

Results

Of the 187 sequences included in this study, 52 ingroup sequences as well as two outgroup sequences were generated in a previous study (Grobler *et al.*, 2006; Chapter 2). The Genbank accession numbers for these previously characterized specimens are AY762267-AY762320. An additional 135 sequences were generated in this study and are deposited in Genbank under accession numbers JF327506-JF327640. Of the 187 sequences used in this study, 131 represent unique haplotypes.

Nested clade analysis

A single haplotype network (Fig. 5), based on the 95 % connection probability (≤ 13 steps), was estimated by TCS 1.13 (Clement *et al.*, 2000). From the nested design presented in Fig. 6, the total cladogram can be divided into two 6-step clades, three 5-step clades, and six 4-step clades at the higher clade level. The six 4-step clades separate from each other by more than five steps, representing missing intermediate haplotypes as indicated in Fig. 7. Nested clade 4-4 which consists of *E. kucheli* (Grobler *et al.*, 2006) was considered basal to the other 4-step clades because time of coalescence of these individuals preceded that of the other clades (Grobler *et al.*, 2006; Chapter 2), and therefore, all higher-step clades were designated with this in mind.

The results of the nested cladistic analysis of geographical distance for the mtDNA data set from the PEIA are shown in Fig. 8. The contingency analyses (Templeton *et al.*, 1995) detected significant geographical associations with the total cladogram, a 6-step clade, two 5-step clades, only one 4-step clade, and a single 1-step clade (Table 1). Significant geographical structure was thus mainly indicated for the higher clade levels. With the exception of nested clade 1-96, no significance at the 1-, 2- and 3-step levels were found. One clade of each of the 4-, 3-, and 2-step clades also had a lower significant geographical association. Table 2 presents the chain of inferences followed when the inference key was applied to the statistical results given in Fig. 7 as well as the resulting inferences obtained.

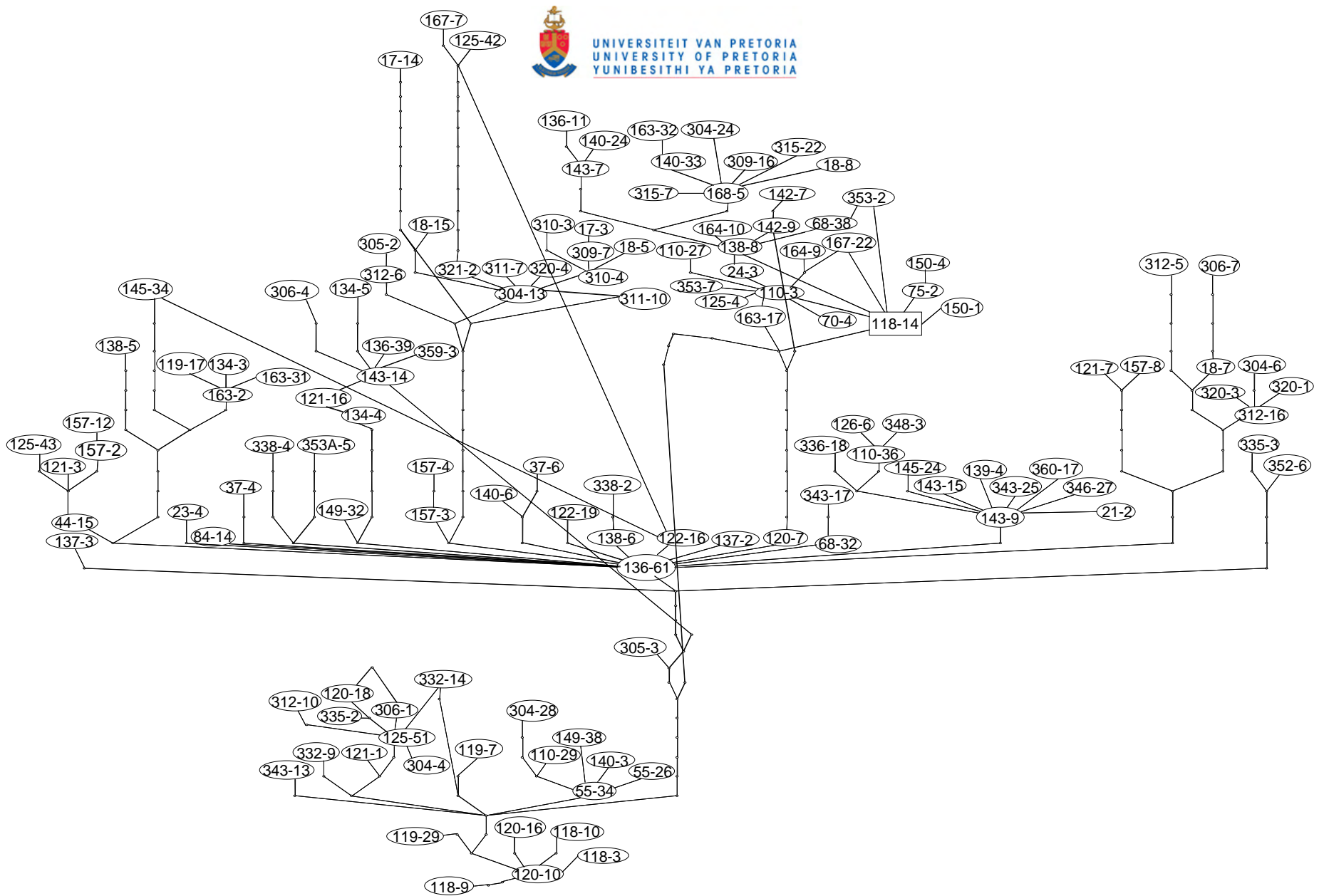


Fig. 5 Haplotype network based on 95% connection probability (≤ 13 steps) as estimated by TCS 1.13 (Clement et al., 2000).

Please find the figure on the disc in the envelope at the back

Fig. 6 Nested design according to previously described nesting rules of Templeton *et al.*, 1987 and Crandall, 1996.

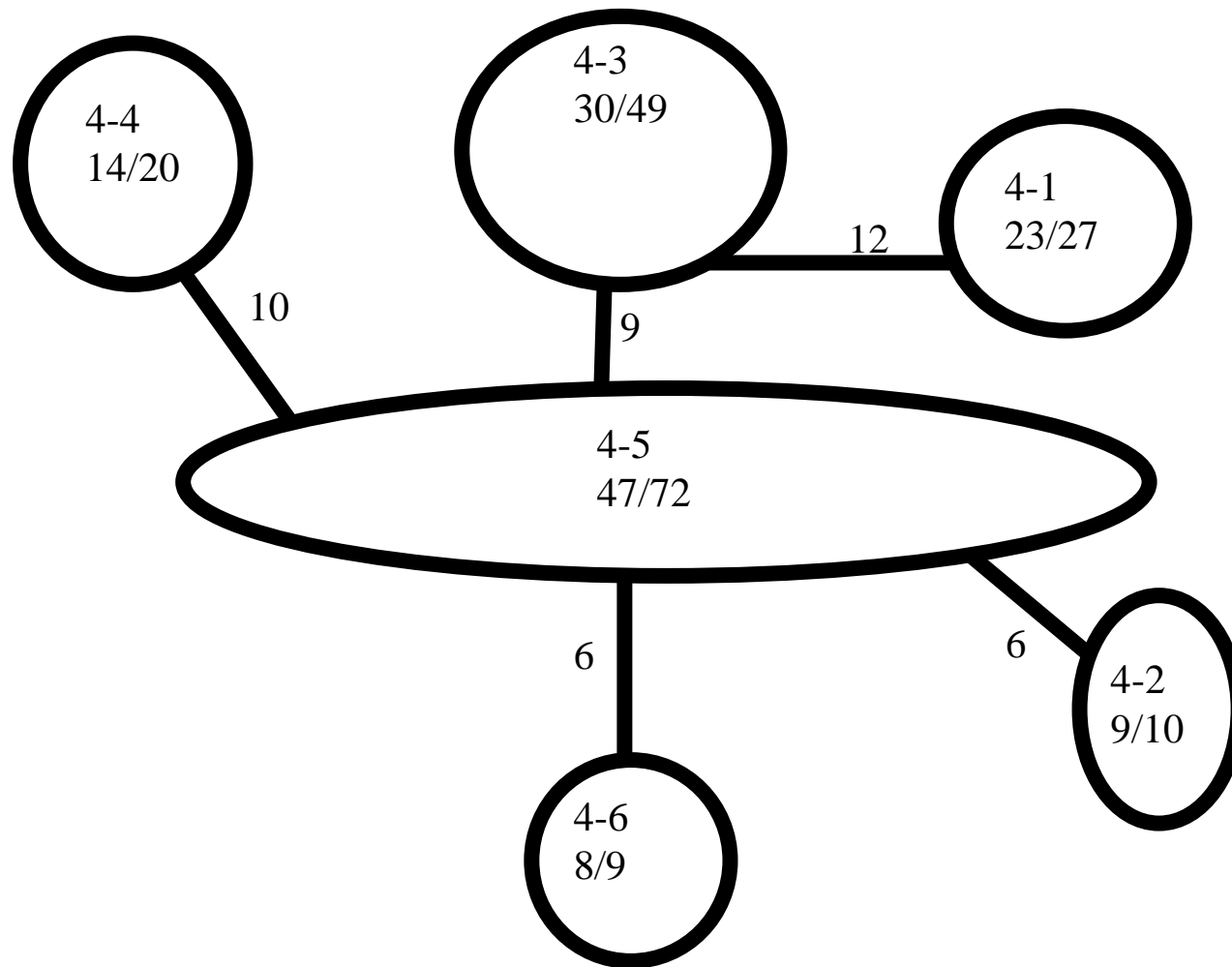


Fig. 7 Haplotype network of 187 *Ectemnorhinus* sequences. Sequences are grouped into six groups separated from each other by the number of missing haplotypes (as indicated on the lines). Within each circle is the designated name for the group followed by the number of unique haplotypes / total number of sequences / specimens characterized for each group.

Please find the figure on the disc in the envelope at the back

Fig. 8 Results of the nested clade analysis of the geographical distance for COI haplotypes of *Ectemnorhinus* weevils from the PEIA. Haplotypes (0-steps) are given at the top in bold. Interior haplotypes/clades are shaded. Higher level clade designations are indicated in bold as one moves down the figure. Boxed groupings indicate the nested structure. Following the number of any given clade is the clade (Dc) and nested clade (Dn) great circle distances. An 's' indicates that the distance is significantly small at the 5 % level and an "L" indicates that the distance is significantly large. For nested groupings that contain both tip and interior clades, the average differences of the distances between interior clades and tip clades within the nested group for clade distances, is indicated by the symbol '(T-1)c', with nested clade distances being indicated by the symbol '(T-1)n'.

Table 1 Nested contingency analyses for geographical association showing the permutation chi-squared probabilities for geographical structure of the clades identified in Fig. 8 from 100 000 resamplings. Clades with a probability value less than 0.05 suggest significant geographical structure and are indicated with *. Clades with a probability value less than 0.1 are indicated with #.

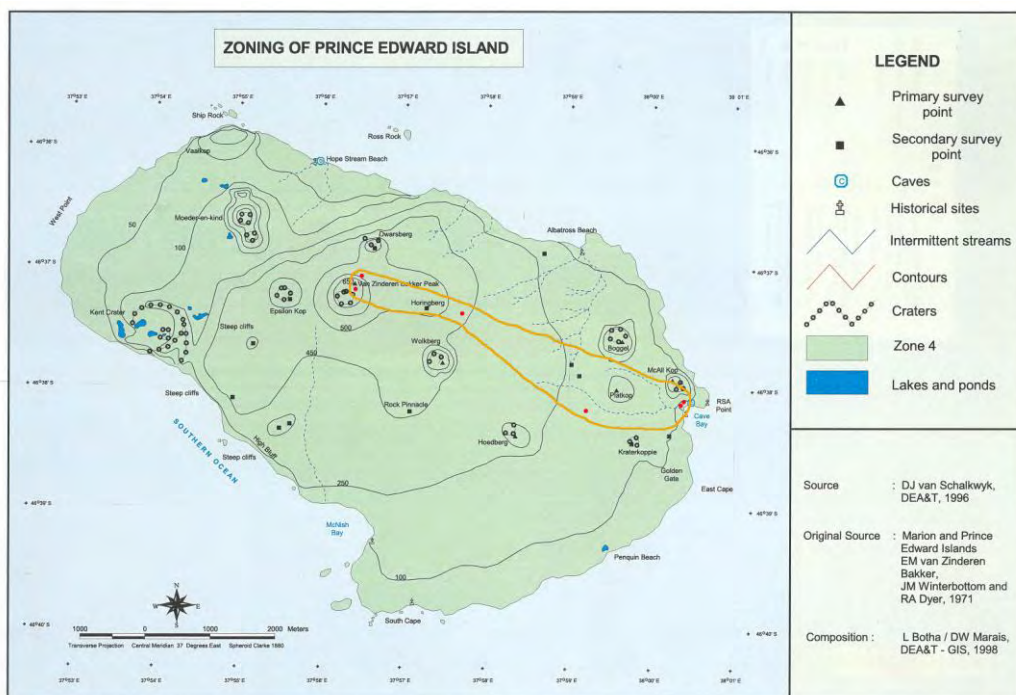
| Clade | Permutational chi-square statistic | Probability |
|-------|------------------------------------|-------------|
| 1-9 | 11.25 | 0.6860 |
| 1-12 | 1.3333 | 1.0000 |
| 1-14 | 15.0000 | 1.0000 |
| 1-15 | 2.0000 | 1.0000 |
| 1-17 | 2.0000 | 1.0000 |
| 1-18 | 34.0000 | 0.8300 |
| 1-20 | 3.0000 | 1.0000 |
| 1-34 | 2.0000 | 1.0000 |
| 1-38 | 0.7500 | 1.0000 |
| 1-43 | 10.0000 | 0.2010 |
| 1-56 | 2.0000 | 1.0000 |
| 1-60 | 1.3333 | 1.0000 |
| 1-67 | 10.0000 | 1.0000 |
| 1-71 | 75.4667 | 0.3210 |
| 1-79 | 2.0000 | 1.0000 |
| 1-80 | 5.9583 | 0.7680 |
| 1-81 | 10.0000 | 0.2900 |
| 1-82 | 2.0000 | 1.0000 |
| 1-83 | 34.0000 | 1.0000 |
| 1-85 | 2.0000 | 1.0000 |
| 1-87 | 3.0000 | 0.3220 |
| 1-89 | 21.3889 | 0.4830 |
| 1-91 | 5.0000 | 0.8370 |
| 1-92 | 2.0000 | 1.0000 |
| 1-93 | 3.0000 | 0.3340 |
| 1-96 | 62.2222 | 0.0410* |
| 1-200 | 2.0000 | 1.0000 |
| 2-1 | 2.9167 | 0.6880 |
| 2-3 | 15.0000 | 1.0000 |
| 2-4 | 2.0000 | 1.0000 |
| 2-5 | 3.0000 | 1.0000 |
| 2-10 | 5.0000 | 0.4300 |
| 2-13 | 3.0000 | 1.0000 |
| 2-15 | 6.0000 | 0.3080 |
| 2-18 | 12.0000 | 0.0730# |
| 2-19 | 4.0000 | 1.0000 |
| 2-20 | 12.0000 | 0.3810 |
| 2-22 | 17.0000 | 0.5780 |
| 2-23 | 2.0000 | 1.0000 |
| 2-27 | 10.0000 | 0.3380 |
| 2-29 | 3.2500 | 0.6120 |
| 2-31 | 18.0000 | 0.4910 |

| Clade | Permutational chi-square statistic | Probability |
|-----------------|------------------------------------|-------------|
| 2-33 | 153.3333 | 0.2680 |
| 2-34 | 30.6000 | 0.3640 |
| 2-35 | 11.2000 | 0.5500 |
| 2-36 | 4.9038 | 0.9120 |
| 2-37 | 11.0714 | 0.9880 |
| 3-2 | 1.2000 | 1.0000 |
| 3-3 | 13.3333 | 0.6430 |
| 3-4 | 19.2000 | 0.1400 |
| 3-6 | 13.9240 | 0.7950 |
| 3-7 | 35.1771 | 0.1100 |
| 3-8 | 0.6857 | 1.0000 |
| 3-9 | 29.8194 | 0.5160 |
| 3-10 | 20.0076 | 0.0950# |
| 3-16 | 149.6400 | 0.4930 |
| 4-1 | 16.5515 | 0.2450 |
| 4-2 | 10.0000 | 0.1060 |
| 4-3 | 24.1412 | 0.1500 |
| 4-4 | 5.9649 | 0.8120 |
| 4-5 | 88.0677 | 0.0000* |
| 4-6 | 3.9375 | 0.7870 |
| 5-1 | 100.0663 | 0.0020* |
| 5-2 | 47.8767 | 0.0002* |
| 6-1 | 71.2705 | 0.0000* |
| Total Cladogram | 101.4970 | 0.0000* |

Table 2 Inference chain for the results of the *Ectemnorhinus* phylogeography as indicated by the nested clade analysis from the results given in Fig. 8 and Table 1. Only the clades that resulted in a rejection of the null hypothesis, i.e. that there are no significant relationships between genetic variation and respective geographical locations of haplotypes, are included. Clades marked with # were not significant at the 0.05 % level but were significant at the 0.1 % level.

| Clade | Chain of inference | Inference |
|-------|--------------------|--|
| 1-96 | 1-2-3-5-6-7 | Insufficient genetic resolution to discriminate between range expansion / colonization and restricted dispersal / gene flow. Restricted gene flow / dispersal but with some long distance dispersal. |
| 2-18# | 1- 2 – 11 - 12 | Continuous range expansion |
| 3-10# | 1- 2 – 11 – 17 - | Inconclusive outcome |
| 4-5 | 1-2-11-17- | Inconclusive outcome |
| 5-1 | 1-2-11-12 | Contiguous range expansion, |
| 5-2 | 1-2-11-17- | Inconclusive outcome |
| 6-1 | 1-2-11-17-4- | Restricted gene flow with isolation by distance |

In order to gain further insight into the geographical structuring, clades were mapped onto the islands according to the sampling localities of the constituent haplotypes. Because of the very large dataset and the fact that the 4-step clades separate from each other by more than three missing haplotypes (Fig. 7) the nested clade diagram was partitioned into the six different 4-step clades to ensure manageable mapping and result presentation. The different maps for the individuals falling in each 4-step clade are presented in Figs. 8-13 with each being divided into six parts.

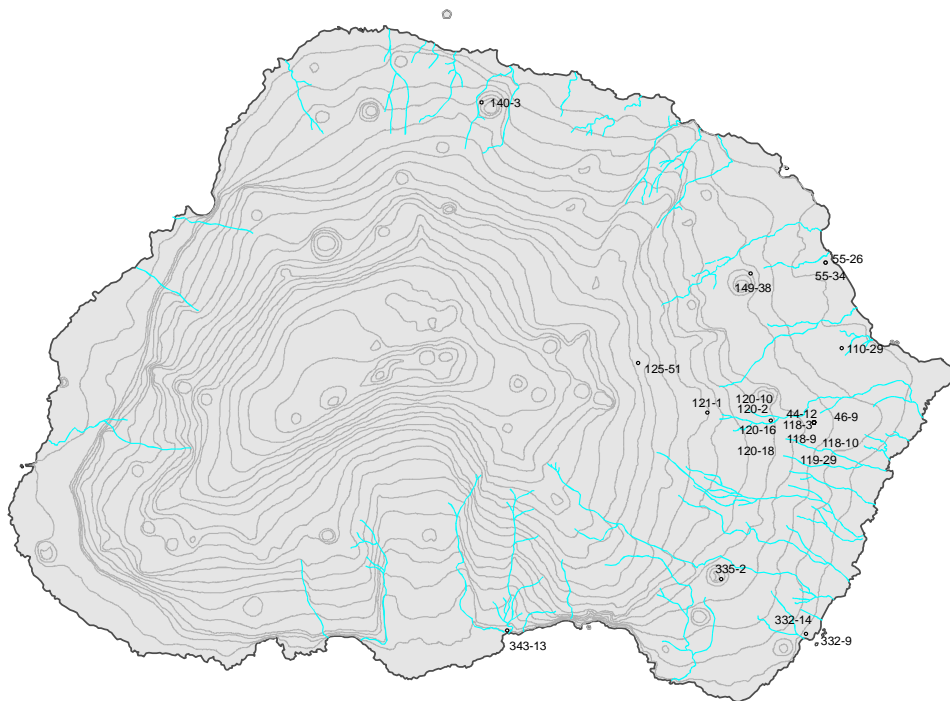
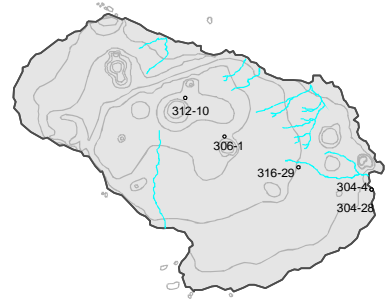
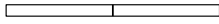


(a)



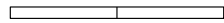
1 *Prince Edward Island*

0 3 6 km



1 *Marion Island*

0 3 6 km



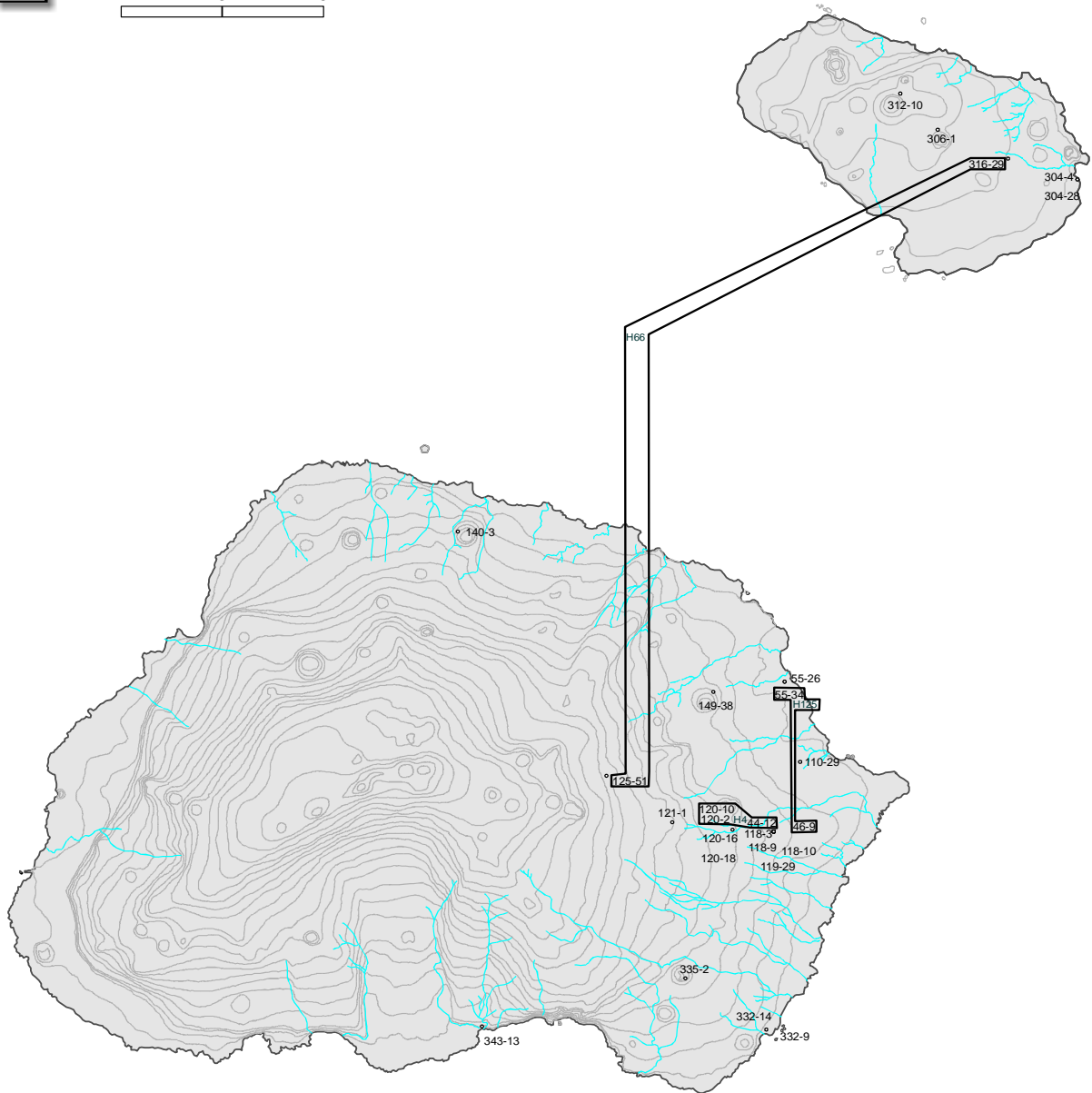
(b)



1

Prince Edward Island

0 3 6 km



1

Marion Island

0 3 6 km

(c)



1 *Prince Edward Island*

0 3 6 km



1 *Marion Island*

0 3 6 km

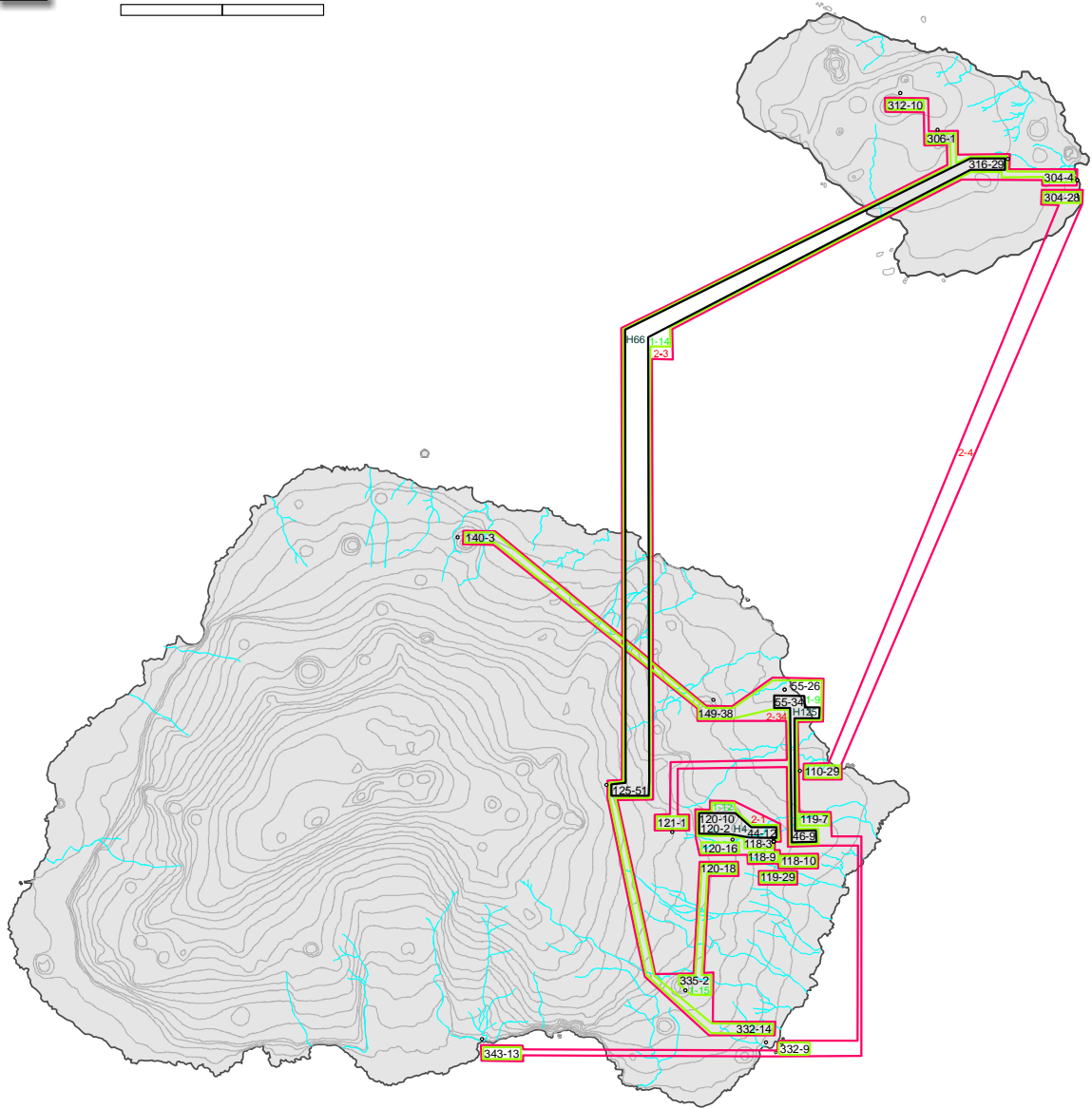
(d)



