# Population genetics of three sympatric springtail species (Hexapoda: Collembola) from the South Shetland Islands: evidence for a common biogeographic pattern

The Linnean Society of London
BIOLOGICAL Journal of the Linnean Society

# Population genetics of three sympatric springtail species (Hexapoda, Collembola) from the South Shetland Islands: evidence for a common biogeographic pattern



| | |
|---|---|
| Journal: | *Biological Journal of the Linnean Society* |
| Manuscript ID | Draft |
| Manuscript Type: | Research Article |
| Date Submitted by the Author: | n/a |
| Complete List of Authors: | Carapelli, Antonio; Universita degli Studi di Siena - Presidio San Miniato, Life Sciences<br>Convey, Peter; British Antarctic Survey, Natural Environment Research Council,<br>Frati, francesco; Universita degli Studi di Siena - Presidio San Miniato, Life Sciences<br>Spinsanti, Giacomo; Universita degli Studi di Siena - Presidio San Miniato, Life Sciences<br>Fanciulli, Pietro; Universita degli Studi di Siena - Presidio San Miniato, Life Sciences |
| Keywords: | Antarctica, Collembola, biogeography, evolutionary origin |
| | |

# Population genetics of three sympatric springtail species (Hexapoda, Collembola) from the South Shetland Islands: evidence for a common biogeographic pattern

**Antonio Carapelli[1*], Peter Convey[2], Francesco Frati[1], Giacomo Spinsanti[1], Pietro Paolo Fanciulli[1]**

[1] Department of Life Sciences, University of Siena, Via Aldo Moro 2, 53100 Siena, Italy.

[2] British Antarctic Survey, High Cross, Madingley Road, Cambridge CB3 0ET, UK.

*Corresponding author

Email: antonio.carapelli@unisi.it

Running title: Biogeography of South Shetlands springtails

**ABSTRACT**

Three sympatric springtail species, from the South Shetland Islands archipelago in the Maritime Antarctica, are here analyzed in a common biogeographic and evolutionary framework. This study is designed to compare their population genetic structure using the same molecular marker. Haplotype data for the mitochondrial *cox1* gene have been obtained for seven populations of *Folsomotoma octooculata* and are compared with data obtained, in previous studies and in the present one, for the sympatric species *Cryptopygus antarcticus antarcticus* and *Friesea grisea*. Molecular data are compatible with the hypothesis that all species were present in the archipelago since before the last glacial maximum (around 20,000 ybp), and that their early diversifications appear to be linked with known interglacials for the region. These springtails may have survived the last glacial cycle in local refugia, from where they have dispersed subsequently to ice-free ground re-exposed during the current interglacial. The populations of the different species diversified at different times, although all of them within the Pleistocene epoch. We propose that the earliest diversification of haplotypes in this archipelago occurred from local refugia in Livingston I., and that some are now distributed across all the South Shetlands Islands populations of these three Antarctic springtails.

ADDITIONAL KEYWORDS: Antarctica – Collembola – biogeography - evolutionary origin.

## INTRODUCTION

The number of collembolan (=springtail) lineages presently occurring in Antarctica is negligible in comparison with ecosystems at lower latitude. Among the 30 described families of Collembola, distributed worldwide in every damp environment, only four, all belonging to the suborder Arthropleona, have Antarctic representatives, with no records of taxa from the two remaining high-rank lineages Neelipleona and Symphypleona. This impoverished composition is the result of past glacial and tectonic events, along with progressively more extreme climatic conditions that, over time, denuded the pre-existing greater biodiversity of invertebrates (e.g. Ashworth & Kuschel, 2003; Lewis *et* al., 2008) and of the atmospheric and oceanic isolation of the continent, completed since the opening of the Drake Passage (≈28 Myr) (see review of Convey *et al.,* 2009).

The distribution of the over twenty known springtail species (Greenslade, 1995; 2010) is almost equally divided between the two major geological elements of the Antarctic continent, in biogeographical terms known as the Maritime and Continental Antarctic, which are broadly equivalent to West (including the Antarctic Peninsula) and East Antarctica. All these species, except that currently classified as *Friesea grisea* (but see Torricelli *et al.,* 2010a,b for evidence of deep molecular differentiation), inhabit one or the other, but not both, sides of an ancient biogeographic boundary known as the Gressitt Line, located at the base of the Antarctic Peninsula, that separates not only the springtails but also most other major contemporary terrestrial invertebrate groups of Antarctica (Chown & Convey, 2007).

In the maritime Antarctic (whose terrestrial habitats encompass the western side of the Antarctic Peninsula, its offshore islands, and the associated Scotia Arc, South Shetland, South Orkney and South Sandwich archipelagoes) the five most common collembolan species, *Archisotoma brucei* (Isotomidae), *Cryptopygus antarcticus antarcticus* (Isotomidae), *Folsomotoma octooculata* (Isotomidae), *Friesea grisea* (Neanuridae) and *Tullbergia mixta* (Tullbergiidae), live in vegetated or ornithogenic soils of coastal ice-free habitats, and display subtly different adaptations to cold and desiccation stress

(Hayward *et al.,* 2004; Russell *et al.,* 2014). Among them, the three isotomid species appear to be more active soil-dwellers, at least during the austral summer. Although each of them has its own distinctive distribution pattern, most (11) of the species known to occur in the maritime Antarctic are present in the South Shetland Islands archipelago, a group of islands (some of geologically recent volcanic origin) located north-west of the Antarctic Peninsula (Greenslade, 2010). Among them, the isotomid *F. octooculata* (previously also known as *Parisotoma octooculata*) is one of the most common, with a distribution associated with the presence of a vegetation cover, soil moisture and organic material (Russell *et al.,* 2014).

*Folsomotoma octooculata* is considered a native species of the South Shetland Islands (*sensu* Greenslade & Convey, 2012), and has congeneric counterparts in South America and Australia continents. Previous studies have primarily focused on its morphological features. Its taxonomic status has been updated on more than one recent occasion, initially by Greenslade (1986), who reassigned some species of *Isotoma* and *Sorensia* to the subgenus *Folsomotoma* and, subsequently, again by Greenslade (1995), assigning *Isotoma octooculata* Willem 1901 to the subgenus *Folsomotoma* (*Isotoma Folsomotoma octooculata*). Finally, Greenslade (2010), following Potapov (2001) who erected the subgenus *Folsomotoma* to generic status, proposed the present name of *Folsomotoma octooculata* for the species. Despite these taxonomic developments, a biogeographic study has never been attempted, unlike some other Antarctic springtails (Stevens *et al.,* 2006a; Hawes, Torricelli & Stevens, 2010; McGaughran *et al.,* 2010).

This study uses a mitochondrial molecular to assess the evolutionary history of the species *F. octooculata* in the South Shetland Islands. These data are also compared with those available for other Antarctic collembolans *Cryptopygus antarcticus antarcticus* and *Friesea grisea* (including additional new data obtained for these species) providing information on the origin and the genetic structure of *F. octooculata* in the archipelago.

