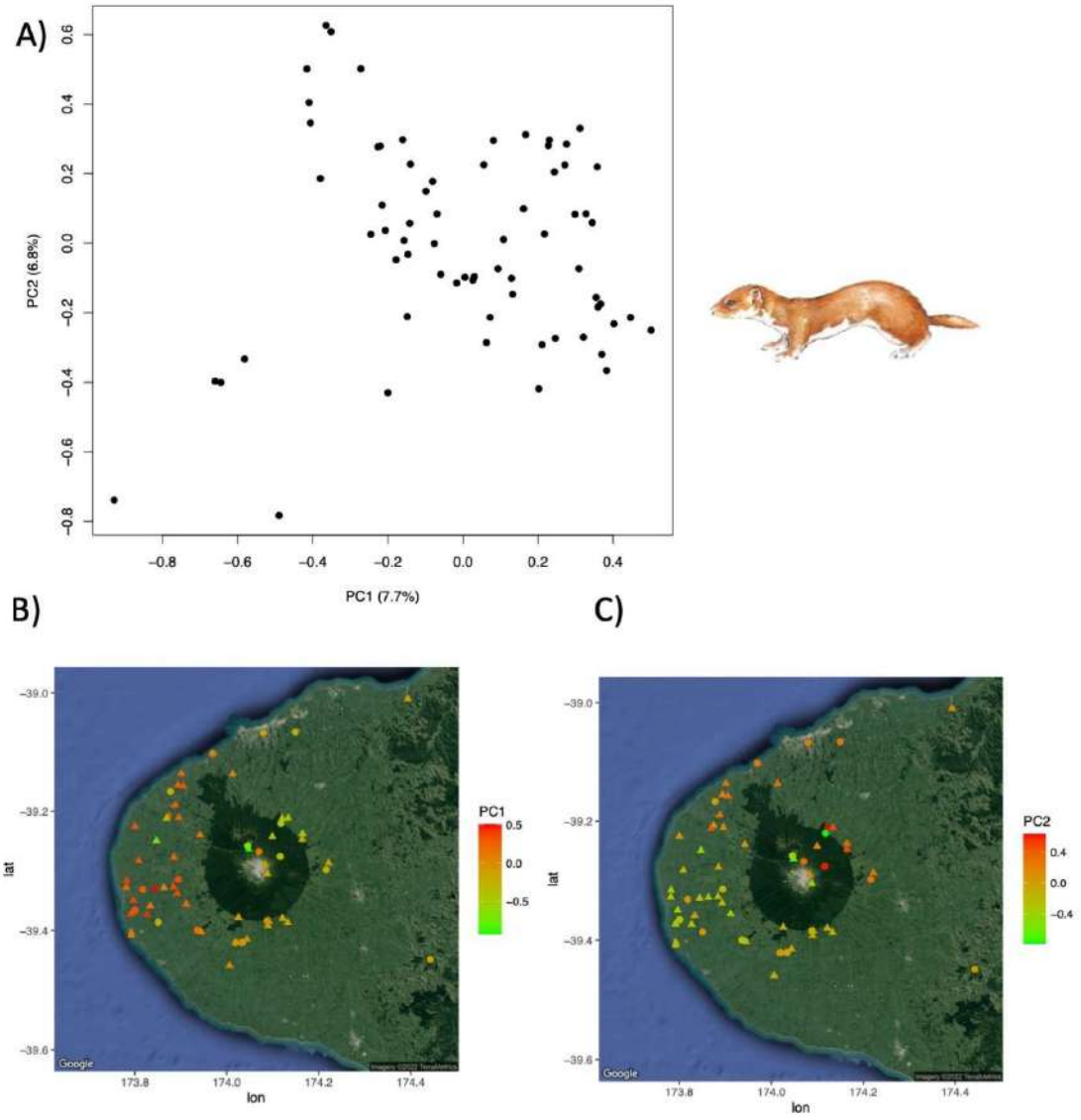

Figure 7. Principal coordinate analyses (PCAs) of genetic distances between weasels. A) PC1 vs PC2; B) PC1 displayed as a gradient over a map of Taranaki; C) PC2 displayed as a gradient over a map of Taranaki.

4.3 Cost-surface analysis

The trapping data consisted of 197 presences and 613 absences for stoats, and 82 presences and 347 absences for weasels. The logistic regression analysis did not identify any statistically meaningful cost-surface, as the best model for both species was barely better than random (stoat AUC = 0.524, weasel AUC = 0.599) meaning that we could not detect any effect of landscape features on trapping rates and hence dispersal. There was also no pattern detectable from the genetic similarity of individuals, as shown in both PCAs. We interpret these findings as indicating that, within the Taranaki landscape, stoat and weasel movement cannot be predicted on the basis of landscape features and, as such, a cost-surface model is not warranted and straight-line distance is the most appropriate dispersal metric.

4.4 Dispersal distances and kernels

No cost-surfaces were successfully fitted to the trapping data so we could only report dispersal in terms of straight-line distances. Only two close-kin pairs, and hence two dispersal events, for weasels were detected (Figure 8), with a male dispersal of 17.427 km and another dispersal of 7.075 km that could have been either sex. With only two dispersal events, a dispersal kernel could not be fitted for weasels. There were 22 (male = 6, female = 6, either = 10) stoat dispersal events detected (Figure 9).

