diversity

MDPI

Article

Molecular Comparison among Three Antarctic Endemic Springtail Species and Description of the Mitochondrial Genome of *Friesea gretae* (Hexapoda, Collembola)

Antonio Carapelli [1], Claudio Cucini [1], Pietro Paolo Fanciulli [1], Francesco Frati [1], Peter Convey [2] and Francesco Nardi [1,*]

[1] Department of Life Sciences, University of Siena, Via A. Moro 2, 53100 Siena, Italy; antonio.carapelli@unisi.it (A.C.); claudio.cucini@student.unisi.it (C.C.); paolo.fanciulli@unisi.it (P.P.F.); francesco.frati@unisi.it (F.F.)

[2] British Antarctic Survey, NERC, High Cross, Madingley Road, Cambridge CB3 0ET, UK; pcon@bas.ac.uk

* Correspondence: francesco.nardi@unisi.it; Tel.: +39-0577-232-128

Received: 6 November 2020; Accepted: 25 November 2020; Published: 27 November 2020

Abstract: Springtails and mites are the dominant groups of terrestrial arthropods in Antarctic terrestrial ecosystems. Their Antarctic diversity includes a limited number of species, which are frequently endemic to specific regions within the continent. Advances in molecular techniques, combined with the re-evaluation of morphological characters and the availability of new samples, have recently led to the identification of a number of new springtail species within previously named, but ill-defined, species entities described in the last century. One such species, the neanurid *Friesea grisea*, originally described from sub-Antarctic South Georgia, was for many years considered to be the only known springtail with a pan-Antarctic distribution. With the recent availability of new morphological and molecular data, it has now been firmly established that the different representatives previously referred to this taxon from the Antarctic Peninsula and Victoria Land (continental Antarctica) should no longer be considered as representing one and the same species, and three clearly distinct taxa have been recognized: *F. antarctica*, *F. gretae* and *F. propria*. In this study, the relationships among these three species are further explored through the sequencing of the complete mtDNA for *F. gretae* and the use of complete mitogenomic as well as cytochrome *c* oxidase I data. The data obtained provide further support that distinct species were originally hidden within the same taxon and that, despite the difficulties in obtaining reliable diagnostic morphological characters, *F. gretae* is genetically differentiated from *F. propria* (known to be present in different locations in Northern Victoria Land), as well as from *F. antarctica* (distributed in the Antarctic Peninsula).

Keywords: invertebrate biodiversity; genetic distances; molecular clock; phylogeny; mitogenomes; Antarctic springtails

1. Introduction

The systematic of Antarctic springtails has recently undergone major reshuffling, with some established species split in new taxonomic entities. This is true also for species of genus *Friesea*, that were lately defined using morphological and molecular data. One of these new taxa, *Friesea gretae* [1] is the main subject of this study, due to the peculiar (and restricted) distribution on protected areas and geographical separation to other congeneric species living in its surroundings.

For more than a century, the neanurid springtail *Friesea grisea* (Schäffer, 1891), originally described from sub-Antarctic South Georgia, was considered to be the only Collembola species with a "pan-Antarctic"

distribution, also present in coastal areas of the western Antarctic Peninsula and the archipelagoes of the Scotia Arc (the maritime Antarctic) and in various locations in Victoria Land in continental Antarctica. New morphological studies have recently demonstrated that this taxon, as identified in the original description, is endemic and restricted to South Georgia [2]. In the latter part of the 20th century, classical taxonomic studies erected two new *Friesea* species from the maritime Antarctic South Orkney and South Shetland archipelagoes (*Friesea woyciechowskii*; [3]) and south-east Alexander Island (*Friesea topo*; [4]). More recently, Greenslade [2,5] also described *Friesea fantaba* from South Georgia and *Friesea eureka* from Enderby Land. Other than *F. topo*, which was described as a newly discovered species from a previously unsurveyed area [6], the remaining species had previously been recorded as *F. grisea* (Figure 1).

Figure 1. Map of the distribution of Antarctic *Friesea* species along major biogeographic regions of the continent. Species, with relative dates of identification, are listed according to their mention along the text as follows: (1) *Friesea gretae* (Greenslade and Fanciulli, 2020); (2) *Friesea grisea* (Schäffer, 1891); (3) *Friesea woyciechowskii* (Weiner, 1890); (4) *Friesea topo* (Greenslade, 1995); (5) *Friesea fantaba* (Greenslade, 2018); (6) *Friesea eureka* (Greenslade, 2018); (7) *Friesea antarctica* (Willem, 1902); (8) *Friesea propria* (Greenslade and Fanciulli, 2020). The "Gressitt" line demarks the discontinuity between the biota of Antarctic Peninsula and that of Continental Antarctica. Data adapted from Greenslade [2] and Carapelli et al. [1].

Following the collection of new morphological and molecular data and their application to the study of the systematics and biogeography of this genus in Antarctica, it became evident that, despite the general uniformity in morphological characters, there was a clear genetic distinction between "*Friesea grisea*" samples collected in maritime and continental Antarctic localities [7,8]. Greenslade [2] first identified the "true identity" of maritime Antarctic material, reinstating the taxonomic status of *Friesea antarctica* (previously described as *Achorutoides antarcticus* by Willem [9]) and then put in synonymy by Wahlgren [10]), and provided a taxonomic key enabling the identification of *Friesea* taxa from the Antarctic Peninsula and Scotia Arc.

With the advent of molecular data applied to the study of the systematics of the genus, and with increasing availability of recently collected samples, analyses of species delimitation have converged with conclusions based on morphological studies. Initially these collectively suggested that *F. antarctica*

is the most common species of the genus in locations along the west coast of the Antarctic Peninsula and that continental Antarctic samples belonged to at least one different species [7,8]. In this respect, two complete mitogenomes were first compared from specimens collected in maritime and continental Antarctica, identifying a remarkable genetic divergence between the two (21.7% genetic distance between protein coding genes). This strongly challenged the taxonomic status of populations from Victoria Land that were still referred to as *F. grisea* at that time.