1

Prince Edward Island

0 3 6 km



1

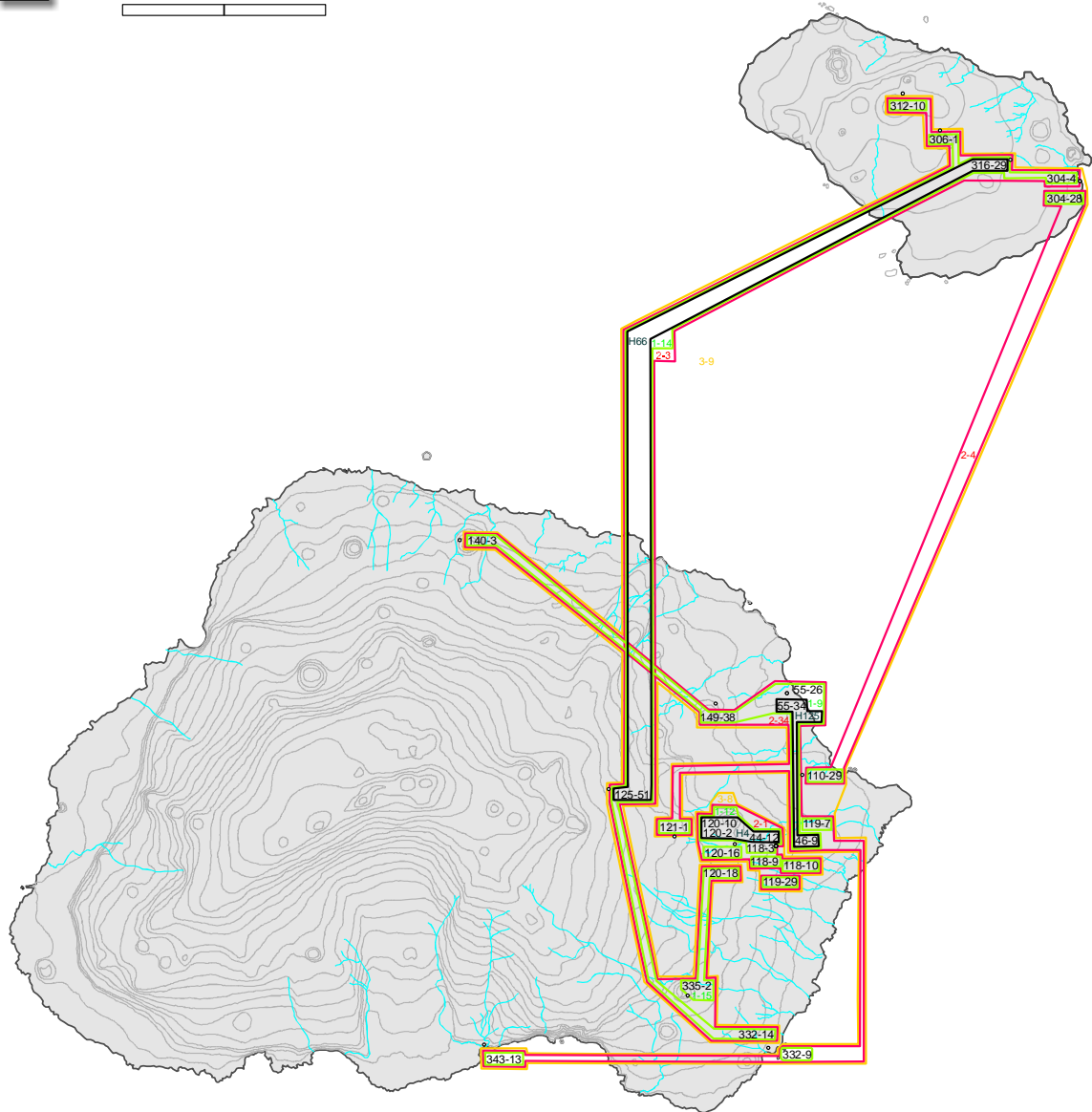
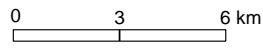
Marion Island

0 3 6 km

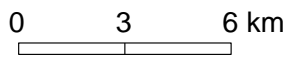
(e)



1 *Prince Edward Island*



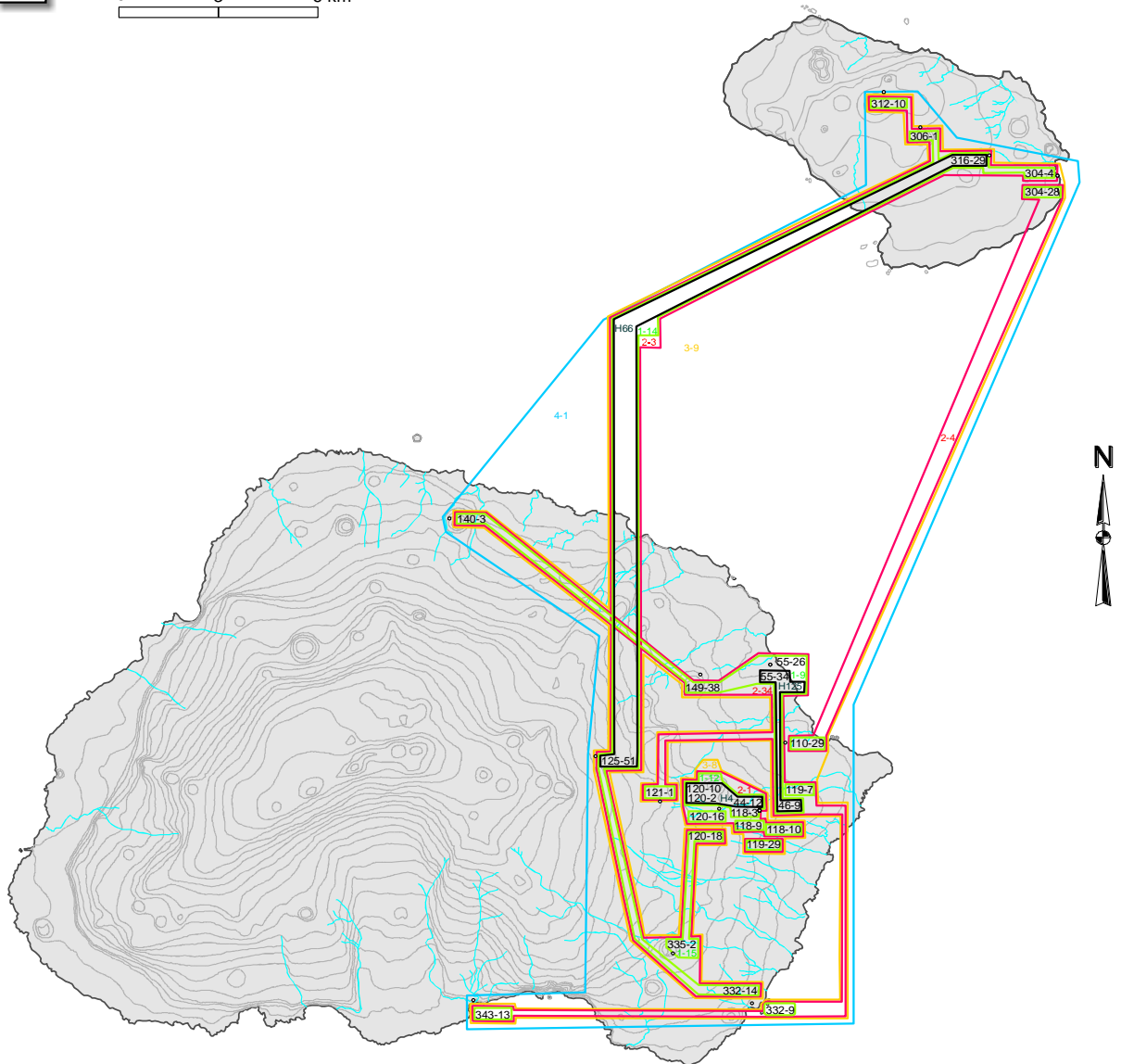
1 *Marion Island*



(f)

1 *Prince Edward Island*

0 3 6 km

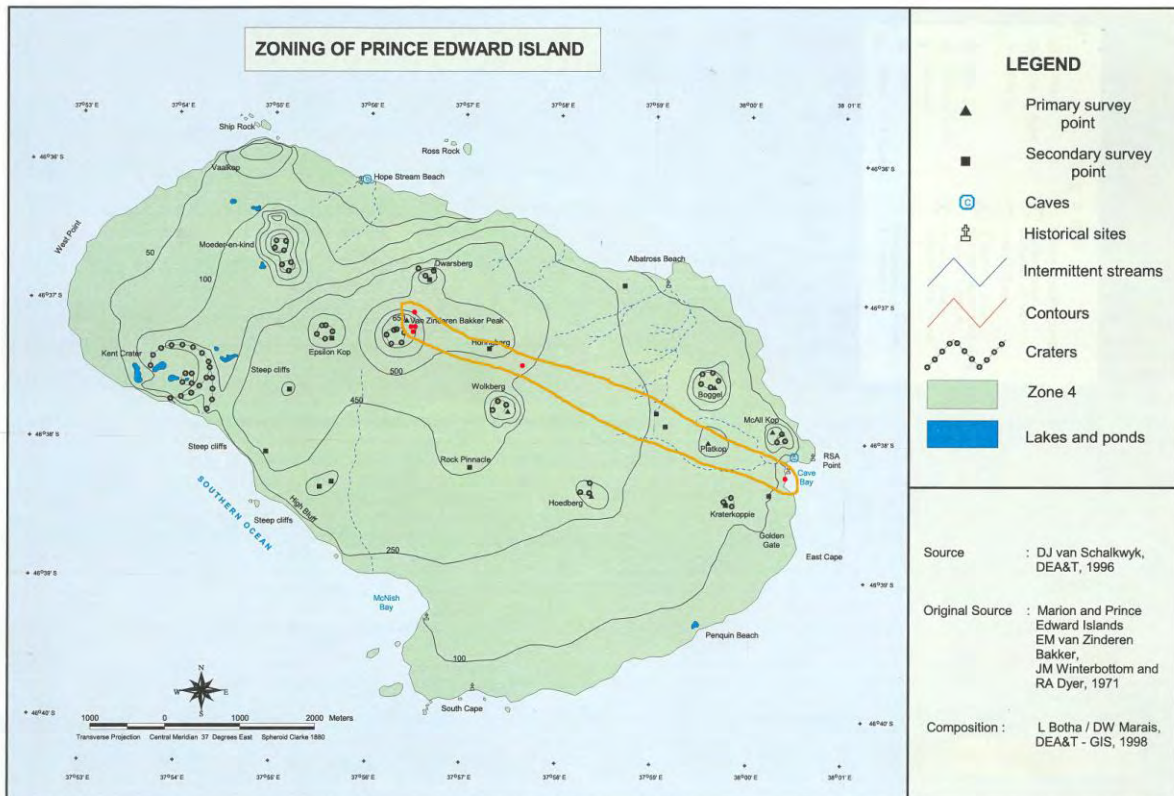


1 *Marion Island*

0 3 6 km

(g)

Fig. 9 *Ectemnorhinus* individuals that group into 4-step clade 4-1 mapped onto the islands according to a) sampling localities with place names for references, samples on Marion Island are indicated in green while those on Prince Edward Island are indicated in red b) sampling localities as used in the nested design c) those that share the same haplotypes nested into 0-step clades, the d) 1-step clades, e) 2-step clades f) 3 step clades and finally g) the complete 4-1 clade as depicted in Fig. 5.



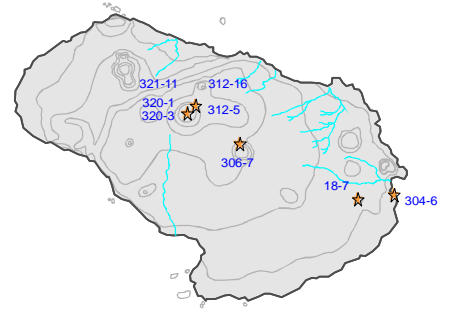
(a)



2

Prince Edward Island

0 3 6 km



2

Marion Island

0 3 6 km

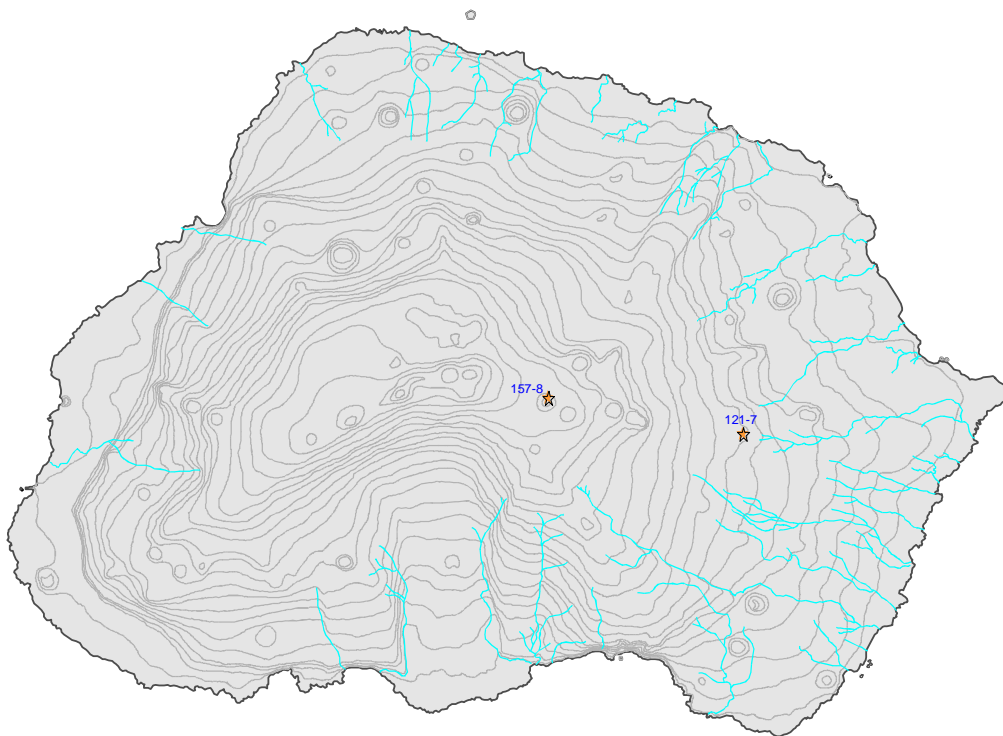
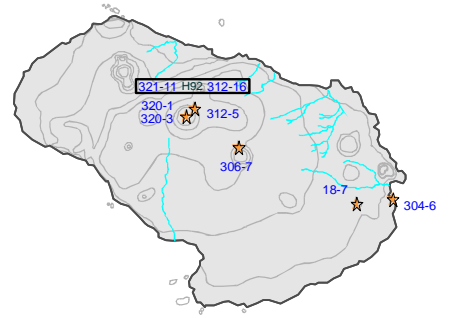
(b)



2

Prince Edward Island

0 3 6 km



2

Marion Island

0 3 6 km

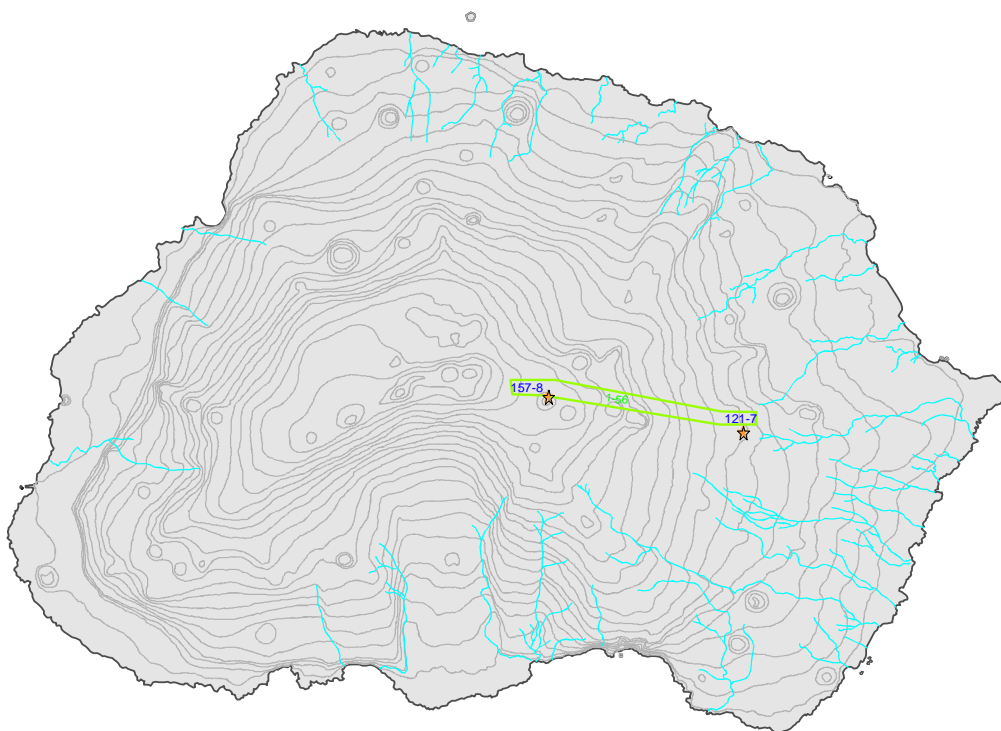
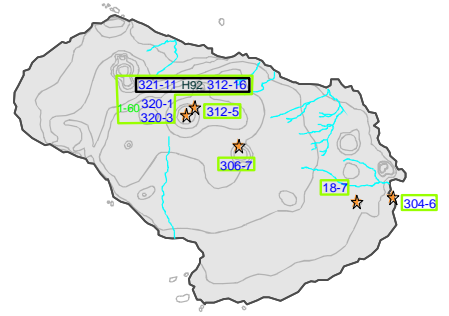
(c)



2

Prince Edward Island

0 3 6 km



2

Marion Island

0 3 6 km

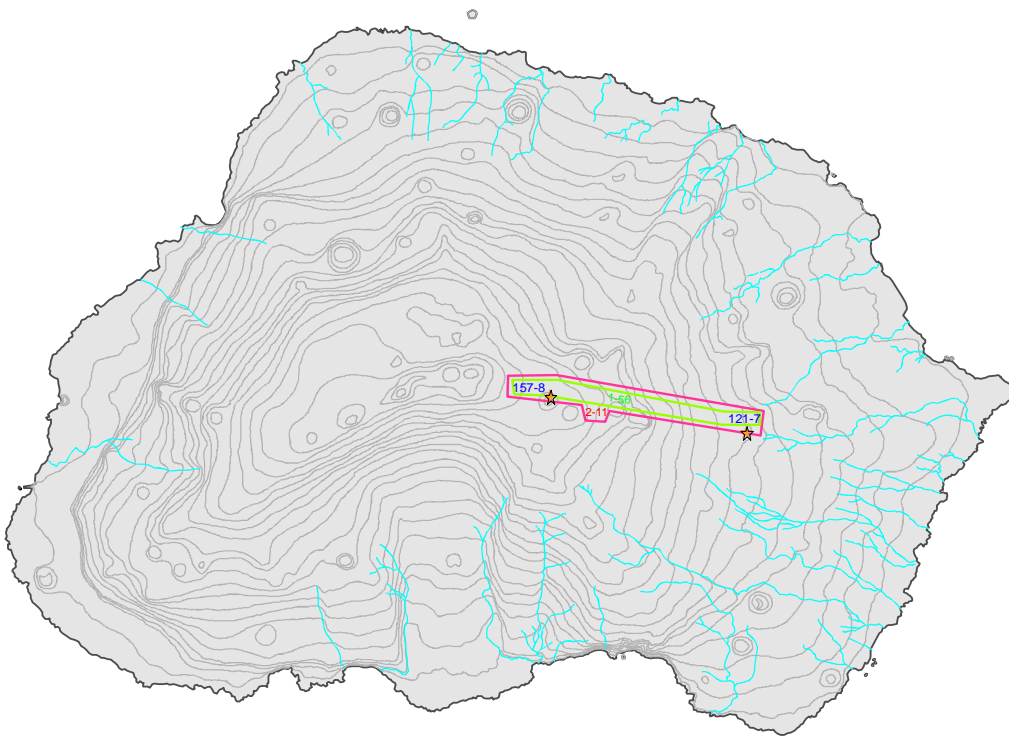
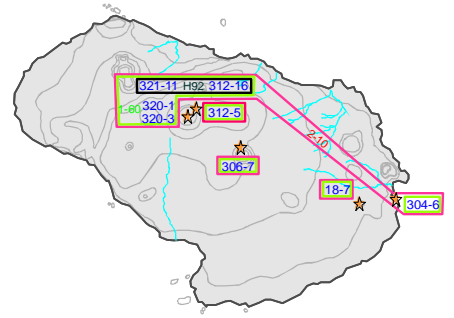
(d)



2

Prince Edward Island

0 3 6 km



2

Marion Island

0 3 6 km

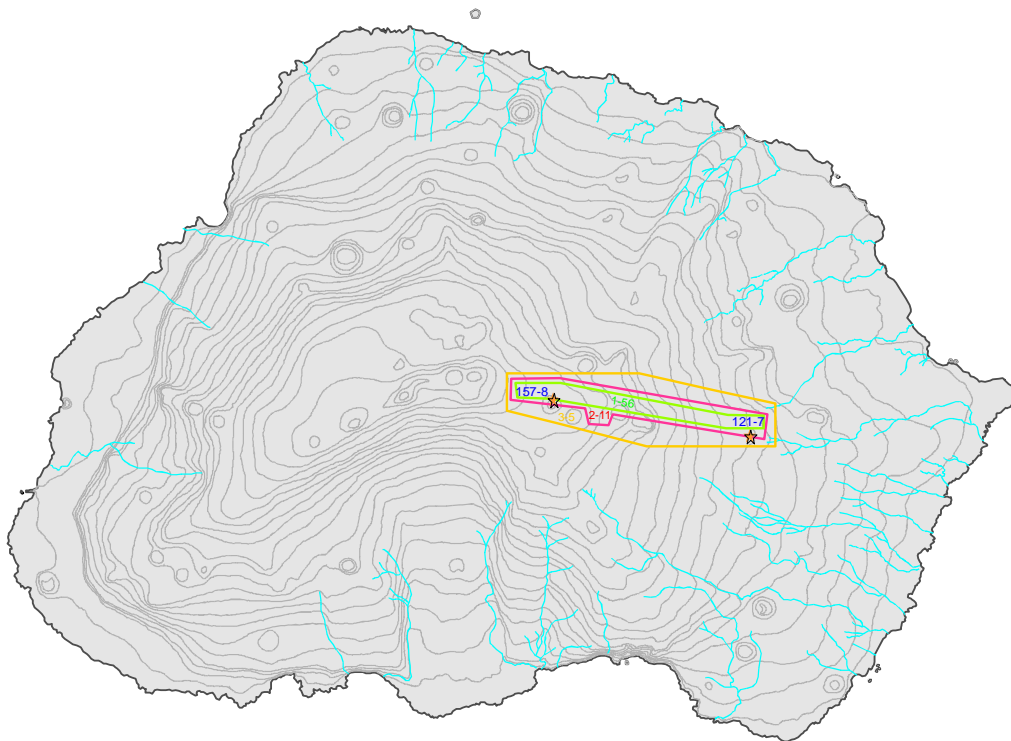
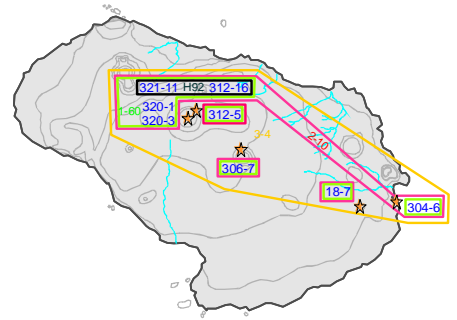
(e)



2

Prince Edward Island

0 3 6 km



2

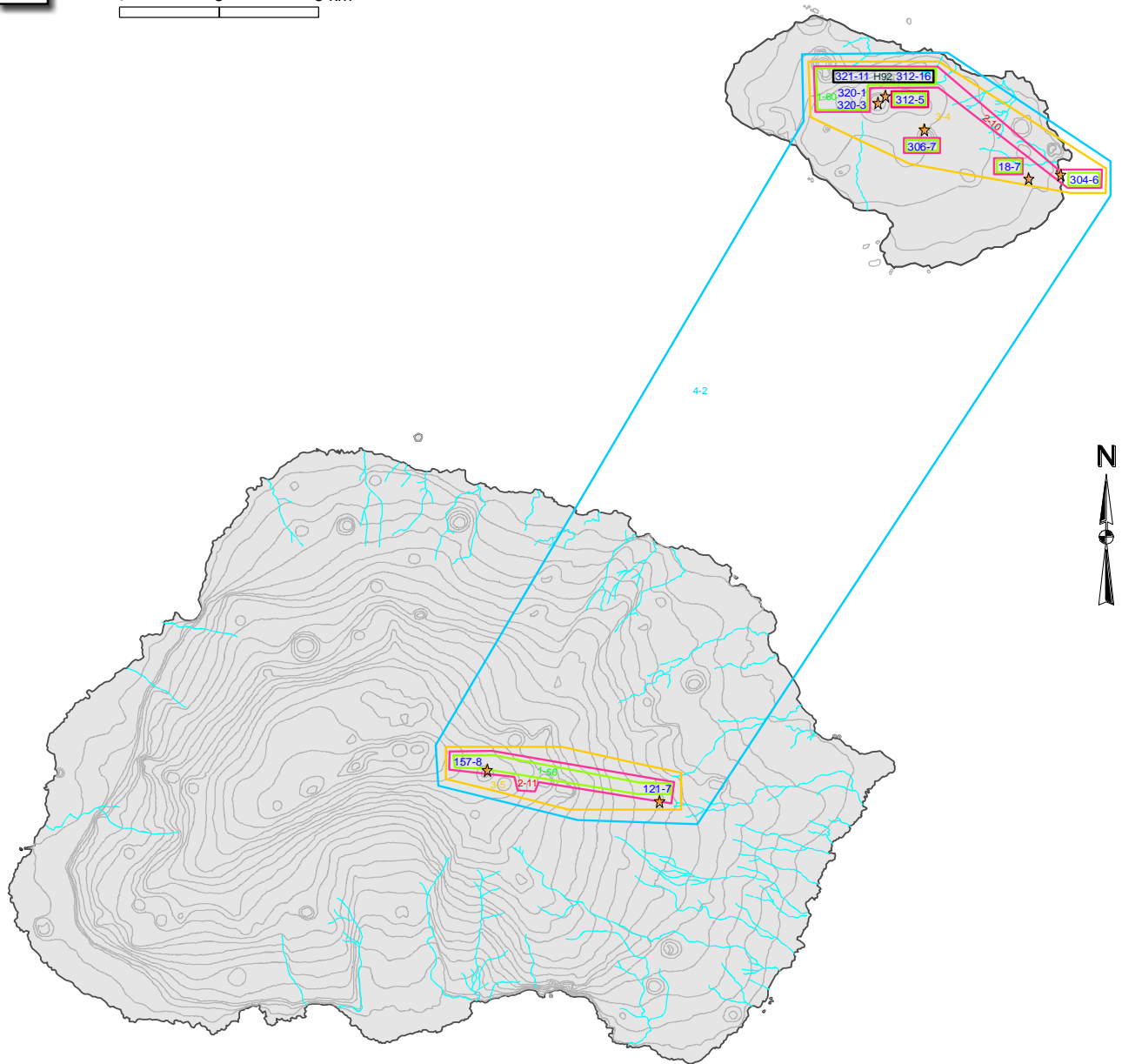
Marion Island

0 3 6 km

(f)

2 *Prince Edward Island*

0 3 6 km

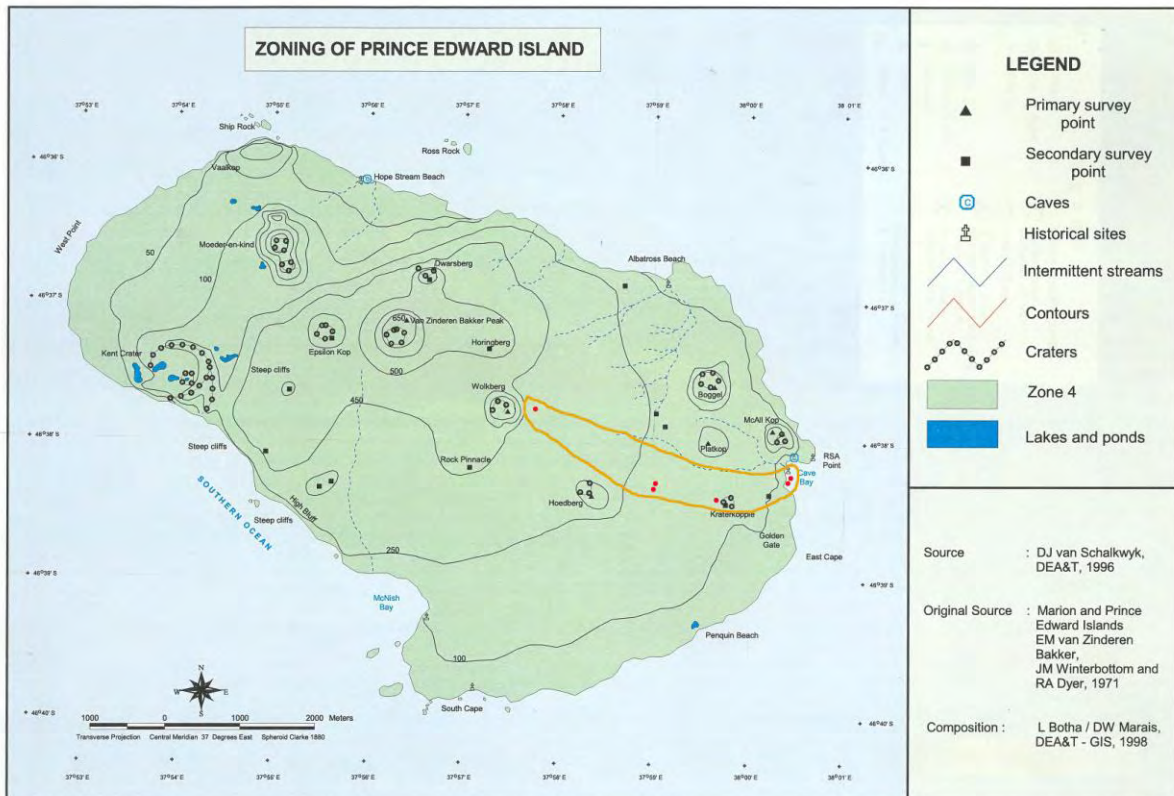


2 *Marion Island*

0 3 6 km

(g)

Fig. 10 *Ectemnorhinus* individuals that group into 4-step clade 4-2 mapped onto the islands according to a) sampling localities with place names for references, samples on Marion Island are indicated in green while those on Prince Edward Island are indicated in red b) sampling localities as used in the nested design c) those that share the same haplotypes nested into 0-step clades, the d) 1-step clades, e) 2-step clades f) 3 step clades and finally g) the complete 4-2 clade as depicted in Fig. 5.