For stoats, the mean geographical distance between close-kin pairs was 7.96 km, the median distance was 3.44 km, and the maximum distance was 38.67 km. There were five close-kin pairs out of 22 that were separated by over 10 km. While the maximum definite male distance of 17.487 km was longer than the maximum female distance of 7.201 km, the Mann-Whitney *U*-test indicated there was no statistical difference between the sexes ($U = 22$, $P = 0.59$). Therefore, a single log-normal dispersal kernel was fitted for all 22 dispersal events (log mean = 8.362, log standard deviation = 1.087, Figure 10).

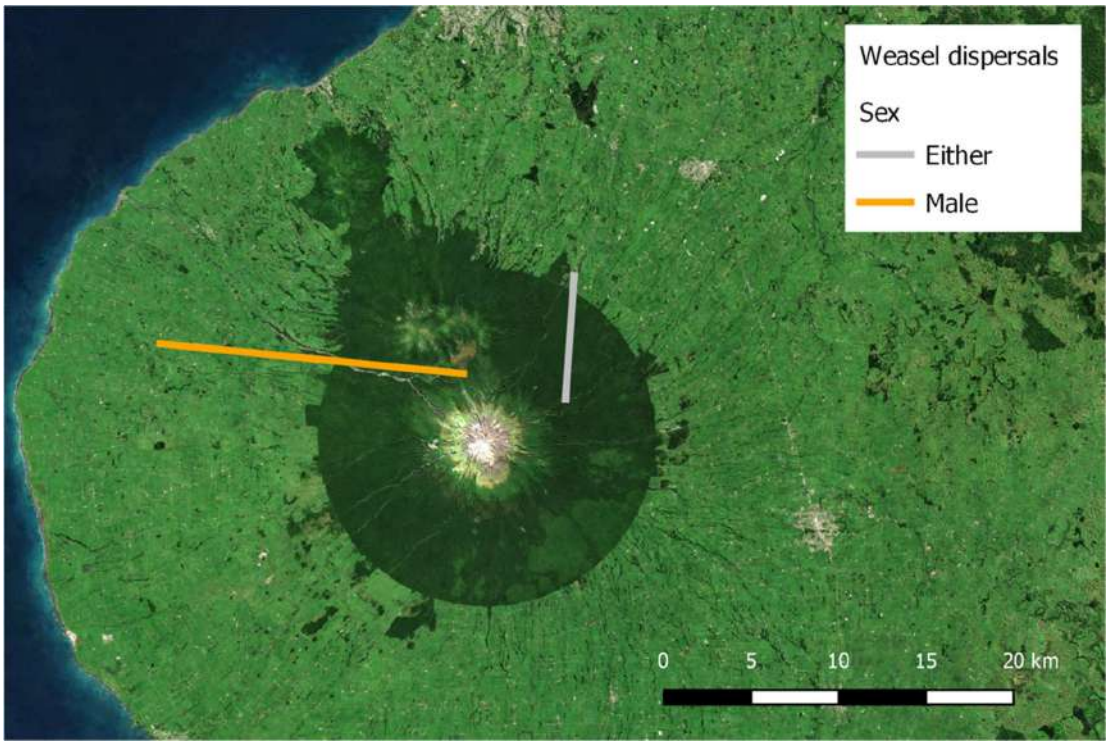

Figure 8. Weasel dispersal events showing lines between capture locations of close-kin pairs.

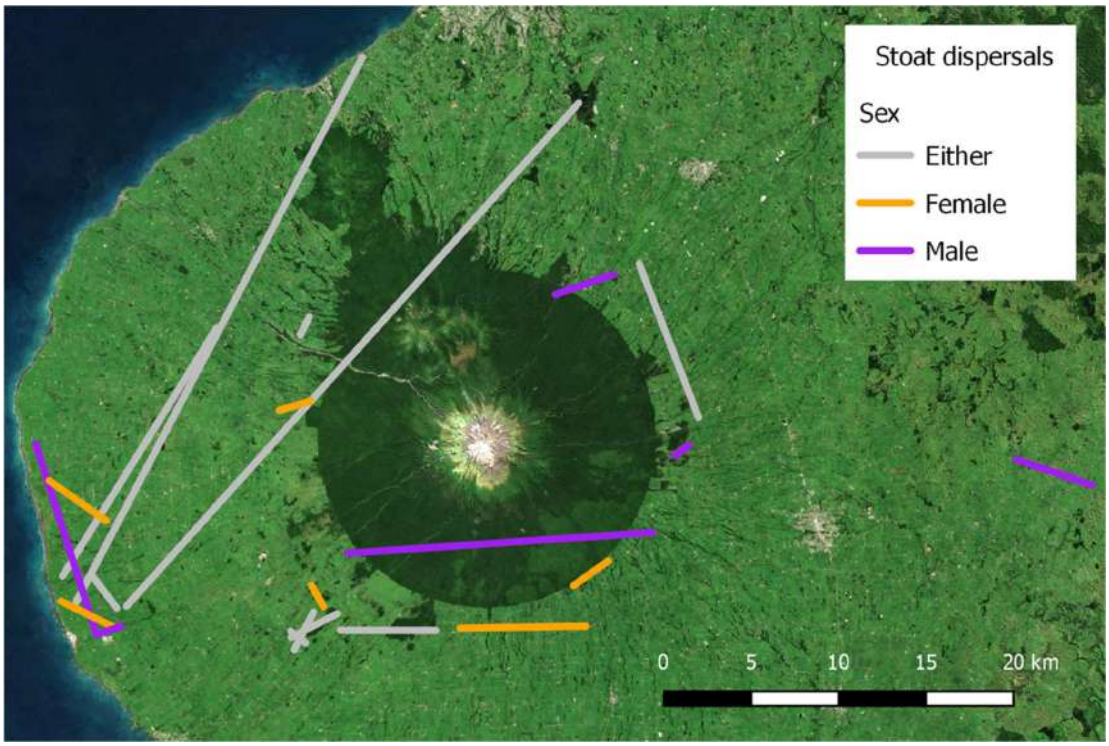

Figure 9. Stoat dispersal events showing lines between capture locations of close-kin pairs.

