Radiation and speciation of the grasshopper genera *Parasphena* and *Gymnobothroides* in the East-African Mountains

Helene LoCascio Sætre

Centre for ecological and evolutionary synthesis (CEES)

Abstract

The East African Mountains can be viewed as continental versions of oceanic archipelagoes, and are therefore great starting point for studying phylogeographic topics. Climatic variations during the past million years have alternately connected and disconnected the forests on these mountains, and this has had great impact on species dependent on montane habitats. The flightless grasshopper species within the genera *Parasphena* and *Gymnobothroides* inhabit different niches in these mountains. *Gymnobothroides* has a wider niche than *Parasphena*, thus, *Gymnobothroides* might have a greater capability to spread than *Parasphena* during the dry phases. Not many genetic analyses have been conducted on these species in the past, and so this study might help shed light on the evolutionary history of these genera.

The main goal of this study was to investigate whether or the species within the grasshopper genera *Parasphena* and *Gymnobothroides* diverged because of climate induced forest fragmentation of the montane zone, or through adaptation novel niches on emerging volcanoes and mountains. Another aim was to find out whether congeneric lineages inhabiting the same mountain, but exploiting different niches were more related to each other than to other members of their genus inhabiting a similar niche, but on a different mountain (adaptive diversification) or whether niche sharing is related to common descent. The last question addressed was to what extent the phylogeny based on molecular markers would match the current classification and taxonomy of *Parasphena* and *Gymnobothroides*.

I sequenced two mitochondrial genes and used Bayesian Inference to investigate the phylogeny of these grasshoppers. Additionally, I used a mismatch distribution approach, calculated the p-distances between the sequences and used a coalescent analysis to further investigate the evolutionary history of these taxa.

This study finds evidence for diversification following a climate induced forest fragmentation for both genera, but certain populations within these genera show signs of having been able to spread and intermingle during periods of putative isolation. The evidence also supports the hypothesis that niche sharing is due to common descent since geographically dispersed

lineages with similar niches were found to be more closely related than were a pair of ecologically differentiated lineages living on the same mountain. Our study adds important data for a revised classification of the target genera. The grasshoppers inhabiting these different mountains are all very similar genetically, and there is little evidence for grouping them into separate species, as is often done in the current classification of these genera.

Acknowledgements

Writing a master thesis is rarely a one-person accomplishment, and for sure this thesis could not have been written were it not for a number of helpful people.

Firstly I would like to thank my supervisor Nils-Christian Stenseth, who made this entire project possible. I am so grateful to him for arranging the project, and connecting me to all the right people within the field. This thesis could certainly not have been completed without my other supervisor, Kjetil Lysne Voje. He was my main go-to-guy during the entire process. He answered all my questions, large or small, assisted me in the lab and during the analyses, proof-read and gave valuable comments each step of the way, and always managed to lower my stress level with his calming voice and reassuring smile. My third supervisor, Øystein Flagstad also deserves huge thanks for taking the time to come to Blindern to brain-storm with me, offering useful advice and comments. He also helped a lot by reading and commenting on the thesis.

Claudia and Andreas Hemp deserve special thanks for taking me in and guiding me around during my stay in Africa. Claudia Hemp collected most of the grasshoppers included in my study, besides those I caught myself, and even those could not have been caught without her. She taught me how to recognize them and where to find them, and she and Andreas drove me to every location I needed to go. Their knowledge extends far beyond grasshoppers and plants of Eastern Africa, and I learned so much during those weeks, ranging from how to shop for gemstones to Tanzanian culture and everyday life, and even some Swahili and German. Huge thanks to both of them for making my stay unforgettable!

I would also like to thank my dad, Glenn-Peter Sætre, for lifting the thesis to a new level through his helpful comments each time we discussed my project, and when he proof-read. He has extensive knowledge of evolution and speciation, and he is a great teacher, and I probably would not have understood my study system and the processes behind my results to the degree I do today were it not for him. He also helped me with the graphics. My mum, Maria LoCascio Sætre and my sister Camilla LoCascio Sætre also proof-read, commented and offered great advice. They also accompanied me down to Africa, making my first trip to that

continent much less scary and way more fun than it otherwise would have been. So in summary, I could not have completed this thesis were it not for my wonderful family and their priceless moral support and proof-reading abilities.

My boyfriend, Tim Dunker also helped me a great deal by proof-reading and commenting on everything I sent his way. He, without any biology background, provided that pair of fresh eyes that my thesis needed, and so helped me understand which parts needed a more thorough explanation. He also deserves thanks for putting up with me the last couple of stressful weeks.

Finally I would like to thank all my wonderful friends for being the wonderful friends that they are. They offered moral support, boosted my confidence and provided distractions when that was needed. I am especially grateful to Ingrid Senje Rasmussen, Idunn Skjetne and Siri Birkeland, who took the time to read and comment on my thesis despite the fact that their own deadlines were approaching.

Table of contents

Introduction

Allopatric speciation, genetic divergence in geographically isolated populations due to different selection pressures and genetic drift, is probably the most common way new species arise (Coyne & Orr 2004). Gene flow between populations will tend to counteract diversification of populations (e.g. Hartl & Clark 2007) and reproductive barriers are therefore less likely to form when geographic barriers are weak. Distance between habitable regions is an important geographic barrier (e.g. Schluter 2000). The theory of island biogeography states that the longer distance the island is from the mainland (or from other islands), the harder it is to colonize, and so the geographic barrier is quite strong, facilitating speciation (MacArthur & Wilson 1967). Island archipelagos are therefore attractive model systems for phylogeographic analyses, and floral and faunal systems on for instance Hawaii, the Canary Islands and the Galapagos have been studied extensively to address a range of evolutionary questions (e.g. Kambysellis *et al*. 1995; Juan *et al*. 2000; Grant & Grant 2008). The montane forest on mountains and volcanoes in East Africa can be understood as a continental version of an island archipelago system, making these forests an excellent model system to study biogeography and speciation (Voje *et al.* 2009). These forests are also well-known biological diversity hotspots and the mountains in the Great Rift Valley and the Eastern Arc in East Africa are rich in endemic species (Myers *et al.* 2000; Matthee *et al.* 2004; Voelker *et al.* 2010). However, hypotheses explaining this high level of biodiversity in East Africa are few and only a small number of phylogeographic studies have been conducted on this continental archipelago system (but see e.g. Bowie *et al.* 2004; Stanley & Olson 2005; Kebede *et al.* 2007; Voje *et al.* 2009).

The mountains vary in age and origin. Some, like Mt. Kilimanjaro, are dead or dormant volcanoes stemming from the formation of the Great Rift (Wilcockson 1956; Hemp 2005), while others like the Taita Hills date back to the Precambrian (Pohl & Horkel 1980). The vegetation along the mountain slopes varies with altitude and precipitation level and can be divided into different climatic zones: wet and arid savannah (up to about 1000 m asl), dry to moist grassland and banana plantations (1000-1600 m asl), forests and forest clearings in the montane zone (1600-3000 m asl), and finally subalpine grass and heathlands up to the afro-

alpine zone (Hemp & Hemp, 2003). Savannah and grassland are the dominant vegetation types on the plains between the mountains. Accordingly, the vegetation zones above the grasslands have a very patchy distribution in the area.

The ability of forest patches to expand and retract under different climatic regimes is an important difference between these continental archipelagos and the oceanic ones, thus making the inhabitants of montane forest islands to a greater extent only temporarily isolated from each other. The East-African montane forests have not always been islands in a sea of dry savannah, although drying seems to be the general trend in Africa in the past million years (deMenocal 1995; Fernandez & Vrba 2006). Climatic fluctuations during the past 40 Myr are associated with alternate spread and retreat of the montane forests, probably forming forest bridges between the mountains during the wetter phases (Burgess *et al.* 2007; Zachos *et al.* 2001; Voje *et al.* 2009). This knowledge is of great value, since such variable conditions may have established alternating opportunities for ecological fragmentation with subsequent genetic isolation of species dependent on forest and grassland habitats. Several paleoclimatic studies have been conducted in East Africa and it is therefore to some extent known when the forests were isolated on mountains and when they interconnected (Fernandez & Vrba 2006; Trauth *et al*. 2005). Forest bridges are believed to have been formed during the wetter phases between 2.7 - 2.5, 1.9 - 1.7 and 1.1 - 0.9 million years ago (Trauth *et al*. 2005). Such bridges may have made differentiated montane populations able to exchange genes at the same time as species dependent on grassland might have been isolated in small patches of favorable habitat. When the forest retracted to higher altitudes during the subsequent dry and cold periods, forest dependent species got isolated at the same time as these conditions eased migration and gene flow between different populations of grassland species. It can therefore be assumed that species dependent on montane habitats and with low migration potential have been isolated from each other since the last time the forests on the mountains were interconnected due to climatic conditions favoring forest in the lowland. A relationship between climatically induced ecological fragmentation of forest habitats and speciation in East Africa has been found for some animal groups (e.g. Bowie *et al.* 2004; Stanley & Olson 2005; 2007, Voje *et al.* 2009).