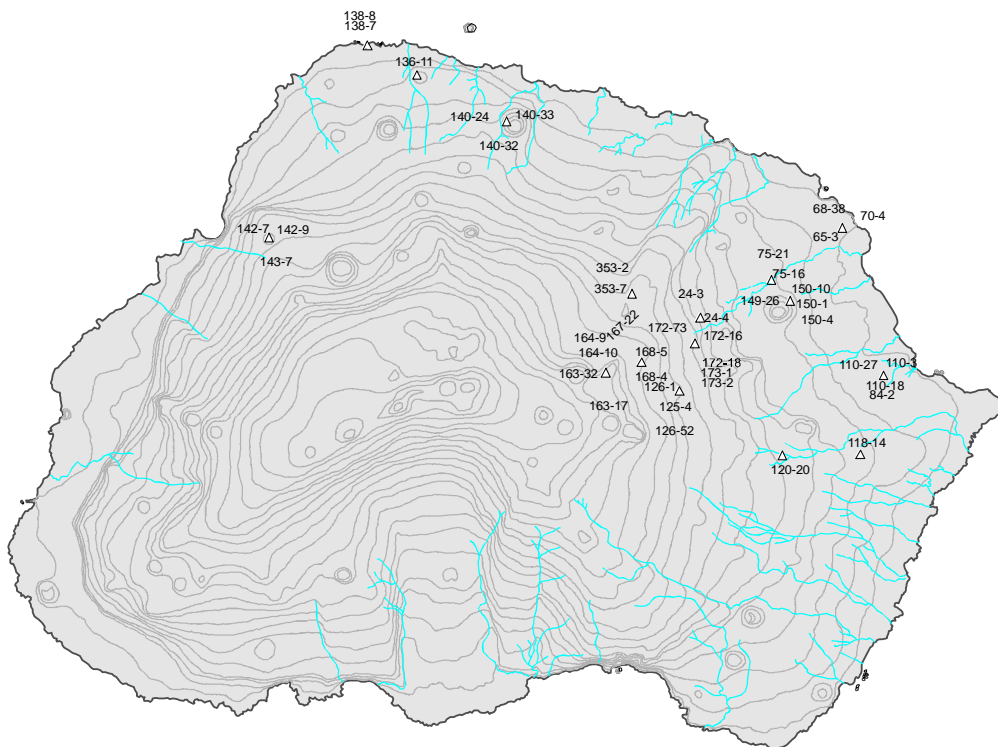
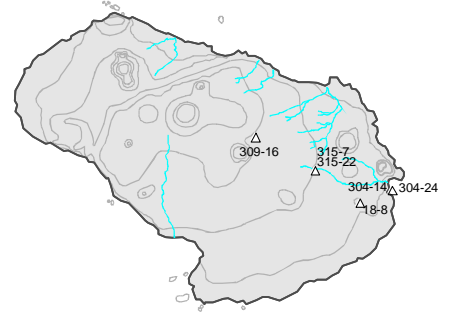
(a)



3

Prince Edward Island

0 3 6 km



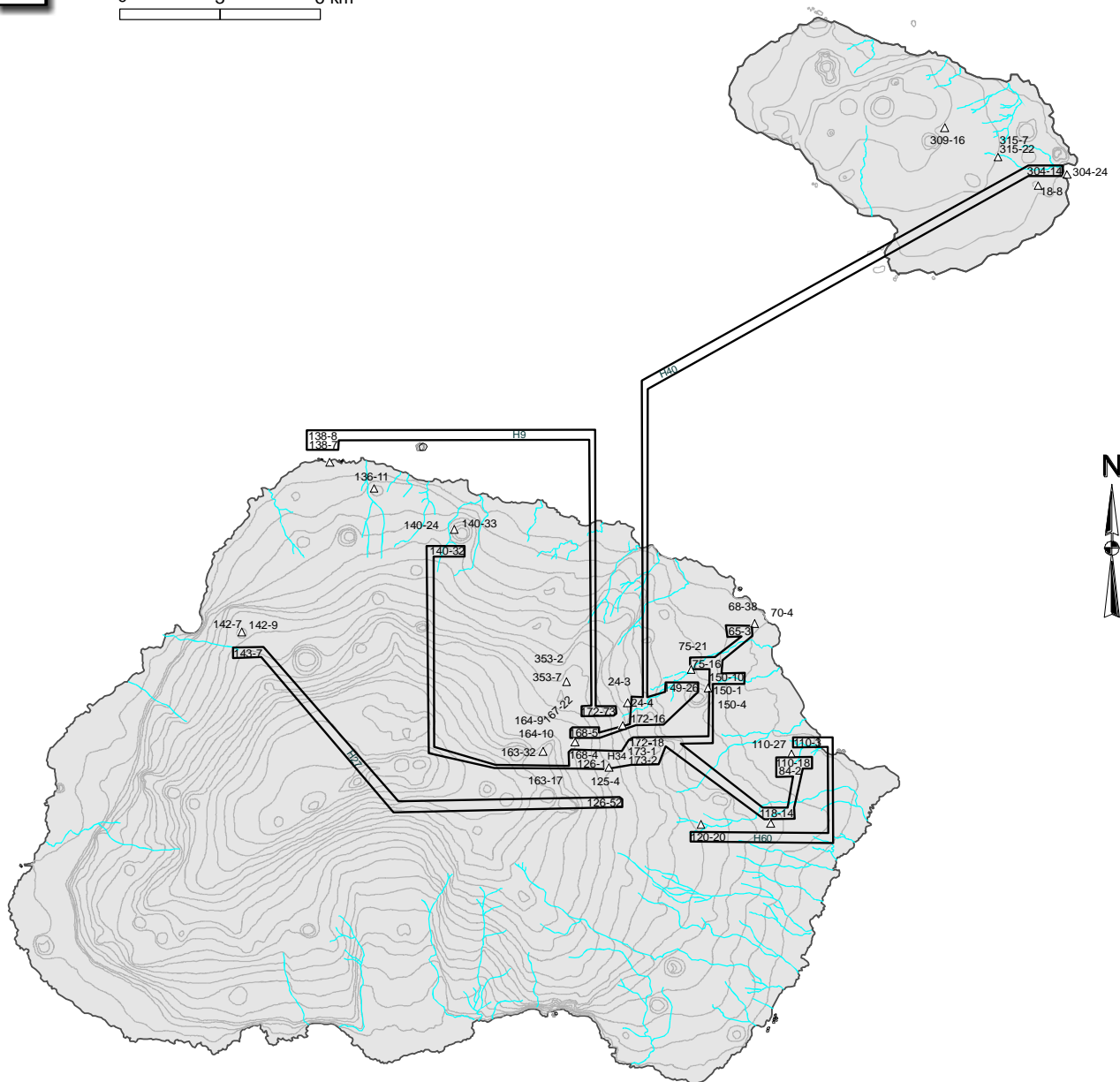
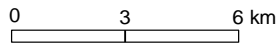
3

Marion Island

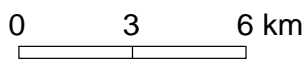
0 3 6 km

(b)

3 *Prince Edward Island*



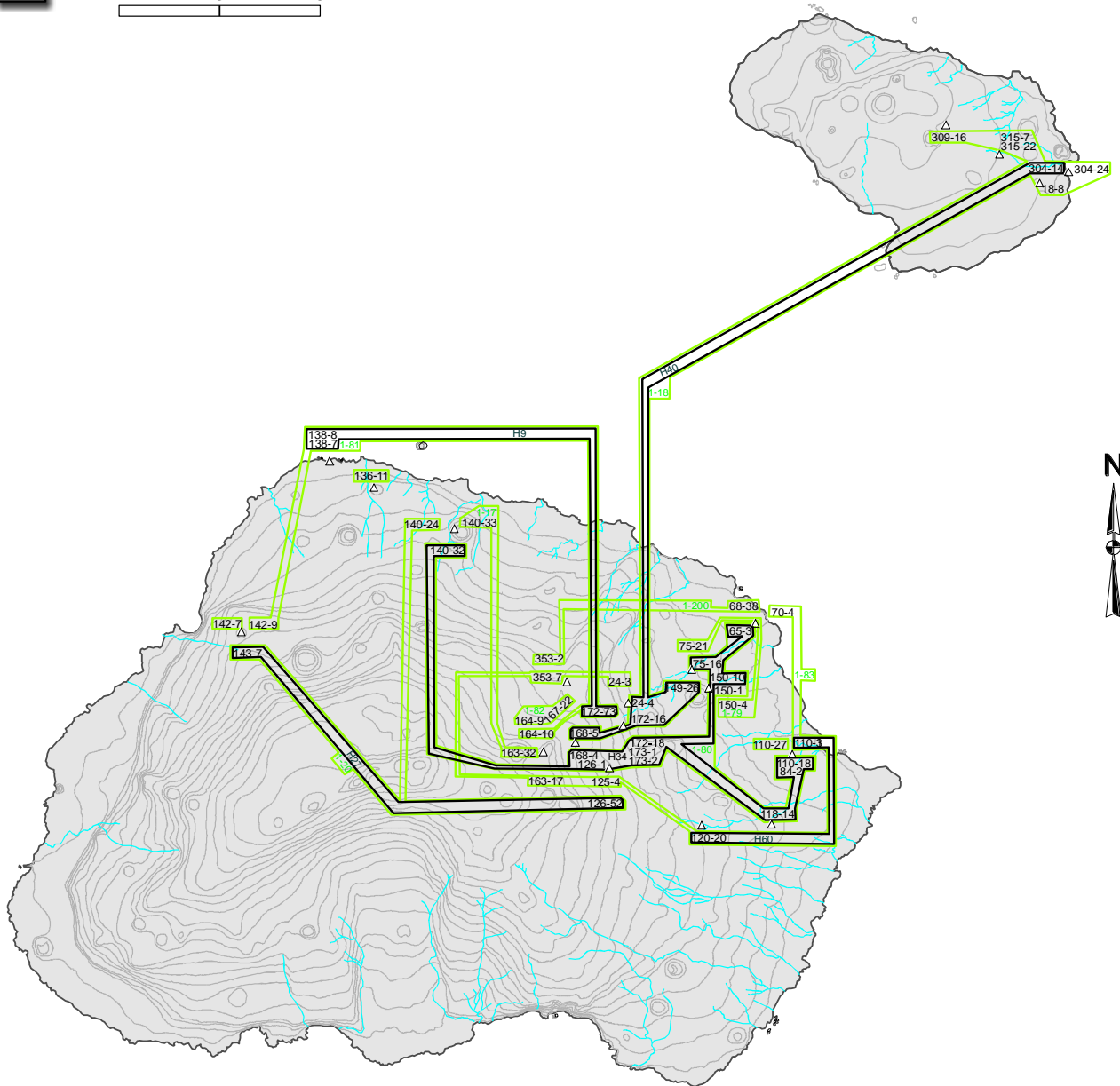
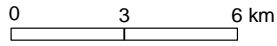
3 *Marion Island*



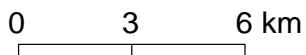
(c)



3 *Prince Edward Island*

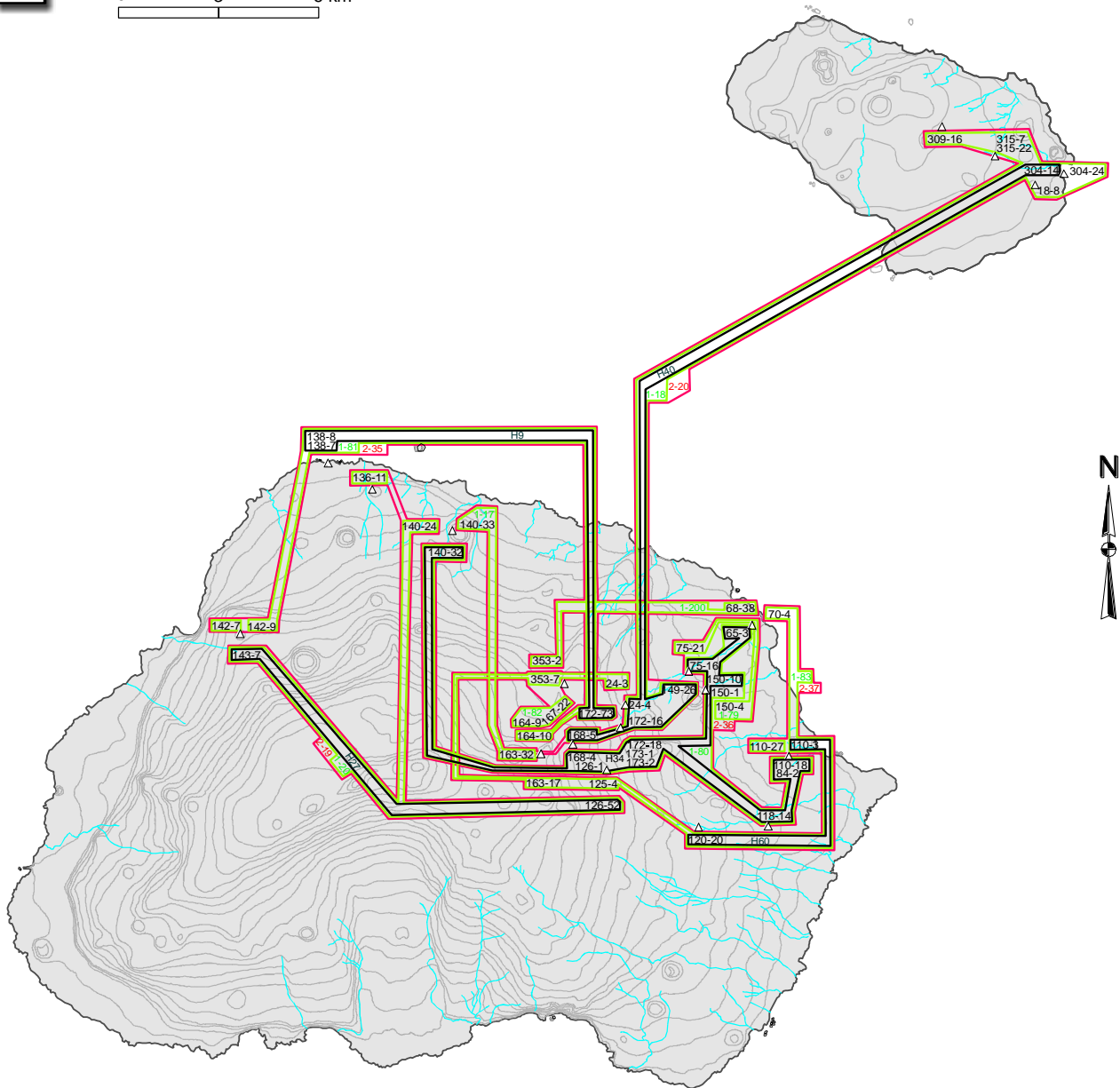
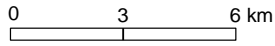


3 *Marion Island*

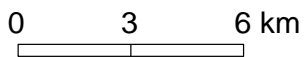


(d)

3 *Prince Edward Island*



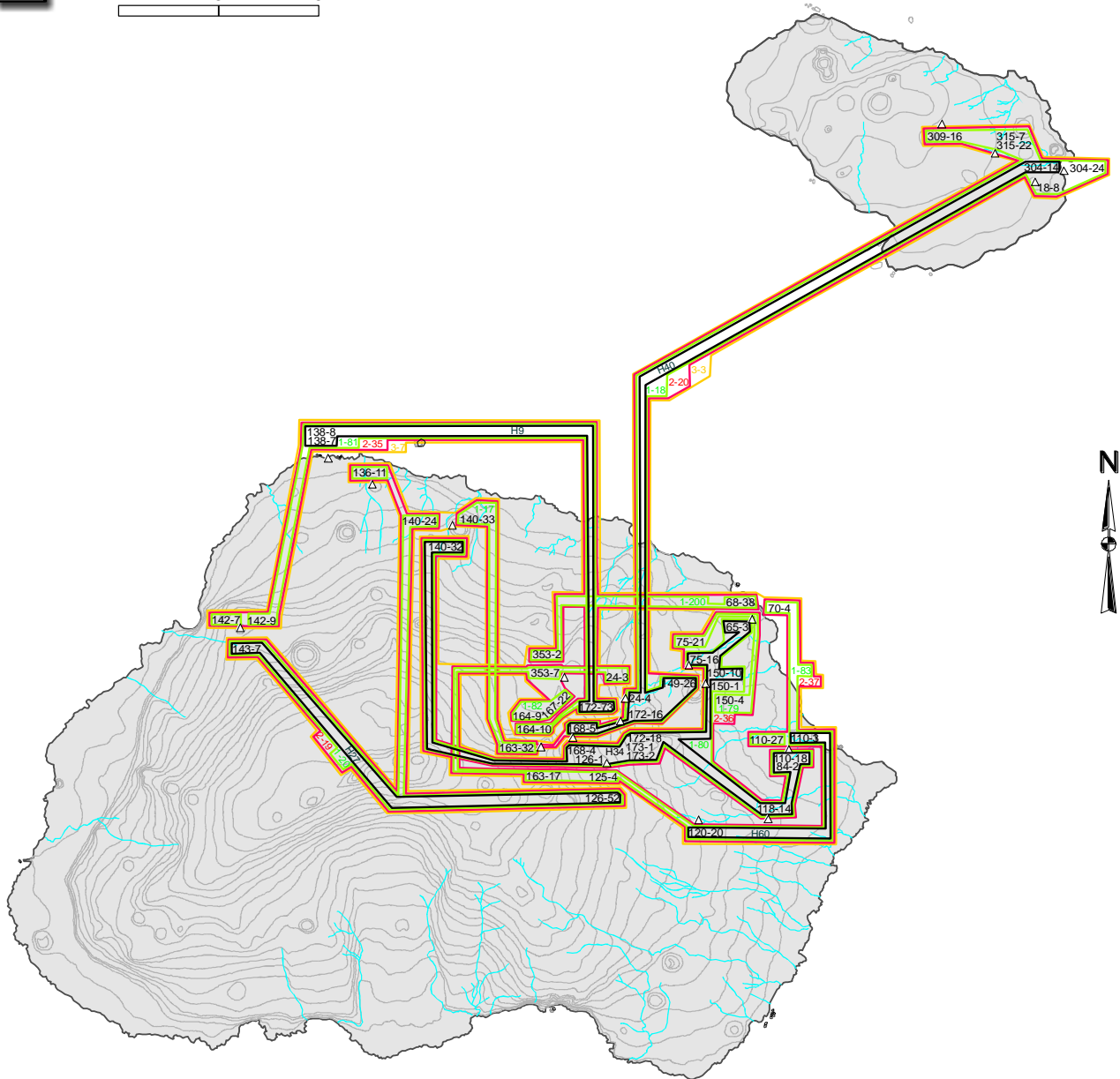
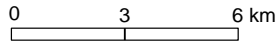
3 *Marion Island*



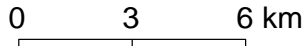
(e)



3 *Prince Edward Island*

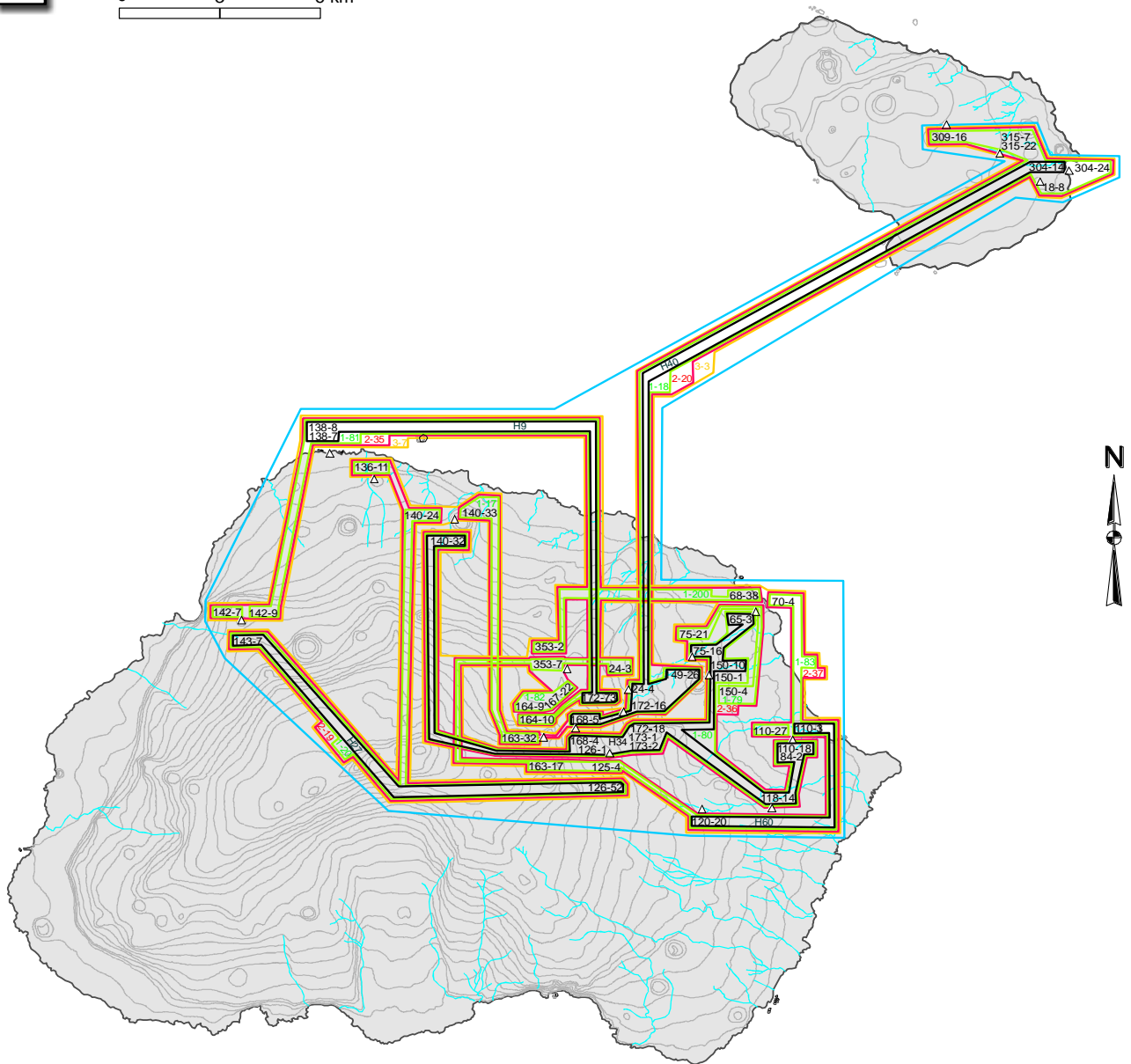
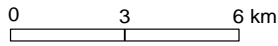


3 *Marion Island*

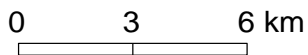


(f)

3 *Prince Edward Island*

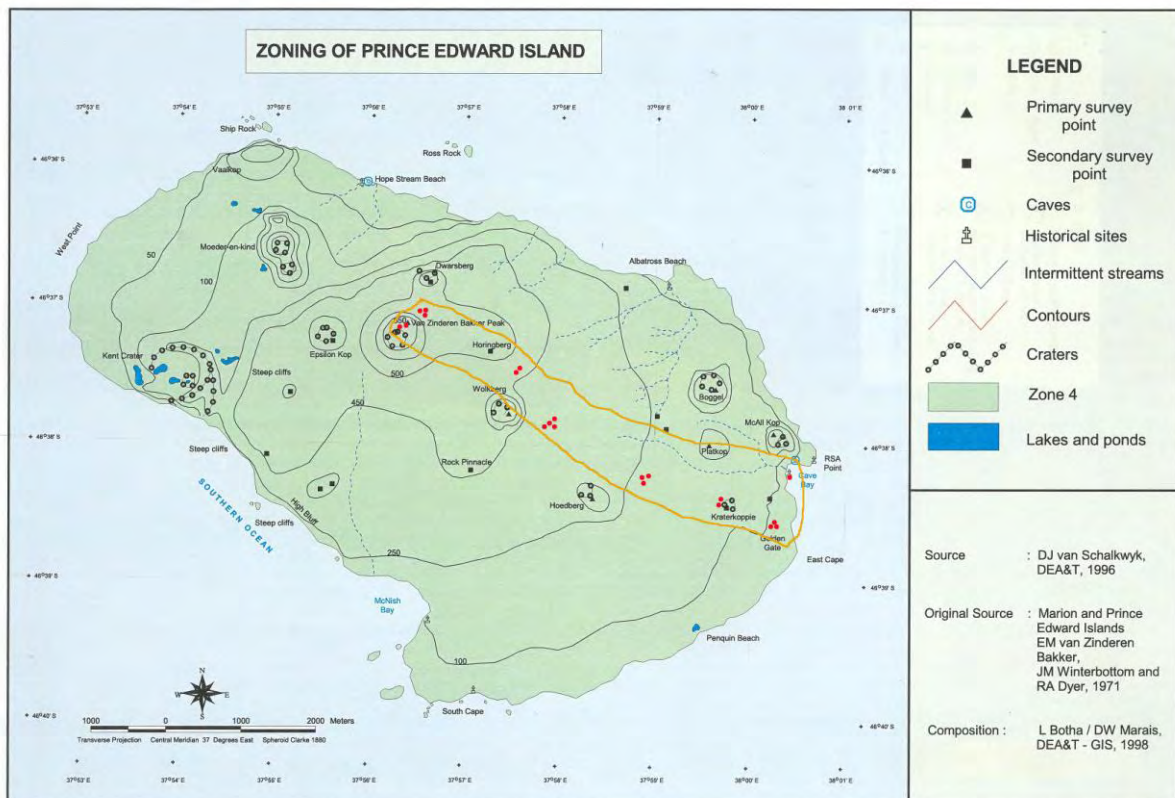


3 *Marion Island*



(g)

Fig. 11 *Ectemnorhinus* individuals that group into 4-step clade 4-3 mapped onto the islands according to a) sampling localities with place names for references, samples on Marion Island are indicated in green while those on Prince Edward Island are indicated in red b) sampling localities as used in the nested design c) those that share the same haplotypes nested into 0-step clades, the d) 1-step clades, e) 2-step clades f) 3 step clades and finally g) the complete 4-3 clade as depicted in Fig. 5.