## MATERIAL AND METHODS

### *SEQUENCING THE DNA FRAGMENTS*

Samples of *F. octooculata* (Willem, 1901) from the South Shetland Islands archipelago were collected from seven locations during a 2002-03 expedition involving a collaboration between the British Antarctic Survey (BAS) and the Italian National Antarctic Program (PNRA) (Fig. 1). The specimens were identified under a stereomicroscope, frozen and preserved at -80°C until their use for molecular analyses. Total DNA was extracted from single individual springtails using the Wizard SV genomic DNA purification system (Promega) and used for amplifications, performed in a GeneAmp PCR System 2700 (Applied Biosystems) thermal cycler. For the PCR amplification of the targeted mitochondrial gene-encoding fragment, corresponding to the almost complete sequence of the cytochrome *c* oxidase subunit I (*cox1*) gene, the following pair of primers (FOC-trnY-1393J: 5'-AAAAATAATTTCTATGATTAAATTTACAG-3'; FOC-trnLuaa-2983N: 5'-GAATTTTAAGTTCATTACACTAATCTG-3') were synthesized, using as reference the complete mtDNA of the species (Carapelli *et al.,* 2014). Their match is located on the two tRNA-encoding genes flanking *cox1* (*trnY* and *trnLuaa*), on the mtDNA molecule. All reactions were performed in a volume of 25 µl containing 2.5 µl of genomic DNA (with a range of concentration between 2.3 and 17 ng $ml^{-1}$), 0.5 mM of each primer, 0.2 mM of each deoxynucleotide, 2.5 mM of $MgCl_2$, 5 µl of Green GoTaq Flexi buffer and 0.625u of GoTaq Flexi DNA Polymerase. PCR conditions were: 35 cycles at 95°C for 1 min, 50°C for 1 min, and 72°C for 90 sec, followed by a final extension step at 72°C for 5 min. PCR products were than purified using the Kit Wizard SV Gel and PCR Clean-up (Promega) and sequenced on both strands at the core facility of the Biofab Research Lab, with the same primers used for the PCR reaction. Furthermore, in order to have a double reading for each nucleotide position of the amplified products, sequencing reactions were also carried out with the internally designed primers: FOC-cox1-2082J (5'-CGTAATTTGAATACATCATTTTTTG-3') and FOC-cox1-2279N (5'-AGTAAATATATGGTGTGCTCAAACG-3'), with numbers in each name corresponding to the position of the 3'-base of the primer in the *F. octooculata* mtDNA (Carapelli *et al.,* 2014).

The final consensus sequences were assembled using Sequencher 4.4.2 (Gene Codes) and deposited in GenBank under the accession numbers: KT008628-KT008644. The

two species *C. a. antarcticus* (Carapelli *et al.,* 2008) and *Orchesella villosa* (Carapelli *et al.,* 2007) were used as cross reference taxa for the genetic distance analysis.

In order to obtain a direct comparison among species that inhabit the same Antarctic localities, a parallel genetic analysis was run for two additional collembolan species that live in sympatry with *F. octooculata*. Therefore, sequences of *C. a. antarcticus* and *F. grisea* were also analyzed. The *F. grisea* data set reported by Torricelli *et al.* (2010a) was enlarged to include the same group of populations and the same number (10) of specimens obtained for *F. octooculata*, adding new sequences (though not increasing the total number of different haplotypes, originally labeled: P1-7) for: DPL (3 x P7), HAL (3 x P3), HPN (3 x P3) and HPL (1 x P3), and sequencing 10 new specimens each for PCK and RPN (both 10 x P3) (abbreviations used for sampling localities listed in Table 1). Methods used for DNA extraction, amplification and sequencing are as described in Torricelli *et al.* (2010a). Similarly, for *C. a. antarcticus* part of the data set analyzed for this species by McGaughran *et al.* (2010), limited to the South Shetland Islands localities, was improved with two new sequences obtained (with the same methods as described by McGaughran *et al.* (2010) for PCK, leading to the addition of one copy for each of H20 and H31 (using the original study's haplotype descriptor for this species). The final data set therefore included 10 springtail specimens for each of the four compared populations. No samples were available for HAL and HPL, although all four individual islands under study were represented for this taxon.

*ASSEMBLING THE DATA SET*

Ten specimens of *F. octooculata* for each of the seven sampling localities (Fig. 1, Table 1) were sequenced for 1533 nucleotides of *cox1* (between positions 1422-2954 of the *F. octooculata* mtDNA). The selected fragment includes all the codon positions of *cox1* (except for the last 2 triplets and the stop codon) and also comprises the last 6 nucleotides at the 3'-end of *trnY* (the gene that foreruns the 5'-end of *cox1*, along the J-strand of mtDNA).

The sequences of *F. octooculata* were manually aligned with MacClade 4.08 (Maddison & Maddison, 2005), resulting in a 1533-bp matrix, with no indels. The 70 sequences of

*F. grisea* and the 40 of *C. a. antarcticus* were also aligned, resulting in matrices of 478- and 618-bp, respectively.

Frequencies of haplotypes for all species (Table 1, Table S1) were obtained using the online tool DNAcollapser (http://users-birc.au.dk/biopv/php/fabox/software), and used in *F. octooculata* for the network clade analysis using TCS 1.21 (Clement, Posada & Crandall, 2000), with the connection limit set to 95%. Haplotype nomenclature for *C. a. antarcticus* and *F. grisea* follows that used in McGaughran *et al.* (2010) and Torricelli *et al.* (2010a), respectively.

Analyses of demographic history of the *F. octooculata* populations, as well as those of *C. a. antarcticus* and *F. grisea*, were performed using 16,000 iterations of the program Arlequin 3.11 (Excoffier, Laval & Schneider, 2005), using the distribution of observed and simulated pairwise differences among haplotypes within each population (Table 2). The time of expansion (t) was calculated using the population demographic parameter tau ($\tau$), and applying the formula $t = \tau/2\mu$, where $\mu$ is the mutation rate per locus per generation (Rogers & Harpending, 1992), and assuming a divergence rate of 1.5-2.3% $Myr^{-1}$ (Brower, 1994) and a generation time of 3 years (McGaughran, Hogg & Stevens, 2008). Parametric bootstrapping was then used to estimate signatures of demographic expansion, using the population demographic parameters tau ($\tau$) and theta ($\Theta$; with $\Theta_0$ at pre- and $\Theta_1$ at post-expansion, respectively). Sum of squared deviations (*SSD*) (Rogers & Harpending, 1992) between observed and expected mismatch patterns, as well as Raggedness (*R*) index, were used to test the model of demographic expansion, assessing the fit of the observed distribution with population expansions chosen as the null hypothesis (Harpending, 1994) (Table 2, Table S2).

Genetic divergence among haplotypes for all species under study, using absolute and pairwise distance methods, was assessed with PAUP* (version 4b10-x86-macosx) (Swofford, 2003), and $F_{ST}$ genetic distances were calculated using Arlequin 3.11 (Excoffier *et al.,* 2005). Matrices of geographical distances among samples (obtained with the program Geographic Distance Matrix Generator; version 1.2.3) and $F_{ST}$ distances were compared (through 16,000 permutations) using Arlequin 3.11 (Excoffier *et al.,* 2005), to assess the significance of the correlated values, as implemented in the Mantel test (Mantel, 1967). The same program was also used to estimate haplotype (*h*) and nucleotide ($\pi$) diversity indices (Nei 1987) and to run "neutrality tests" among

populations, applying Tajima's *D* (Tajima, 1989) and Fu's $F_S$ (Fu, 1997) parameters, with the significance of the values evaluated over 16,000 permutations. AMOVA analysis was performed with Arlequin (Excoffier *et al.,* 2005) to produce estimates of variance between haplotypes at different hierarchical levels. The seven *F. octooculata* populations were therefore tentatively clustered into: a) four groups, each corresponding to the sampled island; b) two groups, associating haplotypes obtained from King George I. and Nelson I. (KN group) and from Livingston I. and Robert I. (LR group). Hierarchical clustering of haplotypes was performed using BAPS 6.0 to assess dependence between unlinked markers under the Bayesian model of clustering method (Cheng *et al.,* 2013). This latter analysis was also applied to the *C. a. antarcticus* and *F. grisea* datasets.

### *PHYLOGENETIC ANALYSIS OF HAPLOTYPES*

Haplotype sequences of *cox1* of the three springtail species (Tables 1 and S1) and of the outgroup species *Onychiurus orientalis* (Cook, Yue & Akam, 2005) were manually aligned (resulting in a matrix of 1554 bp) and used for the phylogenetic analysis. Aligned nucleotides were partitioned in three groups according to their codon position ($1^{st}$, $2^{nd}$ and $3^{rd}$), and examined for the best partitioning strategy and evolutionary model for each partition, as implemented in PartitionFinder 1.0.1 (Lanfear *et al.,* 2012). The resulting partitioning scheme and evolutionary models ($1^{st}$= GTR+Γ; $2^{nd}$= HKY+I; $3^{rd}$= HKY+ Γ) were applied in a bayesian analysis using MrBayes 3.2.1 (Ronquist & Huelsenbeck, 2003). Two parallel runs, each consisting of four chains, were run for 5 million generations, sampling every 1000th generation and removing 20% as burnin, upon stationarity of log-likelihood values. The final consensus tree was used to define the genetic relationships among haplotypes and to visualize the different patterns obtained with the clustering analysis (Fig. 2).