The mountains in the Eastern Arc and the Rift Valley contain a high number of flightless and endemic species of grasshoppers, compared to the more widespread and often fully winged species of grasshoppers on the savannah (Hochkirck 1998). Flightless insects have a reduced ability to disperse and are thus conservative indicators of past biogeographical change. The study organisms of this thesis are two genera of flightless acridid grasshoppers (*Parasphena* and *Gymnobothroides,* see Figures 1 and 2) inhabiting different niches on mountains in East Africa.

Fig. 1 *Gymnobothroides pullus* - Mazumbai. Picture by Claudia Hemp.

Fig. 2 *Parasphena teitensis* – Taita Hills. Picture by Claudia Hemp.

Parasphena are restricted to grasslands in the montane and sub-alpine zones and are not found below the forest belt. *Gymnobothrides* has a wider altitudinal span and can occupy a greater range of ecological niches compared to *Parasphena.* Most species of *Gymnobothroides* are found in sub-montane locations, but some species can also be found at higher altitudes. They can thrive in bush-land, coffee/banana plantations and forest edges, in addition to grasslands. Species within *Gymnobothroides* may therefore be able to spread more easily between mountains compared to the more niche-restricted *Parasphena* species.

The main goal of this study is to investigate whether the patchily distributed grasshoppers within *Parasphena* and *Gymnobothroides* diverged because of climate-induced forest fragmentation of the zones they depend on (vicariant speciation), or through adaptation to novel niches made available on emerging volcanoes (adaptive speciation). These two models of speciation yield different predictions for the timing of speciation events. From the vicariant speciation hypothesis I predict that several speciation events would have occurred within a rather short time period following climate-induced forest fragmentation. Such a pattern is not expected according to the adaptive speciation hypothesis since the mountains

differ in age. From the adaptive speciation hypothesis I rather predict close correlations between the timing of speciation events and the age of different mountains and volcanoes.

Geographically dispersed populations may be ecologically and phenotypically similar because of common descent and/or because they experience similar selection pressures (Schluter 2000). Rundle *et al.* (2000) provided a fascinating example of adaptive diversification, demonstrating repeated events of incipient sympatric speciation caused by parallel adaptation to alternative environments in geographically dispersed populations of three-spined sticklebacks (*Gosterosteus aculeatus*). A related assessment of speciation hypotheses can be done on Mt. Kilimanjaro, where (at least) two ecologically differentiated *Parasphena* species are found. *P. meruensis* occupies a niche in the montane zone, while *P. pulchripes* occupies a different niche in the sub-alpine to the alpine zone (C. Hemp, personal communication). Whether these species are more related to each other (adaptive diversification), or whether they are more closely related to species occupying analogous niches on different mountains (niche sharing due to common descent) is an important question in relation to the speciation mechanisms in this system. I will therefore test the hypothesis of adaptive diversification against the alternative hypothesis of common descent by investigating the relative relatedness among lineages sharing the same ecology or inhabiting the same mountain. Note that Kilimanjaro also houses a third *Parasphena*-species, namely *P. nairobaensis* which otherwise is mainly distributed in Kenya. I do not have information on the ecology of that species, however. Hence, I will not put a similar emphasis on the phylogenetic placement of the latter species.

Finally, I will investigate to what extent this first phylogeny based on molecular markers matches the current morphology and geography-based classification and taxonomy of *Parasphena* and *Gymnobothroides*. Very little work has been done on the genetics and phylogeny of African acrididae (Hemp & Hemp 2003), and the same can also be said for most of the other organisms inhabiting these mountains (Loader *et al.* 2006; Blackburn & Measey 2009). The first survey of the fauna in this area was done during an expedition led by Sjöstedt (Sjöstedt 1909). The current classification and taxonomy of the grasshoppers inhabiting these mountains is mostly based on their morphology and their geographic distribution (e.g. Jago 1971; Grunshaw 1986). The morphological differences between the species within these genera are very small. However, during the field season in 2011, we

found a population of *Parasphena meruensis* on Mt. Meru that had much darker colorings than the other members of that genus. It is therefore interesting to investigate whether such morphological difference will be reflected in our genetic analysis of these taxa.

Methods

2.1 Sampling and laboratory analyses

A total of 52 samples of *Parasphena* and 15 of *Gymnobothroides* grasshoppers were collected between 2006 and 2011 from 19 localities in the East-African Mountains (Fig. 3 and Tab. T4). Twenty-four of them (all *Parasphena* species), were collected by me in 2011. Please note that the specimens called "*G. lineaalba*" have traditionally been classified as "*Gymnobothrus lineaalba*". This was brought to my attention very recently (C. Hemp, personal communication), and I have therefore not had time to investigate if any of the specimens classified as "*G. sp*" also belong to the genus *Gymnobothrus*, and so they will all be analyzed together in this thesis.

Fig. 3 Map over Tanzania and Kenya, showing all locations I have samples from (stars). As a sort of scale, Mt. Meru and Kilimanjaro are approximately 40 km apart.

The grasshoppers were caught using a net, or simply by hand, and then transferred to jars with cyanide in the lid – killing them swiftly. Most samples were stored in 75 - 90% ethanol, but a few were stored dryly prior to extraction of DNA.

DNA was isolated from hind leg tissues using the Mole genetics isolation machine (Mole Genetics AS, Lysaker, Norway) with the Mole kit for tissue samples. The manufacturers' instructions were followed without deviations.

Segments from two mitochondrial genes were amplified using polymerase chain reaction (PCR); the 16s gene, using the primers 16a-M: (5' – CGC CTG TTT ATC AAA AAC AT – 3') and 16b-M: (5' – CCG GTC TGA ACT CAG ATC ACG T – 3') (Simon *et al.* 1994) and the COI gene using the primers HCO2190: (5´- TAA ACT TCA GGG TGA CCA AAA AAT CA - 3´) and LCO1490: (5´ -GGT CAA CAA ATC ATA AAG ATA TTG G - 3´) (Folmer *et al.* 1994).

I added 2 µl template DNA, 2 µl primer, 2 µl buffer, 2 µl nucleotides, 1 µl $MgCl_2$, 0.5 µl taq polymerase and 12.5 µl water, to gain a total reaction volume of 20 µl per sample.

For the 16s-samples a pre-denaturation step for three minutes at 94 ˚C was followed by 35 cycles of 94 ˚C for 45 seconds to denature the DNA chains, 54 ˚C for 45 seconds for primer annealing, and finally 72 ˚C for 1 minute and 20 seconds for extension. The corresponding PCR-profile for COI was three minutes at 95 ˚C, followed by 34 cycles of 94 ˚C (35 sec), 55 ˚C (30 sec) and 72˚C (1 min 30 sec).

PCR products were cleaned for excessive primers and nucleotides using ExoSAP-IT (USB Corporation, Cleveland, OH, USA). The sequencing was performed on an Applied Biosystems (ABI) 3730 high-throughput capillary electrophoresis instrument and was conducted using the same primers as used in the PCR. The sequencing results were usually good, yielding sequences from 400-1000 base pairs, but some sequences were too short or too poor in quality and were consequently omitted from further analysis. Hence, the sample size for the 16s and the COI analysis differs to some extent. The sequences were aligned in Mega version 4 (Tamura *et al*. 2007) using ClustalW. Ambiguous nucleotide positions were determined manually by examining the chromatogram and by comparing the forward and reverse sequences.

2.2 Phylogenetic analyses

Using the JModelTest (Posada 2008), I tested which evolutionary model fitted my sequence data best. This software compares different models of nucleotide evolution in a hierarchic hypotheses testing framework using both log-likelihood values and the Akaike information criterion (AIC). A general time reversible model with invariable sites (GTR+I) fitted all data sets best. Bayesian inference (BI) analyses were completed in MrBayes v.3.1.2 (Huelsenbeck & Ronquist 2001), using an *Amytta* species and two *Ixalidium* species as outgroups for the 16S and COI datasets, respectively. The priors were set to match the GTR+I model, but I did not fix any of the parameters. All the searches were run with between 1.000.000 and 3.000.000 generations, sampling every 1000th generation. The first 75% of the sampled runs were used as the burn-in, yielding between 250 and 750 trees.

MrBayes is a program using Bayesian inference to find the best tree possible. Bayesian inference is a relatively new, but strong method for getting good results fast. The strength and speed of this method is the reason MrBayes was the program of choice in this thesis. Bayes' theorem is as follows:

$$Pr[Tree\ I\ Data] = (Pr[Data\ I\ Tree] \times Pr[Tree])/Pr[Data]$$

It should be read as "the (posterior) probability of the tree given the data is depended on the prior probability of a phylogeny (Pr[Tree]) combined with the likelihood (Pr[Data I tree])" . The posterior probability is then the probability that the tree is correct (Huelsenbeck *et al*. 2001). The Bayesian inference is similar to the maximum likelihood method, which also finds the most likely tree, but the former method is perhaps a faster and more efficient way to find the best trees and the support for these trees are usually higher than those returned by the maximum likelihood bootstrapping method (Holder & Lewis 2003). Instead of bootstrapping, MrBayes finds the posterior probabilities of each clade using Markov chain Monte Carlo (MCMC) method. This method starts with a tree, and then a new tree is proposed following defined rules. The new tree is either accepted or rejected with a certain probability. The posterior probability of a tree is calculated based on how many times that tree was encountered during the MCMC runs, so the essential part is to run the MCMC long enough to make the chains converge on the optimal tree (Huelsenbeck *et al.* 2001). To make sure that the program had run long enough, plots of the generation versus the log-

probability of the data given the parameter values were analyzed. The analysis was regarded to have run sufficiently long when there was no trend in the scatter-plot.