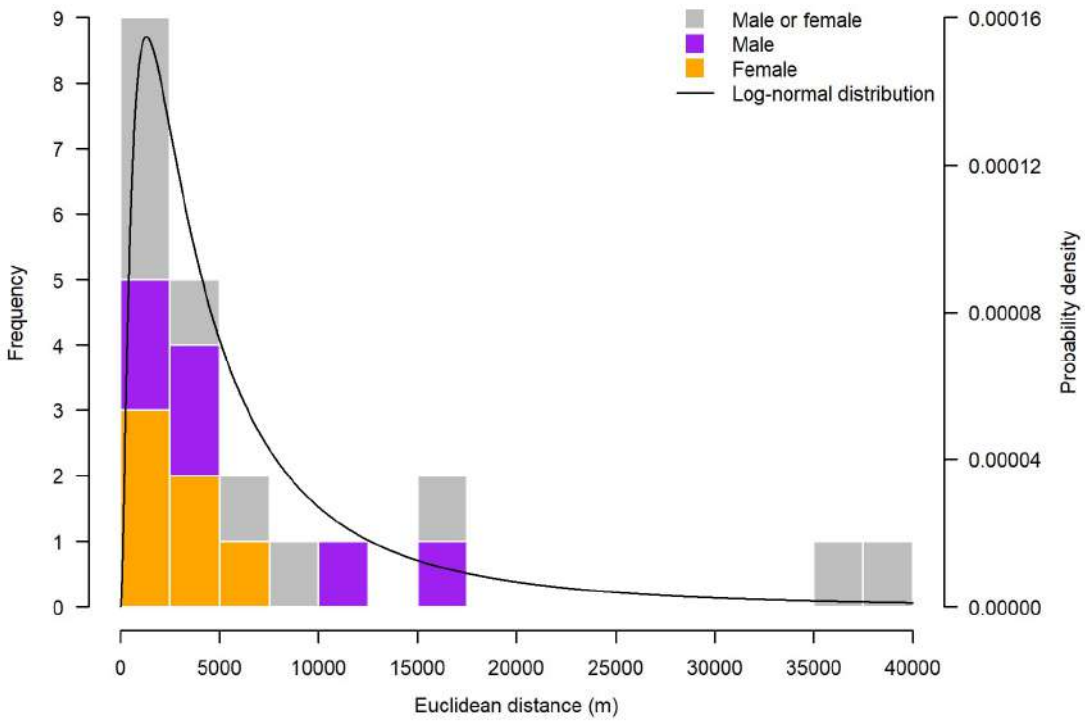

Figure 10. Stoat Euclidean (geographic straight-line) distance dispersal kernel showing the likely distances stoats will disperse.

5 Discussion

For stoats, we were able to create a dispersal model for Taranaki based on genetic data while, for weasels, the limited number of close relatives meant that we were only able to describe a few instances of dispersal rather than producing a complete dispersal model. Given these results and the comparatively few ferrets caught (and their clustered distribution), it would be highly unlikely that we would detect any effect of landscape structure on dispersal for ferrets.

The most appropriate way to model stoat dispersal across this landscape, given the genetic data available, is simply a distance function unrelated to landscape features. This result does not, however, mean that stoats occupy and use this landscape evenly. Instead, our results show that the dispersal capacity of stoats, combined with the patch sizes of landscape features across the Taranaki ring-plain, means that there is little effect detectable from these landscape features. Stoats can travel several kilometres in a day within an established home range, and even further when dispersing as juveniles (King & McMillan 1982; Samson & Raymond 1998). The widths of any less favourable terrain that stoats would have to cross in Taranaki are small enough that they do not constitute barriers or filters to their dispersal and, if stoats were fully removed from a section of the ring-plain, the chance of reinvasion would be determined primarily by the distance from the nearest stoat rather than by the intervening landscape.

Behavioural evidence indicates that stoats prefer to be near cover, and to run along natural linear features (Cameron et al. 2005; Gillies et al. 2007). However, the habitat and microhabitat features that affect stoats are complex and are likely to be specific to local environments. Higher altitude appears to be consistently associated with higher stoat abundance (Christie et al. 2009; Martin et al. 2012; Foster et al. 2021), and this has been noted for Mount Taranaki (Coster 2021), but these effects are likely to be minimal within the ring-plain.

In this farmed environment it appears that stoats can inhabit and cross open farmland and do so often enough that it does not greatly affect their dispersal. This was noted previously with the small amount of data obtained on stoat movement using GPS in the Taranaki ring-plain, reproduced below from Niebuhr & Veale 2020, where the GPS-collared animals were more likely to be detected around forest cover and wetland/scrub, although they crossed open ground.