### *ABBREVIATIONS*

AMOVA, Analysis of molecular variance; Γ, Gamma; GTR, General Time Reversible; HKY, Hasegawa Kishino Yano; *h*, haplotype diversity; I, Invariant; LGM, Last Glacial

Maximum; Myr, million years ago; $\pi$, nucleotide diversity; PCR, Polymerase Chain Reaction; ybp, years before present.

## RESULTS

### *HAPLOTYPE COMPOSITION AND DIVERSITY INDICES*

Screening of the 70 sequences of *F. octooculata* resulted in 17 haplotypes (Table 1) that differed in a total of 22 variable sites (Table S2). Most haplotypes occurred at low frequency: 12 are represented by only one individual, 15 are unique of one single population, while only two of them (A and J) are found in more than one site (Table 1). Haplotype A was present at all locations, while J occurred only in Livingston and Robert Islands (the two southernmost islands of the four investigated) (Fig. 1). Most of the low frequency haplotypes differed from A or J by a single nucleotide substitution. Haplotype P was the most divergent, with 4 or 5 nucleotide differences compared with J and A, respectively (Table S2). The largest number of haplotypes (6) was observed in the Nelson I. population (HPN), whereas that of King George I. (PCK) only hosted individuals with haplotype A. Nelson Island also had the greatest number (8) of the 15 unique haplotypes identified. Nucleotide substitutions among the 70 examined sequences occurred in all three codon positions of *cox1*. The single nucleotide substitutions that differentiate haplotypes C and D from A (at aligned position 1010 and 70, respectively) lead to an amino acid change (Table S2).

Haplotype diversity values were remarkable for four populations (DPL, HPL, HPN and RPN) and lower for the remaining three (CPR, HAL and PCK) (Table 1). Conversely, nucleotide diversity was low ($\pi$<0.0008) for all populations (Table 1), implying recent genetic diversification. In *F. grisea* 7 haplotypes were present (Table S1), with two being very common (P3=55 and P7=10) and 5 unique. P3 was found at all sampled sites with the exception of DPL, whereas the remaining haplotypes were each unique to one location. The DPL population contained only P7, and is the only one where the most frequent haplotype was not represented, suggesting genetic divergence from the most common genetic pool. Intriguingly, in both *F. octooculata* and *F. grisea*, at the Devils Point site (Livingston Island) the overall most frequent haplotype was absent or present

at a very low frequency. In the Livingston Island populations of *F. grisea* five haplotypes were present, including the most frequent overall (P3) and the most frequent of the six unique haplotypes (P7) (Table S1). In *F. grisea*, *h* values were lower than in *F. octooculata*, while values of $\pi$ were similar between the two species, except for those obtained from HAL (Table S1), where the $\pi$ value was the highest amongst all sampling sites, due to the large number of nucleotide changes (from 9 to 11) observed in the comparisons involving P6 with the remaining haplotypes. A large number of haplotypes (28) was observed in total in *C. a. antarcticus*, with number of haplotypes per population ranging between 5 and 9. Only two haplotypes (H15 and H20) were shared between populations (both between CPR and HPN), and these were represented by a limited number (four and two, respectively) of sequences (Table S1). Consequently, high values of *h* and $\pi$ (Table S1) were found in these populations.

*GENETIC DISTANCES*

The matrix of pairwise genetic distances between haplotypes highlights substantial uniformity within populations of *F. octooculata*, which differed an average 0.24% and at maximum by 0.52%. The latter corresponds to 8 changes, and is observed when haplotype P is compared with L, N and Q. Conversely, the lowest estimate (0.06% = one single substitution) was obtained when the following haplotypes were compared: A *vs* B-K, O and M; J *vs* Q, N and L. Values of *p*-distance were considerably higher when haplotypes of *F. octooculata* are compared with the other collembolan species *C. antarcticus antarcticus* (Carapelli *et al.,* 2008) and *O. villosa* (Carapelli *et al.,* 2007), ranging from 20% to 21%. The genetic distances calculated between the seven *F. grisea* haplotypes ranged from 0.21% to 2.30% (corresponding to 1 and 11 nucleotide changes, respectively), giving an average 0.91% divergence. Most of the variability was generated when P6 was compared with the other haplotypes. If P6 is excluded from the comparison, the average *p*-distance value drops to 0.40%.

In *C. a. antarcticus* ranges of nucleotide substitutions (from 1 to 51) and distance estimates among haplotypes were substantially larger than in the other two springtail species, with the highest value (8.25%) observed when H18 was compared with H25. Average *p*-distances calculated between haplotypes in the present data set (2.6%) are

similar to that obtained (2.5%) in a previous analysis based on 14 *cox1* sequences from three Antarctic Peninsula sites (Stevens *et al.,* 2003). However, it should be noted that evaluation of genetic distances calculated for the three different species, although based on the same mitochondrial gene, are not completely comparable, due to the different size of the analyzed fragments. The proportion of nucleotide substitutions (*p*-distance) estimated in *F. octooculata* is based on a longer fragment of *cox1* than that used in *C. a. antarcticus* and *F. grisea*. An overall 59% of the genetic variability of the *cox1* dataset observed in *F. octooculata* (13 out of 22 nucleotide changes) was distributed in the aligned fragment shared between this species and *C. a. antarcticus*.

*MANTEL TEST*

The Mantel test showed a significant correlation between genetic and geographic distances ($r$= 0.487; p= 0.019) in the *F. octooculata* and *F. grisea* populations studied ($r$ = 0.487, p = 0.019 and $r$ = 0.434, p = 0.012, respectively), indicating that neighboring populations are genetically more similar than expected by chance. In contrast, in *C. a. antarcticus* the Mantel test rejected the hypothesis of correlation between genetic and geographical data between the four populations ($r$ = 0.603, p = 0.074).

*ANALYSIS OF MISMATCH DISTRIBUTIONS*

Mismatch analysis (MMD) provided a bimodal distribution of substitution frequencies detected between haplotypes for the complete set of *F. octooculata* populations, as well as for the four collected on Livingston and Robert Islands, suggesting demographic equilibrium; estimates of times of expansion support a pre-LGM expansion only for DPL and HPL (Table 2). For the two populations sampled in Nelson Island (HPN and RPN), calculations of MMD resulted in an unimodal distribution (Table 2), which is a feature of populations that have undergone a recent demographic expansion (Rogers & Harpending, 1992).

Mismatch analysis also provided bimodal or multimodal distributions of substitution frequencies for *C. a. antarcticus*, with an estimated time of expansion for DPL that dates to the mid-Pleistocene. In *F. grisea* populations (Table 2), unimodal distribution

of substitution frequencies would imply demographic equilibrium, with older times of expansion again being indicated in the Livingston Island population. Collectively, p-values for *R* (raggedness) and *SSD* (sum of squared deviations) indexes were not significant, suggesting failure to reject the null hypothesis of demographic expansion of populations.