Trees calculated using MrBayes were opened in Treeview 3.2 (GubuSoft LLC).

In total, four trees were created in MrBayes; one for the *Gymnobothroides* dataset with the 16s gene, one for the *Parasphena* dataset with the 16s gene, one for the *Parasphena* dataset with the COI-gene, and one dataset was made merging the 16s and COI datasets for *Parasphena* together. *Gymnobothroides* was unfortunately not sequenced for the COI-gene, since this genus was included late in the process when there was no time left to do lab work. This means that the genus *Parasphena* will be more thoroughly investigated in this thesis compared to *Gymnobothroides*.

To examine the robustness of the topologies estimated by the BI analysis, I also computed phylogenetic trees with bootstrap values, using the distance-based Neighbor Joining method implemented in MEGA. The model used for these trees was the maximum composite likelihood, with complete deletion of gaps and homogeneity of substitution patterns among lineages.

2.3 Estimation of divergence rate and dating of speciation events

The timing of divergences between species and populations provides an important test of phylogeographic hypotheses. In this respect, the time since the species split apart must be consistent with the putative historical events that are hypothesized to have driven their divergence. Molecular divergence rates applicable to the mitochondrial genome can be obtained from previous studies where divergence time between clades has been calibrated with, for example, geological events (e.g. Brower 1994; Juan *et al.* 1995) and a divergence rate of 1 % per lineage per million years has often been used for the COI-gene in phylogeographic studies on insects (e.g. Juan *et al.* 1995; Trewick & Morgan-Richards 2005; Spooner & Ritchie 2006). However, a recent study by Allegrucci and colleagues (2010) estimated a higher divergence rate (1.6 %) for this gene. It is therefore possible that the divergence rate of 1 % that I use in this study is a slight underestimation. 16S has been less

used in phylogeographic studies on insects and I therefore use the 0.7% divergence rate per million years estimated by Allegrucci *et al.* (2010) for this gene.

In Mega version 4 (Tamura *et al.* 2007), p-distances were calculated between the sequences in all the datasets (appendix, Tab. 1-3). The p-distance is a simple measure of the percentage of differences between each of the sequences, which is an easy way to see how genetically different the various species have become. Only unique haplotypes were used when computing the distance matrices.

A mismatch distribution approach (Slatkin & Hudson 1991; Rogers & Harpending 1992) as implemented in Arlequin 3.5 (Excoffier & Licher 2010) was used to detect sudden population expansions in haplogroups in the phylogenetic trees. In the case of multispecies phylogenies a signal for population expansion can be translated to rapid diversification of lineages or speciation events. The typical phylogenetic signal for such rapid species diversification is their short branch lengths and polytomies. The model assumes that a population at equilibrium at a certain time suddenly expands in size from N_o to N_1, and then returns to an equilibrium state. Arlequin also calculates the raggedness index (Harpending 1994), and the significance (P-value) of this index. The raggedness test evaluates the null-hypothesis that the population has been expanding. If a population has been stable for a long time, the distribution is expected to be highly multimodal and "ragged", giving a high raggedness index. A non-significant raggedness index indicates that there is no support for a stable population, but does not provide absolute evidence for an expanding population either. Low raggedness (and a unimodal mismatch distribution) is however typical for a population which has recently undergone a rapid expansion. The mismatch distribution model implemented in Arlequin estimates three different demographic parameters. One of these is the expected number of differences between two randomly drawn haplotypes at time t since the population expansion T (tau). This parameter can be used to estimate the time since the population expansion by dividing Tau by 2μ, where μ is the expected number of mutations per site per lineage per million years, multiplied by the number of base pairs in the sequence.

Another method used for calculating divergence times and test for gene flow between the populations was a coalescent approach using the MDIV program (Nielsen & Wakely 2001).

The model implemented in MDIV assumes that two populations arose from a single ancestral population at generation t in the past. Bayesian inference is used to estimate three parameters and their posterior probability: Theta (θ, two times the effective population size times the mutation rate), T (the divergence time between two populations) and M (the migration rate between two populations). Both M and T are scaled by population size. I ran three pair-wise comparisons using 4.000.000 generations of MCMC with 1.000.000 burn-in cycles between three *P.meruensis* populations found at Kilimanjaro using the combined 16S- and COI-dataset. The outputs from the runs are three graphs showing the estimated θ, T and M-parameters with their posterior probability. The parameters can be used to calculate the divergence times between the populations by using this formula:

$T_{div} = (T\theta) / (2\mu)$

where μ is the expected number of mutations per site, per lineage, per million years, times the number of base pairs in the sequence. To calculate μ in this dataset containing two genes with different mutation rates, I multiplied each gene's mutation rate per million years with its number of base pairs, and added them up.

Results

3.1 Phylogenetic analyses

The phylogeny of *Gymnobothroides* (Fig. 4) shows little structure, with many polytomies and short branch lengths. Many of the sampled specimens belong to unknown or undescribed species. The two *G. lineaalba* sequences (which traditionally have belonged to the genus *Gymnobothrus*) form a clade with high support, but individuals of another suggested species within this group, *G. levipes*, from different mountains and volcanoes, all appear in a large polytomy with a lower posterior probability.

The phylogeny of *Parasphena* based on the COI-gene can be seen in Figure 5. In this tree the branch lengths are also short, and the phylogeny is somewhat unresolved. The relationship between *P. nairobaensis* and *P. chyulensis* provides one example of such an unresolved phylogenetic group (Fig. 5). The morphologically distinct *P. meruensis* population found at 1700 meters on Mt. Meru forms a separate group and is clustering with the *P. meruensis* from Kidia Kilimanjaro, rather than with the other *P. meruensis* specimens from Mt. Meru. Another interesting aspect of this tree is that the *P. meruensis* from Mkweseko Kilimanjaro form a monophyletic group together with *P. teitensis* collected from the Taita Hills. Kilimanjaro houses (at least) two ecologically divergent species of *Parasphena,* namely *P. meruensis* and *P. pulchripes.* According to this tree, these two species form two well-differentiated lineages where *P. pulchripes* cluster with various species mainly distributed in Kenya, rather than with *P. meruensis*.

The phylogeny of *Parasphena* based on the 16s gene can be seen in Figure 6. This tree has several polytomies, and the branch lengths are generally very short. However, a few interesting clades can be seen: *P. teitensis* from the Taita Hills stand out and cluster with a few *P. meruensis* specimens, and the population of *P. meruensis* from Kidia form a monophyletic group of their own. The morphologically distinct grasshoppers from Mt. Meru (1700 m) do not form a separate clade in this genetic tree, but join the rest of the *P. meruensis* in a large polytomy. Again *P. pulchripes* cluster with various Kenyan species rather than with the *P. meruensis* populations from Kilimanjaro.

I also combined the two genes I had sequenced for *Parasphena* into a large dataset to see if a tree based on these data could give any additional information (Fig. 7). This tree is very similar to the COI-tree, probably because that gene showed the highest level of interspecific variation. *P. meruensis* specimens found at the lowest altitude on Mt. Meru is also in this tree forming a clade with the *P. meruensis* specimens found in Kidia (Kilimanjaro), while the other population on Kilimanjaro (Mkweseko) clusters together with *P. teitensis* from the Taita Hills. An important difference in this tree compared with the other *Parasphena* trees is that this tree resolves the polytomy within the specimens collected at different altitudes on Mt. Meru (M1850, M2150 and M2500), and the specimens from 2150 and 2500 m asl cluster together. This tree contains fewer specimens than the other trees, since this tree only contains those specimens where both a 16s-sequence and a COI-sequence were available. This also accounts for some of the other differences between the trees shown in Figures 5, 6 and 7.

The corresponding Neighbor Joining trees (see the appendix, Figures F1-F4) are generally confirming the topology found in the Bayesian inference analyses.

3.2 Mismatch distribution results

Mismatch distributions and their corresponding raggedness indices were calculated for selected groups that showed phylogenetic signals consistent with rapid species diversification. One for the large polytomy in the *Gymnobothroides* tree, one for the entire in-group in the *Parasphena*-tree based on the COI-gene, and one for the subgroup in the same tree containing the grasshoppers from Kidia (Kilimanjaro) and Mt. Meru. Figure 8 shows the mismatch distribution of the haplogroup within the *Gymnobothroides* tree. The distribution has a unimodal shape and a corresponding low raggedness index. The model (the line in Fig. 8) seems to follow the same general pattern as the bars, and the P-value is non-significant, supporting a recent population expansion. Tau was found to be 4.023, and using the formula

$t = T/2\mu$

where the evolutionary rate of the 16s-gene is assumed to be 0.7 % per site per lineage per million years, and the sequence contains 413 base pairs, the divergence time for this group is calculated to be 0.696 (95 % confidence interval (CI): 0.185 - 1.866) million years ago.