More recently, multilocus and mitogenomic analyses, in combination with morphological studies, provided additional support for this view and suggested the presence of a further additional species. The taxonomic discontinuity between maritime and continental Antarctic taxa became even more undeniable, while at the same time additional diversification was proposed at the local level in continental Antarctica, specifically in different parts of Victoria Land [1]. Two additional species (again, once collectively identified as *F. grisea*) are now recognized in Northern Victoria Land: *F. propria*, distributed between Harrow Peaks (74°06′ S, 164°48′ E) and Crater Cirque (74°06′ S, 164°48′ E), and *F. gretae*, at present reported only from Cape Hallett (72°19′ S, 170°13′ E) and Redcastle Ridge (72°26′ S, 169°56′ E) in the Cape Hallett Peninsula. In this study, we present the complete mitochondrial genome sequence of *F. gretae*, as well as some *cox1* haplotypes from Redcastle Ridge, and analyze these in conjunction with data from the literature in order to assess: (1) if levels of genetic diversification of *F. gretae* mtDNA, with respect to both *F. antarctica* and *Friesea propria*, are consistent with their status as three well-defined species; (2) whether the measured genetic variability, at intra- and inter-specific levels, is consistent with a well-defined "barcode gap" supporting species delimitation; and (3) the timescale over which the three species diverged from each other.

2. Materials and Methods

2.1. Sampling, Sequencing and Genome Annotation

Ten specimens of *F. gretae* were manually collected at Cape Hallett (72°19′ S, 170°13′ E; ASPA No. 106), close to small fresh water ponds in the vicinity of the campsite a few hundred meters from the Adélie penguin rookery, and seven other individuals were sampled from Redcastle Ridge (72°26′ S, 169°56′ E), a neighboring locality along the Hallett Peninsula (Figure 2). Sampling was performed during the XXXIV Italian National Antarctic Program (PNRA) expedition (2018/2019). Individuals were subjected to preliminary morphological identification and stored in alcohol or at −80 °C until further morphological and molecular analyses.

Total DNA was purified from a pool of 10 specimens of *F. gretae* from Cape Hallett and from 7 individual specimens for Redcastle Ridge (the latter used for population genetics screening). The sample was pooled with the DNA of 10 additional arthropod species, chosen that differ at least at the family level to exclude the possibility of cross-assembly, for sequencing. NGS sequencing was carried out at Macrogen Europe using the TruSeq Nano DNA chemistry on an Illumina platform to produce 151 bp paired end reads. Sequences have been deposited on NCBI under BioProject accession PRJNA673074.

NGS sequences were quality controlled in FastQC (ver. 0.11.9; available at https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and trimmed in Trimmomatic (LEADING:28 TRAILING:28 MINLEN:100; ver. 0.39) [11].

Sequences were assembled using NOVOPlasty (ver. 3.8.3) [12] on untrimmed reads (following advice from the manual) using available *cox1* sequences as seed. The full collection of reads was initially enriched for mitochondrial sequences using the filter_reads.pl script from the same package. Final assemblies were in turn produced under two values of k (77 and 101). The resulting circular contigs were compared with assemblies obtained using a different and unrelated method (MEGAHIT ver. 1.2.9 [13]; data not shown) to confirm the correctness of the assembly.

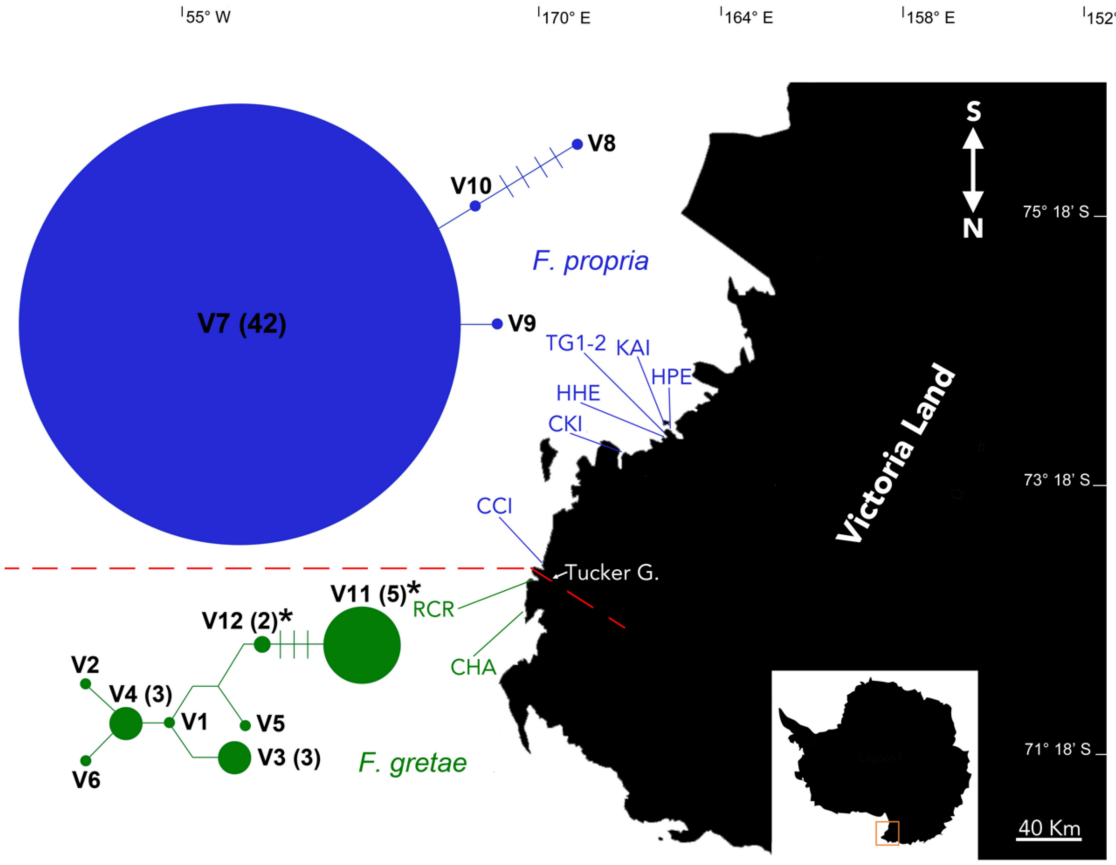

Figure 2. Map of Central and Northern Victoria Land with *Friesea* species distribution. Haplotype networks on the left side: green for *F. propria* and blue for *F. gretae*. Letters with associated numbers correspond to *cox1* haplotypes; haplotypes V1–V10 are obtained from Torricelli et al. [7], whereas V11 and V12 are sequenced for this study (and highlighted with asterisks in the network); numbers within brackets identify their frequency. Species distribution boundary along the Tucker Glacier highlighted with dotted red line. Acronyms identify localities as described in Carapelli et al. [1] and listed in Table S1; colors distinguish species presence in localities and species-specific haplotypes (blue for *F. propria* and green for *F. gretae*).