*AMOVA AND HAPLOTYPE CLUSTERING*

Analysis of molecular variance, performed with *F. octooculata* populations grouped according to the island of origin, suggests that the largest total variation was observed within the populations and secondarily among populations within groups (Table 3). When populations were divided in the two groups KN and LR, total variance was still mostly attributable to the intra-population level (Table 3). This finding suggests that a subdivision into two groups rather than four is more appropriate, with haplotype J representing the molecular signature of the genetic dissimilarities between the KN and LR groups. Bayesian analysis of population structure revealed a nested genetic population subdivision into three clusters ($C_{FO}$ 1-3), with log-marginal likelihood of optimal partition of -197.0363. $C_{FO}$ 1 was represented by haplotypes A-I, K, M and O; $C_{FO}$ 2 by J, L, N and Q; $C_{FO}$ 3 by P. This subdivision exactly corresponds to the three major branches of the haplotype network (Fig. 1) and is also represented in the branching patterns of Fig. 2. Among the four investigated islands, only on Livingston Island were populations with haplotypes of all three clusters identified. Amova analysis for *F. griesea*, whose populations were tentatively clustered in a similar way to *F. octooculata*, generated negative values of variance components suggesting absence of genetic structure. In *C. a. antarcticus*, the more limited number of samples (one for each of the four islands under study) prevented association of the populations into groups. Bayesian clustering of haplotypes in *F. grisea* also led to the identification of three clusters: $C_{FG}$ 1 (P2 and P7), $C_{FG}$ 2 (P1, P3, P4 and P5) and $C_{FG}$ 3 (P6) (log-marginal likelihood of optimal partition: -86.0226) (Fig. 2). King George and Robert Islands had only haplotypes belonging to $C_{FG}$ 2, and Nelson I. to clusters $C_{FG}$ 1 and $C_{FG}$ 2, while all the three clusters were represented in Livingston Island. Clustering of haplotypes for *C. a. antarcticus* resulted in four clusters: $C_{CA}$ 1 (H21-H28), $C_{CA}$ 2 (H6-H8 and H10), $C_{CA}$

3 (H11, H13-H20, H31, H32, H38 and H39) and $C_{CA}$ 4 (H29 and H30) (log-marginal likelihood of optimal partition: -736.9849). CPR, HPN and PCK had only haplotypes of $C_{CA}$ clusters 3, 2 and 1, respectively, whereas at Devils Point (Livingston Island) three ($C_{CA}$ 2-4) of the four clusters were represented (Fig. 2).

*MOLECULAR CLOCK AND TIMING OF POPULATION EXPANSION*

Dating the time of the earliest and most recent diversification between haplotypes of *F. ocotooculata*, assuming a divergence rate of 1.5-2.3% $Myr^{-1}$ (Brower, 1994), led to date range between 104,347 and 160,000 ybp, suggesting an Upper-Middle Pleistocene, but pre-LGM (Last Glacial Maximum), differentiation within the species in the South Shetland Islands. The application of the same rate for the seven *F. grisea* haplotypes, and the use of average divergence values for the populations (0.91%) generated a more ancient differentiation (between 395,652 and 606,666 ybp) than in *F. octooculata*, corresponding to the Middle Pleistocene. The average genetic distance between P6 (the most basal haplotype of the branching pattern in Fig. 2) and the remaining haplotypes was 2.2%, leading to an estimated time of earlier divergence between the $C_{FG}$ 1 and $C_{FG}$ 2 clusters of 0.9 to 1.5 million ybp, within the Lower Pleistocene. However, when the most divergent haplotype (P6) is excluded, the estimated diversification time, calculated from the average *p*-distance value among the remaining haplotypes (0.4% between $C_{FG}$ 1-2) reduced to between 173,913 and 266,667 ybp. In *C. a. antarcticus* the average level of genetic divergence (2.6%) led to Lower Pleistocene dates (1.1–1.7 million ybp), suggesting a 10-times older diversification than in *F. octooculata*.

The calculated time of demographic expansion was very similar in *F. octooculata* and in *F. grisea* for their HPN populations (ranging from 7,432 to 12,343 ybp), whereas in HAL and HPL from Livingston Island, *F. grisea* provided evidence of a more ancient expansion date (from 44,413 to 69,735 ybp). In each possible comparison, the populations of *C. a. antarcticus* appear to have expanded much earlier than the other two species studied. It is noteworthy that inter-specific comparisons identified the populations of Devils Point specifically, and Livingston Island generally, as showing the earliest demographic expansions (Table 2).

*DEMOGRAPHIC ANALYSIS*

All *F. octooculata* populations, except for DPL and HPL, had negative values of Tajima's *D* coefficient, suggesting that the expected average level of variation among haplotypes (i.e. the number of segregating sites) is higher than that observed. Apart from the mono-haplotypic population of PCK, all *D* values were greater than 0, implying departure from neutrality of nucleotide substitutions. These data collectively provide support to the hypothesis that recent expansion of populations may have occurred, although the low number of specimens investigated for each locality suggests caution in drawing conclusions. Similar results were obtained in *F. grisea*, with negative values observed for all populations (except those represented by a single haplotype), whereas in *C. a. antarcticus* PCK was the only population with a negative *D* value. Fu's test of neutrality also suggested recent expansion of the HPN and RPN populations in *F. octooculata*, and for CPR, HPN and PCK in *C. a. antarcticus*. HPN and HPL had negative Fu's values in *F. grisea.* However, none of these values were statistically significant (at $p<0.05$) apart from the Fu's parameter calculated for HPN in *F. octooculata*, for CPR in *C. a. antarcticus* and for HPN in *F. grisea*.

## DISCUSSION

The onset of the most recent deglaciation in the South Shetland Islands was around 11,000 ybp, with the process continuing until 8,400-6,000 ybp (Ingólfsson *et al.,* 1998). More recent re-advance of ice took place in some areas of the archipelago around 5,000 ybp. This recent glaciological history, although with less severe outcomes with respect to that which impacted the invertebrate biota of the Northern Hemisphere (Bergstrom & Chown, 1999), would have permitted the establishment at local scale of ecosystems capable of hosting the invertebrate life typical of contemporary terrestrial ecosystems around 5-7,000 ybp (e.g. Hodgson & Convey, 2005). Despite this, the presence of "relict" springtail (and many other) lineages that must have survived multiple glacial cycles in ice-free refugia in Antarctica is now generally accepted (e.g. Stevens *et al.,* 2003, 2006b; McGaughran *et al.,* 2011), overturning the pre-existing paradigm that the

vast majority of the continent's contemporary terrestrial biota must consist of recent post-LGM dispersers (Convey & Stevens, 2007; Convey *et al.,* 2008).

In the "refugial" scenario, surviving invertebrate taxa would subsequently recolonize habitats made available through glacial retreat, effectively now existing in populations isolated by natural barriers. The genetic parameters of the three species studied here, and the molecular estimate of differentiation times generated from them, support the idea that the diversification of their haplotypes in the South Shetland Islands started within the Pleistocene but well before the LGM. However, our data also suggest that separate differentiation events occurred during different time intervals within this geological period. Thus, the higher numbers of total and intra-population haplotypes observed in *C. a. antarcticus* in comparison with its sympatric counterparts suggest that the species has an evolutionary history in the archipelago that can be traced back at least to the Lower Pleistocene. In contrast, more recent diversification events in the Upper-Middle Pleistocene are indicated for the *F. grisea* and *F. octooculata* populations examined. The high number of haplotypes observed in the latter species also suggest a recent demographic expansion of its populations, with many of the haplotypes obtained at low frequency or from single individuals. Demographic analyses collectively provide support to the hypothesis that recent expansion of populations may have occurred, although the low number of specimens investigated for each locality means that caution in drawing conclusions is required.

The estimate obtained here of the timing of diversification within *F. octooculata* (104-160,000 ybp) is consistent with the Valdivian interglacial period recorded from southern South America (115-130,000 ybp) (Astorga & Pino, 2011; NEEM Community Members 2013). Our data indicate that *F. grisea* diversified in the archipelago earlier (396-607,000 ybp) than *F. octooculata*. While a wider date range, this encompasses at least two recorded interglacial periods in southern South America, the most recent being the Hoxnian (374-424,000 ybp) (Lisiecki & Raymo, 2005). The 1.1-1.7 million ybp divergence estimate for *C. a. antarcticus* is also a wide age range but, given the pre-Pastonian glacial period extended from 0.8-1.3 million ybp, this estimate is consistent with divergence within the preceding Bramertonian interglacial stage of the Pleistocene (1.3-1.55 million ybp) (see Gibbard & van Kolfschoten, 2004 for an overview of glacial/interglacial periods).