Figure 9 shows the mismatch distribution for the entire *Parasphena* in-group in the COI-tree. The observed distribution (the bars) is not a smooth unimodal curve, but is not deviating enough from the expansion model to falsify it since the raggedness index has a non-significant P-value. The estimated time of divergence in this group, based on an evolutionary rate of 1 % per site per lineage per million years, sequences with 524 base pairs and a Tau of 7.766, was calculated to be 0.741 (CI: 0.0037 - 3.45) million years ago. The wide confidence interval reflects the fact that the observed distribution is not strictly unimodal.

Figure 10 shows the mismatch distribution of haplogroup 2, the *Parasphena meruensis* specimens from Mt. Meru and Kidia on Kilimanjaro. The fit between the expansion model and the observed differences between haplotypes is good, and the raggedness index comes out as non-significant. This group had a Tau of 1.256, sequences with 573 base pairs, and the estimated time of divergence was then calculated to be 0.109 (CI: 0 - 0.25) million years ago.

3.3 Coalescent analysis

I used a coalescent analysis to test specifically when the three closely related populations of *P. meruensis* from Mt. Meru and Kilimanjaro diverged from each other. The results of the analysis suggest that they all diverged less than 300.000 years ago (Table 1). The two populations estimated to have split apart most recently are the population found at 1700 meters above sea level on Mt. Meru and the population near Kidia on Kilimanjaro. The population from Kidia and the *P. meruensis* population found at the highest altitudes on Mt. Meru have the oldest estimated divergence time, but not very much older than the split between the two populations on Mt. Meru. Figures 11 to 13 show the posterior probabilities of the three parameters for all three pair wise comparisons. The posterior probabilities overlap quite much between all the comparisons for all the parameters, but note the rather flat curve of posterior probabilities for the divergence time parameter, meaning that both shorter and especially longer divergence times cannot be ruled out.

3.4 Distance matrices

The calculated P-values show that the specimens in the *Gymnobothroides* dataset differ by between 0.1 and 2.4 % (see appendix, Tab. T1 –T 3). The *Parasphena* sequences from the COI-dataset differ by 0.2 - 5.7 %. The distance matrix for the 16s-dataset for both genera combined shows that within-genus variation is lower than 3 %, while nucleotide difference between genera is between 20 - 25 %.

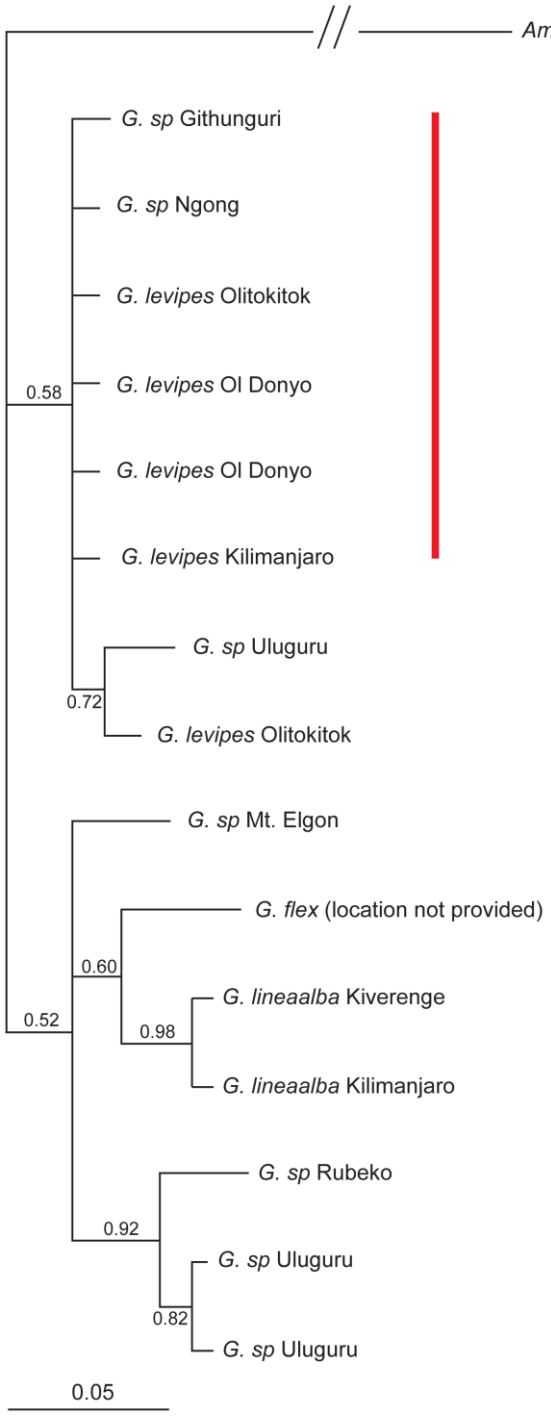

Fig. 4 Bayesian inference of the phylogeny of the genus *Gymnobothroides* based on the 16s-gene. Priors were set to match the GTR+I model. Each sequence is represented with its corresponding species name and the location it was found. *Amytta* sp from Ol Donyo was used as the outgroup. The scale bar represents five changes per 100 nucleotide positions, and the branch support values are the posterior probabilities of each clade. Only nodes that were supported in 50 % or more of the sampled trees are shown. The colored bar represents the haplogroup that was used in the mismatch distribution approach.

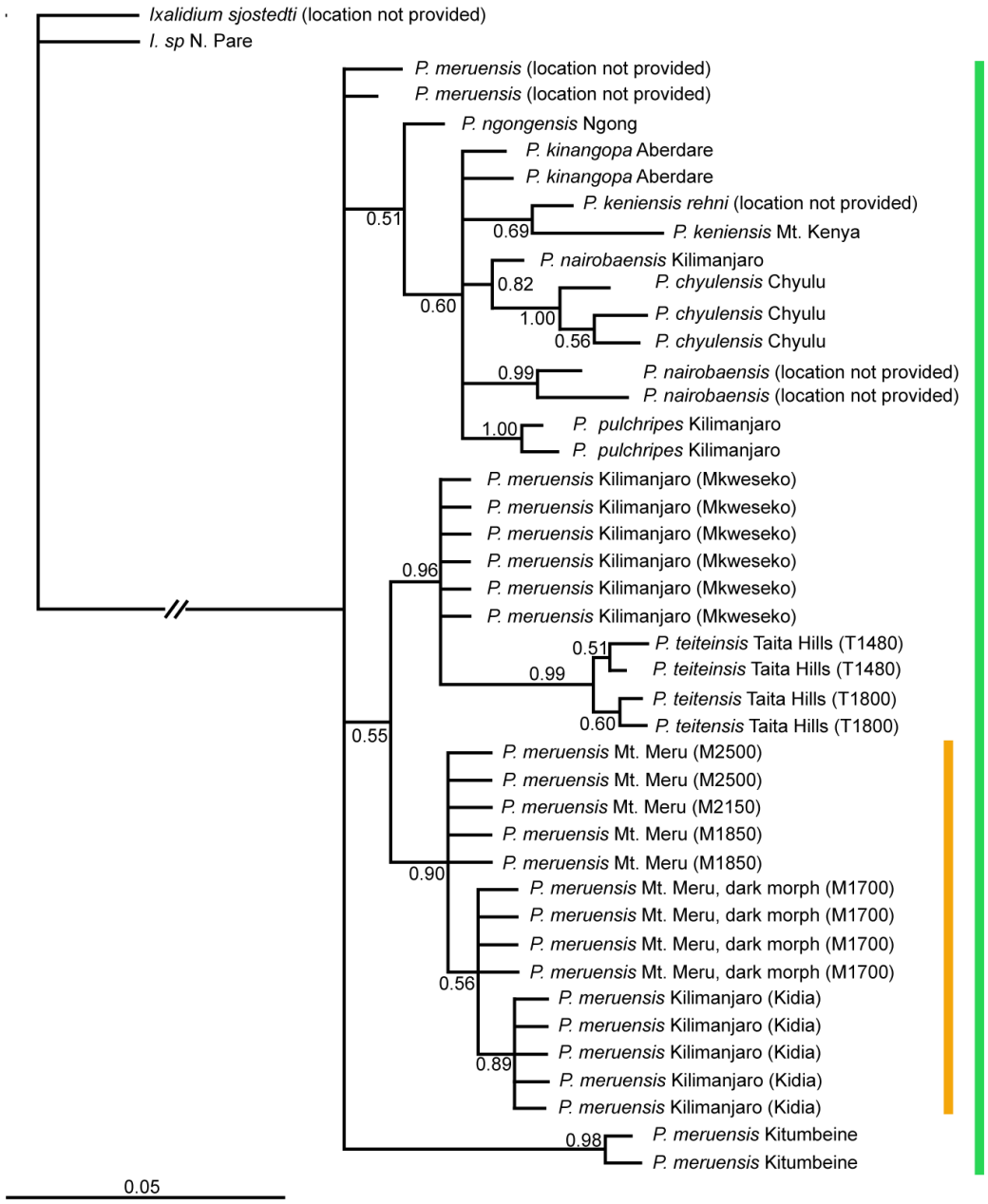

Fig. 5 Bayesian inference of the phylogeny of the genus *Parasphena* based on the COI-gene. Priors were set to match the GTR+I model. Each sequence is represented with its corresponding species name and the location it was found. One *Ixalidium* sp the North Pares, and one *Ixalidium sjostedti were* used as the outgroup. The scale bar represents five changes per 100 nucleotide positions, and the branch support values are the posterior probabilities of each clade. Only nodes that were supported in 50 % or more of the sampled trees are shown. The colored bars represent the haplogroups chosen for further analyses.