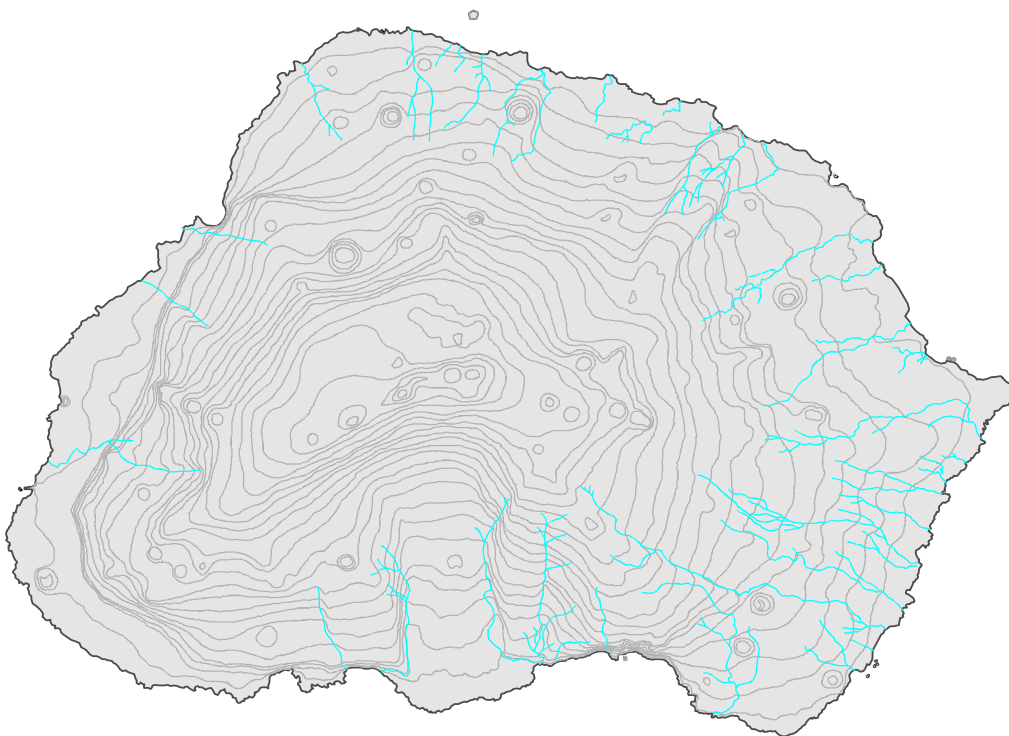
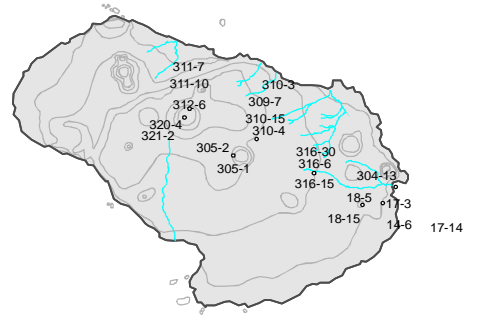
(a)



4

Prince Edward Island

0 3 6 km



4

Marion Island

0 3 6 km

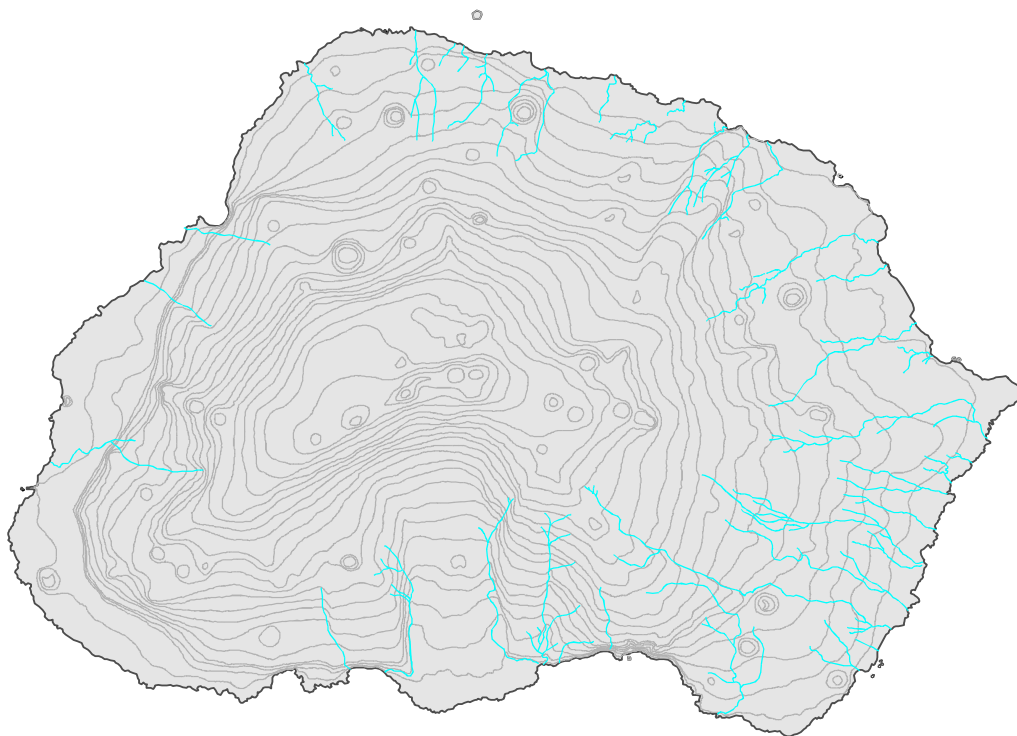
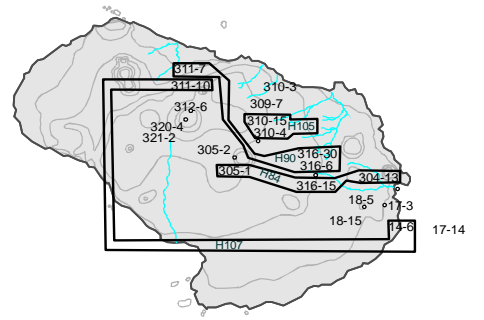
(b)



4

Prince Edward Island

0 3 6 km



4

Marion Island

0 3 6 km

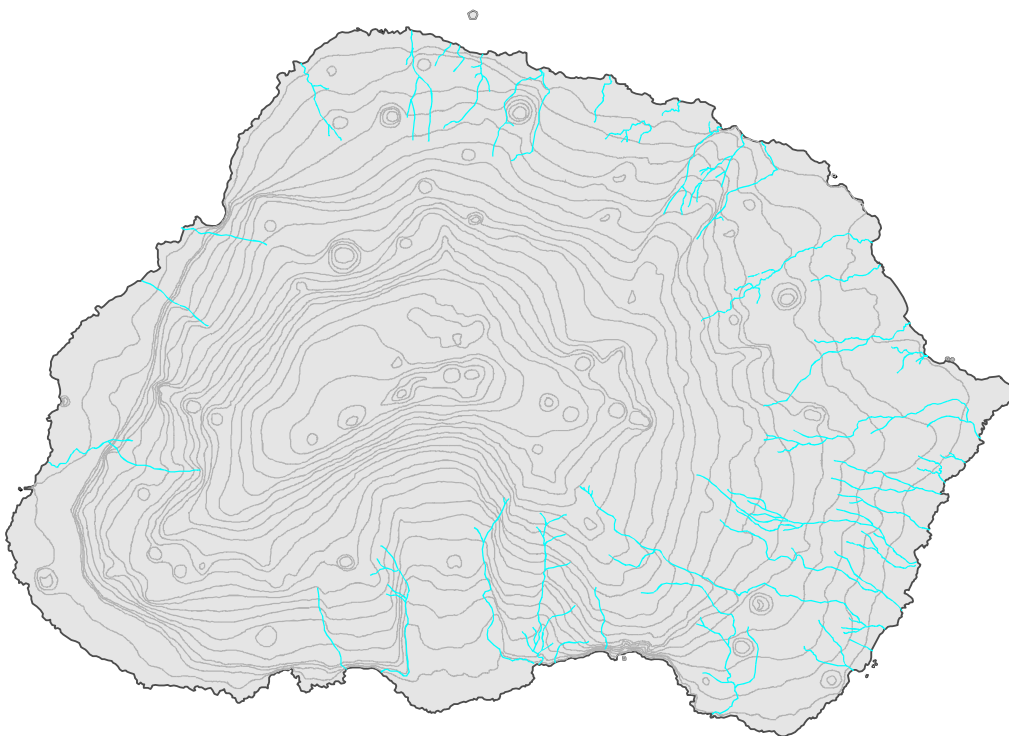
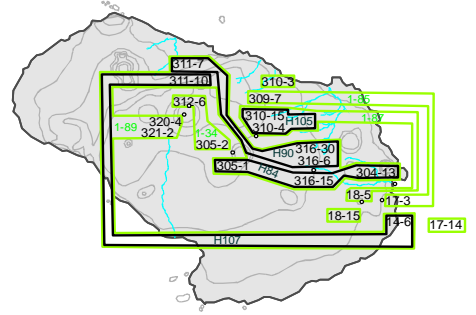
(c)



4

Prince Edward Island

0 3 6 km



4

Marion Island

0 3 6 km

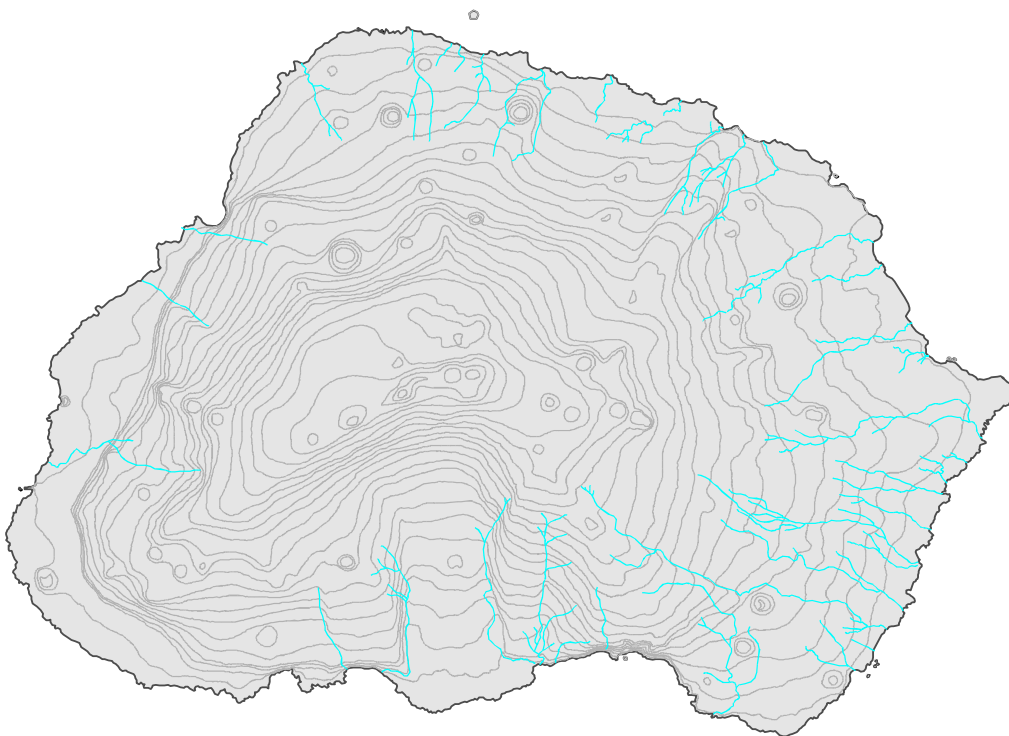
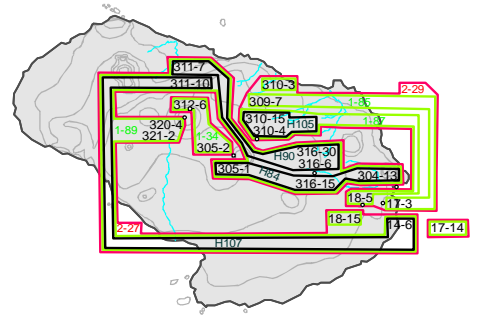
(d)



4

Prince Edward Island

0 3 6 km



4

Marion Island

0 3 6 km

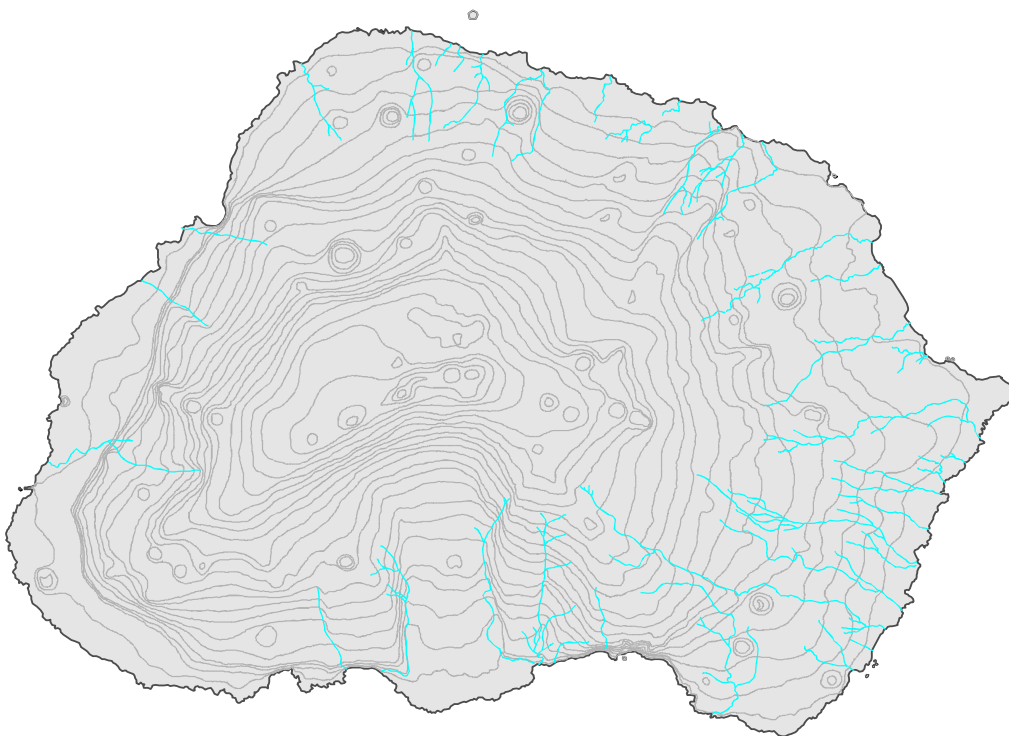
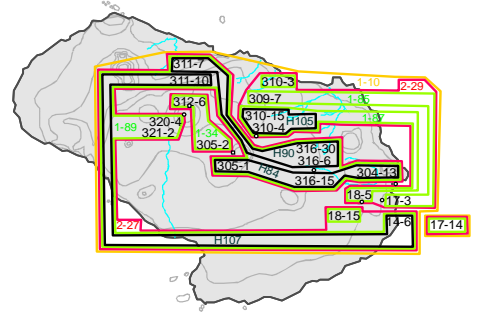
(e)



4

Prince Edward Island

0 3 6 km



4

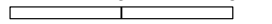
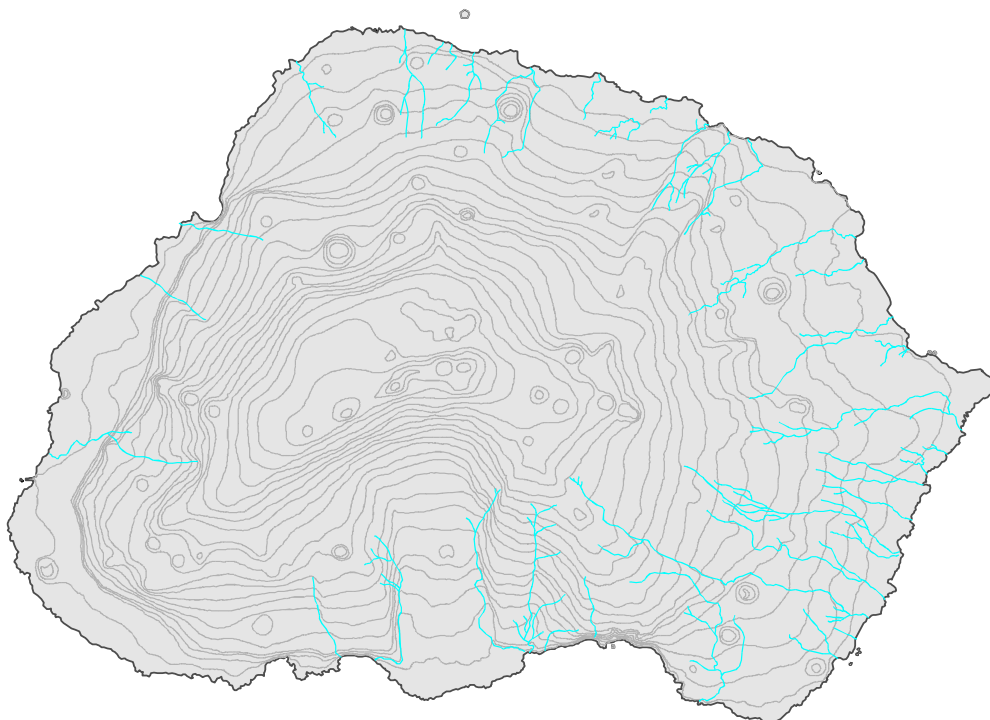
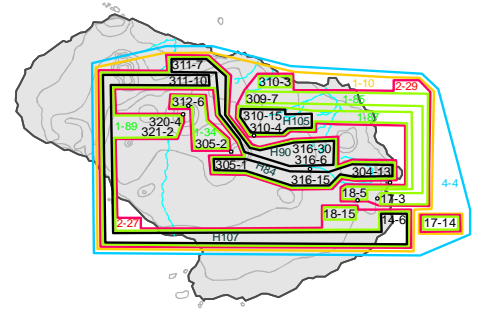
Marion Island

0 3 6 km

(f)

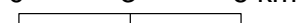
4 *Prince Edward Island*

0 3 6 km

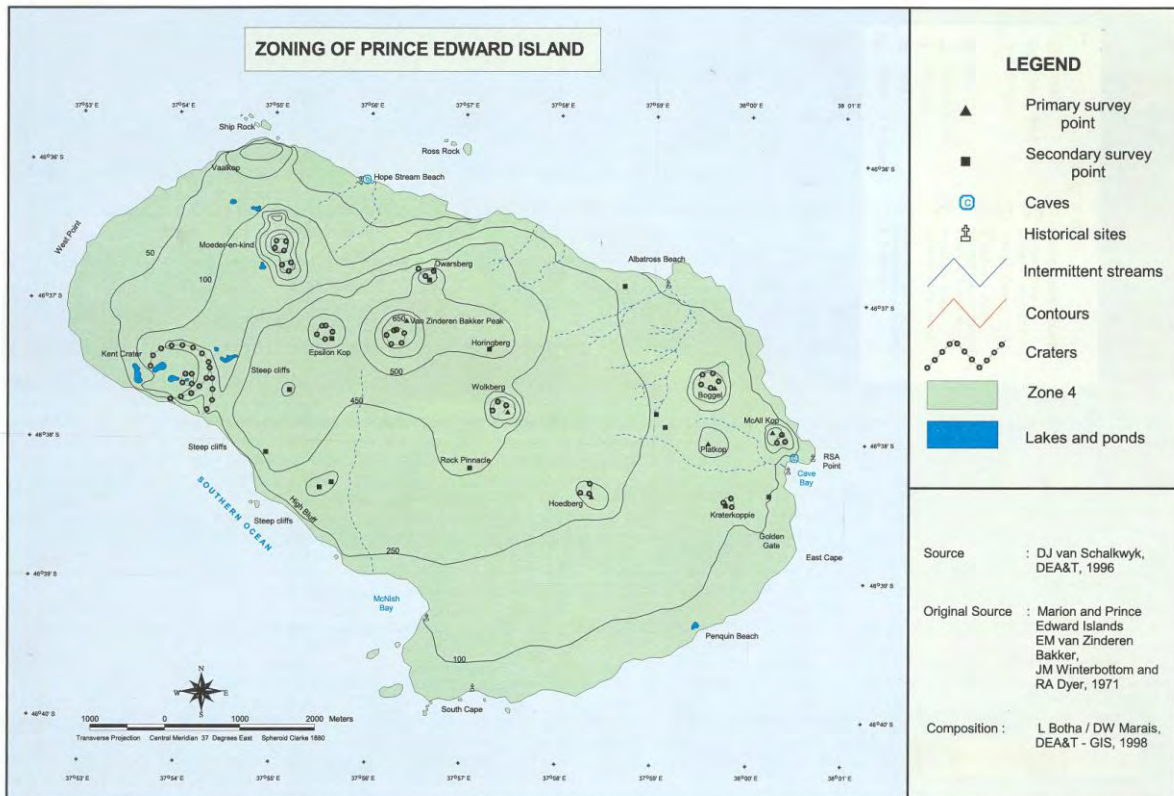
4 *Marion Island*

0 3 6 km



(g)

Fig. 12 *Ectemnorhinus* individuals that group into 4-step clade 4-4 mapped onto the islands according to a) sampling localities with place names for references, samples on Marion Island are indicated in green while those on Prince Edward Island are indicated in red b) sampling localities as used in the nested design c) those that share the same haplotypes nested into 0-step clades, the d) 1-step clades, e) 2-step clades f) 3 step clades and finally g) the complete 4-4 clade as depicted in Fig. 5.

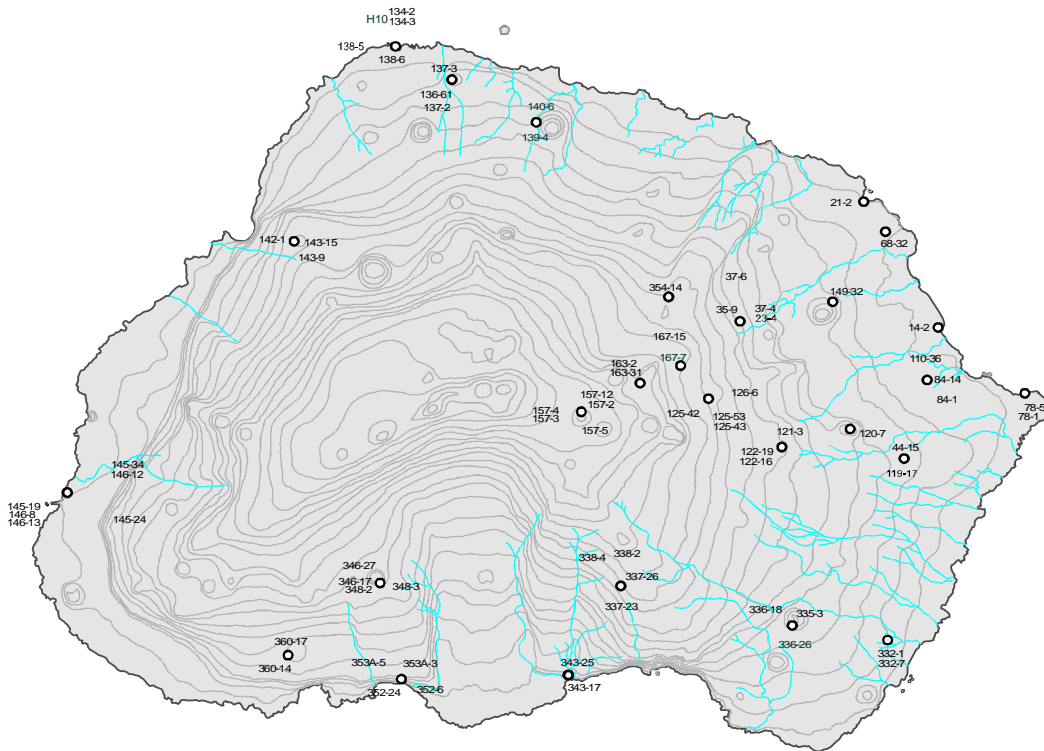
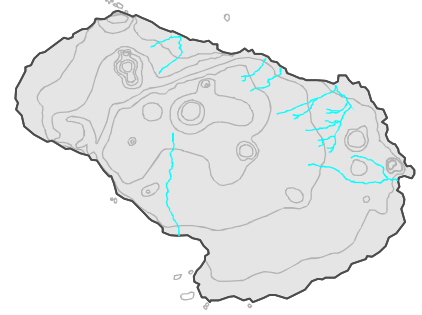
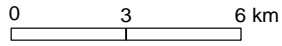


(a)



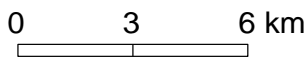
5

Prince Edward Island



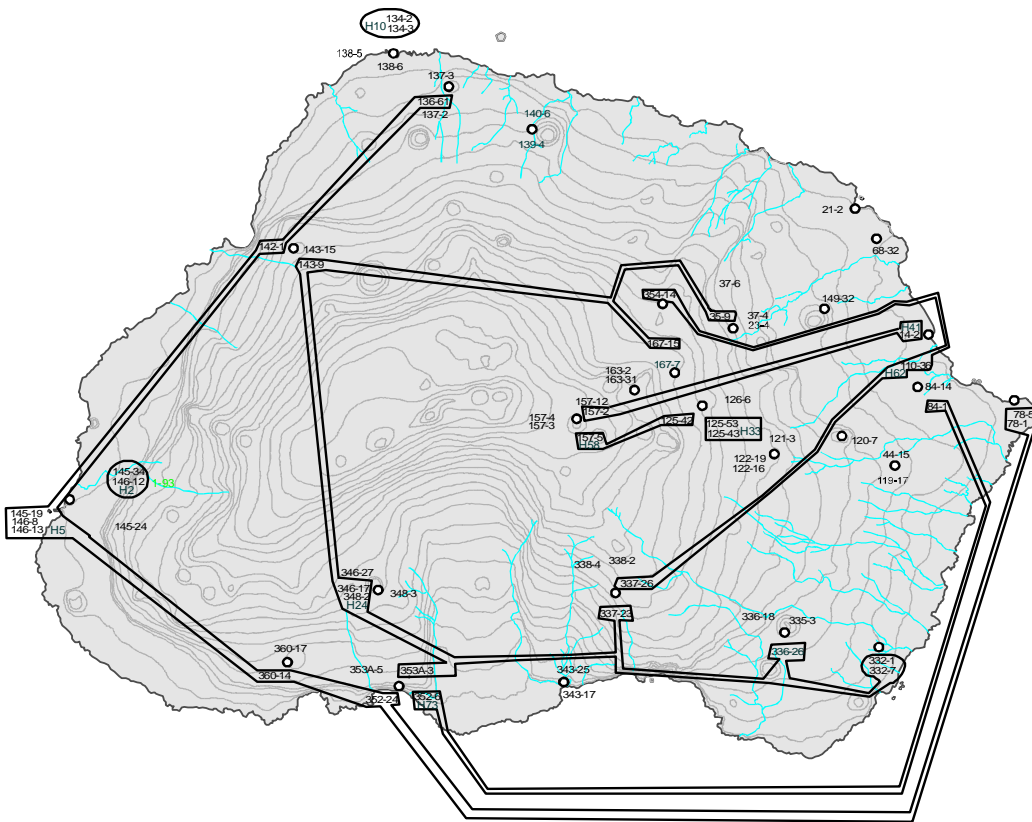
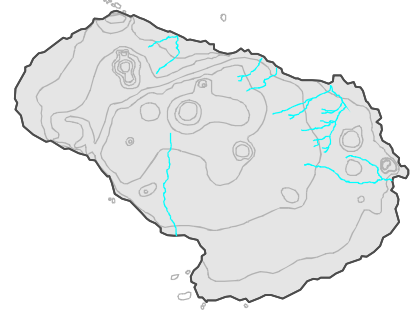
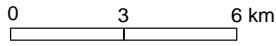
5