The genetic diversity of the South Shetland Islands populations of springtails recorded here is likely to represent only a limited fraction of their total differentiation. Indeed, previous studies performed on haplotype data for populations examined within a larger geographical context have demonstrated much greater levels of intraspecific differentiation (Stevens *et al.,* 2006a; Torricelli *et al.,* 2010a). In *C. a. antarcticus*, the high number of haplotypes found at low frequency and the almost complete absence of linkage between populations suggest recent population expansion (McGaughran *et al.,* 2011), although genetic divergence parameters and MMD analysis date the differentiation of haplotype lineages (clusters) and demographic expansion to deeper in the past. In *F. grisea*, the amount of genetic divergence observed in the South Shetland Islands populations is negligible in comparison with that calculated between populations inhabiting the entire range of distribution of the species along the Antarctic Peninsula (0.91% *vs* 2.7%) and even more so if related to samples obtained from the Victoria Land, in the continental Antarctica (14.4-17.2% divergence between maritime and continental Antarctic haplotypes) (Torricelli *et al.,* 2010a). It should be noted, however, that taxonomic analysis is under way to investigate whether Western and Eastern Antarctic populations of *F. grisea* belong to the same species.

In *F. octooculata* the large number of haplotypes represented by unique sequences, the high values of *h*, the low estimates of $\pi$ and the results of the MMD analysis all suggest recent demographic expansion, at least for the two populations from Nelson Island or which MMD calculations suggest a demographic expansion subsequent to post-LGM glacial retreat. Among the remaining populations, DPL and HPL (two out of the three from Livingston Island), display higher $\pi$ values and include haplotypes belonging to all the groups identified by the cluster analysis.

The pattern of distribution of the haplotypes is most likely dependent on the initial distribution of A over the entire range of suitable environments, followed by differentiation of J in a more restricted area. The abundance of the J haplotype in the southern part of the sampled area now equals or exceeds that of any other haplotype apart from A, although J is the most frequent haplotype only at Devil's Point (Livingston Island) (Table 1). The two populations from Nelson Island (HPN and RPN) have the largest number of low-frequency haplotypes derived from A through a single nucleotide substitution (Fig. 1; Table S2), and are also the only populations with

unimodal distributions of mismatch parameters (Table 2) (implying that they have undergone a recent demographic expansion). Southern populations of *F. octooculata* (i.e. those with either A and J haplotypes; Table 1) are candidate locations that may have ancestrally colonized (or inter-glacially re-colonized) the South Shetland Islands. Specimens with the most frequent haplotypes, A and J (coexisting in Livingston and Robert Islands), may have then dispersed northwards. Only A has (so far) successfully established on Nelson and King George Islands, where it has locally differentiated into several low-frequency haplotypes in the most recently colonised sites (those of Nelson Island).

These factors are consistent with the first haplotype diversification processes taking place on Livingston Island, originating from refugia where the species persisted throughout the LGM period. This process may have initially involved haplotypes from cluster 1 (where the most frequent haplotype A is present), which have locally differentiated to generate clusters 2 and 3 in Livingston and Robert Islands. Cluster 1 is the only cluster whose members are also present on Nelson and King George Islands, supporting a south-to-north route of dispersal. Similarly, in the other two species studied, most (if not all) of the detected haplotype clusters are present in the Livingston Island populations. In addition, the branching pattern obtained in the haplotype phylogenetic tree highlights that $C_{FG}$ 3 and $C_{CA}$ 4 are the most basal clusters (and therefore the ancestral lineages of the remaining groups) for *F. grisea* and *C. a. antarcticus*, respectively, and both groups are unique to Livingston Island. These data again suggest that an earlier diversification occurred on this island, likely initiating from local glacial refugia, with subsequent colonization of the other locations of the Archipelago.

The genetic structure of these three springtail species in the South Shetland Islands is characterized by subdivision of the haplotypes into several groups, the demarcation of which is usually restricted to a small number of nucleotide changes. These data highlight slow rates of molecular differentiation in *F. octooculata*, and is consistent with a recent but pre-LGM evolutionary origin for this species within the South Shetland Islands archipelago. In a wider context, source populations over the different timescales required for all three springtail species considered here are likely to have been elsewhere in the maritime Antarctic region, given their wider contemporary

distributions, but at the present time no appropriate molecular data are available from across these distributions to allow such an assessment to be made.

## ACKNOWLEDGMENTS

This work was funded by the Italian Program of Research in Antarctica (PNRA), and the Italian MIUR (PRIN). The logistic support of the British Antarctic Survey (BAS) and HMS *Endurance* was crucial for the collection of the material. Partial support was also provided by the University of Siena. We also thank Dr. Alessandro Manenti for his assistance in the collection of data. PC is supported by NERC core funding to the BAS 'Biodiversity, Evolution and Adaptation' Team. This paper also contributes to the Scientific Committee on Antarctic Research 'State of the Antarctic Ecosystem' (AntEco) international research programme.

## REFERENCES

**Ashworth AC, Kuschel G. 2003.** Fossil weevils (Coleoptera: Curculionidae) from latitude 85° S Antarctica. *Palaeogeography, Palaeoclimatology, Palaeoecology* **191:** 191-202.

**Astorga G, Pino M. 2011.** Fossil leaves from the last interglacial in Central-Southern Chile: inferences regarding the vegetation and paleoclimate. *Geologica Acta* **9:** 45-54.

**Bergstrom DM, Chown SL. 1999.** Life at the front: history, ecology and change on southern ocean islands. *Trends in Ecology and Evolution* **14:** 472-477.

**Brower AVZ. 1994.** Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences of the United States of America* **91:** 6491–6495.

**Carapelli A, Liò P, Nardi F, van der Wath E, Frati F. 2007.** Phylogenetic analysis of mitochondrial protein coding genes confirms the reciprocal paraphyly of Hexapoda and Crustacea. *BMC Evolutionary Biology* **7:** S8.

**Carapelli A, Comandi S, Convey P, Nardi F, Frati F. 2008.** The complete mitochondrial genome of the Antarctic springtail *Cryptopygus antarcticus* (Hexapoda: Collembola). *BMC Genomics* **9:** 315.

**Carapelli A, Convey P, Nardi F, Frati F. 2014.** The mitochondrial genome of the Antarctic springtail *Folsomotoma octooculata* (Hexapoda; Collembola), and an update on the phylogeny of collembolan lineages based on mitogenomic data. *Entomologia* **2:** 190.

**Cheng L, Connor TR, Sirén J, Aanensen DM, Corander J. 2013.** Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. *Molecular Biology and Evolution* **30:** 1224-1228.

**Chown SL, Convey P. 2007.** Spatial and temporal variability across life's hierarchies in the terrestrial Antarctic. *Philosophical Transactions of the Royal Society B* **362:** 2307–2331.

**Clement M, Posada D, Crandall KA. 2000.** TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9:** 1657-1659.

**Convey P, Stevens MI. 2007.** Antarctic Biodiversity. *Science* **317:** 1877–1878.

**Convey P, Gibson JAE, Hillenbrand C-D, Hodgson DA, Pugh PJA, Smellie JL, Stevens MI. 2008.** Antarctic terrestrial life - challenging the history of the frozen continent? *Biological reviews of the Cambridge Philosophical Society* **83:** 103–117.

**Convey P, Stevens MI, Hodgson DA, Smellie JL, Hillenbrand C-D, Barnes DKA, Clarke A, Pugh PJA, Linse K, Cary SC. 2009.** Exploring biological constraints on the glacial history of Antarctica. *Quaternary Science Reviews* **28:** 3035–3048.

**Cook CE, Yue Q, Akam M. 2005.** Mitochondrial genomes suggest that hexapods and crustaceans are mutually paraphyletic. *Proceedings of the Royal Society B* **272:** 1295–1304.