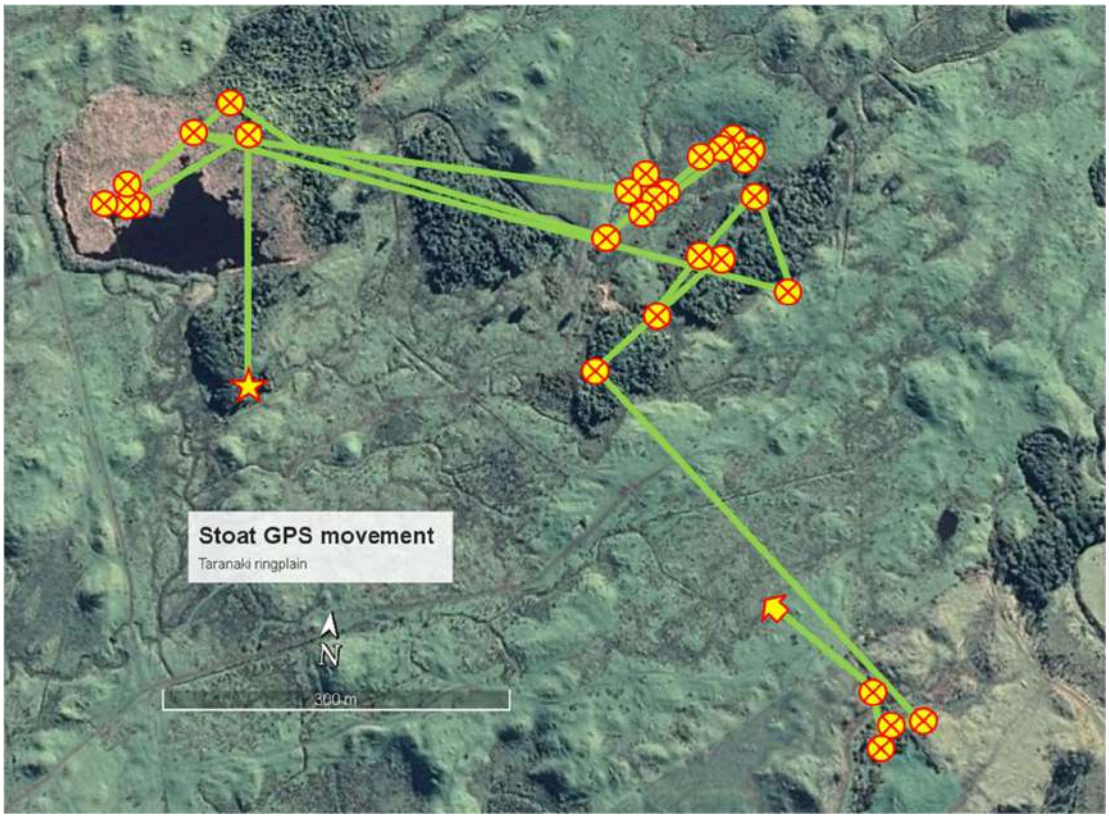

Figure 11. Movement of GPS-collared stoats in the Taranaki ring-plain (reproduced from Niebuhr & Veale 2020).

Given this known habitat preference and our dispersal results, it appears there are enough patches of scrub, riparian margins, and forests in the ring-plain to mean stoats are effectively unaffected by the landscape when dispersing. Traps should still be maintained along these features to maximise catch rates (which appears to be the case based on the trap locations provided). We note that one potential challenge for analysing the trapping data to determine habitat use in the Taranaki ring-plain is that the trapping network has

already been optimised for traps by positioning them along these features, which means trap sites in the middle of open fields are rare.

This is the first study of stoat dispersal in the North Island, and the first study of this type across these habitats (dairy farmland with belts of riparian planting, patches of scrub and North Island broadleaf–podocarp forest). Stoat dispersal in New Zealand has previously been studied only in South Island beech forest and alpine environments (King & McMillan 1982; Murphy & Dowding 1995; Smith & Jamieson 2005). These previous studies (in South Island beech forest) have used ear tagging or radio collars, while our study used relatedness to assess dispersal. For stoats we were able to obtain larger numbers of dispersal instances than in these previous studies, making this the largest such study conducted. A further advantage of this technique was that we did not have the bias of having to find and live-trap animals before they dispersed. In previous studies, dispersal movements are entirely of young of the year, although more work is required to see if adults also disperse, which may occur, particularly for males during the mating season.

While we did not find statistical support for a difference in dispersal ability (or propensity) between sexes, the sample sizes were small. Also, the way we defined sex-specific dispersal meant that there were multiple cases where the sex of the disperser was unknown (if two close kin were different sexes, then we could not know which accounted for the greater proportion of dispersal). The two long-distance dispersal events recorded for stoats involved two males, each caught some distance from more than one close kin, which were each caught closer to each other than to the male in question. Therefore, while we cannot be certain, it seems likely that each of these males (ST_79 and ST_275) was the primary disperser, and their relatives were family members that stayed closer together. The longest confirmed female dispersal recorded was 7.2 km compared to the probable male dispersal of 39 km.

These distances fit broadly with previous dispersal parameters observed in beech forest (King & Veale 2020). One difficulty for all dispersal studies is that we do not know if these animals had finished dispersing when killed: many could have still been in their mother's home range when trapped, or only have recently left their mother's home range and not finished dispersing. Since the animals were killed as part of the sampling, we should conclude that these dispersal distances are specific to how far they disperse in such a trapping grid before being caught, and it seems likely that longer dispersal is more common than recorded here. King and McMillan (1982) found males dispersing between 6 and 24 km, with one individual travelling a minimum of 20.0 km in 5 days (plus a further 4.4 km in the next 2 days). Females in their study generally travelled smaller distances, with one female travelling 5.6 km before returning 3.2 km. While females appear to disperse less than males, the longest-ever dispersal recorded for a stoat is 65 km, recorded in Fiordland (Murphy & Dowding 1995). Overall, stoats in Taranaki appear to disperse similar distances in this trapping grid, although many were killed when still young and near their mother.