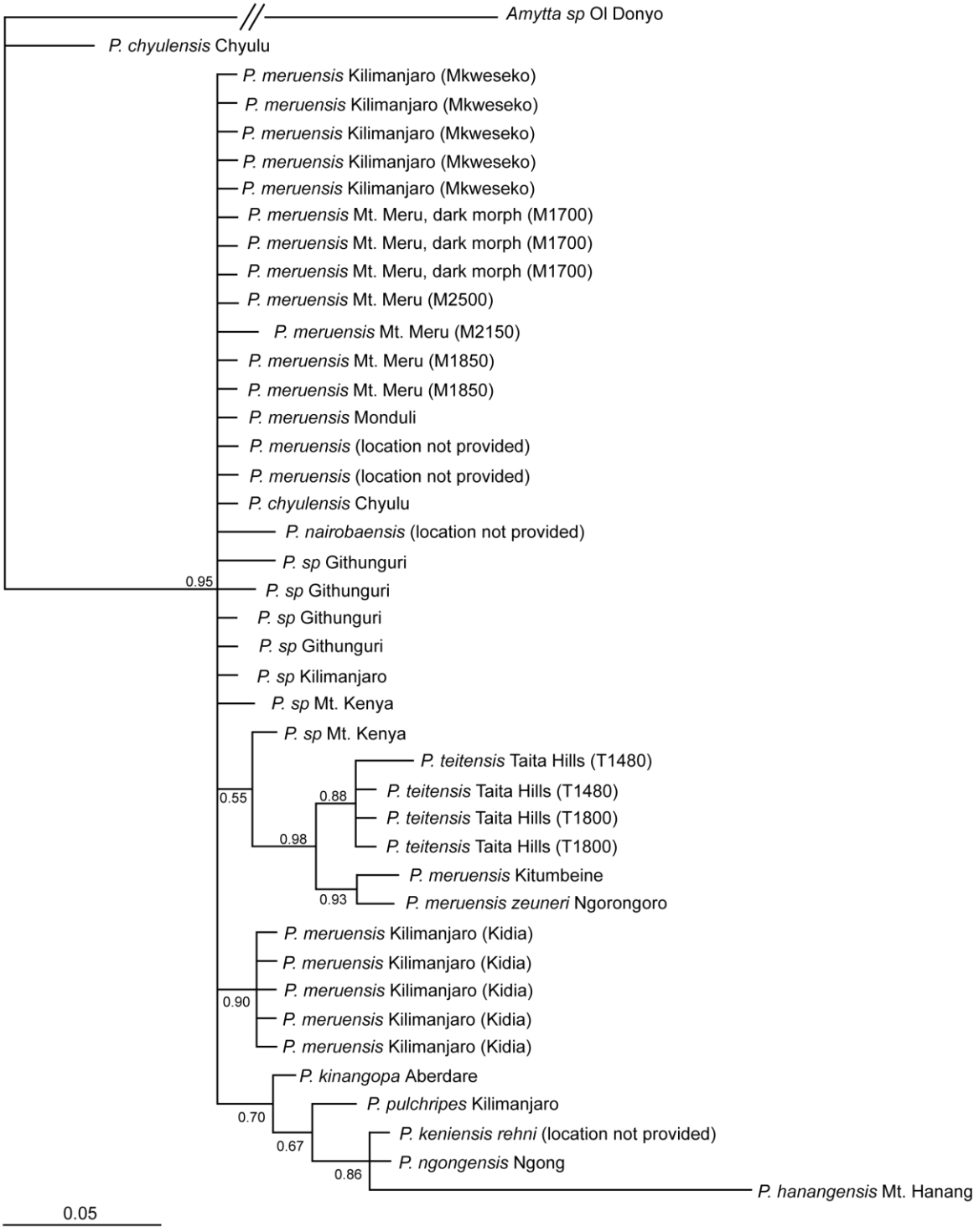

Fig. 6 Bayesian inference of the phylogeny of the genus *Parasphena* based on the 16s-gene. Priors were set to match the GTR+I model. Each sequence is represented with its corresponding species name and the location it was found. *Amytta* sp from Ol Donyo was used as the outgroup. The scale bar represents five changes per 100 nucleotide positions, and the branch support values are the posterior probabilities of each clade. Only nodes that were supported in 50 % or more of the sampled trees are shown.

Fig. 7 Bayesian inference of the phylogeny of the genus *Parasphena* based on both the 16s-gene and the COI-gene. Priors were set to match the GTR+I model. Each sequence is represented with ots corresponding species name and the location it was found. One *Ixalidium sp* from the North Pares and one *Ixalidium sjostedti* were used as the outgroup. The scale bar represents five changes per 100 nucleotide positions, and the branch support values are the posterior probabilities of each clade. Only nodes that were supported in 50 % or more of the sampled trees are shown.

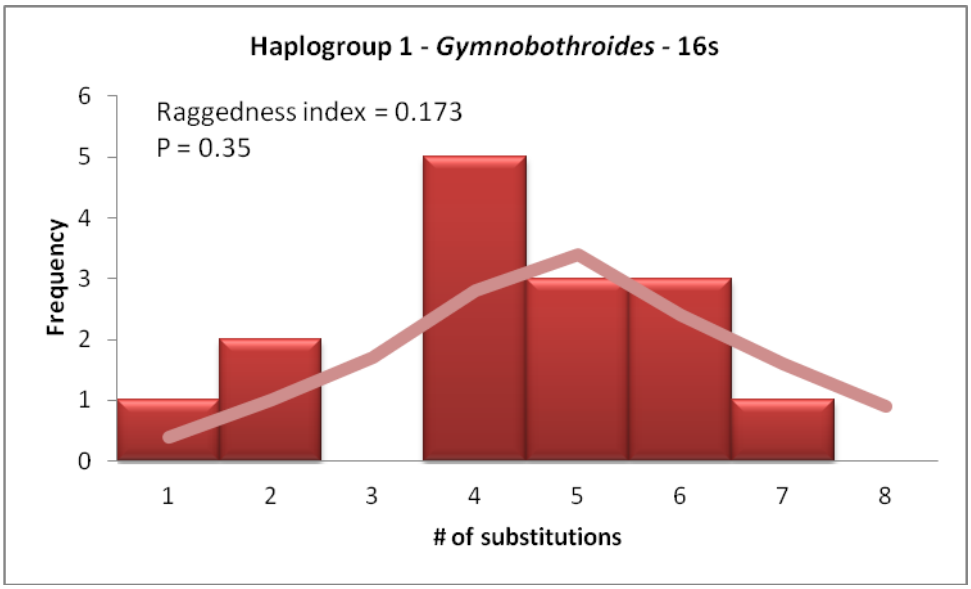

Fig. 8 Mismatch distribution of the sequences within haplogroup 1 from the *Gymnobothroides* tree (see Fig. 4). The line represents the expected values given the sudden expansion model implementedin Arlequin, while the bars are the observed values. The raggedness index and its P-value indicates whether or not the null-hypothesis (a stable, non-expanding population) can be rejected.