Single base ambiguities (~20, not unexpected given that sequences were obtained from a pool of individuals) were resolved on a majority rules basis by remapping reads on the regions in question using bbmap (ver. 38.84; sourceforge.net/projects/bbmap/) and visualizing alignments in IGV (ver. 2.8.2) [14]. The final genome sequence was preliminarily annotated using Mitos [15] and manually curated based on a comparison with the mitochondrial genomes of the congeneric *F. propria* and *F. antarctica* to produce final annotations.

The *cox1* "barcode" marker was sequenced in six specimens of *F. gretae* from Redcastle Ridge following the methods described by Torricelli et al. [7]. The haplotypes obtained were compared with those previously deposited in GenBank, corresponding to samples from Victoria Land, following Carapelli et al. [16]. A matrix of genetic distances was calculated using PAUP* 4b10 [17] and a haplotype network constructed using TCS1.21 [18].

2.2. Genomic and Genetic Features

Start and stop codons of PCGs were identified and gene boundaries were inspected for the presence of gene overlaps. tRNA structures, provided by Mitos, were similarly identified. Basic sequence statistics and genetic distances were calculated using PAUP* 4b10 [17]. Strand asymmetry was described using the formulas: AT-skew = [A − T]/[A + T] and CG-skew = [C − G]/[C + G] [19].

2.3. Single-Locus Phylogenetic Relationships of Friesea Species

All *cox1* sequences available in Bold and GenBank for 15 *Friesea* species were downloaded. Most of them were discarded as duplicates or due to their short length. The remaining 29 sequences, from 12 *Friesea* species (Table S2), were retained for analysis and collated with the two new haplotypes sequenced from *F. gretae* sampled in Redcastle Ridge and that of *Bilobella aurantiaca*, the latter used as outgroup in the phylogenetic analyses. Sequences were aligned with Clustal Omega [20] to produce an alignment of 478 bp. The latter was partitioned in three subsets according to the three nucleotide codon positions.

The most appropriate partitioning scheme and associated evolutionary model for each charset were determined using PartitionFinder (v. 2.1.1) [21] using the Akaike Information Criterion. Bayesian phylogenetic analyses were conducted on the complete dataset using MrBayes (v. 3.2.7a) [22]. Two parallel runs of four chains were conducted for 100 million generations, and the resulting trees and posterior probabilities were summarized. Convergence was assessed in Tracer (v. 1.7.1) [23], and the initial 20% of generations was excluded as burn-in. The resulting tree was visualized using Figtree (v. 1.4.4) [24].

2.4. Mitogenomic Genetic and Phylogenetic Relationships of Friesea Species

Genome comparisons were conducted between the three complete mtDNAs available for *Friesea* species. In detail: (a) the mitochondrial genome of *F. gretae* collected in Cape Hallett, Victoria Land (continental Antarctica) and submitted to GenBank under the accession number MT644085; (b) the sequence previously described by Torricelli et al. [8] as belonging to "*F. grisea*" collected at Hannah Point, Livingston Island, South Shetland Islands (maritime Antarctica: 62°39′16″ S, 60°36′48″ W) and here referred to as *F. antarctica* (GenBank accession number EU016196); (c) the sequence previously described by Torricelli et al. [8] as belonging to *F. grisea* collected in Kay Island, Victoria Land (continental Antarctica: 74°04′ S, 165°18′ E) and here referred to as *F. propria* (GenBank accession number EU0124719). The reassignment of taxon names is discussed in Carapelli et al. [1].

The newly determined sequence of *F. gretae* was collated into the dataset described by Leo et al. [25] and a dating analysis was performed as therein described (Figure 2 in Leo et al. [25]). Briefly, first and second codon positions of alignable portions of all PCGs were used in conjunction with a Log normal relaxed clock and four (two in the ingroup, two in the outgroup) calibration points to estimate nodal ages in BEAST (ver. 1.10.4) [24]. In addition, all PCGs of the three available mitogenomes, for each of the species *F. antarctica*, *F. gretae* and *F. propria*, were aligned and concatenated into a matrix of 14,995 bp aligned sequences that was used for comparisons in genome divergence.

3. Results

3.1. Genome Features

The mtDNA of *F. gretae* was 15,439 bp in length, 14 nucleotides longer and 3 nucleotides shorter that those of *F. antarctica* and *F. propria*, respectively. The gene order is identical to that most frequently observed in springtails [25] and in Pancrustacea [26]. It includes 13 PCGs required for the assembly of protein complexes of the mitochondrial oxidative phosphorylation system: subunits 6 and 8 of the ATPase (atp6 and atp8), cytochrome *c* oxidase subunits 1 to 3 (cox1–3), apo cytochrome b (cob) and NADH dehydrogenase subunits 1 to 6 and 4L (nad1–6 and nad4L). In addition, 2 genes encoding for ribosomal RNAs (small and large subunits [rrnS and rrnL]) and 22 genes encoding for transfer RNAs (trnXs) are also present. Twenty-three genes are oriented on the J- or (−) strand and 14 on the N- (+) strand (Table S3). The so-called A + T-rich region in *F. gretae*, as in the majority of arthropods, is located between *rrnS* and *trnI*. The nucleotide composition of the *F. gretae* J-strand is biased towards a higher number of Ts (40%), whereas the A + T content (71%) is identical to that of *F. propria* and slightly lower than in *F. antarctica* (72.2%). Parameters of nucleotide skew, calculated for all three *Friesea* species, are negative for the AT-skew (~−0.13) and slightly positive for the CG-skew (up to 0.025) on

the J-strand. The limited number of nucleotide differences among the three *Friesea* species considered here are mainly confined in non-coding, trnX and rDNA sequences, although several dissimilarities in length also occur within PCGs.