Marion Island

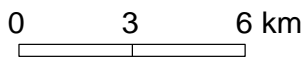


(b)

5 *Prince Edward Island*



5 *Marion Island*

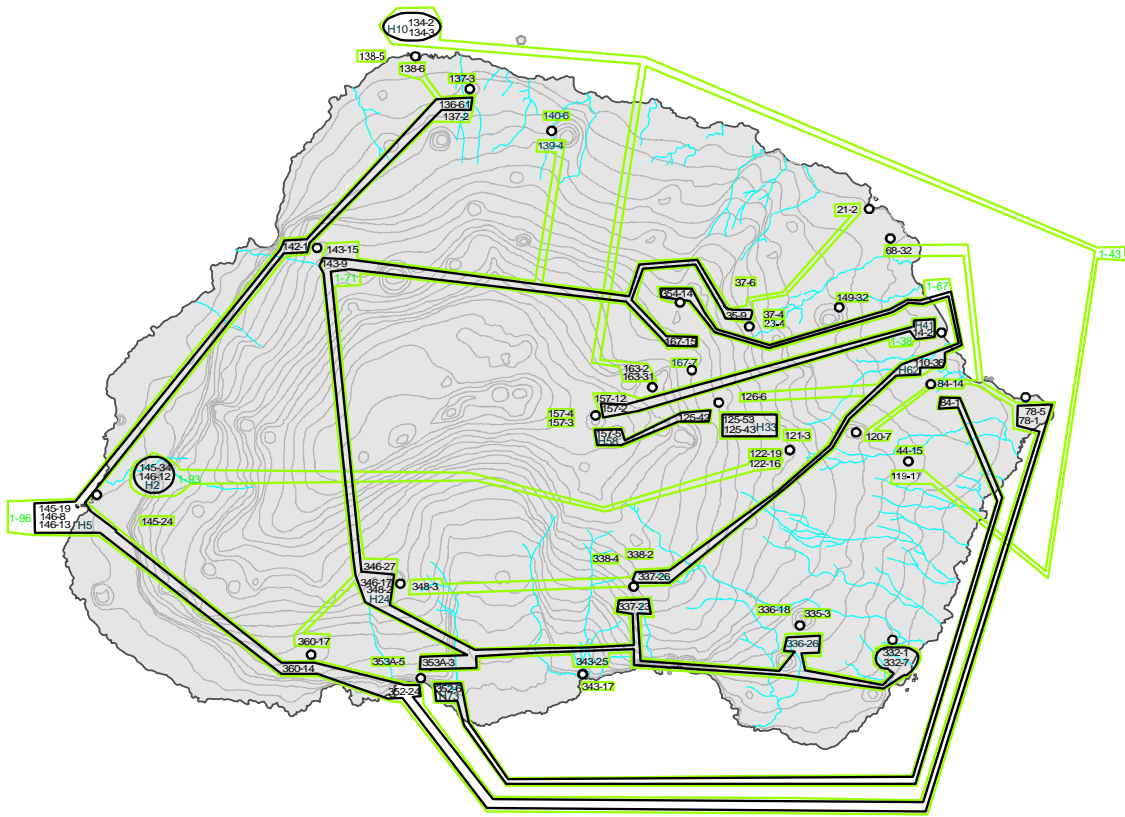
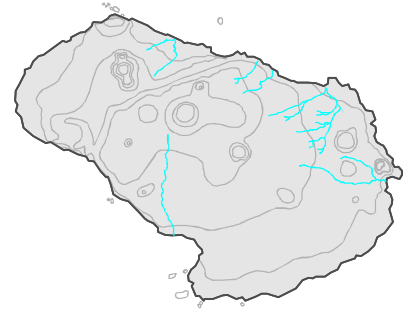
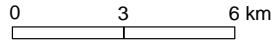


(c)



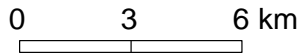
5

Prince Edward Island



5

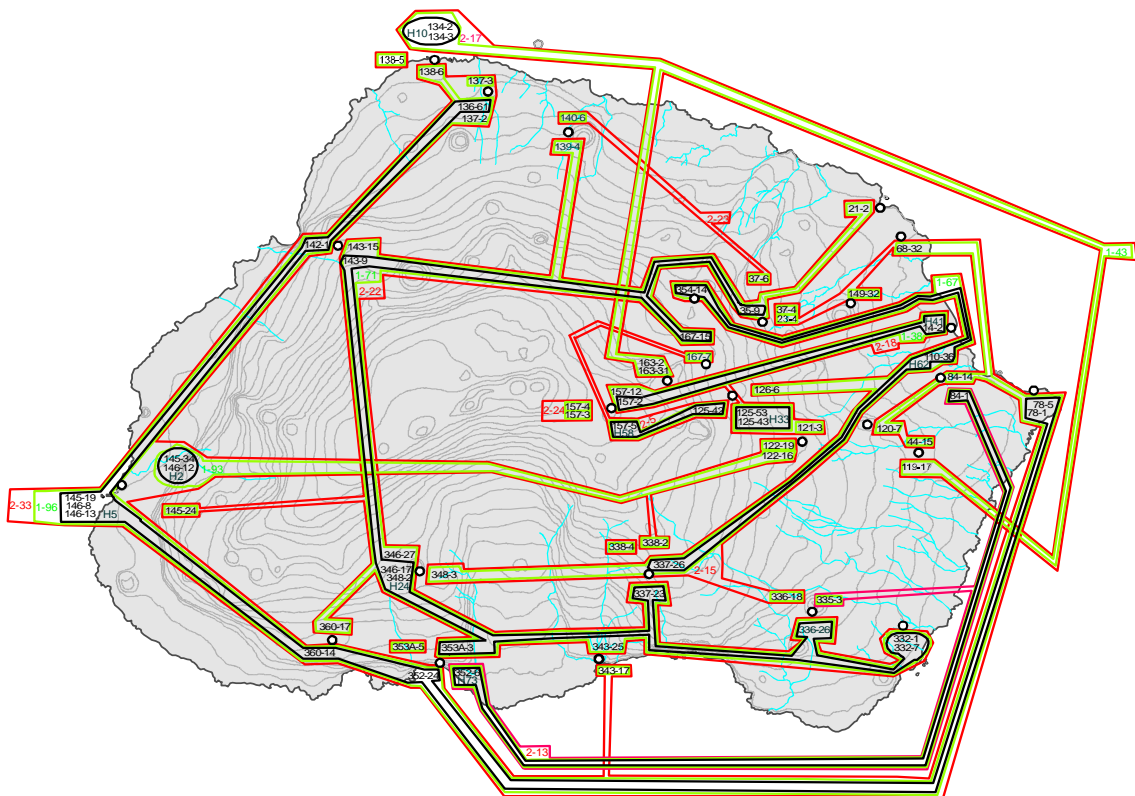
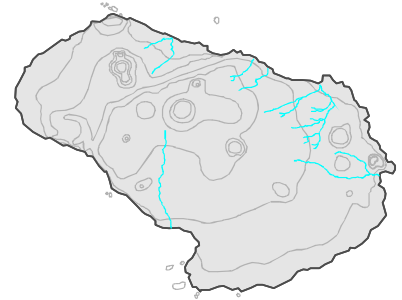
Marion Island



(d)

5 *Prince Edward Island*

0 3 6 km




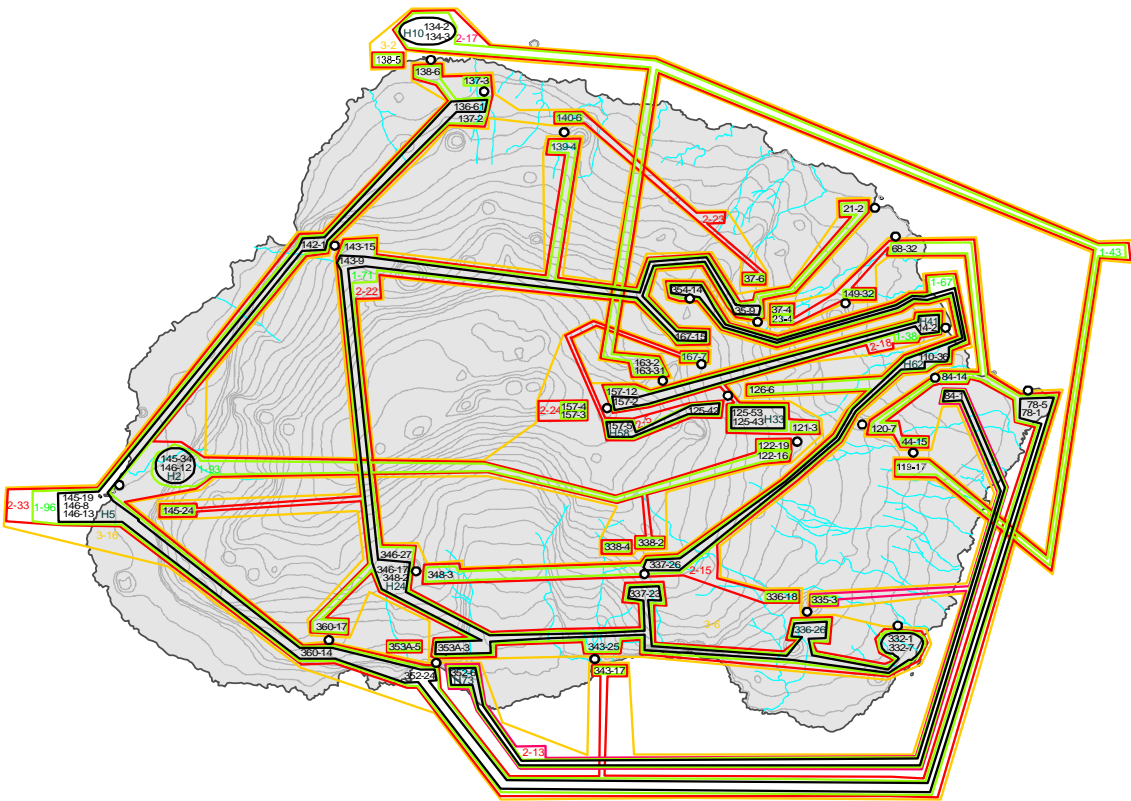
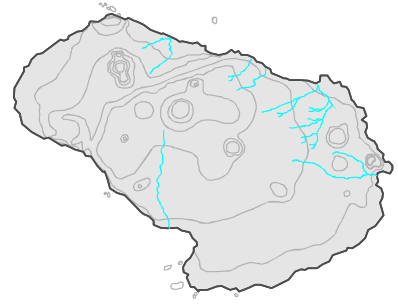
5 *Marion Island*

0 3 6 km

(e)

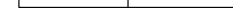
5 *Prince Edward Island*

0 3 6 km

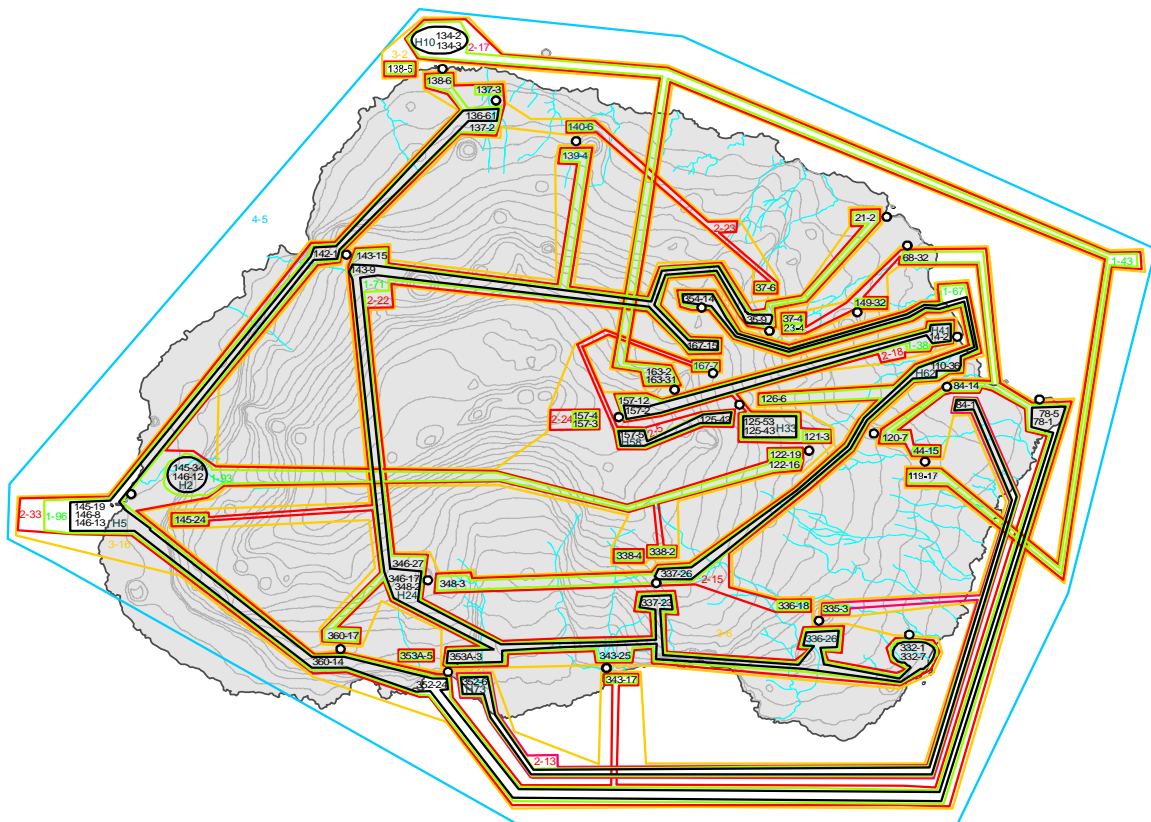
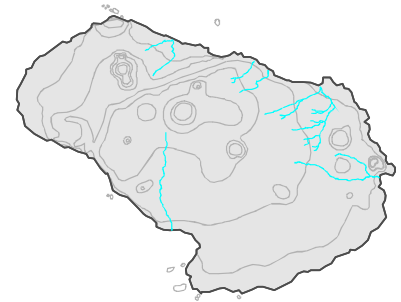
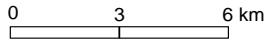
5 *Marion Island*

0 3 6 km

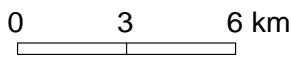



(f)

5 *Prince Edward Island*

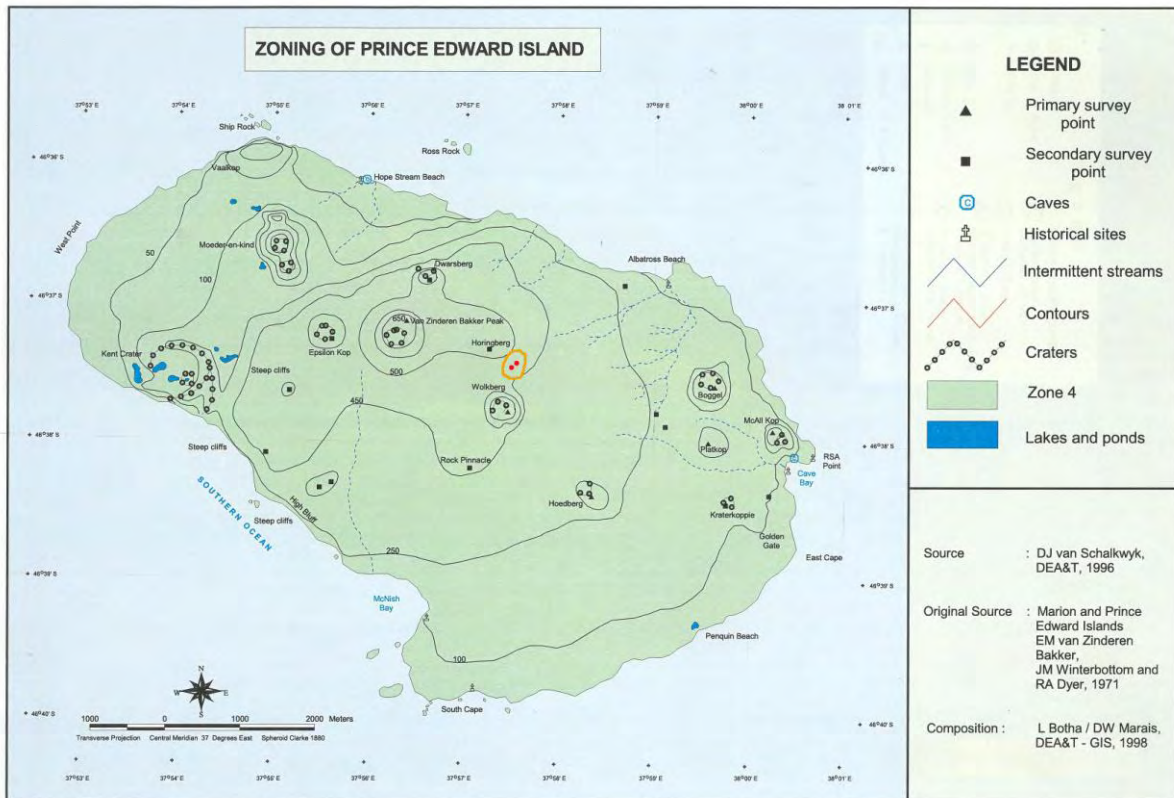


5 *Marion Island*



(g)

Fig. 13 *Ectemnorhinus* individuals that group into 4-step clade 4-5 mapped onto the islands according to a) sampling localities with place names for references, samples on Marion Island are indicated in green while those on Prince Edward Island are indicated in red b) sampling localities as used in the nested design c) those that share the same haplotypes nested into 0-step clades, the d) 1-step clades, e) 2-step clades f) 3 step clades and finally g) the complete 4-5 clade as depicted in Fig. 5.



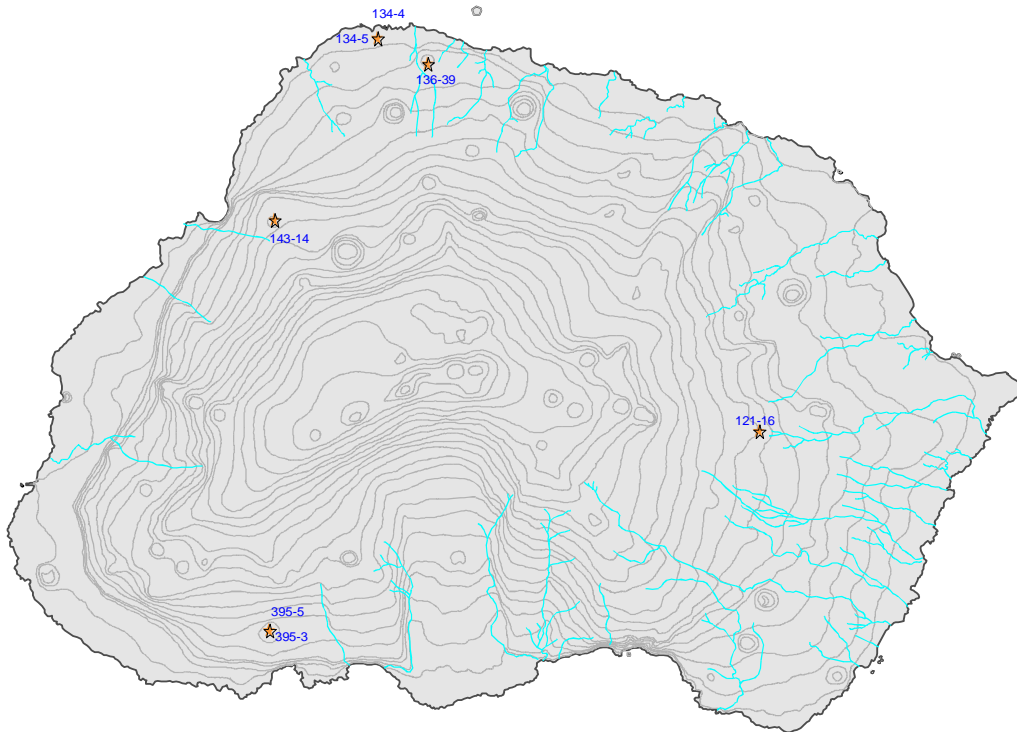
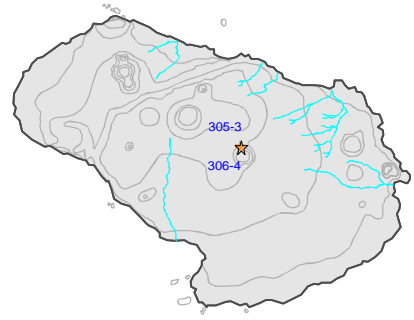
(a)



6

Prince Edward Island

0 3 6 km



6

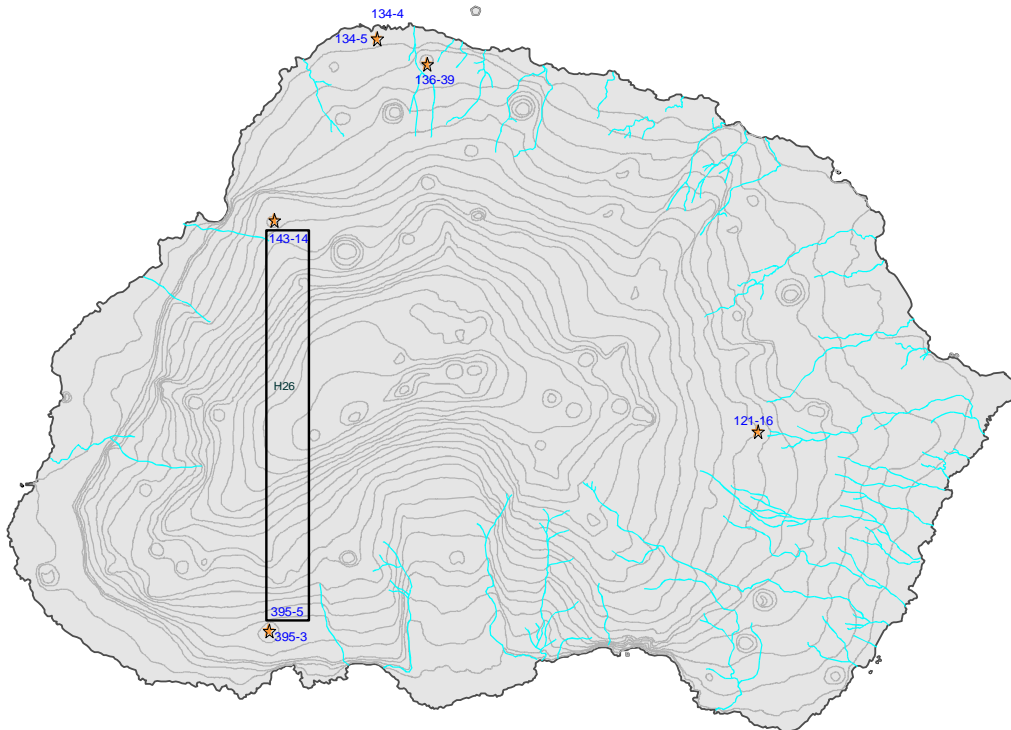
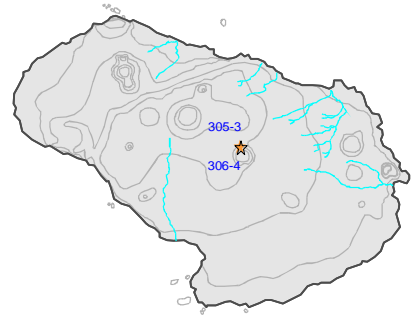
Marion Island

0 3 6 km

(b)

6 *Prince Edward Island*

0 3 6 km



6 *Marion Island*

0 3 6 km

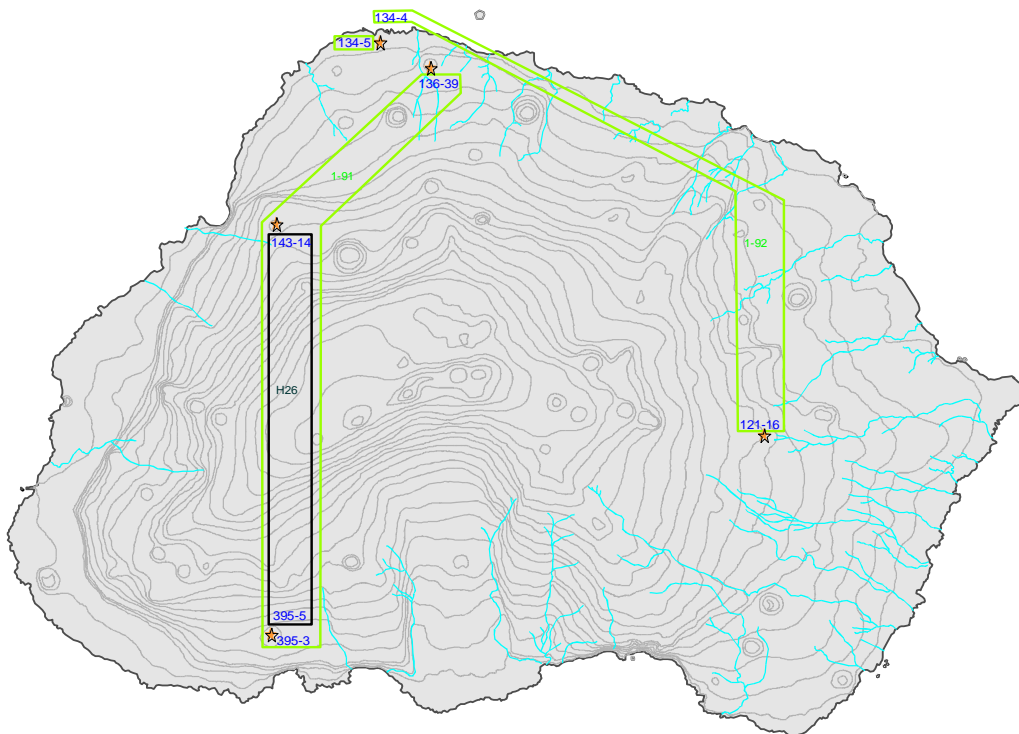
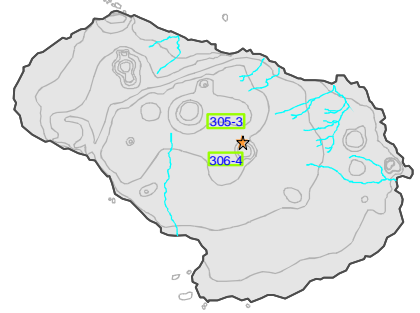
(c)