The sex ratio of trapping programmes for stoats is one of the best ways to evaluate the effectiveness of the trapping network, because female stoats are harder to trap due to their smaller home ranges and high levels of trap aversion in winter and spring (King & Powell 2006). The fact that the sex ratio is indistinguishable from 1:1 means that trap

network coverage is sufficient to have at least one trap in each female's home range – very few, if any, do not encounter traps. This means that recruitment into the control area is either from stoats that encountered but avoided traps, or from immigration from outside the trapping area. The genetic sexing method developed in this study provides an accurate way to do this sex-ratio assessment, even on highly degraded or fragmentary specimens, and ensured complete accuracy in this assessment.

The highly male-biased sex ratio of weasels is unsurprising, as all studies of trapped weasels in New Zealand have similarly male-biased sex ratios (King & Murphy 2020). These male-biased sex ratios have multiple causes, including:

- the greater ranges of males, leading to their higher chances of finding a trap
- the differences between the sexes in activity and behaviour
- in particular, the lighter weight of females, which lowers their chances of setting off a trap.

While all of these factors may contribute towards the low female weasel catch, the fact that the average female weasel is c. 57 g (King & Murphy 2020), well below the trigger weight set for DOC 200s used by Taranaki Regional Council, means this is likely to be the major contributing factor. It seems likely that most female weasels remain uncaught, so this species will not be controlled. The weasel population in this region may even increase beyond what it was before control started, as stoats are likely to suppress weasels.

The weasel population sample had only two close-kin pairs compared to the estimated 49 close-kin pairs for stoats (correcting for sample size) which strongly indicates that the proportion of the weasel population caught in the trapping programme is much lower than that of the stoat population. We have already described the low capture rates for female weasels, but this level of difference indicates that male weasels are also being caught at considerably lower rates than stoats. Male weasels (particularly juveniles) may also be too light to trigger DOC200s. The home-ranges of weasels are poorly understood in New Zealand (King & Murphy 2020), and it is possible that in some environments (particularly with high mouse numbers) male weasels could have smaller home ranges than female stoats, and so they could be missed by the trapping grid. It is also possible that their behaviour means they are less likely to go into the traps than stoats, or that the lures used are not optimised for weasels. Modelling using this measure in future may provide more insight into the precise proportions of each species caught.

On Secretary and Resolution Islands, stoat populations have been maintained despite intensive trapping over 10 years. Genetic evidence has shown that while there is some immigration from swimming stoats, most recruitment every year is from residents that have avoided traps. The genetic evidence in Taranaki indicates this is probably true tere too; there is a gentle pattern of isolation by distance in the population, and family groups are still largely being caught close together.

All these results mean that, while long-distance dispersal is occurring, most of the stoats contributing to recruitment in the control area are survivors that have avoided traps. We say 'avoided' rather than 'not encountered' because the even sex ratio means that most stoats must be encountering traps (if the network were not dense enough, females with

their smaller home ranges would be trapped less frequently). Therefore, the primary management implication is that lowering stoat numbers to reach elimination must focus on removing these recalcitrant (difficult-to-catch) animals (with better or more varied traps, detection dogs, cameras, etc.). Immigration must be managed, but it is only likely to account for a small percentage of catches currently.

6 Conclusions

The lack of association between trapping and landscape features suggests that mustelids are moving freely through this landscape, seemingly unaffected by landscape features. Their dispersal ability, combined with the habitat patch sizes in Taranaki, means that if there are habitats that are less favourable for habitation or dispersal, they are small enough not to have a significant effect on the distances stoats can disperse.

The few dispersal events detected indicate that mustelid dispersal in this landscape can easily be tens of kilometres. Dispersal is essentially random within the landscape and over very long distances.

The sex ratio of stoats across this trapping programme indicates that the trapping grid is dense enough to have at least one trap within most stoats' home-ranges. The sex ratio of weasels indicates a significant proportion of female weasels are not being caught, and the low number of detected close kin for weasels indicates a substantial proportion of the population is remaining untrapped.

7 Recommendations

To achieve elimination all stoats must be removed from the region. Stoats that are caught now either (1) have not encountered traps, (2) have avoided traps, or (3) are immigrants from outside the area. Our results indicate that there will be very few of #1, and we interpret this to mean that the majority of stoats will be #2: recruitment from animals that have avoided traps. Our results also indicate that long-distance dispersal occurs, and that immigration must therefore be managed to achieve a stoat-free state across the region, although these immigrants are probably less common than the residents that avoid traps.

A trapping network with the current trap density needs to be very wide (over10 km, or perhaps over 40 km) to achieve zero immigration. Most stoats, particularly females, will disperse shorter distances than this, so if elimination cannot be achieved, a narrower buffer will provide some control of immigrants. Currently, the trapping network is dense enough that most stoats will have a chance of encountering traps, and therefore the network should be maintained until stoats are eliminated. Weasels are under-managed, however, and new traps are required to investigate how best to control them in this environment.