The number of nucleotide differences between tRNA encoding genes (1335 aligned nucleotides) was 38 between *F. gretae* and *F. propria*, and 195 between *F. antarctica* and both *F. gretae* and *F. propria*. In the rRNA-encoding genes (1973 aligned characters), 440 nucleotide changes were present between *F. antarctica* and *F. propria*, 436 between *F. antarctica* and *F. gretae,* and 144 between *F. gretae* and *F. propria*. The A + T-rich regions (1335 aligned nucleotides) differed by 85 nucleotides between *F. antarctica* and *F. propria*, 73 nucleotides between *F. antarctica* and *F. gretae*, and 42 nucleotides between *F. gretae* and *F. propria*.

Apart from start and stop codons, most of the length differences of PCGs were between *F. antarctica* vs. *F. gretae/F. propria*, with one inserted/deleted codon in each of *atp8*, *nad1* and *nad5*. For *nad4L*, in *F. propria* this gene was one codon longer than in the other species, whereas *nad6* in *F. propria* and *F. gretae* were two codons shorter and three codons longer compared with *F. antarctica*, respectively.

In the mtDNA of the three *Friesea* species under study, canonical initiation codons (ATA or ATG, both encoding for methionine) were present in most PCGs, with few exceptions: in *atp8* and *cox2* of all species the first codon was ATT or ATC (isoleucine), while ATT was used as starting codon in *F. antarctica* and *F. propria* for *nad2*, and in *F. gretae* for *nad6* (Table S3).

3.2. Genetic Distances

Genetic distances (p-distances), calculated for the *cox1* dataset among 12 *Friesea* species (Table S4), ranged between 9% (*F. propria* HQ315682 vs. *F. gretae* (several comparisons)) and 23% (*F. najtae* KY753925 vs. *F. claviseta* KT808354/*F. millsi* KM620437, GU657122).

Amongst the Antarctic species, *F. antarctica*, *F. gretae*, *F. propria* and *F. topo*, the highest interspecific genetic divergence (18%) was that between *F. topo* MN957905 and *F. antarctica* HQ315674, whereas the lowest (9%) was between *F. propria* and *F. gretae*. Intraspecific variability within *F. antarctica*, *F. gretae* and *F. propria* (highlighted in grey in Table S4) ranged between 0.2% and 2.3%. Interspecific variability between these same three species varied from 8.8% to 9.8% between *F. gretae* and *F. propria*, and from 14.4% to 17.4% between *F. antarctica* vs. *F. gretae* and *F. antarctica* vs. *F. propria*. Comparison of intra- and inter-specific variability among haplotypes of *F. antarctica*, *F. gretae* and *F. propria* highlighted a marked barcode gap (between 2% and 9%) clearly demarking the range of differentiation among within and between species genetic variability. The barcode gap calculated for all *Friesea* species (Figure 3) ranges between 2.3% to 8.8%, with intraspecific and interspecific differentiation only overlapping when the two *F. flaviseta* specimens are compared. When complete mtDNA sequences of Antarctic *Friesea* species (three whole genomes currently available) were compared, estimates of genetic variability between *F. antarctica* vs. *F. propria* and *F. gretae* vs. *F. antarctica* were 23% (*p*-distances), and that between *F. propria* and *F. gretae* was 10%.

3.3. Haplotype Screening

Sequencing of Redcastle Ridge specimens identified two new *F. gretae* haplotypes that correspond to genetic variants V11 (two sequences) and V12 (five sequences) (the latter deposited on GenBank under the accession numbers MW136278 and MW136279, respectively). The remaining V1–V10 haplotypes correspond to sequences of *F. gretae* from Cape Hallett, whereas V7–V10 are those that have been sampled in other Victoria Land localities and belong to populations of *F. propria*. Gene genealogies represented by network estimation of all known haplotypes from Victoria Land localities provided a graphical view demarking the clear separation between Cape Hallett + Redcastle Ridge samples (both of *F. gretae*) with respect to those from other Victoria Land localities (from *F. propria*), a genetic and taxonomical discontinuity between the two species corresponding, geographically, to the position of the Tucker Glacier (Figure 2). The two sub-networks are separated by an estimated 42–47 mutational steps.

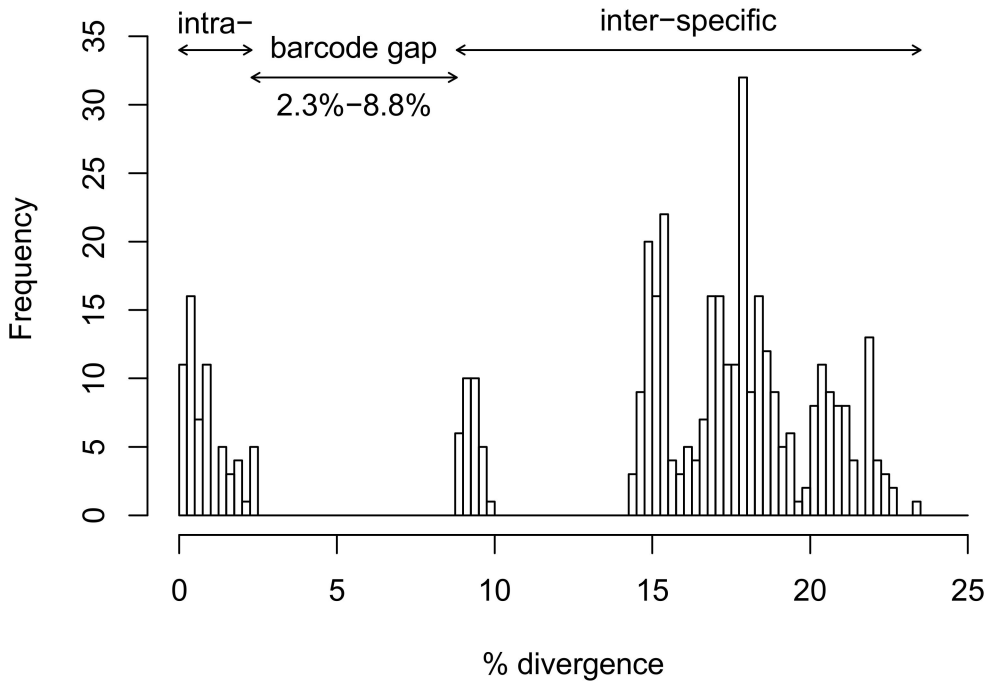

Figure 3. Histograms of p-distance comparisons between the *Friesea* species under study with highlighted intraspecific and interspecific ranges and barcode gap.