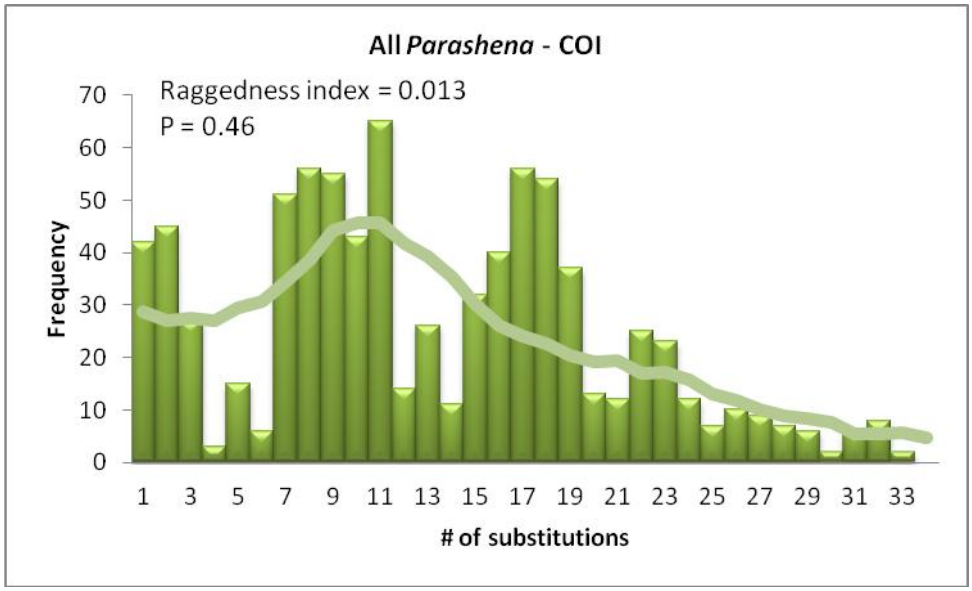

Fig. 9 Mismatch distribution of all the species except the outgroup from the *Parasphena* COI-tree (see Fig. 6). The line represents the expected values given the sudden expansion model, while the bars are the observed values. The raggedness index and its P-value indicates whether or not the null-hypothesis (a stable, non-expanding population) can be rejected.

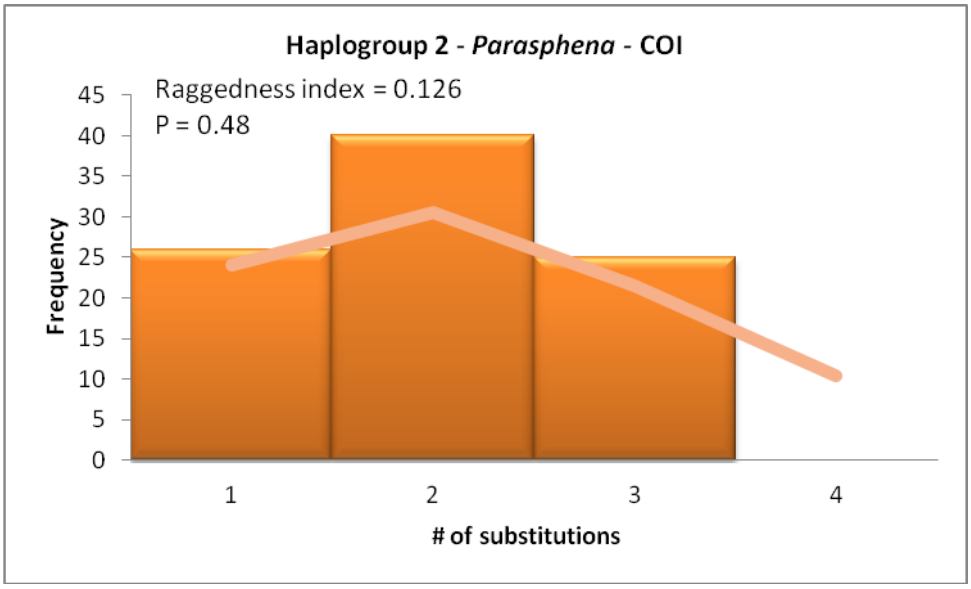

Fig. 10 Mismatch distribution of the sequences within haplogroup 2 from the *Parasphena* COI- tree (see Fig. 6). The line represents the expected values given the sudden expansion model, while the bars are the observed values. The raggedness index and its P-value indicates whether or not the null-hypothesis (a stable, non-expanding population) can be rejected.

Tab. 1 Estimated divergence times in years before present (ybp) between three populations of *P. meruensis* based on the parameters T and Theta estimated in the coalescent analysis implemented in MDIV.

	Divergence time
Mt. Meru 1700 m.asl. and Kidia Kilimanjaro	56 228 (ybp)
Mt. Meru 1700 m.asl. and Mt. Meru 1800-2500 m.asl.	196 829 (ybp)
Mt. Meru 1800-2500 m.asl. and Kidia Kilimanjaro	259 093 (ybp)

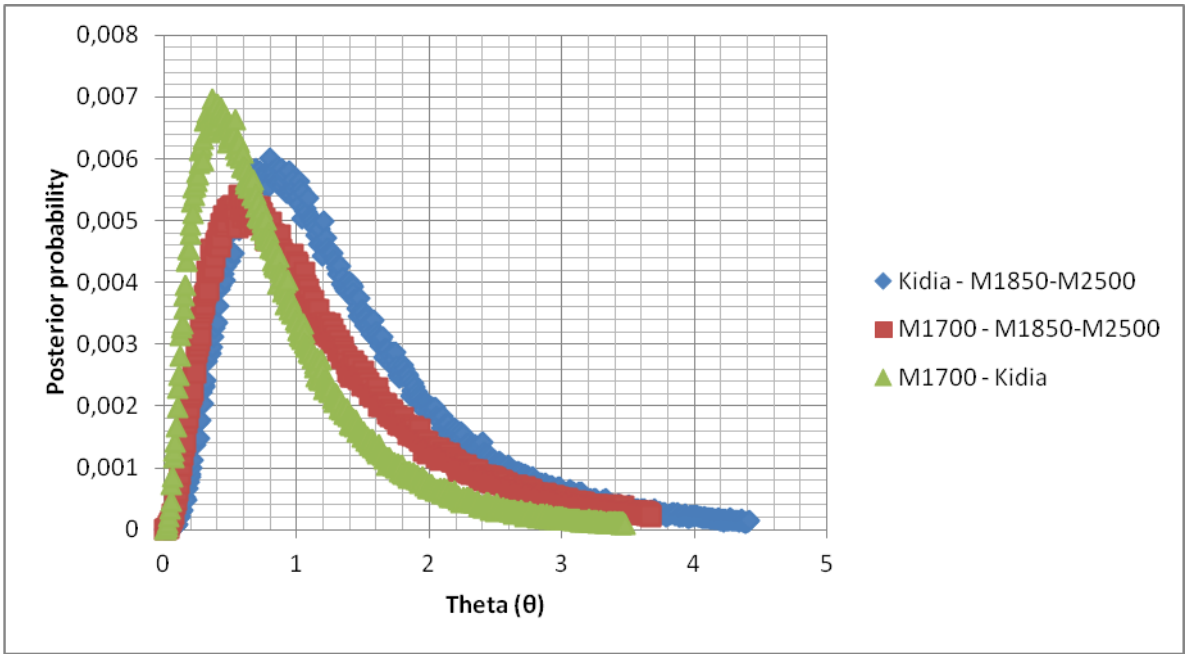

Fig. 11 The posterior probabilities of theta (θ) for pairwise comparisons between three populations of *P. meruensis*. Comparison of the population from Mt. Meru (1700 m asl) and the population from Kidia on Kilimanjaro in green, the comparison of the two populations on Mt. Meru in red, and the comparison of the populations on Mt. Meru from the highest altitudes and the population from Kidia on Kilimanjaro in blue.

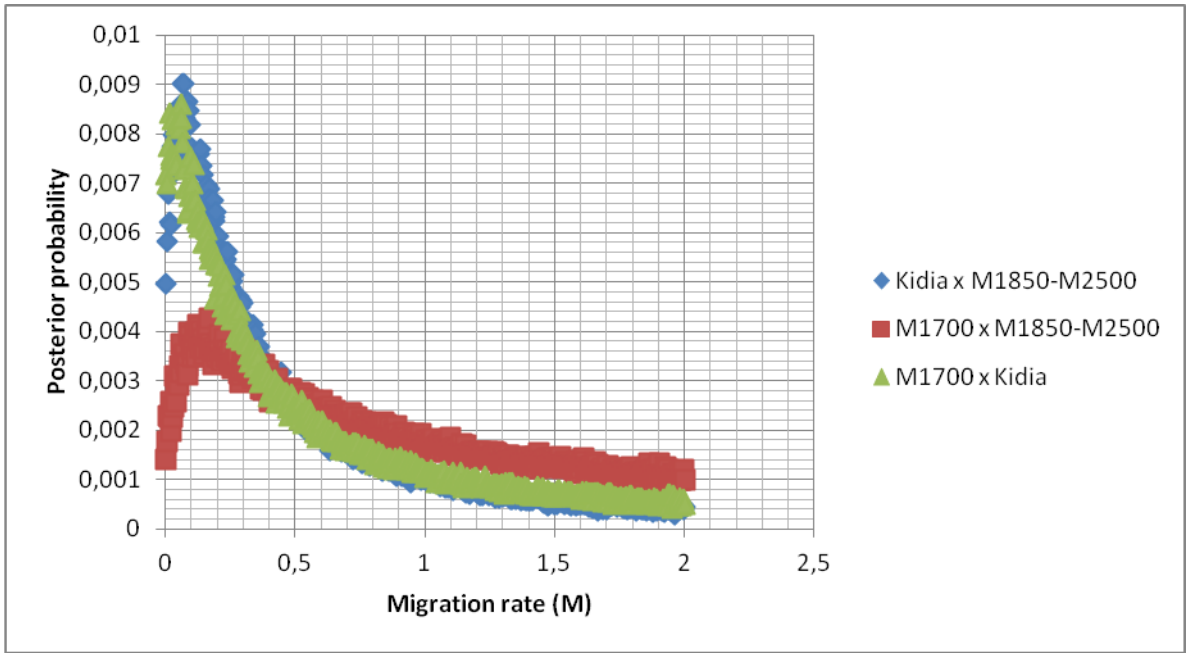

Fig. 12 The posterior probabilities of the migration rates (M) for all pairwise comparisons between three populations of *P. meruensis*. Comparison of the population from Mt. Meru (1700 m asl) and the population from Kidia on Kilimanjaro in green, the comparison of the two populations on Mt. Meru in red, and the comparison of the populations on Mt. Meru from the highest altitudes and the population from Kidia on Kilimanjaro in blue.