6

Prince Edward Island

0 3 6 km



6

Marion Island

0 3 6 km

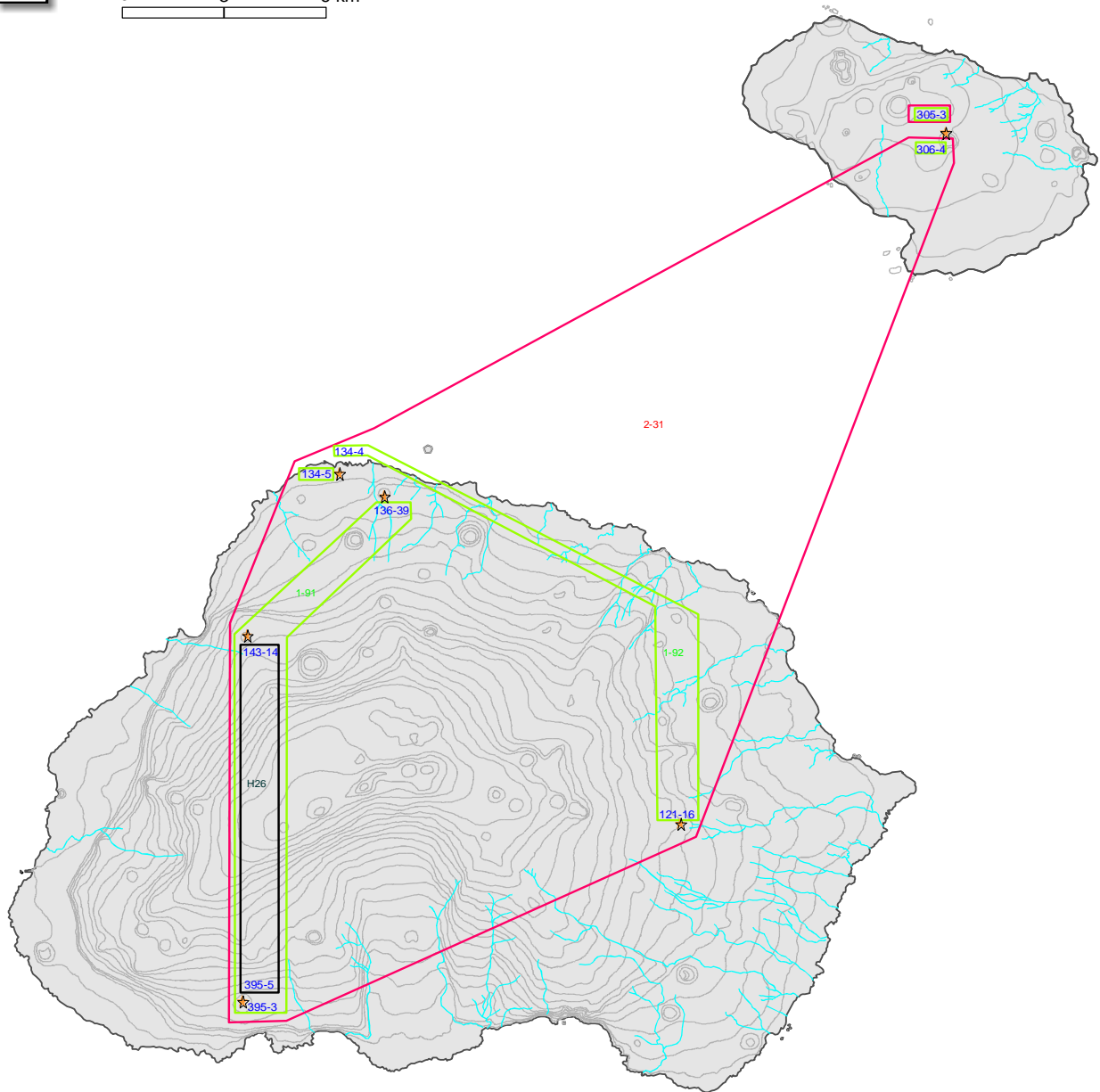
(d)



6

Prince Edward Island

0 3 6 km



6

Marion Island

0 3 6 km

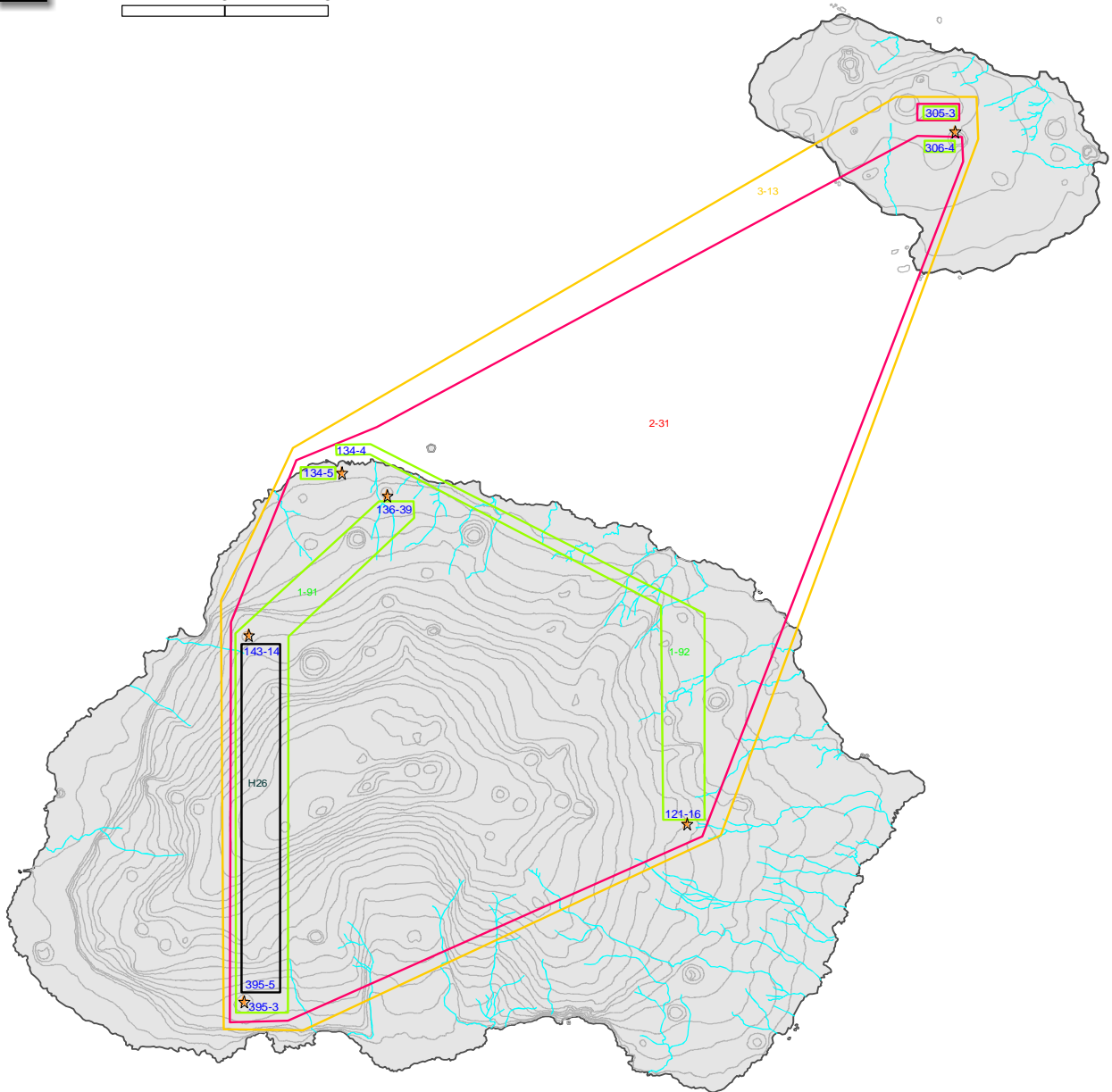
(e)



6

Prince Edward Island

0 3 6 km



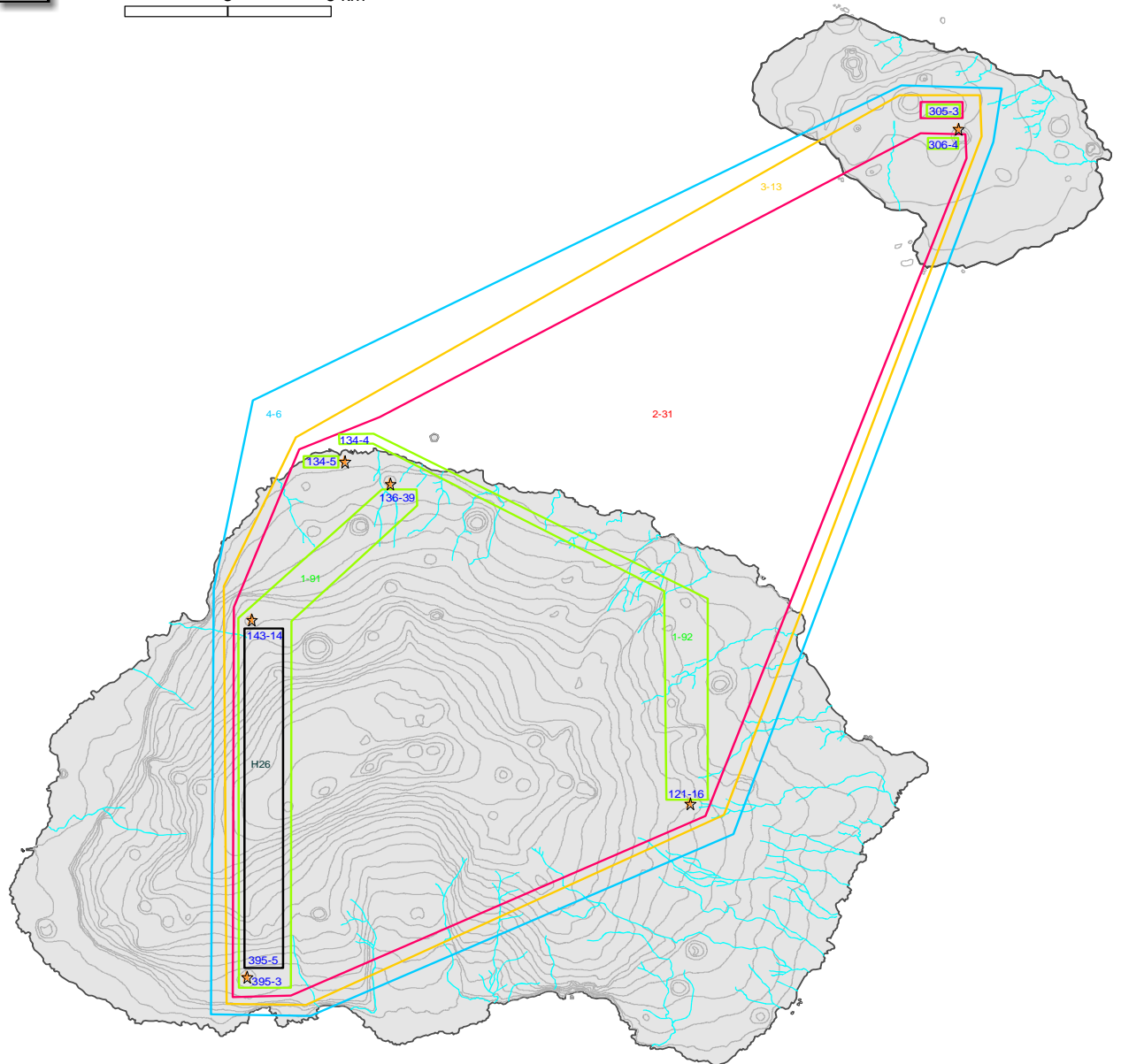
6

Marion Island

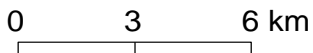
0 3 6 km

(f)

6 *Prince Edward Island*



6 *Marion Island*



(g)

Fig. 14 *Ectemnorhinus* individuals that group into 4-step clade 4-6 mapped onto the islands according to a) sampling localities with place names for references, samples on Marion Island are indicated in green while those on Prince Edward Island are indicated in red b) sampling localities as used in the nested design c) those that share the same haplotypes nested into 0-step clades, the d) 1-step clades, e) 2-step clades f) 3 step clades and finally g) the complete 4-6 clade as depicted in Fig. 5.

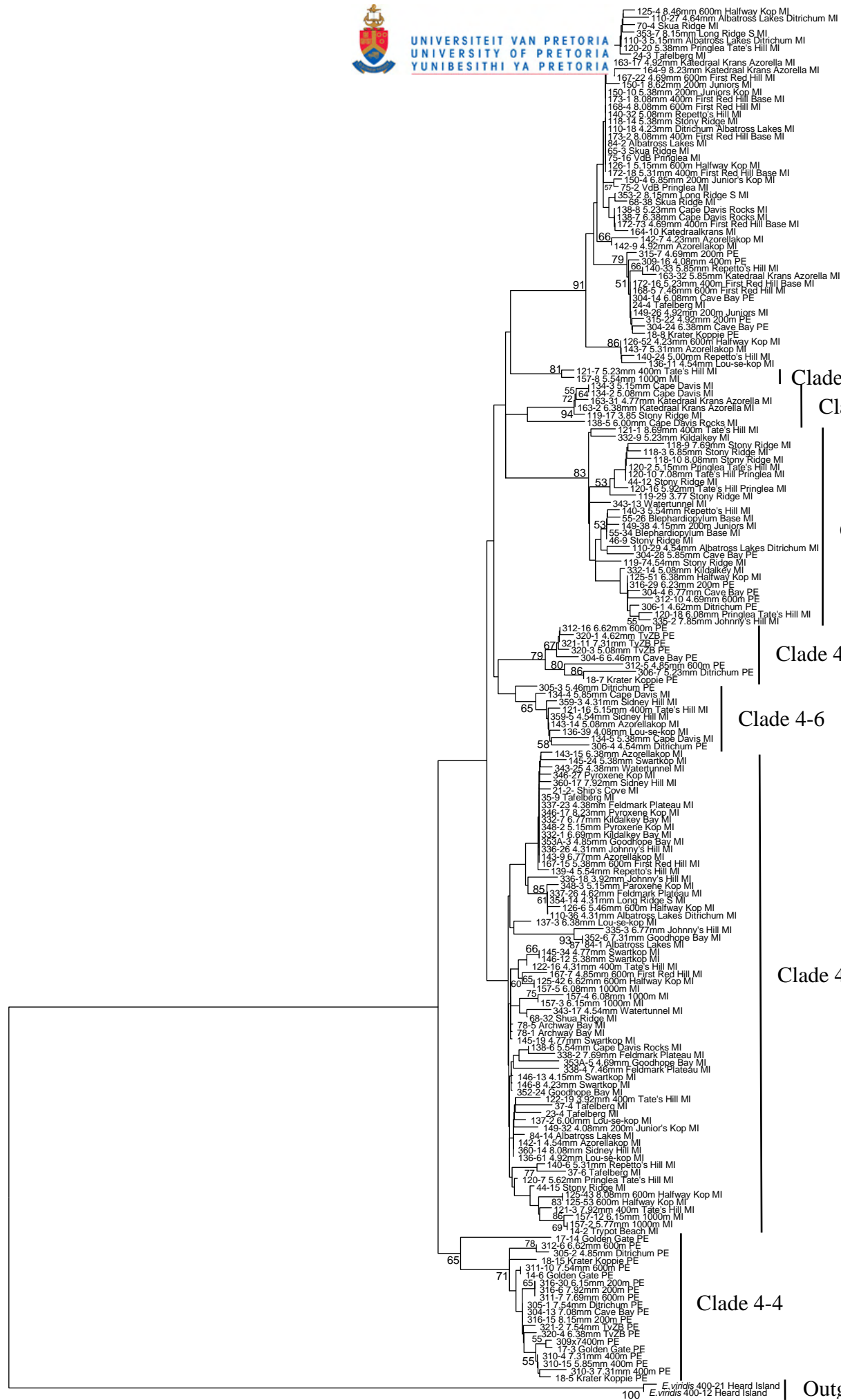
Fig. 9a shows that the individuals included in nested clade 4-1 are located mainly to the eastern part of MI with some individuals occurring on PEI. Although 23 haplotypes were detected in clade 4-1 only three are represented by more than one individual (Fig. 9b) and only one of these haplotypes is shared between the two islands. These individuals group into 15 one-step clades (Fig. 9c), five two-step clades (Fig. 9d) and two 3-step clades (Fig. 9e). Nested clade 4-2 consists of just ten individuals the vast majority of which were collected from PEI (Fig. 10a). Within these ten individuals, nine different haplotypes were identified, with the two individuals sharing a haplotype occurring on PEI (Fig. 10b). The individuals that make up nested clade 4-2 group into six nested one-step clades (Fig. 10c), five nested 2-step clades (Fig. 10d) and two nested 3-step clades (Fig. 10e). It is important to note that the 3-step clades also separate the MI individuals from the PEI individuals. Nested clade 4-3 consists of 49 individuals that map mainly to the north-eastern side of MI, again with a few individuals occurring on PEI (Fig. 11a). In Fig. 11b, 30 different haplotypes are apparent with haplotype H34 occurring in 13 individuals. The nested design unfolds further into 12 one-step clades (Fig. 11c), five two-step clades (Fig. 11d) and two three-step clades (Fig. 11e). Fig. 12a shows that all 20 *E. kucheli* (Grobler *et al.*, 2006) individuals of nested clade 4-4 map to PEI while Fig. 12b shows that they represent 14 different haplotypes. Seven 1-step clades (Fig. 12c), four 2-step clades (Fig. 12d) and two 3-step clades (Fig. 12e) all combine into nested clade 4-4 (Fig. 11f). Fig. 13a shows that nested clade 4-5 maps to the whole of MI and as the largest of the 4-step clades, consists of 72 individuals. Nested clade 4-5 has no representatives on PEI and the individuals therein can be grouped into 47 different haplotypes (Fig. 13b), 29 1-step clades (Fig. 13c), 12 2-step clades (Fig. 13d) and three 3-step clades (Fig. 13e). Four of the nine individuals making up nested clade 4-6 map to the north-western side of MI, two to the south-western side of MI, one to the central eastern side of MI and the last two to PEI (Fig. 14a). Only one of the eight haplotypes that occurred in more than one individual was shown to link the individuals from the north-western side of MI to those of the south-western side of MI (Fig. 14b). The remaining subclades of nested clade 4-6 consist of five one-step clades (Fig. 14c), two 2-step clades (Fig. 14d) and a single 3-step clade (Fig. 14e).

Phylogenetic analyses

Table 3 shows the jModel Test results obtained for the whole data set, the two islands individually, as well as the different 4-step nested clades. Two phylogenetic trees containing all the individuals that are represented in the nested clade analyses are presented. Fig. 15 shows the tree obtained from the Minimum Evolution analyses while Fig. 16 shows the tree obtained from the Bayesian analyses. All the four-step nested clades are retained in the trees, except nested clade 4-2 and nested clade 4-5, which were each divided into ‘a’ and ‘b’ subclades. Individuals grouping in nested clades 4-1, 4-3 and 4-4 also group together in both trees with high levels of support. Individuals grouping together in nested clade 4-6 also group together in both trees. Although clade 4-6 does not have high bootstrap support in the Minimum Evolution tree (Fig. 15), it has 93 % support in the Bayesian tree (Fig. 16). Clade 4-2 in the nested design shows that sequences 121-7 and 157-8 (clade 2b) are from MI while the rest of the sequences (clade 2a) are from PEI and that the individuals from the different islands separates from each other even at the 3-step level (Fig. 10e). Both clades 2a and 2b have high levels of support in both trees (Figs. 15& 16). Nested clade 4-5 separates in both the Minimum Evolution (Fig. 15) and the Bayesian trees (Fig. 16) into clades 4-5a and 4-5b. Although neither clades 4-5a nor 4-5b have bootstrap support of more than 50% in the Minimum Evolution tree (Fig. 15), clade 4-5a has 94 % support and clade 4-5b has 85 % support in the Bayesian tree (Fig.16). The nested design (Fig. 6) shows that individuals grouping in clade 4-5b separate from those grouping in clade 4-5a by at least six missing haplotypes. Other 3-step nested clades are also easily recognisable in the trees. As the minimum evolution tree did not recover any of the internal nodes with high levels of support, this analysis gave no information about the relationship between the major clades. The Bayesian analyses (Fig. 16) on the other hand, clarifies the relationship between the major clades, where individuals grouping together in nested clade 4-4 are basal to the rest of the tree, as suggested by Grobler *et al.* (2006), and all the other clades also group together with 95% support.

Table 3. jModelTest results obtained for the different datasets. In each case the outgroup was included. I = Proportion of invariable sites and G= Gamma distribution shape parameter

| Dataset | Model selected | Base frequencies | | | | Substitution model | | | | | | I | G |
|------------------|----------------|------------------|--------|--------|--------|--------------------------------------|-----------|----------|---------|-----------|--------|--------|---------------------------|
| | | A | C | G | T | A-C | A-G | A-T | C-G | C-T | G-T | | |
| Total Cladogram | TIM3+I+G | 0.3111 | 0.1671 | 0.1296 | 0.3921 | 2.7910 | 46.0397 | 1.0000 | 2.7910 | 15.6917 | 1.0000 | 0.4090 | 0.2410 |
| MI dataset | TIM3+I+G | 0.3129 | 0.1668 | 0.1291 | 0.3912 | 2.5888 | 38.6352 | 1.0000 | 2.5888 | 13.5356 | 1.0000 | 0.4360 | 0.2480 |
| PEI dataset | TrN+I | 0.3098 | 0.1682 | 0.1309 | 0.3911 | 1.0000 | 37.0083 | 1.0000 | 1.0000 | 10.9496 | 1.0000 | 0.8220 | Equal rates for all sites |
| Nested clade 4-1 | TIM1+I | 0.3027 | 0.1671 | 0.1390 | 0.3913 | 1.0000 | 92.6399 | 5.0722 | 5.0722 | 29.0701 | 1.0000 | 0.8470 | Equal rates for all sites |
| Nested clade 4-2 | TPM2uf+I | 0.3016 | 0.1653 | 0.1405 | 0.3927 | 9.4493 | 47.3177 | 9.4493 | 1.0000 | 47.3177 | 1.0000 | 0.7550 | Equal rates for all sites |
| Nested clade 4-3 | TPM1uf+I+G | 0.3110 | 0.1597 | 0.1409 | 0.3883 | 1.0000 | 885.2974 | 59.7997 | 59.7997 | 885.2974 | 1.0000 | 0.6660 | 0.2320 |
| Nested clade 4-4 | TPM2uf+G | 0.3014 | 0.1663 | 0.1426 | 0.3897 | 172.1505 | 1175.5851 | 172.1505 | 1.0000 | 1175.5851 | 1.0000 | - | 0.1560 |
| Nested clade 4-5 | HKY+I+G | 0.3095 | 0.1672 | 0.1462 | 0.3772 | Transitions/ Transversions = 14.0814 | | | | 8.9150 | 1.0000 | 0.8540 | 0.4590 |
| Nested clade 4-6 | TPM2uf+G | 0.3040 | 0.1655 | 0.1427 | 0.3878 | 216.8281 | 1008.4432 | 216.8281 | 1.0000 | 1008.4432 | 1.0000 | - | 0.1210 |



Clade 4-3

Clade 4-2b

Clade 4-5b

Clade 4-1

Clade 4-2a

Clade 4-6

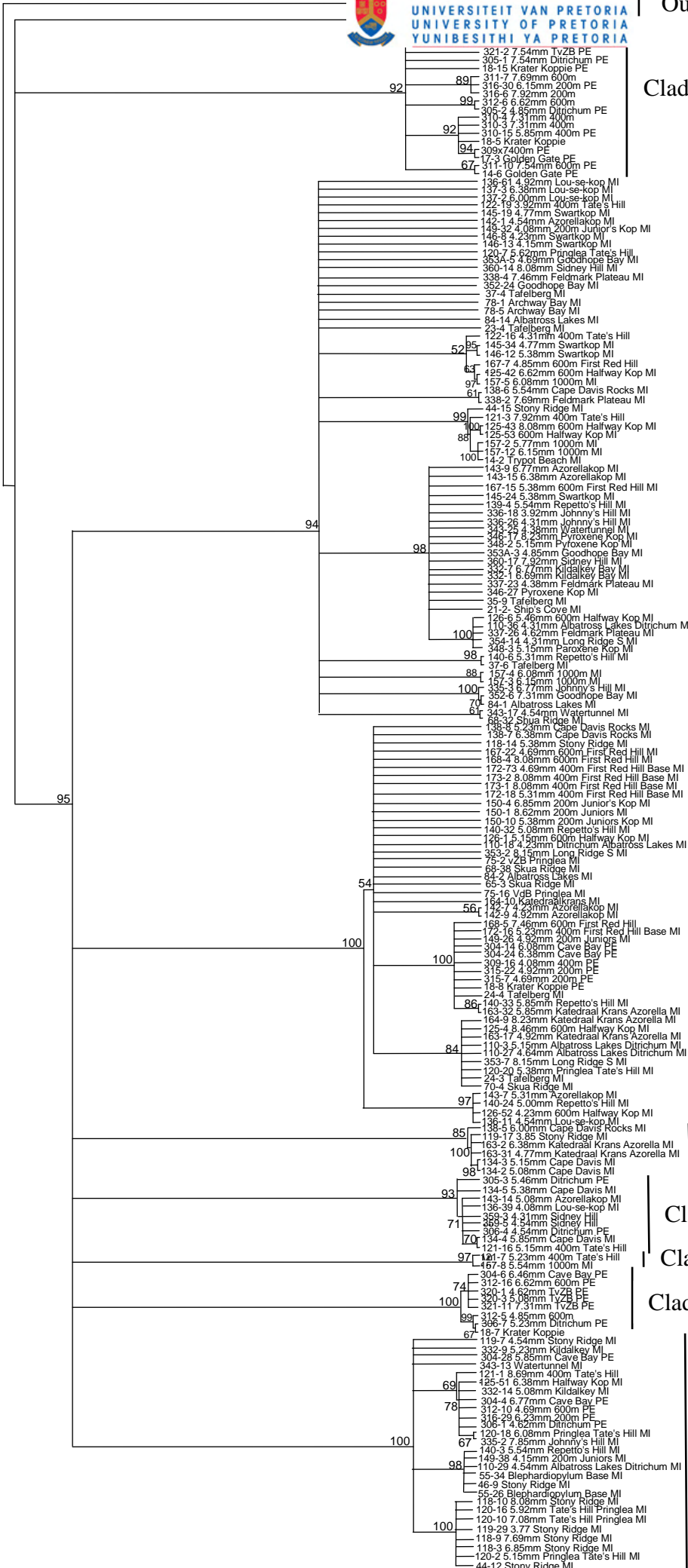
Clade 4-5a

Clade 4-4

Outgroup

Fig. 15 Minimum Evolution (ME) tree inferred using the Tamura-Nei distance correction algorithm. In each case, the sample number is given followed by the body length measurement, vegetation type it was collected from, locality it was collected from, sex and island of origin. ‘MI’ denotes Marion Island samples whilst Prince Edward Island samples are indicated by ‘PEI’. Nodal support obtained following 100 000 bootstrap replications above 50% is indicated. TvZB indicates samples collected at the top of Van Zinderen Bakker Peak at an elevation of 672 m above sea level. (On page 109)

Fig. 16 Bayesian phylogenetic analyses. For each taxon, the sample number is given followed by the body length measurement, vegetation type it was collected from, locality it was collected from, sex and island of origin. ‘MI’ denotes Marion Island samples whilst Prince Edward Island samples are indicated by ‘PEI’. TvZB indicates samples collected at the top of Van Zinderen Bakker Peak at an elevation of 672 m above sea level. (On page 111)