8 Acknowledgements

We would like to thank Jordan Lasenby, Toby Shanley, and all the trappers who collected the samples and recorded everything so well for us. Also, many thanks to John McEwan, Rudi Brauning, and the rest of the team at AgResearch, who helped process the samples. Thanks to Grant Norbury and Chris Niebuhr for reviewing the document, and to Chris Niebuhr for ongoing discussions on mustelid movement research.

9 References

Benestan LM, Ferchaud A-L, Hohenlohe PA, Garner BA, Naylor GJP, Baums IB et al. 2016. Conservation genomics of natural and managed populations: building a conceptual and practical framework. Molecular Ecology 25: 2967–2977.

Cameron BG, van Heezik Y, Maloney RF, Seddon PJ, Harraway JA 2005. Improving predator capture rates: analysis of river margin trap site data in the Waitaki Basin, New Zealand. New Zealand Journal of Ecology 29(1): 117–128.

Christie JE, Brown DJ, Westbrooke I, Murphy EC 2009. Environmental predictors of stoat (Mustela erminea) and ship rat (Rattus rattus) capture success. DOC Research & Development Series 305.

Coster K 2021. Gut content analysis of trapped mustelids and trend analysis for ship rats and stoats over six trapping seasons at Mount Taranaki: a thesis submitted in partial fulfilment of the requirements for the degree of Master of Science at Lincoln University.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. 2011. 1000 Genomes Project Analysis Group: the variant call format and VCFtools. Bioinformatics 1,27,(15): 2156–2158.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. 2021. Twelve years of SAMtools and BCFtools. Gigascience 16,10,(2): giab008.

Dodds KG, McEwan JC, Brauning R, Anderson RM, van Stijn TC, Kristjánsson T, et al. 2015. Construction of relatedness matrices using genotyping-by-sequencing data. BMC Genomics 16: 1047.

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLOS ONE 6. doi: 10.1371/journal.pone.0019379

Escoda L, Fernández-González Á, Castresana J 2019. Quantitative analysis of connectivity in populations of a semi-aquatic mammal using kinship categories and network assortativity. Molecular Ecology Resources 19: 310–326.

Etherington TR 2016. Least-cost modelling and landscape ecology: concepts, applications, and opportunities. Current Landscape Ecology Reports 1: 40–53.

Etherington TR, Perry GLW, Cowan PE, Clout MN 2014. Quantifying the direct transfer costs of common brushtail possum dispersal using least-cost modelling: a combined cost-surface and accumulated-cost dispersal kernel approach. PLOS ONE 9: e88293.

Fielding AH, Bell JF 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24: 38–49.

Flanagan SP, Jones AG 2019. The future of parentage analysis: from microsatellites to SNPs and beyond. Molecular Ecology 28: 544–567.

Foster NJ, Maloney RF, Seddon PJ, Recio MR, Khan MS, van Heezik Y 2021. Altitudinal distribution of the entire invasive small mammal guild in the eastern dryland zone of New Zealand's Southern Alps. Biological Invasions 23(6): 1837–1857.

Galla SJ, Moraga R, Brown L, Cleland S, Hoeppner MP, Maloney RF, et al. 2020. A comparison of pedigree, genetic and genomic estimates of relatedness for informing pairing decisions in two critically endangered birds: implications for conservation breeding programmes worldwide. Evolutionary Applications 13: 991–1008.

García-Díaz P, Niebuhr CN 2019. Buffer design to protect Taranaki Mounga from mustelid inflow. Manaaki Whenua – Landcare Research Report LC3497 prepared for Taranaki Regional Council.

Gillies CA, Graham PJ, Clout MN 2007. Home ranges of introduced mammalian carnivores at Trounson Kauri Park, Northland, New Zealand. New Zealand Journal of Zoology 34(4): 317–333.

Gregory S 1978. Statistical methods and the geographer. London, OK, Longman.

Heinrich V, Kamphans T, Mundlos S, Robinson PN, Krawitz PM 2017. A likelihood ratio-based method to predict exact pedigrees for complex families from next-generation sequencing data. Bioinformatics 1,33,(1): 72–78.

Kahle D, Wickham H 2013. ggmap: spatial visualization with ggplot2. The R Journal 5(1): 144–161.

King CM, McMillan CD 1982. Population structure and dispersal of peak-year cohorts of stoats (Mustela erminea) in two New Zealand forests, with especial reference to control. New Zealand Journal of Ecology 1: 59–66.

King CM, Murphy EM 2020. Weasel. In: King CM, Forsyth DM eds, The handbook of New Zealand mammals. 3rd edn. Dunedin: Otago University Press.

King CM, Powell RA 2006. The natural history of weasels and stoats: ecology, behaviour & management 2nd edn. Oxford: Oxford University Press.

King CM, Veale AJ 2020. Stoat. In: King CM, Forsyth DM eds. The handbook of New Zealand mammals. 3rd edn. Dunedin: Otago University Press.

Lu F, Lipka AE, Glaubitz J, Elshire R, Cherney JH, Casler MD, et al. 2013. Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. PLOS Genetics 9(1): e1003215.

Martin M 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17(1): 10–12.

Martin RD 2012. Geospatial and temporal patterns of stoat (Mustela erminea) activity in Tongariro Forest, central North Island, New Zealand. Master's thesis, Massey University.

McCulloch AF, Jauregui R, Maclean PH, Ashby RL, Moraga RA, Laugraud A, et al. 2018. An entropy-reducing data representation approach for bioinformatic data. Database 2018. doi: 10.1093/database/bay029.