3.4. Molecular Clock

Molecular comparison performed using the concatenated set of PCGs analyzed by Leo et al. [25], plus the newly obtained sequences for *F. gretae*, were used to estimate divergence times between *F. antarctica*, *F. gretae* and *F. propria*. The resulting dated tree conforms to that depicted by Leo et al. [25] both in topology and timing. A fully supported cluster is observed including the three *Friesea* sequences (i.e., *F. antarctica*, *F. gretae* and *F. propria*) (Figure S1). The basal diversification appears to be between *F. antarctica* from the Antarctic Peninsula and the other two sequences from Victoria Land. The node is dated at 79 Mya (71–107 95% HPD). Within Victoria Land, the diversification between *F. propria* and *F. gretae* is dated at 16 Mya (14–25 95% HPD). The differentiation between *Cryptopygus antarcticus* (NC_010533) and *C. terranovus* (NC_037610) that, although not phylogenetically related, display a similar geographic distribution in comparison with *F. antarctica* and *F. propria*, is dated at 120 Mya (94–147 95% HPD).

3.5. Phylogenetic Analysis of Friesea Species

The phylogenetic tree obtained from the Bayesian analysis of the *cox1* data set is not conclusive for deeper relationships, probably as a consequence of sequence saturation and the limited length of the analyzed fragment, with several basal nodes not supported by sufficient statistical values (posterior probabilities) (Figure 4). Nevertheless, all taxonomically recognized species represented by more than one sequence formed monophyletic clusters, with the possible exception of *Friesea claviseta*, which remained unresolved.

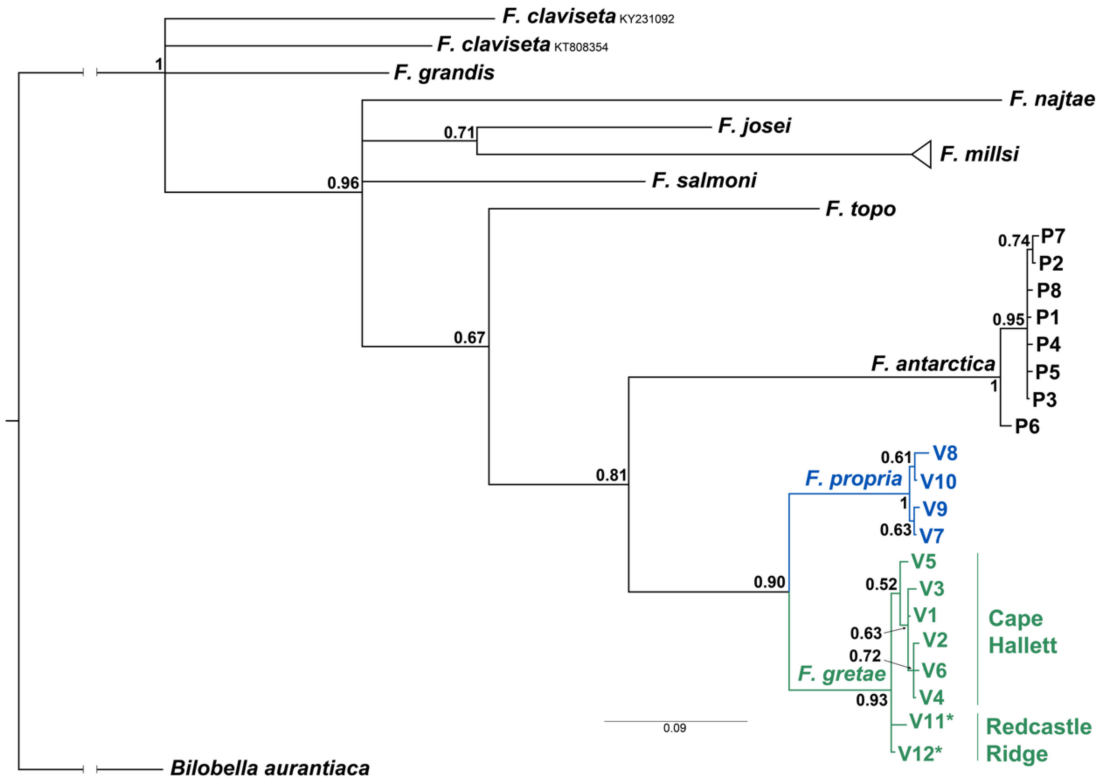

Figure 4. Phylogenetic tree of *Friesea* species based on *cox1* sequences and inferred with Bayesian method. Numbers at nodes indicate values of posterior probabilities.

4. Discussion

The close relationship between *F. gretae* and *F. propria* (both from Victoria Land) is in agreement with the multilocus phylogenetic study performed by Carapelli et al. [1]. Among the Antarctic representatives of *Friesea* that have been subjected to molecular studies, *F. topo* appears to be basal to all others, although the node supporting the four Antarctic species as a monophyletic group shows only marginal support.

From a biodiversity and conservation standpoint, Cape Hallett is a significant area (ASPA No. 106) and hosts the richest terrestrial biodiversity in this region of the Ross Sea coastline, with about 46 lichen species, 9 bryophytes and 11 microarthropods [27,28]. Springtails are the dominant microarthropod group, being present at densities two to three orders of magnitude greater than mites [29]. The neanurid springtail *F. gretae* is commonly observed in the vicinity of small freshwater ponds at Cape Hallett, as well as on the volcanic rocks of Redcastle Ridge. Based on currently available records, the species appears to be endemic to these two specific parts of the Hallett Peninsula.

Friesea gretae is morphologically distinguishable from *F. antarctica* (Antarctic Peninsula) and *F. propria* (Victoria Land, south of the Tucker glacier). The main differences are in body size, as it is significantly smaller than *F. antarctica* and larger than *F. propria*, and the shape of the body, which is more cylindrical and slenderer than in the other two species. Other finer scale characters useful for species identification are the shape of the a1 chaeta on Abd. VI and the number of chaetae on the ventral tube (see Table 4 in [1] for a complete list of diagnostic characters).

The evolutionary history of Antarctic springtails has been studied in increasing detail using molecular tools over the last two decades [30–33] and a growing consensus has been obtained for their long-term persistence in Antarctica and survival in local refugia during glacial maxima [34,35]. Present-day populations are the descendants of lineages that colonized suitable environments up to many millions of years ago, subsequently persisting through multiple and intense periods of glaciation [36,37].