Clade 4-4

Clade 4-5a

Clade 4-3

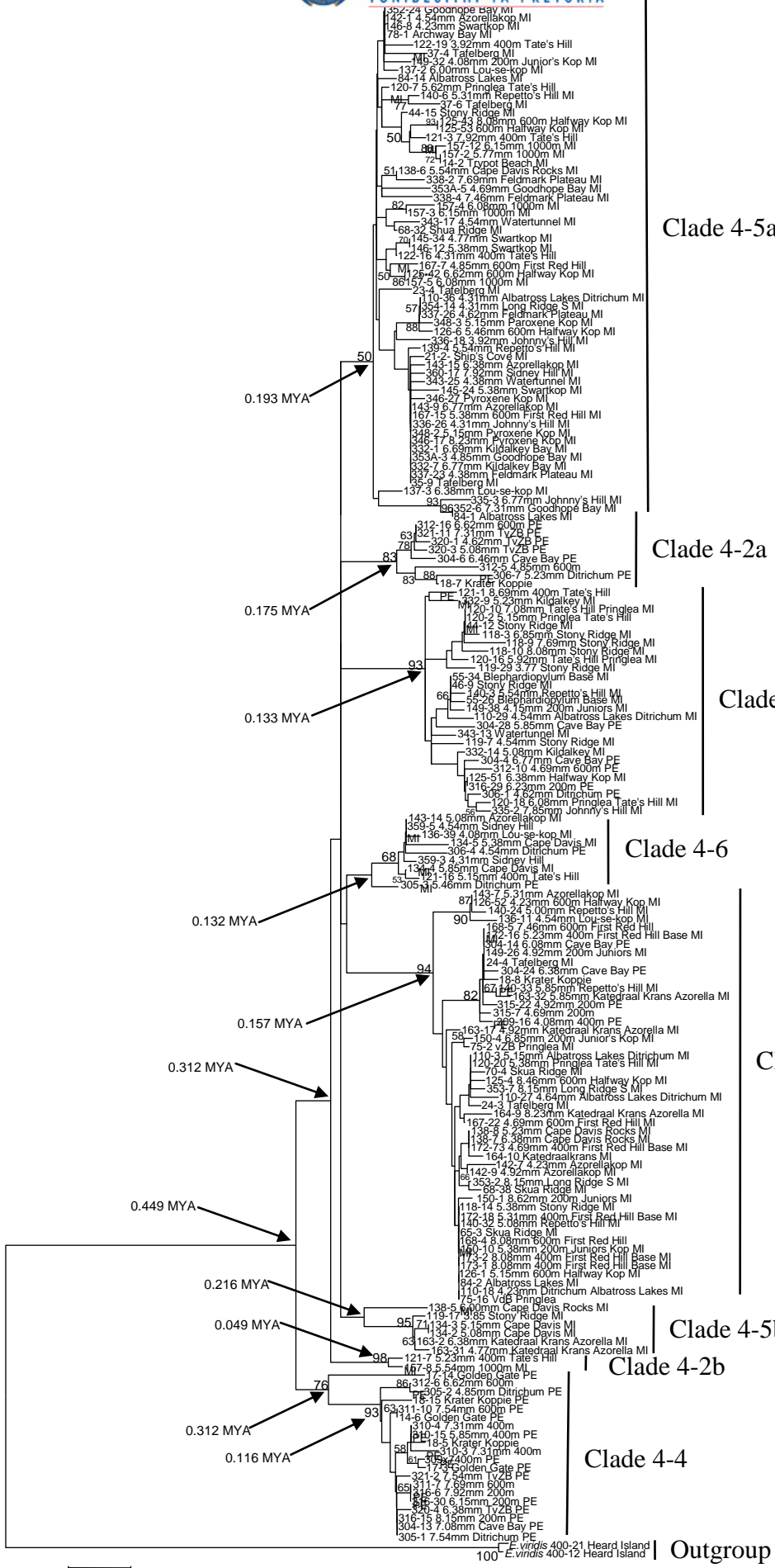
Clade 4-5b

Clade 4-6

Clade 4-2b

Clade 4-2a

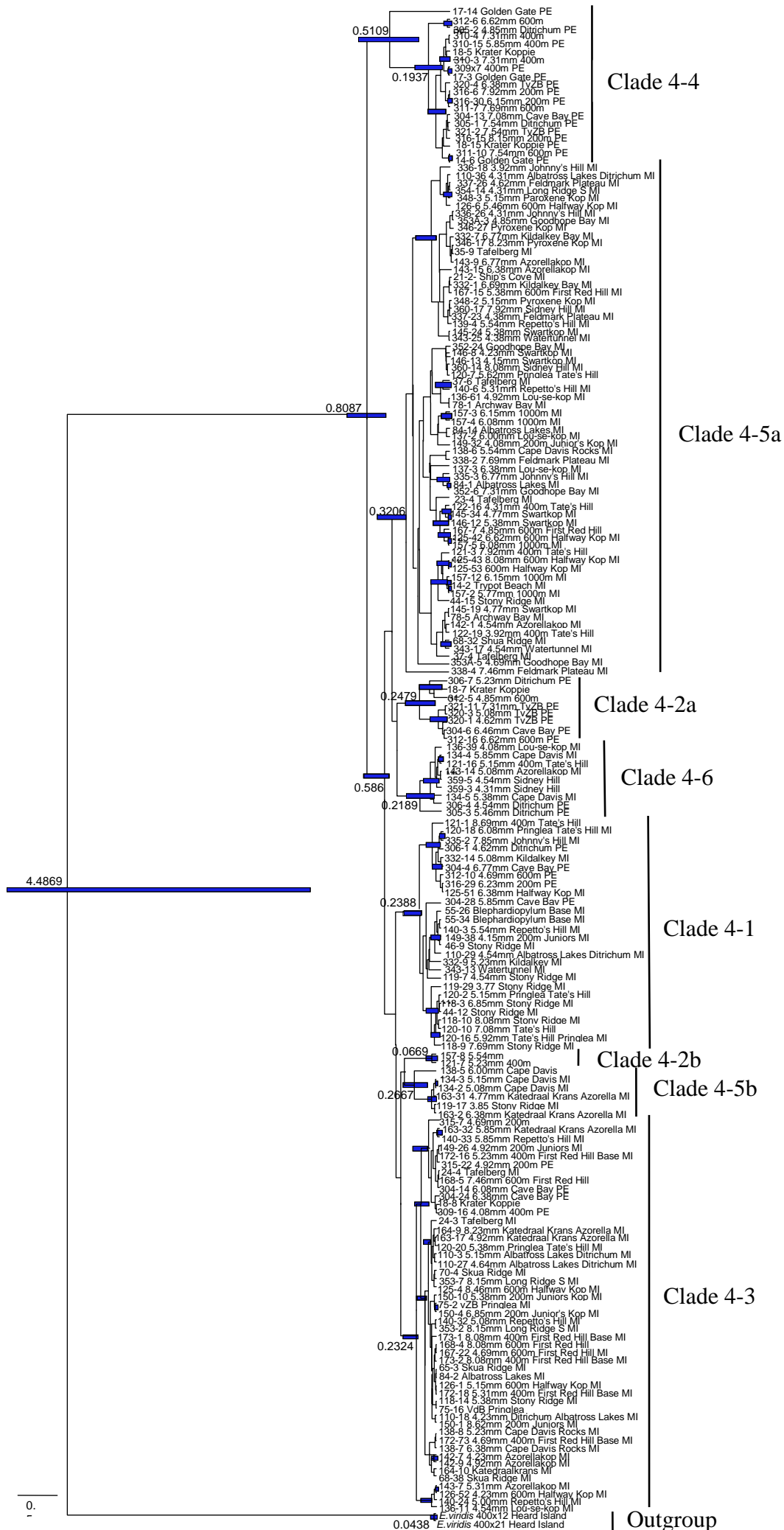
Clade 4-1



0.005

Fig. 17 Neighbour Joining (NJ) tree inferred using uncorrected p-distance values. For each taxon, the sample number is followed by the body length measurement, vegetation type it was collected from, locality it was collected from, sex, and its island of origin. ‘MI’ denotes Marion Island samples whilst Prince Edward Island samples are indicated by ‘PEI’. Nodal support obtained following 100 000 bootstrap replications above 50% is indicated next to the relevant nodes. Time to lineage coalescence was estimated using a 2.3 % nucleotide sequence divergence per million years molecular clock calibration, based on the arthropod mtDNA survey of Brower (1994). Estimated time of coalescence is indicated by arrows. (On page 112)

Fig. 18 Ultrametric tree obtained with BEAST 1.6.1 (Drummond & Rambaut 2007) (On page 114). The numbers in the nodes correspond to the estimated age in MY. Blue bars correspond to the 95 % confidence interval. A clock rate of 2.3 % sequence divergence per million years (MY), based on the arthropod mtDNA survey of Brower (1994), was used. The BEAST 1.6.1 (Drummond & Rambaut, 2007) statistics as analyzed with Tracer v1.5 (Drummond & Rambaut, 2007) are summarised in Table 4.



A neighbour joining p-distance tree with *E. viridis* as outgroup is presented in Fig. 17. The molecular clock, with a clock rate of 2.3 % sequence divergence per million years (MY) based on the arthropod mtDNA survey of Brower (1994), was utilised to determine the time when the individuals in each 4-step clade last shared a common ancestor so as to give a rough indication of when each of the 4-step clades originated on the PEIA. A Neighbour-Joining p-distance tree (Fig. 17) as well as an ultrametric tree obtained with BEAST (Fig. 18) to estimate times of coalescence were utilised. The results obtained are presented in Table 5. All BEAST estimates are older than those estimated by Neighbour-Joining uncorrected p-distance values. In each case, clade 4-4, representing *E. kucheli* (Grobler *et al.*, 2006), exhibits to oldest coalescence times with nested clade 4-5, consisting of clades 4-5a and 4-5b, exhibiting the next oldest coalescence times. This supports the theory that nested clade 4-5 was the first to colonize Marion Island.

Table 4 BEAST 1.6.1 (Drummond & Rambaut, 2007) statistics as analyzed with Tracer v1.5 (Drummond & Rambaut, 2007).

| Statistic | Mean | ESS |
|----------------------|-----------|-----------|
| posterior | -3768.93 | 961.232 |
| prior | -401.868 | 288.1 |
| likelihood | -3367.062 | 1592.394 |
| treeModel.rootHeight | 4.593 | 13816.543 |
| tmrca(4-1) | 0.225 | 2918.564 |
| tmrca(4-2a) | 0.235 | 6186.949 |
| tmrca(4-2b) | 5.668E-2 | 11594.391 |
| tmrca(4-3) | 0.218 | 3832.055 |
| tmrca(4-5a) | 0.309 | 2837.015 |
| tmrca(4-5b) | 0.268 | 2842.287 |
| tmrca(4-6) | 0.21 | 4599.945 |
| tmrca(Outgroup) | 3.245E-2 | 4334.886 |
| tmrca(4-4) | 0.514 | 8447.389 |
| constant.popSize | 3.057 | 5423.428 |
| ac | 0.163 | 3522.169 |
| ag | 2.888 | 1715.709 |
| at | 8.15E-2 | 3099.201 |
| cg | 0.195 | 1447.834 |
| gt | 1.82E-2 | 2102.082 |
| alpha | 1.01 | 916.967 |
| pInv | 0.681 | 494.606 |
| clock.rate | 2.3E-2 | - |
| treeLikelihood | -3367.062 | 1592.394 |
| coalescent | -396.494 | 5675.378 |

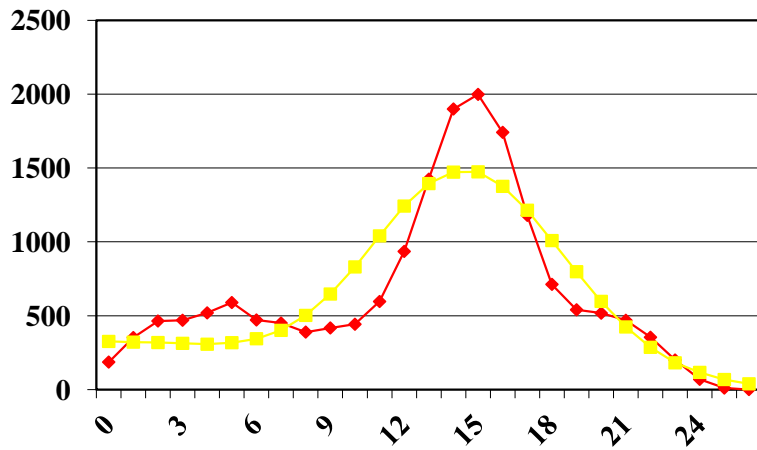
Table 5 Estimates of lineage coalescence for the major clades as determined by Neighbour-Joining (NJ) uncorrected p-distance values (Fig. 17) and BEAST (Fig. 18) using a clock rate of 2.3 % sequence divergence per million years (MY) based on the arthropod mtDNA survey of Brower (1994).

| Clade | NJ estimates | BEAST estimates |
|---------------------------------|---------------------|------------------------|
| Clade 4-1 | 0.133 MYA | 0.2388 MYA |
| Clade 4-2a | 0.175 MYA | 0.2479 MYA |
| Clade 4-2b | 0.049 MYA | 0.0669 MYA |
| Clade 4-3 | 0.157 MYA | 0.2324 MYA |
| Clade 4-4 (<i>E. kucheli</i>) | 0.312 MYA | 0.5109 MYA |
| Clade 4-5a | 0.193 MYA | 0.3206 MYA |
| Clade 4-5b | 0.216 MYA | 0.2667 MYA |
| Clade 4-6 | 0.132 MYA | 0.2189 MYA |

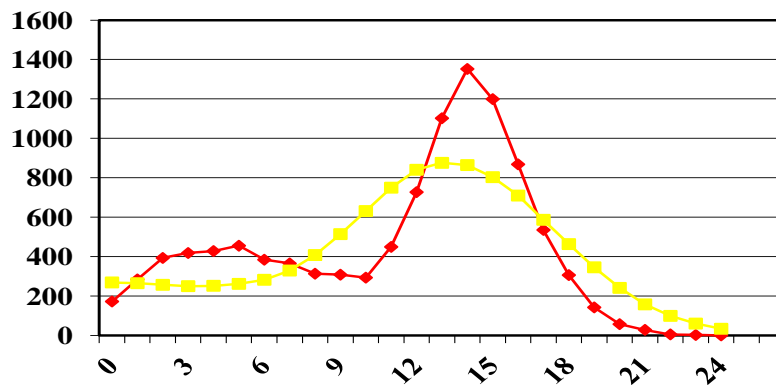
Mismatch analysis and population expansion

Fig. 19 a-j comprises of mismatch distributions under the sudden expansion model for the archipelago as a whole, each island on its own as well as the mismatch distributions of each of the eight discrete clades as identified in the phylogenetic trees (Figs. 15 & 16). Fig. 20 a-j, on the other hand, comprises the mismatch distributions under the spatial expansion model, of the archipelago as a whole, each island on its own as well as the mismatch distributions of each of the eight discrete clades as identified in the phylogenetic trees (Figs. 15 & 16). Both SSD and raggedness index values are shown in Table 6 under the sudden expansion model and in Table 7 under the spatial expansion model. None of these values shows any significant deviation from either the sudden expansion model or the spatial expansion model ($P > 0.05$), indicating a recent population expansion for each island as well as for each group on its own. In each case, demographic expansions predate range expansions, except in the cases of clades 4-2b and 4-6. The earliest demographic expansion for the different clade populations is estimated to have occurred more than 100 000 years ago (clades 4-1, 4-3 & 4-5a) with the most recent demographic expansion occurring more than 40 000 years ago (clades 4-b & 4-6). The earliest range expansions occurred approximately 170 000 years ago in clade 4-2b and the most recent approximately 40 000 years ago in clade 4-5b. Under the assumption of a purely demographic expansion, the populations represented by each clade were inferred to have expanded from small effective female population sizes to large effective female population sizes. Clades 4-5a & 4-6 needed larger initial effective female population sizes in order to explain demographic expansion. Under the assumption of a spatial expansion, populations represented in clades 4-5a, 4-3 & 4-2a required large effective female population sizes and no genetic exchange with neighboring populations to explain the inferred spatial expansion, while the population that is represented by clade 4-6 required a large effective female population size and very little genetic exchange with neighboring populations. The population represented by clade 4-5b required a small effective female population size and extensive genetic exchange with neighboring populations to explain the inferred spatial expansion. This is presumably due to individual 138-5 that may be from another population that was poorly sampled. Clade 4-5b also nested with clade 4-5a. Clade 4-4 also required a degree of genetic exchange with neighboring populations to explain the inferred spatial expansion, and this may possibly also be explained by individual 17-14 that is from another poorly sampled population.

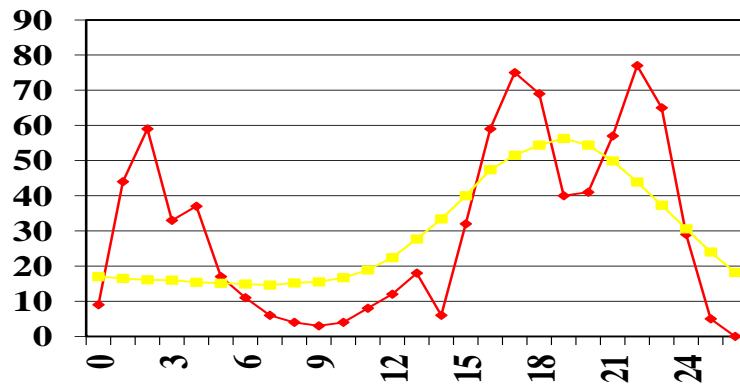
(a)



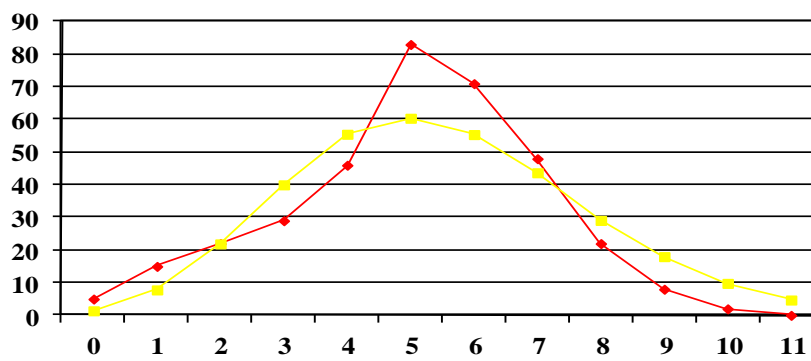
(b)



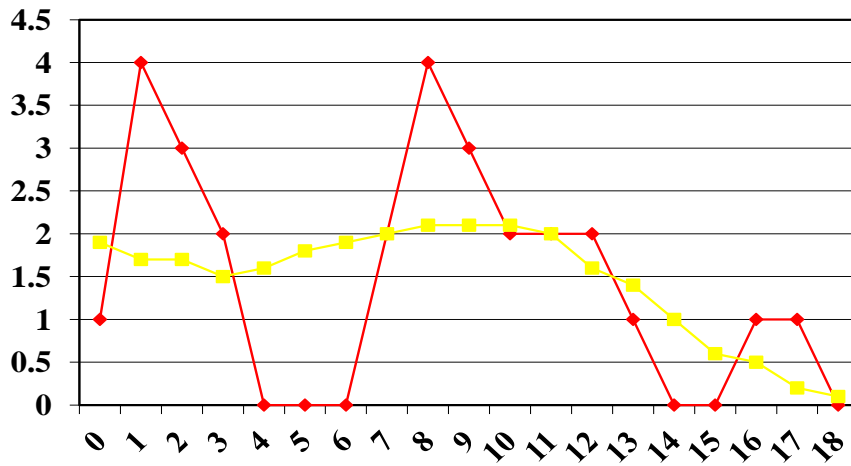
(c)



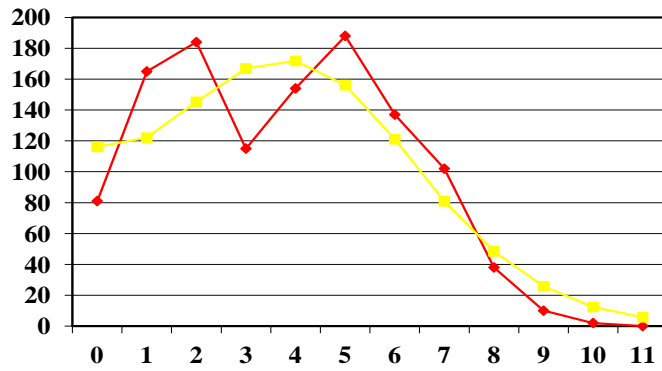
(d)



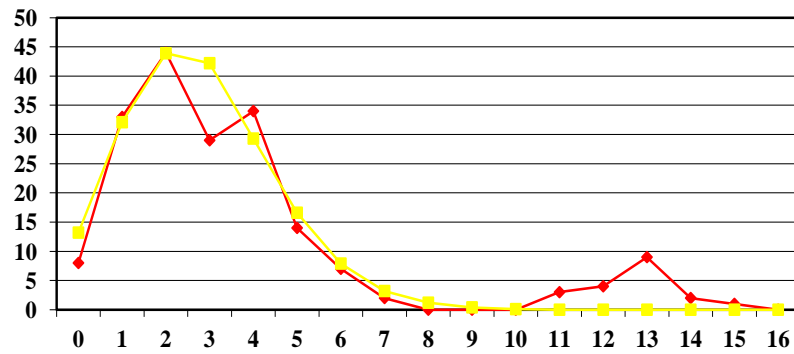
(e)



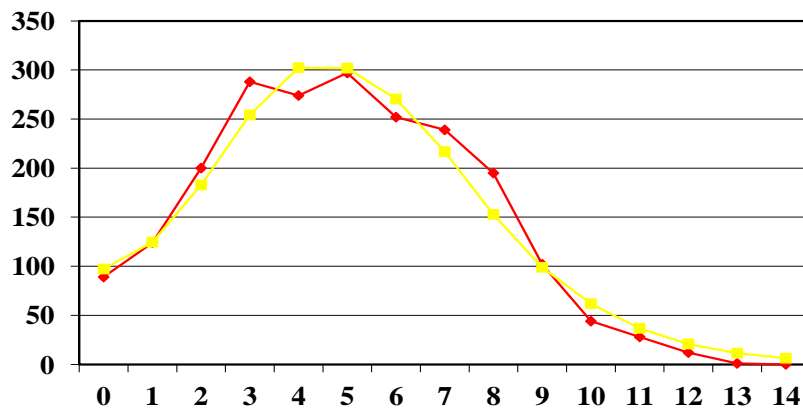
(f)



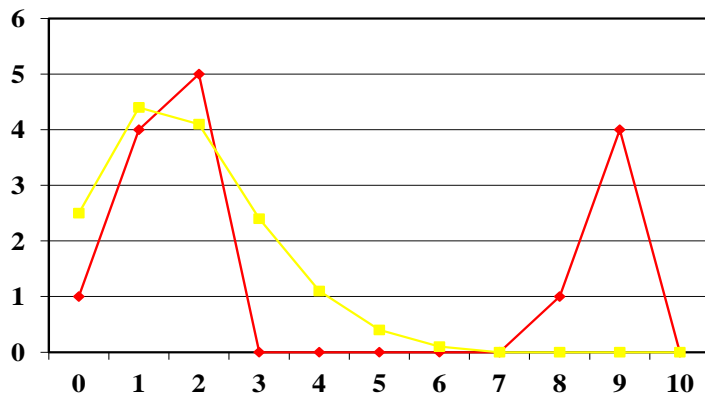
(g)



(h)



(i)



(j)

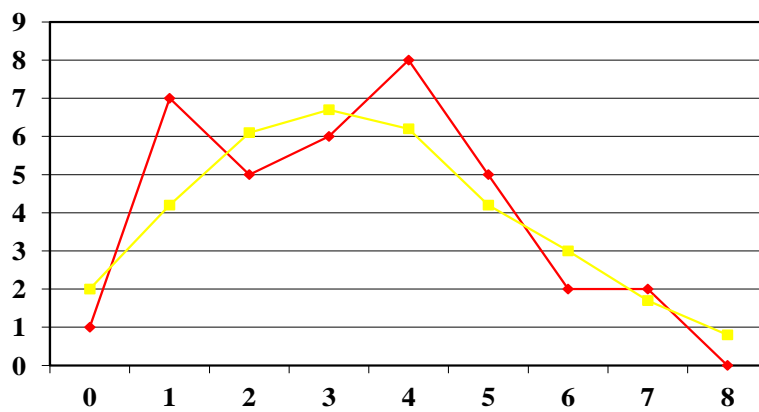
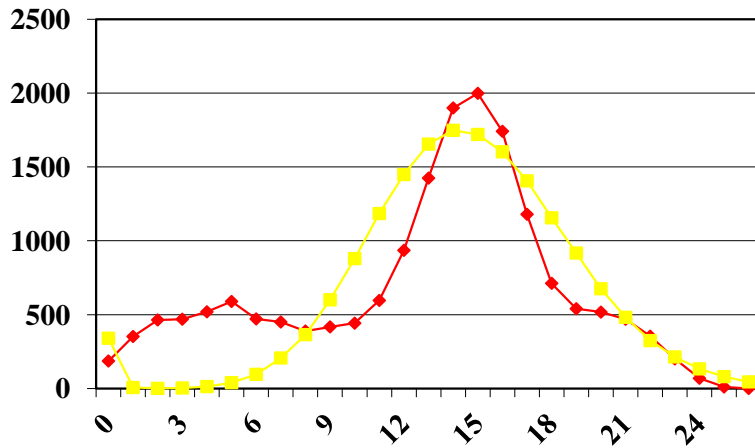


Fig. 19 Mismatch distribution estimated under the sudden expansion model for the different data sets: a) Prince Edward Islands Archipelago (PEIA), b) Marion Island (MI), c) Prince Edward Island (PEI), d) nested clade 4-1, e) clade 4-2a, f) nested clade 4-3, g) nested clade 4-4, h) clade 4-5a and i) clade 4-5b j) nested clade 4-6. The yellow line represents the simulated stepwise expansion model and the red line represents the observed data. No mismatch analyses were possible for clade 4-2b because of a small sample size.