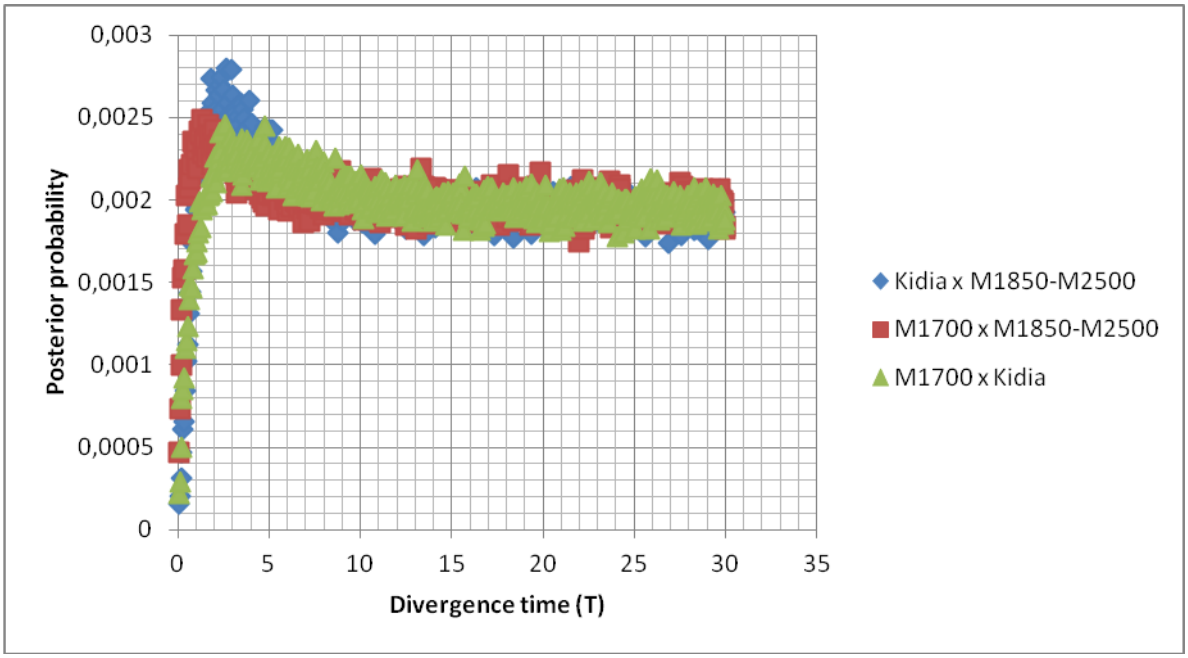

Fig. 13 The posterior probabilities of the divergence time (T) for all pairwise comparisons between three populations of *P. meruensis*. Comparisons of the population from Mt. Meru (1700 m asl) and the population from Kidia on Kilimanjaro in green, the comparison of two populations on Mt. Meru in red, and the comparison of the population on Mt. Meru from the highest altitudes and the population from Kidia on Kilimanjaro in blue.

Discussion

The main goal of this study was to investigate whether the genera *Parasphena* and *Gymnobothroides* diverged because of climate induced forest fragmentation of the montane zone or through adaptation to novel niches made available on emerging volcanoes.

The results of my study mainly support the fragmentation hypothesis. All of the mismatch distributions failed to reject the stable population hypothesis, and they pointed towards a sudden expansion consistent with the climatically induced habitat fragmentation hypothesis. All the grasshoppers seem to be genetically very similar based on the distance matrices and the phylogenetic trees, with short branch lengths and frequent polytomies, as expected if many new lineages emerged more or less at the same time. In further support of the speciation-through-fragmentation hypothesis is the fact that all the estimated divergence times are much more recent than the emergence of the mountains and volcanoes the species inhabit. Many of the East-African volcanoes emerged in the past millions of years. For instance, Mt. Kenya is approximately 4 million years old, while Kilimanjaro and Mt. Meru are about 1 million years old. Many of the non-volcanic mountains in the area are a lot older; like the Taita Hills which arose in the Precambrian. The estimated times of divergence of *Gymnobothroides* and *Parasphena* is on the other hand matching reasonably well with the last wet phase that ended about 0.9 million years ago (Trauth *et al.* 2005). Although having large confidence intervals, the estimated divergence events for the *Gymnobothroides* haplogroup (0.7 Mya), and the entire *Parasphena* genus (0.74 Mya) are very similar and both just a few hundred thousand years after the last wet maximum. Based on the fact that *Gymnobothroides* occupy a wider niche than *Parasphena,* I expected to find a less clear link between the climatic events and the divergence events within this genus, than for the other genus. On the contrary, it seems that the savannah is a strong barrier for gene flow for *Gymnobothroides*, and that this genus also underwent a sudden expansion some time after the retraction of the forests about 900.000 years ago. However, the supposedly niche-constricted *Parasphena* contain haplogroups where different populations seem to have intermingled despite fragmentation of their favored habitat. The less unimodal shape of the mismatch distribution of *Parasphena* also supports this hypothesis. More than one peak can

be seen in the observed distribution of the pair-wise differences, indicating that some diversification within *Parasphena* probably happened both before and after the estimated radiation about 700.000 to 800.000 years ago. It is for example possible that different populations of *Parasphena* were trapped in islands of grassland during the wet phase where the forests spread where they started diverging due to genetic drift and perhaps adaptation to their habitat. Indications that some populations of *Parasphena* seem to have intermingled during times of fragmentation of their favored habitat also comes from the analyses of the third haplogroup examined. *P. meruensis* from Mt. Meru and Kidia (Kilimanjaro) seem to have diverged around 100.000 years ago. This haplogroup was also examined with a coalescent approach. The results from this method fit the divergence times found with the mismatch distribution approach quite well: all three populations within this haplogroups diverged less than 300.000 years ago according to the latter estimate. These two mountains are quite close to each other (see Fig. 3), so it appears that it has not been impossible for these grasshoppers to intermingle even though the climate has been quite dry the past 900.000 years. This indicates that the savannah is not as strong a barrier for these grass-living grasshoppers as for the forest-dwelling grasshoppers previously studied (Voje *et al.* 2009). Dry savannah landscape is obviously a barrier, since we find a radiation consistent with the climate induced speciation hypothesis for the entire group, but apparently it is not an impossible barrier to cross if the distance is short enough. It is also interesting to see that the *P. meruensis* population living the farthest down on Mt. Meru (the dark morph) is genetically more similar to the population on Kilimanjaro than other populations on Mt. Meru. These two populations cluster together in the combined dataset-tree and the COI-tree, and they were estimated to have split apart only about 50.000 – 60.000 years ago.

P. teitensis and *P. meruensis* seem to be very closely related based on all the phylogenetic trees and the distance matrices. One population of *P. meruensis* (the one found in Mkweseko on Kilimanjaro) is actually more similar to *P. teitensis* than to other *P. meruensis*-populations. The Taita Hills and Kilimanjaro are of different origin and different age, and so the ancestors of these species must have intermingled, or migrated from one mountain to the other long after the formation of the mountains.

Another set of hypotheses investigated in this thesis is whether geographically dispersed lineages occupying the same niches are more related to each other than species that inhabit

the same mountains or whether niche separation on the same mountain is due to recent adaptive diversification in situ. *P. meruensis* and *P. pulchripes* both live on Mt. Kilimanjaro, but occupy different niches. *P. meruensis* inhabits grasslands in the montane zone, while *P. pulchripes* lives in the sub-alpine zone. The two Kilimanjaro taxa should cluster together according to the adaptive differentiation hypothesis. I found that *P. pulchripes* cluster with several of the species found in Kenya, rather than to the *P. meruensis* occupying the same mountain. This indicates that geographically dispersed groups occupying the same niche have a common descent, and have managed to spread to new mountains and finding the same niche there, without intermingling with other species in other niches on the way. Differentiated habitat preferences may constitute a significant pre-mating barrier in itself (Coyne & Orr 2004). Additionally, habitat choice may be connected to other differences; such as timing of breeding or song further reducing gene flow (see e.g. Sætre & Sæther 2010 for a case of multiple reproductive barriers).