**Enderlein G. 1903.** Die Insekten und Arachnoiden der Kerguelen. *Valvidia* **3:** 199-249.

**Excoffier L, Laval G, Schneider S. 2005.** Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1:** 47–50.

**Fu YX. 1997.** Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147:** 915-925.

**Gibbard P, van Kolfschoten T. 2004.** The Pleistocene and Holocene Epochs. In: Gradstein FM, Ogg JG, Smith A. Gilbert, eds. *A Geologic Time Scale 2004*. Cambridge: Cambridge University Press, 441-453.

**Greenslade P. 1986.** Additions to the Collembola of Heard Island. *Records of the South Australian Museum* **19:** 91–96.

**Greenslade P. 1995.** Collembola from the Scotia Arc and Antarctic Peninsula including description of two new species and notes on biogeography. *Polskie Pismo Entomologiczne* **64:** 305–319.

**Greenslade P. 2010.** South Shetlands Collembola fauna revisited. *Antarctic Science* **22:** 233–242.

**Greenslade P, Convey P. 2012.** Exotic Collembola on subantarctic island: patways, origin and biology. *Biological Invasions* **14:** 405-417.

**Harpending HC. 1994.** Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology* **66:** 591-600.

**Hawes TC, Torricelli G, Stevens MI. 2010.** Haplotype diversity in the Antarctic springtail *Gressittacantha terranova* at fine spatial scales - a Holocene twist to a Pliocene tale. *Antarctic Science* **22:** 766-773.

**Hayward SAL, Worland MR, Convey P, Bale JS. 2004.** Effects of moisture on the local distribution of the Antarctic Collembola *Cryptopygus antarcticus* and *Friesea grisea*. *Soil Biology and Biochemistry* **36:** 927-934.

**Hodgson DA, Convey P. 2005.** A 7000 year record of the oribatid mite communities on a maritime-Antarctic island: responses to climate change. *Arctic Antarctic and Alpine Research* **37:** 239-245.

**Ingólfsson Ó, Hjort C, Berkman P, Björck S, Colhoun E, Goodwin ID, Hall B, Hirakawa K, Melles M, Möller P, Prentice M. 1998.** Antarctic glacial history since the Last Glacial Maximum: an overview of the record on land. *Antarctic Science* **10:** 326-344.

**Lanfear R, Calcott B, Ho SYW, Guindon S. 2012.** Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution* **29:** 1695-1701.

**Lewis A, Marchant D, Ashworth A, Hedenäs L, Hemming S, Johnsong J, Leng M, Newton A, Raine J, Willenbring J, Williams M, Wolfem A. 2008.** Mid-Miocene cooling and the extinction of tundra in continental Antarctica. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 1–5.

**Lisiecki LE, Raymo ME. 2005.** A Pliocene-Pleistocene stack of 57 globally distributed benthic d18O records. *Paleoceanography* **20:** 1-17.

**Maddison DR, Maddison WP. 2005.** MacClade 4: Analysis of phylogeny and character evolution. Version 4.08a. Available at: http://macclade.org.

**Mantel N. 1967.** The detection of disease clustering and a generalized regression approach. *Cancer Research* **27:** 209–220.

**McGaughran A, Hogg ID, Stevens MI. 2008.** Patterns of population structure for springtails and mites in southern Victoria Land, Antarctica. *Molecular Phylogenetics and Evolution* **46:** 606–618.

**McGaughran A, Redding GP, Stevens MI, Convey P. 2010.** Patterns of temporal and spatial metabolic rate variation in an Antarctic springtail. *Journal of Insect Physiology* **56:** 57–64.

**McGaughran A, Stevens MI, Hogg ID, Carapelli A. 2011.** Extreme glacial legacies: A synthesis of the Antarctic springtail phylogeographic record. *Insects* **2:** 62-82.

**NEEM community members (2013).** Eemian interglacial reconstructed from a Greenland folded ice core. *Nature* **493:** 489–94.

**Nei M. 1987.** *Molecular Evolutionary Genetics*. Columbia University Press: New York.

**Potapov M. 2001.** Synopses on Palaearctic Collembola vol. 3: Isotomidae. *Abhandlungen und Berichte des Naturkundemuseums Görlitz* **73:** 1-603.

**Ronquist F, Huelsenbeck JP. 2003.** MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19:** 1572-1574.

**Rogers A, Harpending H. 1992.** Population growth makes waves in the distribution of pairwise differences. *Molecular Biology and Evolution* **9:** 552-569.

**Russell DJ, Hohlberg K, Potapov M, Bruckner A, Otte V, Christian A. 2014.** Native terrestrial invertebrate fauna from the northern Antarctic Peninsula: new records, state of current knowledge and ecological preferences – Summary of a German federal study. *Soil Organisms* **86:** 1-58.

**Stevens MI, Hogg ID. 2003.** Long-term isolation and recent range expansion from glacial refugia revealed for the endemic springtail *Gomphiocephalus hodgsoni* from Victoria Land, Antarctica. *Molecular Ecology* **12:** 2357–2369.

**Stevens MI, Greenslade P, Hogg ID, Sunnucks P. 2006a.** Examining Southern Hemisphere springtails: could any have survived glaciation of Antarctica? *Molecular Biology and Evolution* **23:** 822–874.

**Stevens MI, Fjellberg A, Greenslade P, Hogg ID, Sunnucks P. 2006b.** Redescription of the Antarctic springtail *Desoria klovstadi* using morphological and molecular evidence. *Polar Biology* **29:** 820–830.

**Swofford DL. 2003.** PAUP* Phylogenetic analysis using parsimony (* and other methods), Version 4. Sinauer Associates, Sunderland, MA.

**Tajima F. 1989.** Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123:** 585–595.

**Torricelli G, Frati F, Convey P, Telford M, Carapelli A. 2010a.** Population structure of *Friesea grisea* (Collembola, Neanuridae) in the Antarctic Peninsula and Victoria Land: evidence for local genetic differentiation of pre-Pleistocene origin. *Antarctic Science* **22:** 757–765.

**Torricelli G, Carapelli A, Convey P, Nardi F, Boore JL, Frati F. 2010b.** High divergence across the whole mitochondrial genome in the 'pan-Antarctic' springtail *Friesea grisea*: evidence for cryptic species? *Gene* **449:** 30–40.

**Willem V. 1901.** Les Collemboles recueillis par l'expédition Antarctique belge. *Annales de la Société Entomologique de Belgique* **45:** 260-262.

Table 1. Statistics of population genetic parameters in *F. octooculata*. *n*, number of individuals; $N_H$, number of haplotypes; *h*, haplotype diversity; $\pi$, nucleotide diversity.

| | | | *Folsomotoma octooculata* | | | | |
|---|---|---|---|---|---|---|---|
| ***Collecting site*** | | ***Coordinate*** | ***n*** | ***$N_H$*** | ***h*** | ***π*** | ***Cluster*** |
| Potter Cove, King George I. | PCK | 62°14'S, 58°42'W | 10 | A(10) | - | - | $C_{FO}1$ |
| Rip Point, Nelson I. | RPN | 62°15'S, 58°59'W | 10 | A(4), B(4), C(1), D(1) | 0.7333 ± 0.1005 | 0.000609 ±0.000515 | $C_{FO}1$ |
| Harmony Point, Nelson I. | HPN | 62°19'S, 59°10'W | 10 | A(3), E(1), F(2), G(1), H(1), I(2) | 0.8889 ± 0.0754 | 0.000855 ± 0.000657 | $C_{FO}1$ |
| Coppermine Peninsula, Robert I. | CPR | 62°23'S, 59°42'W | 10 | A(8), J(1), K(1) | 0.3778 ± 0.1813 | 0.000652 ± 0.000540 | $C_{FO}1$+$C_{FO}2$ |
| Hannah Point, Livingston I. | HAL | 62°39'S, 60°36'W | 10 | A(8), J(1), O(1) | 0.3778 ± 0.1813 | 0.000652 ± 0.000540 | $C_{FO}1$+$C_{FO}2$ |
| Devils Point, Livingston I. | DPL | 62°40'S, 61°11'W | 10 | A(2), J(5), P(2), Q(1) | 0.7333 ± 0.1199 | 0.002334 ±0.001463 | $C_{FO}1$+$C_{FO}2$+$C_{FO}3$ |
| Hurd Peninsula, Livingston I. | HPL | 62°41S, 60°23'W | 10 | A(4), J(3), L(1), M(1), N(1) | 0.8000 ± 0.1001 | 0.001841 ± 0.001198 | $C_{FO}1$+$C_{FO}2$ |
| Total | | | 70 | 17 | | | |