In general, molecular data suggest an evolutionary scenario characterized by limited dispersal of Antarctic springtails within the continent, due to their poor vagility and the occurrence of insurmountable geographical barriers, to the extent that, even at local scales of as little as tens of kilometers, some conspecific populations appear to be highly genetically differentiated [34]. In this respect, the geographically small distance (28 km between Crater Cirque and Redcastle Ridge) between different *Friesea* populations collected on either side of the Tucker Glacier contrasts markedly with the sizable genetic differentiation between specimens from these two localities, leading to the conclusion that the glacier is a significant barrier preventing dispersal in this region [1,34]. Similar levels of genetic divergence (~9–10%) were estimated based on *cox1* sequences and of the complete mitochondrial genome between *F. propria* and *F. gretae*. Notably, this level of divergence is almost half of that observed between *F. antarctica* and *F. gretae/F. propria*, groups that are separated by almost four thousand kilometers of ice shelves.

High levels of cryptic diversity have been recognized in various species of springtails [38–40] including some occurring in the Antarctic Continent [30]. The three congeneric *Friesea* species analyzed in this study have genetic interspecific divergence ranges (9% to 17.4%) consistent with values reported for other congeneric springtail species (8% to 22%) [31,40–42] and values of genetic divergence ≤2% and ≥9% in intraspecific and interspecific comparisons, respectively (i.e., a 7% barcode gap).

Levels of genetic divergence with Antarctic springtail species of <5% have been interpreted as being compatible with isolation occurring within the last 2 Mya [35], whereas between species comparisons suggest more ancient diversification, as is apparently the case in *Friesea* species. The dating analysis carried out in this study generated two independent dates for divergence within separate congeneric species pairs present in the Antarctic Peninsula and Victoria Land (i.e., in *Cryptopygus* and *Friesea*). These dates point to a mid-Cretaceous differentiation. The differentiation between the two species of *Friesea* within Victoria Land is somewhat more recent in absolute terms, from the early Tertiary, but is still ancient considering that two sibling species are being compared. The finding of deep divergences between sibling species is not unusual in Collembola, with other such events being documented between two lineages of *Parisotoma notabilis* [25] and in southern European *Lepidocyrtus* [43].

Based on molecular clock data, the split between *F. antarctica* and the ancestor of *F. gretae/F. propria* occurred around 79 Mya. In a paleoclimatic context, this corresponds to a geological period when Antarctica was characterized by a semi-tropical climate and long before the formation of continental ice sheets. The separation between *F. propria* and *F. gretae* took place more recently (around 16 Mya). By this time, the Antarctic continent had become geographically isolated from the other southern continents, ice sheets had started to form and the region's cooling had been accelerated by the formation of the Antarctic Circumpolar Current and atmospheric circulations [44].

It is also not long before the very extensive glaciations in the Miocene, around 12–14 Mya, which are associated with the final extinction of remaining "tundra" communities in the Transantarctic Mountains [37,45].

Whatever the macro-evolutionary processes responsible for generating species diversity among geographical regions in *Friesea*, it appears evident that: (1) lineage diversification took place over extended geological periods and under very different environmental conditions; (2) species divergence events were not necessarily connected with large geographical distances (e.g., the ~28 km between Redcastle Ridge and Crater Cirque, the latter being the northern-most locality where *F. propria* has been recorded).

All dating estimates are subject to different sources of uncertainty and should be considered with caution. One possible source of uncertainty, in closely related species, is lineage sorting. However, given that the three species considered here, based on a larger mitochondrial (*cox1*) dataset, do not share mitochondrial haplotypes and that the difference between intra- and interspecific divergence values (barcode gap) is substantial, lineage sorting does not appear likely to apply here. An alternative, systematic source of uncertainty may lay in the use of calibration points that are very distant, both in terms of time and taxonomy, from the group of interest. This cannot be discounted in the current analysis,

as available calibration points within Collembola are limited to two Palaeozoic nodes [37,46–48] for a compilation of fossil Collembola that may be deployed provided adequate taxon sampling is available). Future dating analyses specifically focused on Antarctic divergences are clearly required, although the associated difficulties (e.g., the availability of appropriate calibration points that do not themselves derive from phylogeographic interpretations, considering that the resulting dates are to be used to test phylogeographic hypotheses) at present preclude this possibility.

Supplementary Materials: The following are available online at http://www.mdpi.com/1424-2818/12/12/450/s1. Figure S1: Dated tree based on first/second positions of concatenated PCGs of mtDNA sequences, see Leo et al. [25] for complete prior information. Numbers highlight node ages expressed in Mya. Colors different from black indicate the three *Friesea* species analyzed in this study. In bold, dates for nodes connecting *Friesea* species, Table S1: Overview of sampling localities, with respective label (applied in Figure 2) and geographic coordinates; the haplotypes detected in each population are shown with capital letter "V" from Victoria Land and numbers, Table S2: List of *Friesea* species used for genetic distance calculation and phylogenetic inference using the *cox1* data set with their Accession number and correspondence to haplotype symbol, Table S3: Annotation, nucleotide composition and other features of the mitochondrial genome of *Friesea gretae*, Table S4: genetic distances (*p*-distances) calculated for the *cox1* data set. In grey, comparisons within Antarctic *Friesea* species. Boxed values highlight interspecific genetic distances among *F. antarctica*, *F. gretae* and *F. propria*.

Author Contributions: Conceptualization, A.C., F.N. and P.C.; formal analysis, A.C., C.C. and F.N.; resources, A.C., F.F. and P.C.; data curation, A.C., F.N., P.P.F., C.C. and P.C.; writing—original draft preparation, A.C., F.N., P.C. and P.G.; writing—review and editing A.C., P.C. and F.N.; visualization, A.C., C.C., F.N. and P.P.F.; supervision, A.C., F.F., P.C. and P.P.F.; funding acquisition, A.C., F.F., F.N. and P.C. All authors approved the final manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Italian Program of Research in Antarctica (PNRA16_00234). Partial support was also provided by the University of Siena. P.C. is supported by NERC core funding to the British Antarctic Survey's 'Biodiversity, Ecosystems and Adaptation' Team. The paper also contributes to the SCAR 'State of the Antarctic Ecosystem' (AntEco) international program.