Murphy EC, Dowding JE 1995. Ecology of the stoat in Nothofagus forest: home range, habitat use and diet at different stages of the beech mast cycle. New Zealand Journal of Ecology 19: 97–108.

Niebuhr CN 2021. Investigations into novel technologies for assessing mustelid movement in the landscape: progress report. Manaaki Whenua – Landcare Research Report Contract Report LC4010 prepared for Taranaki Regional Council.

Niebuhr CN, Veale AJ 2020. Mustelid movement in the Taranaki ring plain: update 2020. Manaaki Whenua – Landcare Research Report LC3821 prepared for Taranaki Regional Council.

Samson C, Raymond M 1998. Movement and habitat preference of radio tracked stoats, Mustela erminea, during summer in southern Quebec. Mammalia 62: 165–174.

Smith DH, Jamieson IG 2005. Lack of movement of stoats (Mustela erminea) between Nothofagus valley floors and alpine grasslands, with implications for the conservation of New Zealand's endangered fauna. New Zealand Journal of Ecology 1: 45–52.

Storfer A, Patton A, Fraik AK 2018. Navigating the interface between landscape genetics and landscape genomics. Frontiers In Genetics 9: 68. doi: 10.3389/fgene.2018.00068.

Zeller KA, McGarigal K, Whiteley AR 2012. Estimating landscape resistance to movement: a review. Landscape Ecology 27: 777–797.

Appendix

Genomic analyses

A total of 554 mustelid tissue samples were obtained from both the ring-plain trapping (491 samples) conducted by Taranaki Regional Council, and trapping conducted by the Department of Conservation in Egmont National Park (63 samples). These samples consisted of 357 stoats (*Mustela erminea*), 121 weasels (*Mustela nivalis*), and 70 ferrets (*Mustela putorius*), along with a handful of unidentified mustelid samples. While most tissue samples were ears, preserved in 70% ethanol, some were tails, and a few of the more degraded samples were just collections of miscellaneous unidentified tissue and fur.

We used a DNeasy blood and tissue extraction kit on a QiaCube to extract the DNA from c. 0.5 cm^2 of tissue, with an overnight digest using proteinase K according to the manufacturer's protocols (QIAGEN, Hilden, Germany). DNA was eluted into 200 µl of Buffer AE and then stored at –20ºC. The quality and quantity of the DNA were evaluated using a denovix DS-11 nucleic acid spectrophotometer, examining the 260/230 and 260/280 ratios to determine if there was any contamination. Any sample that did not meet the criteria for purity (260/280 = 1.7–2.1, 260/230 = 1.9–2.2) or that did not have a concentration greater than 40 ng/µl was excluded and DNA extraction was repeated at least once.

All DNA extractions were then diluted to a uniform 50 ng/µl (with a concentration step using a SpeedVac for samples that had low concentration), with 1 µg of DNA sent for each sample for GBS sequencing. This GBS was performed at Genomnz Animal Genomics Group (AgResearch, New Zealand). Procedures followed Dodds et al. 2015 after Elshire et al. 2011, with the following modifications. Briefly, genomic DNA was digested with PstI and MspI restriction enzymes (NEB R140L and R0106L, New England Biolabs, Ipswich, United States). We chose enzymes based on bioanalyser traces (2100 Bioanalyser, Agilent Technologies, Santa Clara, United States) showing an even digestion pattern with no evidence of repeat sequences through the region of interest. Following ligation to barcoded adapters, the uniquely barcoded individuals were pooled into five multiplexed libraries of 94 samples. Libraries post-pooling were run through PCR in multiples of four and pooled again before column clean-up, then each library was further purified and size selected (193–500 bp) using a Pippin (SAGE Science, Beverly, United States; 2% agarose, dye-free with internal standards CDF2050, Marker L CDF2010). We then sequenced each library on an Illumina HiSeq2500 using single-end reads, with 101 cycles in high-output mode (v4 chemistry). Each of the five libraries was sequenced separately on a single lane on a single flow cell.

Quality checks and adapter removal followed Dodds et al. 2015. Raw fastq files were quality checked using FastQC v. 0.10.1 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Barcodes and adapters were removed using cutadapt (Martin 2011), then a random 15,000 reads were checked for contamination using BLAST+ against the NT database (https://blast.ncbi.nlm.nih.gov), with the following settings: blastn -query - -task blastn - num_threads 2 -db nt -evalue 1.0e-10 -dust '20 64 1' -max_target_seqs 1 -outfmt '7 qseqid sseqid pident evalue staxids sscinames scomnames sskingdoms stitle'. The other approach to checking for contamination utilised kmers to provide a high-level overview of

the sequence composition, as described in McCulloch et al. 2018. We then produced a catalogue of SNP loci (single nucleotide polymorphisms, or single base variants within genes) following Dodds et al. 2015 and the general guidelines of Benestan et al. 2016.

It became clear that some individuals had their species misidentified due to the high divergence of some individuals, and so all individuals were amalgamated into a single *de novo* assembly in UNEAK Tassel v. 3.0.170 (Lu et al. 2013). We then reclassified individuals to species based on their principal coordinates. From these species-level assignments we then created two separate assemblies – one for weasels and one for stoats. The pipelines for these two assemblies were identical.

After trimming adapters with cutadapt (Martin 2011), reads were assembled with a modified reference-based pipeline in bcftools (Danecek et al. 2021) using the recently completed stoat genome (mMusErm1.Pri: GCA_009829155.1). All following data filtering steps were then conducted using vcftools (Danecek et al. 2011).