Table 2. Analysis of Mismatch distribution in the *F. octooculata* populations, and for comparison in *C. a. antarcticus* and *F. grisea*. $\tau$, coefficient tau; $T_{1.5\%}$, estimated time of expansion using 1.5% divergence $Myr^{-1}$; $T_{2.3\%}$, estimated time of expansion using 2.3% divergence $Myr^{-1}$; $\theta_0$, theta ($4N_e\mu$) at pre-expansion; $\theta_1$, theta at post-expansion; *SSD*, sum of squared deviations; *R*, raggedness index; Demographic parameters for each population represented with the estimated distribution (modality).

***Folsomotoma octooculata***

| *Statistics* | ***PCK*** | ***RPN*** | ***HPN*** | ***CPR*** | ***HPL*** | ***HAL*** | ***DPL*** |
|---|---|---|---|---|---|---|---|
| $\tau$ | - | 1.186 | 1.703 | 3.0 | 5.469 | 3.0 | 7.021 |
| $T_{1.5\%}$ | - | 8,596 | 12,343 | 21,743 | 39,639 | 21,743 | 50,888 |
| $T_{2.3\%}$ | - | 5,606 | 8,050 | 14,181 | 25,851 | 14,181 | 33,188 |
| $\theta_0$ | - | - | 0.00176 | - | - | - | 0.00352 |
| $\theta_1$ | - | 99999.0 | 99999.0 | 0.53115 | 5.66680 | 0.53115 | 5.78003 |
| *SSD* | - | 0.03099 | 0.05453 | 0.03901 | 0.07891 | 0.03901 | 0.08465 |
| *R* | - | 0.22222 | 0.30667 | 0.28543 | 0.15457 | 0.28543 | 0.19457 |
| *Modality* | - | *unimodal* | *unimodal* | *bimodal* | *bimodal* | *bimodal* | *bimodal* |

***Cryptopygus antarcticus***

| *Statistics* | ***PCK*** | ***RPN*** | ***HPN*** | ***CPR*** | ***HPL*** | ***HAL*** | ***DPL*** |
|---|---|---|---|---|---|---|---|
| $\tau$ | 2.9 | - | 4.7 | 2.9 | - | - | 43.2 |
| $T_{1.5\%}$ | 52,139 | - | 84,502 | 52,139 | - | - | 776,699 |
| $T_{2.3\%}$ | 30,004 | - | 55,109 | 30,004 | - | - | 506,543 |
| $\theta_0$ | 253.125 | - | 11.56641 | - | - | - | - |
| $\theta_1$ | 51.64062 | - | 34.82109 | 99999.0 | - | - | 329.619 |
| *SSD* | 0.01232 | - | 0.02970 | 0.01733 | - | - | 0.25489* |
| *R* | 0.03753 | - | 0.02914 | 0.08000 | - | - | 0.37284 |
| *Modality* | *multimodal* | - | *multimodal* | *bimodal* | - | - | *bimodal* |

***Friesea grisea***

| *Statistics* | ***PCK*** | ***RPN*** | ***HPN*** | ***CPR*** | ***HPL*** | ***HAL*** | ***DPL*** |
|---|---|---|---|---|---|---|---|
| $\tau$ | - | - | 0.49023 | - | 2.92969 | 3.0 | - |
| $T_{1.5\%}$ | - | - | 11,395 | - | 68,101 | 69,735 | - |
| $T_{2.3\%}$ | - | - | 7,432 | - | 44,413 | 45,479 | - |
| $\theta_0$ | - | - | - | - | - | - | - |
| $\theta_1$ | - | - | 99999.0 | - | 3.60000 | 0.49873 | - |
| *SSD* | - | - | 0.00579 | - | 0.33101 | 0.04225 | - |
| *R* | - | - | 0.18272 | - | 0.40000 | 0.28543 | - |
| *Modality* | - | - | *unimodal* | - | *unimodal* | *bimodal* | - |

Table 3. Hierarchical analysis of molecular variance (AMOVA). Components of differentiation calculated with alternative clustering into individual islands (four groups) or, as discussed along the text, gathering these latter into KN and RL groups.

| population compositions | | Component of differentiation | | |
|---|---|---|---|---|
| | | Among groups $\Phi_{CT}$ | Among populations within groups $\Phi_{SC}$ | Among all populations $\Phi_{ST}$ |
| 4 groups | Variance component | 0.05931 (0.43206±0.01412) | 0.24119 (0.00880±0.00288) | 0.76032 (0.00) |
| | % variation | 5.59 | 22.74 | 71.67 |
| 2 groups | Variance component | 0.17425 (0.09384±0.00602) | 0.18963 (0.00293±0.00164) | 0.76032 (0.00) |
| | % variation | 15.50 | 16.87 | 67.63 (0.00) |

SUPPLEMENTARY MATERIAL

Table S1. Statistics of population genetic parameters in *C. a. antarcticus* and *F. grisea*. *n*, number of individuals; $N_H$, number of haplotypes; *h*, haplotype diversity; $\pi$, nucleotide diversity.

| Collecting site | | Coordinate | *Friesea grisea* | | | | | *Cryptopygus antarcticus* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | *n* | $N_H$ | *h* | $\pi$ | Cluster | *n* | $N_H$ | *h* | $\pi$ | Cluster |
| Potter Cove, King George I. | PCK | 62°14'S, 58°42'W | 10 | P3(10) | - | - | $C_{FG}2$ | 10 | H21(1), H22(2), H23(1), H24(1), H25(1), H26(1), H27(2), H28(1), | 0.9556 ± 0.0594 | 0.008846 ± 0.005260 | $C_{CA}1$ |
| Rip Point, Nelson I. | RPN | 62°15'S, 58°59'W | 10 | P3(10) | - | - | $C_{FG}2$ | - | - | - | - | - |
| Harmony Point, Nelson I. | HPN | 62°19'S, 59°10'W | 10 | P1(1), P2(1), P3(8) | 0.3778 ± 0.1813 | 0.000837 ± 0.000953 | $C_{FG}1$+$C_{FG}2$ | 10 | H6(1), H7(1), H8(1), H11(1), H15(1), H16(1), H18(1), H20(2), H44(1) | 0.9778 ± 0.0540 | 0.017332 ± 0.009759 | $C_{CA}2$+$C_{CA}3$ |
| Coppermine Peninsula, Robert I. | CPR | 62°23'S, 59°42'W | 10 | P3(10) | - | - | $C_{FG}2$ | 10 | H13(1), H14(1), H15(1), H17(1), H19(1), H20(2), H31(2), H32(1) | 0.9556 ± 0.0594 | 0.004207 ± 0.002777 | $C_{CA}3$ |
| Hannah Point, Livingston I. | HAL | 62°39'S, 60°36'W | 10 | P3(8), P5(1), P6(1) | 0.3778 ± 0.1813 | 0.004603 ± 0.003136 | $C_{FG}2$+$C_{FG}3$ | - | - | - | - | - |
| Devils Point, Livingston I. | DPL | 62°40'S, 61°11'W | 10 | P7(10) | - | - | $C_{FG}1$ | 10 | H10(1), H29(5), H30(1), H38(2), H39(1) | 0.7556 ± 0.1296 | 0.038799 ± 0.021109 | $C_{CA}2$+$C_{CA}3$+$C_{CA}4$ |
| Hurd Peninsula, Livingston I. | HPL | 62°41S, 60°23W | 10 | P3(9), P4(1) | 0.2000 ± 0.1541 | 0.000418 ± 0.000637 | $C_{FG}2$ | - | - | - | - | - |
| Total | | | 70 | 7 | | | | 40 | 28 | | | |