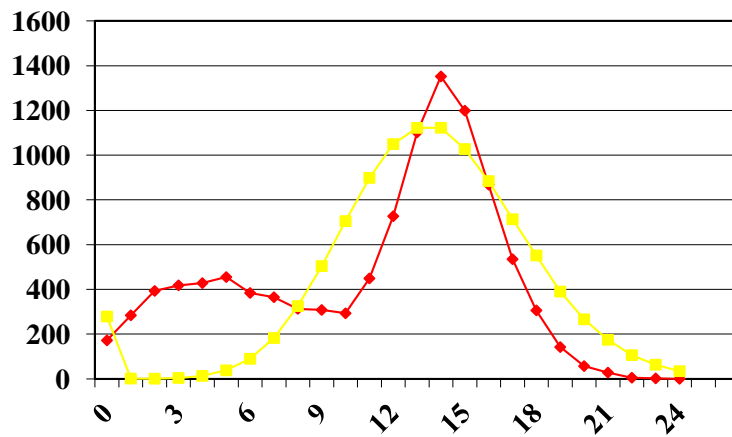
Table 6 Mismatch parameters estimated under the sudden expansion model: τ is the expansion parameter or date of the Growth or Decline measured in units of mutational time with $\tau = 2ut$ where u is the mutation rate and t is the time in generations. The timing of the most important expansion in each group ($t_{\text{divergence}}$) was calculated based on the equation $\tau = 2\mu t$. θ_0 and θ_1 are the substitution rates before and after expansion. SSD is the test of the validity of a stepwise expansion model based on the sum of the square deviations between the observed and the expected mismatch. The raggedness index quantifies the smoothness of the observed pairwise differences distribution. PEIA = Prince Edward Islands Archipelago; PEI = Prince Edward Island; MI = Marion Island. The significance of both SSD and the raggedness index are estimated with a parametric bootstrap approach (probability values: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$).

| Population | Mean number of differences | Mismatch observed variance | τ (alpha = 0.100 confidence values) | $t_{\text{divergence}}$ (years) | $\theta_0 / N_{f_t=0}$ | $\theta_1 / N_{f_t=0}$ | SSD | P(Sim. SSD \geq Obs. SSD) | Raggedness index | P(Sim. Rag. \geq Obs. Rag.) |
|------------|----------------------------|----------------------------|--|---------------------------------|------------------------|-------------------------------|-----------|-----------------------------|------------------|-------------------------------|
| PEIA | 12.838 | 30.069 | 15.448 (10.840-18.299) | 379464 (266273-449496) | 0.004/ 98 | 52.266/ 1283861 | 0.0049468 | 0.179 | 0.00453067 | 0.367 |
| PEI | 14.609 | 59.701 | 20.406 (13.926-24.506) | 501252 (342078-601965) | 0.002/ 49 | 48.091/ 1181306 | 0.0127341 | 0.095 | 0.01198245 | 0.297 |
| MI | 11.087 | 25.953 | 14.458 (9.697-17.828) | 355146 (238197-437926) | 0.002/ 49 | 38.125/ 936502 | 0.0088584 | 0.131 | 0.00598689 | 0.414 |
| 4-1 | 5.057 | 3.894 | 5.494 (4.232-6.676) | 134954 (103954-163989) | 0.000/ 0 | >99999.0/ >2 $\times 10^9$ | 0.0103789 | 0.112 | 0.02792997 | 0.177 |
| 4-2a | 7.179 | 23.411 | 10.586 (4.113-15.930) | 260034 (101031-391304) | 0.000/ 0 | 13.398/ 329108 | 0.0290171 | 0.781 | 0.03698980 | 0.941 |
| 4-2b | 2.000 | 0.0000 | - | - | - | - | - | - | - | - |
| 4-3 | 3.728 | 5.216 | 4.984 (2.090-7.809) | 122426 (51338-191820) | 0.032/ 786 | 9.736/ 239154 | 0.0071099 | 0.599 | 0.01708562 | 0.822 |
| 4-4 | 3.742 | 11.272 | 2.867 (1.539-3.842) | 70424 (37803-94374) | 0.002/ 49 | 69.531/ 1707958 | 0.0090510 | 0.313 | 0.04321330 | 0.465 |
| 4-5a | 4.912 | 6.891 | 4.570 (2.711-8.467) | 112257 (66592-207983) | 1.106/ 27167 | 23.008/ 565168 | 0.0012460 | 0.865 | 0.00700866 | 0.950 |
| 4-5b | 3.867 | 13.410 | 1.771(0.293-3.455) | 43502 (7197-84868) | 0.000/ 0 | >99999.0/ >2 $\times 10^9$ | 0.1221409 | 0.119 | 0.27111111 | 0.252 |
| 4-6 | 3.278 | 3.349 | 3.033 (1.330-6.082) | 47502 (32670-149398) | 0.643/ 15794 | 31.055/ 762834 | 0.0126028 | 0.643 | 0.05169753 | 0.721 |

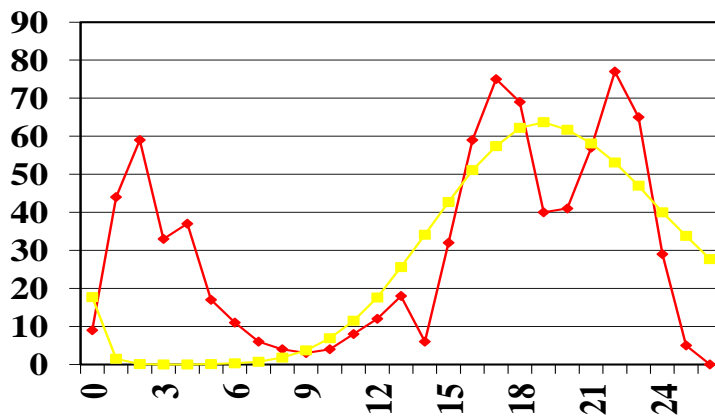
(a)



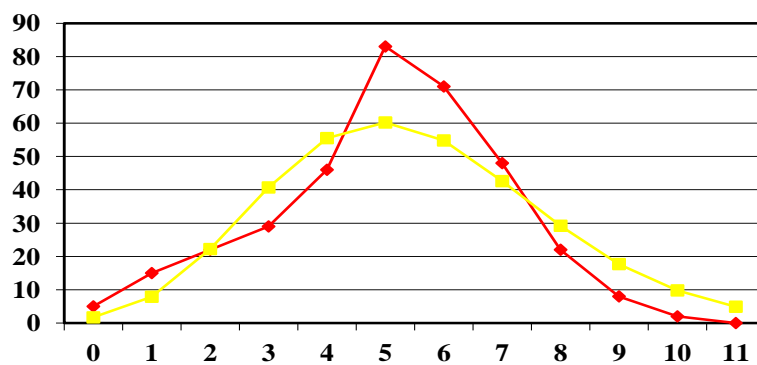
(b)



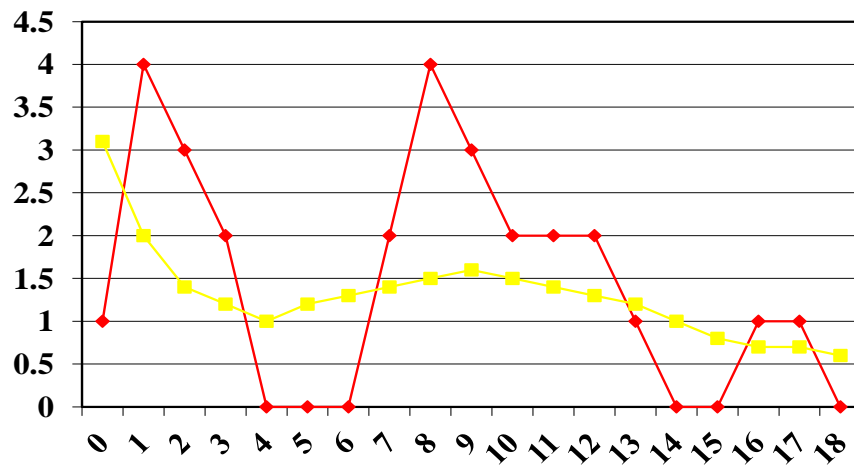
(c)



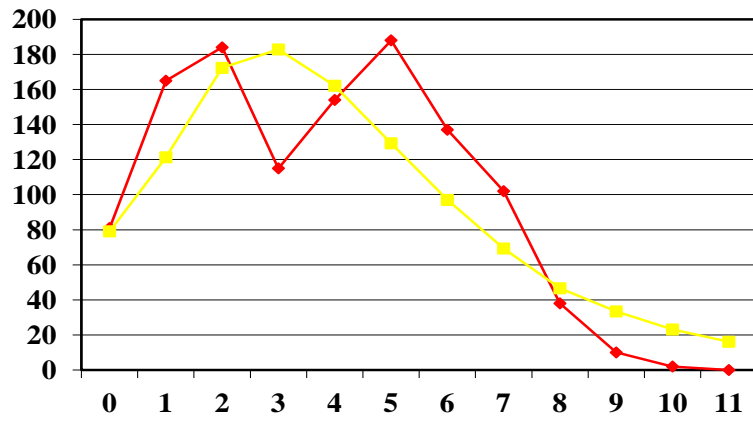
(d)



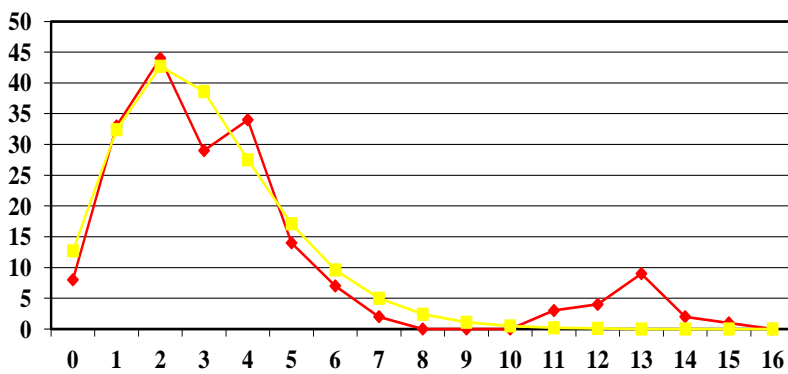
(e)



(f)



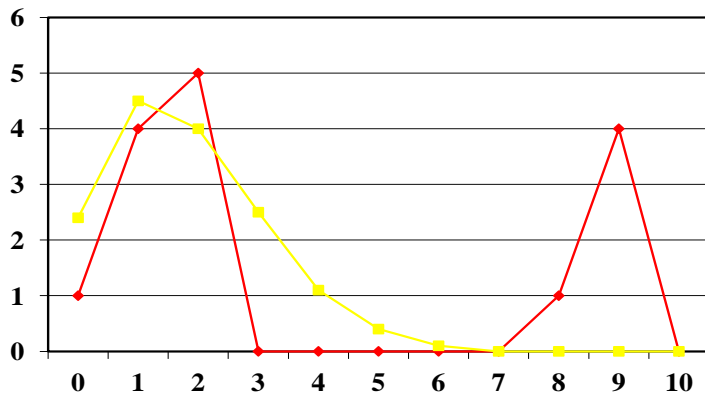
(g)



(h)



(i)



(j)



Fig. 20 Mismatch distribution estimated under the spatial expansion model for the different data sets: a) Prince Edward Islands Archipelago (PEIA), b) Marion Island (MI), c) Prince Edward Island (PEI), d) nested clade 4-1, e) clade 4-2a, f) nested clade 4-3, g) nested clade 4-4, h) clade 4-5a and i) clade 4-5b j) nested clade 4-6. The yellow line represents the simulated stepwise expansion model and the red line represents the observed data. No mismatch analyses were possible for clade 4-2b because of a small sample size.

Table 7 Mismatch parameters estimated under the spatial expansion model: τ is the expansion parameter or date of the Growth or Decline measured in units of mutational time with $\tau = 2ut$ where u is the mutation rate, and t is the time in generations. The timing of the most important expansion in each group ($t_{\text{divergence}}$) was calculated based on the equation $\tau = 2\mu t$. θ is ancestral population size and $M = 2Nm$ is an indication of migration. m is the rate at which the sampled deme would exchange migrants with a unique population of infinite size. The effective female population size $N_f = \theta/2u$. SSD is the test of the validity of a stepwise expansion model based on the sum of the square deviations between the observed and the expected mismatch. The raggedness index quantifies the smoothness of the observed pairwise differences distribution. PEIA = Prince Edward Islands Archipelago; PEI = Prince Edward Island; MI = Marion Island. The significance of both SSD and the raggedness index are estimated with a parametric bootstrap approach (probability values: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$).

| Population | Mean number of differences | Mismatch observed variance | τ (alpha = 0.100 confidence values) | $t_{\text{divergence}}$ (years) | θ / N_f | M / m | SSD | P(Sim. SSD \geq Obs. SSD) | Raggedness index | P(Sim. Rag. \geq Obs. Rag.) |
|------------|----------------------------|----------------------------|--|---------------------------------|---------------------|----------------------|-------|-----------------------------|------------------|-------------------------------|
| PEIA | 12.838 | 30.069 | 14.238 (11.912-16.261) | 349742 (292606-399435) | 0.884/ 21714.566 | 49.775/ 0.0011 | 0.009 | 0.007** | 0.00453067 | 0.610 |
| PEI | 14.609 | 59.701 | 17.413 (14.268-21.705) | 427732 (350478-533161) | 3.507/ 86145.910 | 40.687/ 0.0002 | 0.019 | 0.000*** | 0.01198245 | 0.475 |
| MI | 11.087 | 25.953 | 13.889 (11.028-15.340) | 341169 (270891-376811) | 0.157/ 3856.546 | 37.661/ 0.0049 | 0.015 | 0.004** | 0.00598689 | 0.685 |
| 4-1 | 5.057 | 3.894 | 5.494 (3.301-6.565) | 134954 (81085-161262) | 0.001/ 24.564 | 1003.383/ 20.4238 | 0.010 | 0.095 | 0.02792997 | 0.200 |
| 4-2a | 7.179 | 23.411 | 6.946 (0.741-20.289) | 170621 (18201-498378) | 6.053/ 148685.83 | 2.053/ 0.000006 | 0.034 | 0.718 | 0.03698990 | 0.980 |
| 4-2b | 2.000 | 0.0000 | - | - | - | - | - | - | - | - |
| 4-3 | 3.728 | 5.216 | 1.964 (0.810-5.255) | 48243 (19896-129083) | 2.608/ 64062.884 | 28.924/ 0.0002 | 0.011 | 0.271 | 0.01708562 | 0.740 |
| 4-4 | 3.742 | 11.272 | 2.025 (1.284-4.047) | 49742 (31540-99410) | 0.034/ 835.176 | 99999/ 59.867 | 0.008 | 0.339 | 0.04321330 | 0.473 |
| 4-5a | 4.912 | 6.891 | 3.562 (2.332-6.298) | 87496 (57283-154704) | 1.923/ 47236.551 | 28.069/ 0.0003 | 0.002 | 0.747 | 0.00700866 | 0.966 |
| 4-5b | 3.867 | 13.410 | 1.770 (0.732-3.267) | 43478 (17980-80250) | 0.001/ 24.564 | 99999/ 2035.479 | 0.122 | 0.057 | 0.27111111 | 0.222 |
| 4-6 | 3.278 | 3.349 | 2.414 (1.214-5.125) | 59297 (29820-125890) | 1.215/ 29845.247 | 99999/ 1.675 | 0.013 | 0.590 | 0.05169753 | 0.732 |

Table 8 Population expansion statistics showing Fu's F_s values (probability values: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$), Tajima's D as well as Tajima's estimations for θ . PEIA = Prince Edward Islands Archipelago; PEI = Prince Edward Island; MI = Marion Island.

| Population | Fu's F_s test | P(sim_ F_s \leq obs_ F_s) | Fu's estimations for θ | Tajima's D | P(sim_ D \leq obs_ D) | Tajima's estimations for θ |
|---------------|-----------------|----------------------------------|-------------------------------|--------------|------------------------------|-----------------------------------|
| Ectemnorhinus | | | | | | |
| PEIA | -23.98497*** | 0.00100 | 12.83831 | -1.38277 | 0.04600 | 12.83831 |
| PE | -13.56464*** | 0.00100 | 14.60854 | -0.33924 | 0.44300 | 14.60854 |
| Marion | -24.22080*** | 0.00100 | 11.08654 | -1.40322 | 0.04300 | 11.08654 |
| 4-1 | -17.73110*** | 0.00000 | 5.05698 | -1.37102 | 0.06600 | 5.05698 |
| 4-2a | -1.11105 | 0.21200 | 7.17857 | -1.49761 | 0.05900 | 7.17857 |
| 4-2b | 0.69315 | 0.36800 | 2 | 0 | 1 | 2 |
| 4-3 | -24.46672*** | 0.00000 | 3.72789 | -1.55217 | 0.04200 | 3.72789 |
| 4-4 | -6.90308*** | 0.00100 | 3.74211 | -1.62575 | 0.04000 | 3.74211 |
| 4-5a | -25.57643*** | 0.00000 | 4.91235 | -2.10023 | 0.00200 | 4.91235 |
| 4-5b | -0.80758 | 0.22100 | 3.86667 | -1.19369 | 0.13500 | 3.86667 |
| 4-6 | -4.03444** | 0.01000 | 3.27778 | -1.74560 | 0.02700 | 3.27778 |

Fu's F_s statistic (Table 8) showed significant large negative values for most groups. Only clades 4-2a, 4-2b and 4-5b have no significant large negative values. A significant large negative F_s value is an indication of population expansion. The F_s values for clades 4-2a and 4-5b are still negative and thus show a tendency for population growth (Fu, 1997). The fact that clades 4-2a and 4-5b do not show significant F_s values may be due to some individuals in these clades being separated by up to seven missing haplotypes (Fig. 6). Clade 4-2b is the only clade with a positive F_s value, however as it comprises of only two individuals, it is not possible to compute any reliable statistics. Tajima's (1983) estimate for θ (Table 8) is based on the calculation of the mean number of pairwise differences of the sequences. Tajima's estimate of θ puts more weight on ancient mutations, therefore reflecting ancient population events (Fu, 1997).

Statistical analyses

Haplotype diversity (h) and nucleotide diversity (π) values for the PEIA as a whole, each island, and the different discrete clades as identified in the phylogenetic trees (Fig. 15 and Fig. 16) are summarised in Table 9. In every case, h is very close to 1. The number of haplotypes and parsimony informative sites are also summarised in Table 9. The number of haplotypes in the different clades as identified in the phylogenetic trees (Figs. 15 & 16) varied between 2 and 42, while the number of parsimony informative sites varied between 1 and 31.

Table 10 summarizes the population pairwise distance F_{ST} values as computed in Arlequin version 3.5.1.2 (Excoffier & Lischer, 2010) as well as the absolute number of migrants exchanged between two populations. This was computed between each of the eight discrete clades as identified in the phylogenetic trees (Figs. 15 & 16). F_{ST} assesses the variation in the subpopulations relative to that in the total population and has values between 0 and 1.0. An F_{ST} of zero means that all the subpopulations have the same gene frequencies while an F_{ST} of 1.0 means that the subpopulations have completely non-overlapping sets of alleles. Natural panmictic populations tend to have F_{ST} values that range between near zero up to just greater than 0.5. Values of F_{ST} above approximately 0.2 are considered "high". Consequently, all the F_{ST} values in Table 10 can be considered "high" with the lowest value being 0.578 between genetically distinct clades 4-5a and 4-6. The only other F_{ST} values below 0.6 were those between genetically distinct clades 4-5a and 4-5b ($F_{ST} = 0.580$) and between genetically distinct clades 4-2a and 4-5a ($F_{ST} = 0.583$). In each case, the number of migrants exchanged between the distinct clades was less than one migrant per generation

with the estimate of 0.366 migrants per generation, between clades 4-5a and 4-6, being the highest estimate obtained. Thus, the higher the F_{ST} value the lower the migration rate. The Hudson *et al.*'s (1992) F_{ST} values as estimated in DNASP 5.1 (Librado & Rozas, 2009) and presented in Table 11 show a similar trend to the F_{ST} values presented in Table 10. All Hudson *et al.*'s (1992) F_{ST} values can again be considered high with the lowest F_{ST} values observed between clades 4-2a and 4-5a ($F_{ST} = 0.56502$) and between clades 4-5a and 4-5b ($F_{ST} = 0.59530$). The N_{ST} values in Table 12 (F_{ST} to which the correction of Jukes & Cantor (1969) has been applied) almost mirror the F_{ST} values in Table 11, with the differences being negligible. The number of pairwise differences between populations as presented in Table 13 reveal that the number of pairwise differences between populations correlates with the F_{ST} values found in the previous tables: The lower the F_{ST} values between two genetically distinct clades, the lower the number of pairwise differences between them. This is to be expected as fewer differences show a closer resemblance between populations whereas high F_{ST} values are an indication of differentiation between populations. Table 14 shows that the variation between the genetically distinct clades is more than double the variation within the clades. The overall fixation index of 0.73 is also very high. All F_{ST} values suggest that there are a low number of overlapping alleles between clades with less than one migrant between clades per generation. Given that these distinct clades overlap in geographical range, it may be possible that there is either some barrier that prevents interbreeding between individuals from different genetically identified clades, or that insufficient time has passed to eradicate genetic signatures of past isolation coupled with differentiation. Table 15 summarizes the tau values as well as their P values as calculated by Arlequin version 3.5.1.2 (Excoffier & Lischer, 2010) for divergence times between the different genetically identified clades. None of the tau values have P values under the 0.05 confidence level. Divergence times between clade 4-4, representing *E. kucheli* (Grobler *et al.*, 2006), and the other clades are in each instance the oldest.

Table 9 Summary of mitochondrial DNA diversity. NH = number of unique haplotypes, h = haplotype diversity and π = nucleotide diversity. PEIA = Prince Edward Islands Archipelago; PEI = Prince Edward Island; MI = Marion Island.

| Dataset (N) | NH | h | π | Parsimony informative sites |
|-----------------------------|-----|--------|---------|-----------------------------|
| PEIA (<i>n</i> = 187) | 131 | 0.9893 | 0.01451 | 84 |
| MI (<i>n</i> = 146) | 99 | 0.984 | 0.01253 | 69 |
| PEI (<i>n</i> = 41) | 34 | 0.989 | 0.01651 | 43 |
| Clade 4-1 (<i>n</i> = 27) | 23 | 0.986 | 0.00571 | 11 |
| Clade 4-2a (<i>n</i> = 8) | 7 | 0.964 | 0.00577 | 4 |
| Clade 4-2b (<i>n</i> = 2) | 2 | 1 | 0.00226 | 0 |
| Clade 4-3 (<i>n</i> = 49) | 30 | 0.931 | 0.00421 | 14 |
| Clade 4-4 (<i>n</i> = 20) | 14 | 0.958 | 0.00423 | 9 |
| Clade 4-5a (<i>n</i> = 66) | 42 | 0.959 | 0.00552 | 31 |
| Clade 4-5b (<i>n</i> = 6) | 5 | 0.933 | 0.00437 | 1 |
| Clade 4-6 (<i>n</i> = 9) | 8 | 0.972 | 0.00370 | 1 |

Table 10 Population pairwise distance F_{ST} values as computed in Arlequin version 3.5.1.2 (Excoffier & Lischer, 2010) using Tamura Nei (1993) distances and gamma as estimated by jModelTest 0.1.1 (Posada, 2008; Guindon & Gascuel, 2003). Pairwise F_{ST} values can be used as short-term genetic distances between populations (Reynolds, *et al.*; 1983; Slatkin, 1995). All P values estimated for F_{ST} values are significant at the 0.05 level. The value after the “/” is the M value (Slatkin, 1991) obtained in Arlequin version 3.5.1.2 (Excoffier & Lischer, 2010) where M is the absolute number of migrants exchanged between two populations.

| | 4-2a | 4-2b | 4-3 | 4-4 | 4-5a | 4-5b | 4-6 |
|------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 4-1 | 0.70819 / 0.20602 | 0.71085 / 0.20339 | 0.76985 / 0.14948 | 0.79697 / 0.12737 | 0.67993 / 0.23537 | 0.70290 / 0.21134 | 0.70031 / 0.21397 |
| 4-2a | | 0.62887 / 0.29508 | 0.78374 / 0.13797 | 0.78003 / 0.14100 | 0.58263 / 0.35818 | 0.67785 / 0.23763 | 0.64613 / 0.27384 |
| 4-2b | | | 0.74562 / 0.17058 | 0.78705 / 0.13529 | 0.60757 / 0.32295 | 0.70664 / 0.20757 | 0.72504 / 0.18962 |
| 4-3 | | | | 0.84003 / 0.09522 | 0.71118 / 0.20305 | 0.76549 / 0.15317 | 0.75907 / 0.15870 |
| 4-4 | | | | | 0.71450 / 0.19979 | 0.79462 / 0.12923 | 0.80109 / 0.12415 |
| 4-5a | | | | | | 0.57977 / 0.36241 | 0.57758 / 0.36567 |
| 4-5b | | | | | | | 0.68816 / 0.22657 |

Table 11 Hudson *et al.*'s (1992) F_{ST} values as estimated in DNASP 5.1 (Librado & Rozas, 2009). An average F_{ST} value of 0.72745 was estimated as well as an average Nm (Number of migrants) of 0.09. Nm is based on the island model of population structure: $F_{ST} = 1 / (1 + 2Nm)$ (Wright, 1951).

| | 4-2a | 4-2b | 4-3 | 4-4 | 4-5a | 4-5b | 4-6 |
|------|---------|---------|---------|---------|---------|---------|---------|
| 4-1 | 0.68663 | 0.77317 | 0.74400 | 0.78477 | 0.65725 | 0.70255 | 0.70519 |
| 4-2a | | 0.70232 | 0.74327 | 0.75402 | 0.56502 | 0.67169 | 0.63053 |
| 4-2b | | | 0.78770 | 0.82056 | 0.69697 | 0.74857 | 0.75515 |
| 4-3 | | | | 0.82590 | 0.70037 | 0.74500 | 0.75371 |
| 4-4 | | | | | 0.71837 | 0.78261 | 0.79474 |
| 4-5a | | | | | | 0.59430 | 0.60340 |
| 4-5b | | | | | | | 0.66741 |

Table 12 Lynch and Crease's (1990) N_{ST} / Number of migrants (N_m) as estimated in DNASP 5.1 (Librado & Rozas, 2009). An average N_{ST} value of 0.72941 was estimated as well as an average N_m of 0.09. N_m is based on the island model of population structure: $N_{st} = 1 / (1 + 2N_m)$ (Wright, 1951).

| | | | | | | | |
|------|---------|---------|---------|---------|---------|---------|---------|
| | 4-2a | 4-2b | 4-3 | 4-4 | 4-5a | 4-5b | 4-6 |
| 4-1 | 0.68902 | 0.77505 | 0.74630 | 0.78712 | 0.65951 | 0.70452 | 0.70724 |
| 4-2a | | 0.70382 | 0.74548 | 0.75615 | 0.56682 | 0.67339 | 0.63219 |
| 4-2b | | | 0.78924 | 0.82205 | 0.69848 | 0.74978 | 0.75654 |
| 4-3 | | | | 0.82797 | 0.70238 | 0.74672 | 0.75553 |
| 4-4 | | | | | 0.72031 | 0.78437 | 0.79656 |
| 4-5a | | | | | | 0.59564 | 0.60496 |
| 4-5b | | | | | | | 0.66869 |

Table 13 Above diagonal: Average number of pairwise differences between genetically distinct clades ($PiXY$). Diagonal elements (indicated with grey shading): Average number of pairwise differences within genetically distinct clades (PiX). Below diagonal: Corrected average pairwise difference ($(PiXY - (PiX + PiY)) / 2$). All values were calculated in Arlequin version 3.5.1.2 (Excoffier & Lischer, 2010) according to Nei & Li (1979) using Tamura Nei (1993) distances and gamma as estimated by jModelTest 0.1.1 (Posada, 2008; Guindon & Gascuel, 2003).

| | | | | | | | | |
|------|--------|--------|--------|--------|--------|--------|--------|--------|
| | 4-1 | 4-2a | 4-2b | 4-3 | 4-4 | 4-5a | 4-5b | 4-6 |
| 4-1 | 5.270 | 18.179 | 17.098 | 19.121 | 23.111 | 16.19 | 16.816 | 15.675 |
| 4-2a | 12.847 | 5.393 | 12.863 | 19.240 | 19.992 | 12.461 | 15.043 | 12.263 |
| 4-2b | 13.433 | 9.136 | 2.061 | 14.606 | 17.689 | 12.242 | 12.640 | 11.541 |
| 4-3 | 14.556 | 14.613 | 11.646 | 3.860 | 24.300 | 15.819 | 16.663 | 15.589 |
| 4-4 | 18.503 | 15.321 | 14.685 | 20.397 | 3.947 | 16.751 | 19.457 | 18.907 |
| 4-5a | 10.995 | 7.195 | 8.642 | 11.320 | 12.208 | 5.139 | 11.708 | 11.111 |
| 4-5b | 12.112 | 10.277 | 9.540 | 12.664 | 15.415 | 7.069 | 4.139 | 11.846 |
| 4-6 | 11.348 | 7.875 | 8.819 | 11.968 | 15.242 | 6.849 | 8.085 | 3.383 |

Table 14 AMOVA analyses using Tamura Nei (1993) distances and gamma as estimated by jModelTest 0.1.1 (Posada, 2008; Guindon & Gascuel, 2003) as calculated by Arlequin version 3.5.1.2 (Excoffier & Lischer, 2010). The eight genetically distinct clades were used as different populations. V_a and F_{ST} : $P(\text{random value} > \text{observed value}) = 0.00000$, $P(\text{random value} = \text{observed value}) = 0.00000$ and $P(\text{random value} \geq \text{observed Value}) = 0.00000+0.00000$.

| Source of variation | Degrees of freedom | Sum of squares | Variance components | Percentage of variation |
|-----------------------------------|--------------------|----------------|---------------------|-------------------------|
| Among clades | 7 | 907.362 | 6.19684 V_a | 73.04 |
| Within clades | 179 | 409.447 | 2.28742 V_b | 26.96 |
| Total | 186 | 1316.810 | 8.48425 | |
| Fixation Index F_{ST} : 0.73039 | | | | |

Table 15 Divergence times allowing for unequal population sizes (τ) / tau P values. All values were calculated in Arlequin, version 3.5.1.2 (Excoffier & Lischer 2010) using Tamura Nei (1993) distances and gamma as estimated by jModelTest 0.1.1 (Posada, 2008, Guindon & Gascuel, 2003). Time in years (t) was determined using the equation $\tau = 2ut$, where $u = m_T\mu$ (m_T is the number of nucleotides used and μ is the mutation rate per generation), and is shown under the grey-shaded diagonal.

| | 4-1 | 4-2a | 4-2b | 4-3 | 4-4 | 4-5a | 4-5b | 4-6 |
|------|--------|------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| 4-1 | | 8.887 / 0.478 | 10.073 / 0.414 | 10.649 / 0.482 | 14.243 / 0.535 | 7.414 / 0.541 | 8.471 / 0.312 | 7.931 / 0.397 |
| 4-2a | 218300 | | 6.126 / 0.654 | 10.666 / 0.387 | 11.279 / 0.420 | 4.487 / 0.437 | 6.928 / 0.570 | 5.056 / 0.541 |
| 4-2b | 247433 | 150478 | | 8.909 / 0.336 | 11.789 / 0.212 | 5.777 / 0.424 | 6.974 / 0.470 | 6.459 / 0.368 |
| 4-3 | 261581 | 261999 | 218840 | | 16.597 / 0.457 | 7.872 / 0.6115 | 9.267 / 0.596 | 8.835 / 0.481 |
| 4-4 | 349864 | 277057 | 289584 | 407688 | | 8.612 / 0.536 | 11.740 / 0.601 | 11.826 / 0.468 |
| 4-5a | 182117 | 110218 | 141906 | 193367 | 211545 | | 4.472 / 0.453 | 4.312 / 0.483 |
| 4-5b | 208081 | 170179 | 171309 | 227634 | 288381 | 109850 | | 5.456 / 0.487 |
| 4-6 | 194816 | 124195 | 158658 | 217022 | 290493 | 105919 | 134012 | |