Sex determination

An initial data set for determining the sex of individuals was created with the following filtering parameters:

- minimum mean depth per locus (--min-meanDP 0.3)
- maximum mean depth per locus (--max-meanDP 20).

These were chosen because any more stringent filtering would result in the potential loss of Y-chromosome or X-chromosome markers. From this assembly we then output each of the chromosomes separately and calculated the average sequencing depth per locus for each chromosome.

To determine sex we used two different measures: the ratio of the average sequencing depth on the X chromosome vs the 21 autosomes, and the ratio of the average depth on loci on the Y chromosome vs the 21 autosomes. Males only have one X chromosome while females have two, so the read depth will differ between them, and only males will have a Y chromosome (though any sequencing error or barcode jumping could add a few reads that map to it). We found that, particularly for the Y chromosome for weasels (since the assembly was based on the stoat genome), there were issues with sexing, so we just used a small section (248 SNPs) that appeared to best match the X-chromosome data.

Assembling the final data set to calculate relatedness

Once sex was determined for all samples, we created filtered data sets, removing the two sex chromosomes along with any SNPs not assigned to the 21 autosomes, and conducted further filtering with the following parameters:

- individuals with a lower average coverage per locus than 2 were removed
- a maximum percentage missing data per locus of 20% (--max-missing 0.8) was allowed
- minimum mean depth per locus (--min-meanDP 0.3)

- maximum mean depth per locus (--max-meanDP 20)
- Hardy-Weinberg filtering (--hwe 0.001)
- minor allele frequency (0.03).

These combined filtering parameters should remove all low-quality data and noise and ensure that accurate precise relationships between individuals are obtained. A range of additional quality control diagnostics were applied via the software package KGD (kinship using GBS with depth adjustment; https://github.com/AgResearch/KGD) (Dodds et al. 2015). Allele frequencies and sequence depths were evaluated, and SNP call rates and their minor allele frequencies (MAF)s were calculated. To identify repeat regions and filter out SNPs that were not behaving properly, we used a fin-plot of MAFs versus Hardy-Weinberg (H-W) disequilibrium (observed frequency of the reference allele homozygote minus its expected value). The fin-plot highlights SNP average depth and identifies SNPs with non-Mendelian inheritance. High depth and low H-W disequilibrium (often high MAF) are assumed to represent genome duplication or repetitive regions (Dodds et al. 2015). Consequently, SNPs with H-W disequilibrium below –0.05 were removed to avoid the inclusion of SNPs not properly assembled.

We then used KGD (Dodds et al. 2015) to calculate pairwise relatedness between all individuals. This matrix of relatedness values was then compared to the pairwise geographical distances between samples. KGD provides a relatedness index that goes from 0 to 1, with 0 being unrelated for many generations and 1 being identical (either the same individual sampled twice or identical twins). Parent/offspring pairs and full sibling pairs will average 0.5 relatedness, with a recommended cut-off threshold for identifying close kin of >0.35 (Dodds et al. 2015).

To help determine the true pedigree of the samples, we refiltered the data for a second program that requires different parameters. KGD benefits from using high numbers of loci and considers the quality of these loci, which means the estimates obtained for relatedness are reliable. However, KGD cannot easily differentiate parent/offspring pairs from full siblings without age data. The program VCF2LR (Heinrich et al. 2017) can attempt to determine these pedigree relationships without age data but requires significantly higher depth and certainty for each locus. Therefore, we refiltered the data for use in VCF2LR with the following parameters:

- --minDP 9
- --maxmissing 0.7
- --hwe 0.001
- --maxmissing indiv 0.8.

To translate our data into a format usable in VCF2LR we used the custom Python script from Escoda et al. 2019.

Spatial patterns for genetic similarity

For both stoats and weasels we defined kinship groups based on clustering in KGD (Dodds et al. 2015). For stoats these consisted of family groups (which most likely originated from the same litter) having a kinship coefficient of greater than 0.35. For weasels, given the small number of individuals that had these levels of kinship we lowered the threshold to display patterns of lower kinship (cousins, grandparents, etc.).

In KGD we also conducted a principal coordinate analysis (PCA) of kinship for each species, and mapped the first two principal coordinates and displayed these on maps of Taranaki using the package ggmaps (Kahle & Wickham 2013) in R.

Database cleaning

From the database supplied by Taranaki Regional Council identifying the mustelid samples, there were a number of samples with an incorrect species identification. Given the degraded nature of some of the samples, this confusion – generally between weasels and stoats – is understandable. Identifying individuals to species was extremely easy with the genomic data set; however, each species needs to have the genetic reads assembled separately otherwise problems arise when calculating relative relatedness.

We performed a PCA, which clearly clustered most samples to their correct species, highlighting those that had been misidentified. Therefore, after an initial assembly we used the first principal component (PC1) values and clustering to determine which species each of the samples came from, then re-ran all analyses separating the two species correctly. These samples with altered IDs were as follows: sample 83 was confirmed as a stoat and not a weasel, and the following samples listed as stoats were actually weasels: 19-4, 31, 80, 85, 136, 175.

Some samples were identified as replicates of the same individual rather than independent individuals (most of these were listed as having some confusion in the database): 7,60,55; 9,65; 10,50, 19-1,19-20, 19-4,19-17). We provide these specific individual numbers here in case any future analyses are performed on the data set (such as habitat modelling) so that Taranaki Regional Council can correct their metadata files.