Acknowledgments: We wish to thank the Department of Biotechnology, Chemistry and Pharmacy of the University of Siena for granting the use of the 'Marie' computational cluster, and Andrea Bernini for continuous and competent assistance with the bioinformatic analysis. We also thank two anonymous reviewers for their insightful comments.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

ASPA	Antarctic Specially Protected Area
HPD	Highest Posterior Density
mtDNA	mitochondrial DNA
PCGs	protein encoding genes
Mya	million years ago
bp	base pairs
NGS	Next Generation Sequencing

References

1. Carapelli, A.; Greenslade, P.; Nardi, F.; Leo, C.; Convey, P.; Frati, F.; Fanciulli, P.P. Evidence for cryptic diversity in the "pan-Antarctic" springtail *Friesea antarctica* and the description of two new species. *Insects* **2020**, *11*, 141. [CrossRef]
2. Greenslade, P. An Antarctic biogeographical anomaly resolved: The true identity of a widespread species of Collembola. *Polar Biol.* **2018**, *41*, 969–981. [CrossRef]
3. Weiner, W.M. *Friesea woyciechowskii* sp. n. (Collembola, Neanuridae) from King George Island South Shetland Islands, Antarctica. *Pol. Pismo Entomol.* **1980**, *50*, 491–495.
4. Greenslade, P. Collembola from the Scotia Arc and Antarctic Peninsula including descriptions of two new species and notes on biogeography. *Pol. Pismo Entomol.* **1995**, *64*, 305–319.

5. Greenslade, P. A new species of *Friesea* (Collembola: Neanuridae) from the Antarctic Continent. *J. Nat. Hist.* **2018**, *52*, 2197–2207. [CrossRef]
6. Convey, P.; Smith, R.I.L. The terrestrial arthropod fauna and its habitats in northern Marguerite Bay and Alexander Island, maritime Antarctic. *Antarct. Sci.* **1997**, *9*, 12–26. [CrossRef]
7. Torricelli, G.; Frati, F.; Convey, P.; Telford, M.; Carapelli, A. Population structure of *Friesea grisea* (Collembola, Neanuridae) in the Antarctic Peninsula and Victoria Land: Evidence for local genetic differentiation of pre-Pleistocene origin. *Antarct. Sci.* **2010**, *22*, 757–765. [CrossRef]
8. Torricelli, G.; Carapelli, A.; Convey, P.; Nardi, F.; Boore, J.L.; Frati, F. High divergence across the whole mitochondrial genome in the 'pan-Antarctic' springtail *Friesea grisea*: Evidence for cryptic species? *Gene* **2010**, *449*, 30–40. [CrossRef] [PubMed]
9. Willem, V. Collemboles Expédition Antarctique Belge Résultants du Voyage du S Y Belgica 1897–1899. *Rapports Scientifique Zoologie* **1902**, *9*, 1–19.
10. Wahlgren, E. Antarktische und subantarktische Collembolen gessammelt von der schwedischen Sudpolar expedition. In *Wissenschaft Ergeb Schwed Sudpolar–Expedition 1901–1903*; Lithographische Institut des Generalstabs: Stockholm, Sweden, 1908; pp. 1–22.
11. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [CrossRef] [PubMed]
12. Dierckxsens, N.; Mardulyn, P.; Smits, G. NOVOPlasty: De novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* **2017**, *45*, e18. [CrossRef] [PubMed]
13. Li, D.; Liu, C.M.; Luo, R.; Sadakane, K.; Lam, T.W. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **2015**, *31*, 1674–1676. [CrossRef] [PubMed]
14. Robinson, J.T.; Thorvaldsdóttir, H.; Winckler, W.; Guttman, M.; Lander, E.S.; Getz, G.; Mesirov, J.P. Integrative Genomics Viewer. *Nat. Biotechnol.* **2011**, *29*, 24–26. [CrossRef] [PubMed]
15. Bernt, M.; Donath, A.; Jühling, F.; Externbrink, F.; Florentz, C.; Fritzsch, G.; Pütz, J.; Middendorf, M.; Stadler, P.F. MITOS: Improved de novo metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.* **2013**, *69*, 313–319. [CrossRef] [PubMed]
16. Carapelli, A.; Convey, P.; Frati, F.; Spinsanti, G.; Fanciulli, P.P. Population genetics of three sympatric springtail species (*Hexapoda*: *Collembola*) from the South Shetland Islands: Evidence for a common biogeographic pattern. *Biol. J. Linn. Soc.* **2017**, *120*, 788–803. [CrossRef]
17. Swofford, D.L. *PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4*; Sinauer Associates: Sunderland, MA, USA, 2003.
18. Clement, M.; Posada, D.; Crandall, K.A. TCS: A computer program to estimate gene genealogies. *Mol. Ecol.* **2000**, *9*, 1657–1660. [CrossRef]
19. Perna, N.T.; Kocher, T.D. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J. Mol. Evol.* **1995**, *41*, 353–358. [CrossRef]
20. Madeira, F.; Park, Y.M.; Lee, J.; Buso, N.; Gur, T.; Madhusoodanan, N.; Basutkar, P.; Tivey, A.R.; Potter, S.C.; Finn, R.D.; et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **2019**, *47*, W636–W641. [CrossRef]
21. Lanfear, R.; Frandsen, P.B.; Wright, A.M.; Senfeld, T.; Calcott, B. PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* **2017**, *34*, 772–773. [CrossRef]
22. Ronquist, F.; Teslenko, M.; van der Mark, P.; Ayres, D.L.; Darling, A.; Hohna, S.; Larget, B.; Liu, L.; Suchard, M.A.; Huelsenbeck, J.P. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **2012**, *61*, 539–542. [CrossRef]
23. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior eummarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **2018**, *67*, 901–904. [CrossRef]
24. Suchard, M.A.; Lemey, P.; Baele, G.; Ayres, D.L.; Drummond, A.J.; Rambaut, A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **2018**, *4*, vey016. [CrossRef]
25. Leo, C.; Carapelli, A.; Cicconardi, F.; Frati, F.; Nardi, F. Mitochondrial genome diversity in Collembola: Phylogeny, dating and gene order. *Diversity* **2019**, *11*, 169. [CrossRef]
26. Boore, J.L. Animal mitochondrial genomes. *Nucleic Acids Res.* **1999**, *27*, 1767–1780. [CrossRef]