Table S2. List and positions (along the aligned sequences) of nucleotide and amino acid changes of haplotypes B to Q, with respect to A. Each substitution categorized according to the corresponding codon position. Numbers within brackets give the frequency of each haplotype across the samples sequenced.

| | Haplotype | Variable sites | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A (39) | G | C | T | C | C | A | T | C | T | C | G | T | T | T | G | G | A | C | T | A | C | T |
| 2 | B (4) | . | . | . | . | . | . | . | . | . | . | . | C | . | . | . | . | . | . | . | . | . | . |
| 3 | C (1) | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | G | . | . | . | . |
| 4 | D (1) | A | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| 5 | E (1) | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | T | . |
| 6 | F (2) | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | C | . | . | . |
| 7 | G (1) | . | . | . | . | . | . | . | . | . | . | . | . | . | . | A | . | . | . | . | . | . | . |
| 8 | H (1) | . | . | . | . | . | . | . | . | . | . | . | . | C | . | . | . | . | . | . | . | . | . |
| 9 | I (1) | . | T | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| 10 | J (10) | . | . | . | . | T | . | . | . | . | . | . | . | . | C | . | A | G | . | . | . | . | . |
| 11 | K (1) | . | . | . | . | . | . | C | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| 12 | L (1) | . | . | . | . | T | . | . | . | . | T | . | . | . | C | . | A | G | . | . | . | . | . |
| 13 | M (1) | . | . | C | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| 14 | N (1) | . | . | . | . | T | . | . | . | . | . | . | . | . | C | . | A | G | . | . | G | . | T |
| 15 | O (1) | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | C |
| 16 | P (2) | . | . | . | T | . | G | . | T | . | . | A | . | . | . | . | A | . | . | . | . | . | . |
| 17 | Q (1) | . | . | . | . | T | . | . | . | C | . | . | . | . | C | . | A | G | . | . | . | . | . |
| | Aligned position | 70 | 106 | 111 | 115 | 192 | 345 | 348 | 399 | 533 | 597 | 612 | 621 | 681 | 742 | 819 | 951 | 954 | 1010 | 1089 | 1110 | 1255 | 1491 |
| | Codon position | I | I | III | III | III | III | III | III | III | III | III | III | III | I | III | III | III | II | III | III | I | III |
| | aa change | V→I | | | | | | | | | | | | | | | | | A→G | | | | |
| | | Haplotype D | | | | | | | | | | | | | | | | | Haplotype C | | | | |

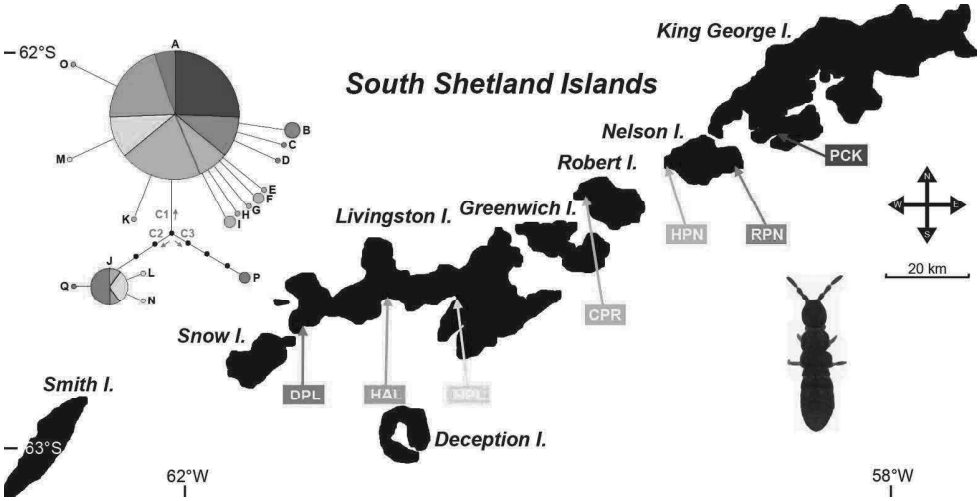


Figure 1. Map of the South Shetland Islands, with sampling localities indicated using a three-letter code (listed in Table 1). Top left, haplotype network observed between the seven analyzed populations. Haplotype nomenclature with one capital letter as in Table 1.
242x122mm (300 x 300 DPI)

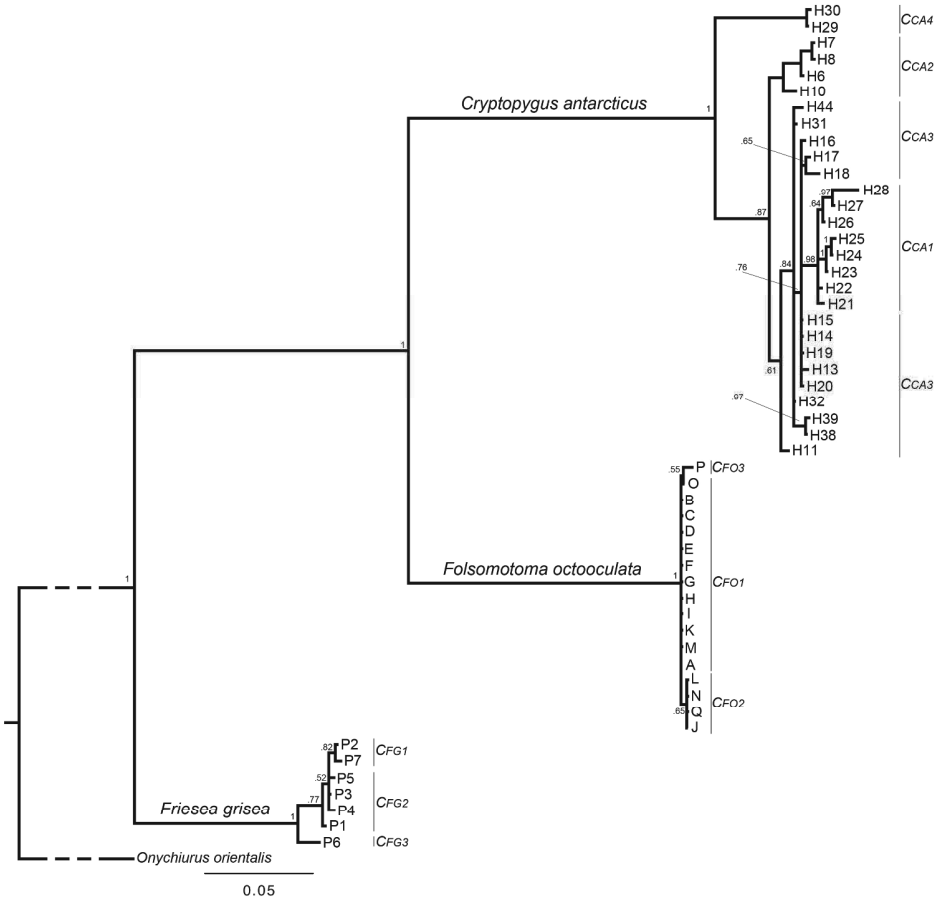


Figure 2. Phylogenetic tree, obtained using the Bayesian method, inclusive of all haplotypes obtained from the South Shetland Islands populations for the sympatric species C. a. antarcticus, F. octooculata and F. grisea. Haplotypes grouped according to the corresponding clusters, as obtained under the Bayesian model of clustering method (Cheng et al., 2013). Tree drawing obtained using FigTree, vers. 1.4.2 (Available at: http://tree.bio.ed.ac.uk/software/figtree/).
237x225mm (300 x 300 DPI)