27. Gressitt, J.L. *Introduction*; Antarctic Research Series; American Geophysical Union: Washington, DC, USA, 1967; Volume 10, pp. 1–33.
28. Green, T.G.A.; Seppelt, R.D.; Brabyn, L.R.; Beard, C.; Türk, R.; Lange, O.L. Flora and vegetation of Cape Hallett and vicinity, northern Victoria Land, Antarctica. *Polar Biol.* **2015**, *38*, 1825–1845. [CrossRef]
29. Sinclair, B.; Scott, M.; Klok, C.; Terblanche, J.; Marshall, D.; Reyers, B.; Chown, S. Determinants of terrestrial arthropod community composition at Cape Hallett, Antarctica. *Antarct. Sci.* **2006**, *18*, 303–312. [CrossRef]
30. Stevens, M.I.; Greenslade, P.; Hogg, I.D.; Sunnucks, P. Examining Southern Hemisphere springtails: Could any have survived glaciation of Antarctica? *Mol. Biol. Evol.* **2006**, *23*, 822–874. [CrossRef]
31. McGaughran, A.; Torricelli, G.; Carapelli, A.; Frati, F.; Stevens, M.I.; Convey, P.; Hogg, I.D. Contrasting phylogeographical patterns for springtails reflect different evolutionary histories between the Antarctic Peninsula and continental Antarctica. *J. Biogeogr.* **2010**, *37*, 103–119. [CrossRef]
32. McGaughran, A.; Terauds, A.; Convey, P.; Fraser, C.I. Genome-wide SNP data reveal improved evidence for Antarctic glacial refugia and dispersal of terrestrial invertebrates. *Mol. Ecol.* **2019**, *28*, 4941–4957. [CrossRef]
33. Carapelli, A.; Fanciulli, P.; Frati, F.; Leo, C. Mitogenomic data to study the taxonomy of Antarctic springtail species (Hexapoda: Collembola) and their adaptation to extreme environments. *Polar Biol.* **2019**, *42*, 715–732. [CrossRef]
34. Collins, G.E.; Hogg, I.D.; Convey, P.; Barnes, A.D.; McDonald, I.R. Spatial and Temporal Scales Matter When Assessing the Species and Genetic Diversity of Springtails (Collembola) in Antarctica. *Front. Ecol. Evol.* **2019**, *7*, 76. [CrossRef]
35. Collins, G.E.; Hogg, I.D.; Convey, P.; Sancho, L.G.; Cowan, D.A.; Lyons, W.B.; Adams, B.J.; Wall, D.H.; Allan Green, T.G. Genetic diversity of soil invertebrates corroborates timing estimates for past collapses of the West Antarctic Ice Sheet. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 22293–22302. [CrossRef]
36. Fraser, C.I.; Terauds, A.; Smellie, J.; Convey, P.; Chown, S.L. Geothermal activity helps life survive glacial cycles. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 5634–5639. [CrossRef]
37. Convey, P.; Biersma, E.M.; Casanova-Katny, A.; Maturana, C.S. Refuges of Antarctic diversity. In *Past Antarctica: Paleoclimatology and Climate Change*; Oliva, M., Ruiz-Fernández, J., Eds.; Elsevier: Amsterdam, The Netherlands, 2020; pp. 181–200. [CrossRef]
38. Deharveng, L. Recent advances in Collembola systematics. *Pedobiologia* **2004**, *48*, 415–433. [CrossRef]
39. Carapelli, A.; Frati, F.; Fanciulli, P.P.; Nardi, F.; Dallai, R. Assessing species boundaries and evolutionary relationships in a group of south-western European species of *Isotomurus* (Collembola, Isotomidae) using allozyme data. *Zool. Scr.* **2005**, *34*, 71–79. [CrossRef]
40. Porco, P.; Bedos, A.; Greenslade, P.; Janion, C.; Skaržynski, D.; Stevens, M.I.; Jansen van Vuuren, B.; Deharveng, L. Challenging species delimitation in Collembola: Cryptic diversity among common springtails unveiled by DNA barcoding. *Invertebr. Syst.* **2012**, *26*, 470–477. [CrossRef]
41. Hogg, I.D.; Hebert, P.D.N. Biological identification of springtails (Hexapoda: Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes. *Can. J. Zool.* **2004**, *82*, 749–754. [CrossRef]
42. Stevens, M.I.; Hogg, I.D. Contrasting levels of mitochondrial DNA variability between mites (*Penthalodidae*) and springtails (*Hypogastruridae*) from the Trans-Antarctic Mountains suggest long-term effects of glaciation and life history on substitution rates, and speciation processes. *Soil Biol. Biochem.* **2006**, *38*, 3171–3180. [CrossRef]
43. Cicconardi, F.; Nardi, F.; Emerson, B.C.; Frati, F.; Fanciulli, P.P. Deep phylogeographic divisions and long-term persistence of forest invertebrates (Hexapoda: Collembola) in the North-Western Mediterranean basin. *Mol. Ecol.* **2010**, *19*, 386–400. [CrossRef]
44. Convey, P.; Stevens, M.I.; Hodgson, D.A.; Smellie, J.L.; Hillenbrand, C.-D.; Barnes, D.K.A.; Clarke, A.; Pugh, P.J.A.; Linse, K.; Cary, S.C. Exploring biological constraints on the glacial history of Antarctica. *Quat. Sci. Rev.* **2009**, *27–28*, 3035–3048. [CrossRef]
45. Lewis, A.R.; Marchant, D.R.; Ashworth, A.C.; Hedenäs, L.; Hemming, S.R.; Johnson, J.V.; Leng, M.J.; Machlus, M.L.; Newton, A.E.; Raine, J.I.; et al. Mid-Miocene cooling and the extinction of tundra in continental Antarctica. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 10676–10680. [CrossRef]
46. Rick, E.F. An entomobryid collembolan (Hexapoda: Collembola) from the Lower Permian of Southern Africa. *Paleontol. Afr.* **1976**, *19*, 141–143.

47. Whalley, P.; Jarzembowski, E. A new assessment of *Rhyniella*, the earliest known insect, from the Devonian of Rhynie, Scotland. *Nature* **1981**, *291*, 317. [CrossRef]
48. Sánchez-García, A.; Engel, M. Springtails from the Early Cretaceous Amber of Spain (Collembola: Entomobryomorpha), with an Annotated Checklist of Fossil Collembola. *Am. Mus. Novit.* **2016**, *3862*, 1–47. [CrossRef]