Unfortunately, ecological differentiation has not been investigated between other species or populations of *Parasphena* in any detail. It is possible, however, from the results I have found regarding *P. meruensis* on Mt. Meru, that the population living closest to the mountain foot is genetically more similar to a population on another mountain. This might represent a parallel to the pattern found in the ecologically differentiated *P. meruensis* and *P. pulchripes* on Kilimanjaro, as described above. Kidia on Kilimanjaro is also at a quite low altitude, so it is possible that this is a similar habitat as the 1700 m asl area on Mt. Meru. Possibly these populations have adapted to a warmer climate, or to some other ecological characteristic of these altitudes that reduces gene flow with other populations on their respective mountains. The fact that the population farthest down on Mt. Meru is morphologically distinct from the other populations on that mountain can be a sign that pre-zygotic barriers have evolved. Similarly, *P. nairobaensis* also resides on Kilimanjaro, but clusters with Kenyan conspecifics (and other species) rather than with the other *Parasphena* taxa on Kilimanjaro. The presence of such genetically differentiated lineages on the same mountain suggests that distinct ecotypes have evolved, of which some may constitute distinct species.

The pattern of common descent of lineages occupying different mountains, and differentiated lineages living on the same mountain, could be explained by two not mutually

exclusive hypotheses. First, as mentioned earlier some diversification of the *Parasphena* genus may have started before the last episode of forest fragmentation, as is suggested by the peak around 17 differences in the mismatch distribution for this group.

Second, grasshoppers may have been able to disperse between mountains despite forest fragmentation. As discussed above, recent gene flow between Kilimanjaro and Mt. Meru does seem likely. However, dispersal between Kilimanjaro and Kenyan mountains seems less likely simply because of the long distance. Hence I suggest that some of the diversification of *Parasphena* actually may be more ancient than my divergence estimates would suggest.

I also set out to investigate whether the phylogeny based on molecular markers matched the current classification and taxonomy of *Parasphena* and *Gymnobothroides*. The traditional morphology- and geography-based classification of these grasshoppers (e.g. Jago 1971; Grunshaw 1986) deviates from the results of my study. For instance, the grasshoppers living in the Taita hills definitely seem to be a part of the *P. meruensis* species complex, and should perhaps rather have a sub-species status. I also conclude that there is probably no need to distinguish between the genera *Gymnobothroides* and *Gymnobothrus.* They cluster together in the phylogenetic tree without any obvious split that would indicate two separate genera, and the distance matrix shows that the number of nucleotide differences between the proposed genera is about as large as the number of nucleotide differences within each genus.

As for the rest of the species, it is virtually impossible to draw any firm conclusions on species status or phylogenetic relations now, partly because the sample sizes are relatively small. However, it appears that all the species within both genera are genetically very similar, and it might be that some of the genetic differences would break down if the forests were to venture down again. According to the biological species concept (Mayr 1963), these would not be different species if that were the case. However, as discussed above in connection with the genetically distinct lineages found on Kilimanjaro, reproductive isolation might have developed between some of the lineages. It has also been shown that the calling songs of genetically very similar species of bush-crickets are different (Hemp *et al*. 2010).

In conclusion, it seems that these grasshopper genera diverged in a large degree due to the climatic shift that took place about 900.000 years ago, but gene flow has occurred since that

time. Niche dependency appears to keep grasshoppers genetically separated despite them sharing the same mountain. A thorough investigation of the taxonomy for these East-African grasshoppers appears to be needed as the current classification does not match the molecular one depicted in this study.

There is a long way yet to go before we fully understand the processes driving speciation in this system. I would recommend further studies on ecological differentiation and potential reproductive barriers, for instance using common garden experiments in order to properly revise their taxonomy and in order to learn more about speciation events and their phylogeography. Additionally, more individuals from all the localities should be included in future phylogenetic studies. My study indicates that the well-sampled localities provide more in-depth knowledge about the phylogeny and evolutionary history of the taxa. We see the clearest pattern where I had the most samples: namely the Taita Hills, Mt. Kilimanjaro and Mt. Meru. Finally, additional traits, including nuclear genes and traits that may be important for reproductive isolation, such as morphology of genitalia, habitat preferences, song, and timing of breeding, should be included in future studies on these fascinating animals.

References

Allegrucci G, Trucchi E, Sbordoni V (2010) Tempo and mode of species diversification in *Dolichopoda* cave crickets (Orthoptera, Rhapidophoridae). *Molecular Phylogenetics and Evolution* **60**, 108-121.

Blackburn DC, Measey GJ (2009) Dispersal to or from an African biodiversity hotspot? *Molecular Ecology* **18**, 1904-1915.

Bowie RCK, Fjeldså J, Hackett SJ, Crowe TM (2004) Systematics and biogeography of double-collared sunbirds from the Eastern Arc Mountains, Tanzania. *Auk*, **121**, 660–681.

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 Science USA*, **91**, 6491-6495.

Burgess ND, Butynski TM, Cordeiro NJ (2007) The biological importance of the Eastern Arc Mountains of Tanzania and Kenya. *Biological Conservation*, **134**, 209–231.

Coyne JA, Orr HA (2004). *Speciation*. Sinauer & Associates, Sunderland Massachusetts.

deMenocal PB (1995) Plio-Pleistocene African Climate. *Science*, **270**, 53-59.

Excoffier, L. and H.E. L. Lischer (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564-567.

Fernandez MH, Vrba ES (2006) Plio-pleistocene climatic change in the Turkana Basin (East Africa): evidence from large mammal faunas. *Journal of Human Evolution*, **50**, 595–626.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**, 294–299.

Grant PR, Grant BR (2008) *How and Why Species Multiply: The Radiation of Darwin's Finches.* Princeton University Press, NJ.

Grunshaw JP (1986) Revision of the East African grasshopper genus Kassongia with a description of a new, closely related taxon, Labidioloryma gen.n. (Orthoptera: Acrididae: Hemiacridinae). *Systematic Entomology*, **11**, 33-51.

Harpending H (1994) Signature of ancient population growth in a low-resolution mitocondrial DNA mismatch distribution. *Human Biology*, **66**, 591-600.

Hartl DL, Clark AG (2007) *Principles of Population Genetics 4th ed.* Sinauer Associates Inc. MA.

Hemp A, Hemp C (2003) Saltatoria coenoses of high-altitude grassland on Mt. Kilimanjaro, Tanzania (Orthoptera: Saltatoria). *Ecotropica*, **9**, 71-97.

Hemp A (2005). Climate change-driven forest fires marginalize the impact of ice cap wasting on Kilimanjaro. *Global Change Biology*, **11**, 1013-1023.

Hemp C, Heller KG, Kehl S, Warchalowska-Sliwa E, Wägele W, Hemp A (2010) The *Plesirthes* complex (Orthoptera, Tettigoniidae, Conocephalinae, Conocephalini) rewiewed: integrating morphological, molecular, chromosomal and bioacoustic data. *Systematic Entomology*, **35**, 554-580.

Hochkirck A (1998) A Comparison of the Grasshopper Fauna (Orthoptera: Acridoidea & Eumastacoidea) of the Uluguru Mountains and the East Usambara Mountains, Tanzania. *Journal of East African Natural History*, **87**, 221-232.

Holder M, Lewis PO (2003) Phylogeny estimation: traditional and Bayesian approaches. *Nature reviews*, **4**, 275-284.

Huelsenbeck JP, Ronquist F (2001) MrBayes: bayesian inference of phylogenetic trees. *Bioinformatics*, **17**, 754–755.

Huelsenbeck JP, Ronquist F, Nielsen R, Bollback, JP (2001) Bayesian Inference of Phylogeny and Its Impact on Evolutionary Biology. *Science*, **294**, 2310.

Jago ND (1971) A review of the Gomphocerinae of the world with a key to the genera (Orthoptera, acrididae). *Proceedings of the Academy of Natural Sciences of Philadelphia*, **123**, 205-343.

Juan C, Oromi P, Hewitt GM (1995) Mitochondrial DNA phylogeny and sequential colonization of Canary Islands by darkling beetles of the genus *Pimelia* (Tenebrionidae). *Proceedings of the Royal Society of London, Series B*, **261**, 173-180.

Juan C, Emerson BC, Oromi P, Hewitt GM (2000) Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. *Trends in Ecology and Evolution*, **15**, 104-109.

Kambysellis MP, HO KF, Craddock EM, Piano F, Parisi M, Cohen J (1995) Pattern of Ecological shifts in the diversification of Hawaiian *Drosophila* inferred from a molecular phylogeny. *Current Biology*, **5**, 1129-1139.

Kebede M, Ehrich D, Taberlet P, Nemomissa S, Brochmann C (2007) Phylogeography and conservation genetics of a giant lobelia (Lobelia giberroa) in Ethiopian and Tropical East African mountains. *Molecular Ecology*, **16**, 1233–1243.

Loader SP, Channing A, Menegon M, Davenport TRB (2006) A new species of *Probreviceps* (Amphibia: Anura) from the Eastern Arc mountains, Tanzania. *Zootaxa*, **1237**, 45–60.

MacArthur RH, Wilson EO (1967) *The Theory of Island Biogeography*. Princeton University Press, NJ.

Matthee CA, Tilbury CR, Townsend T (2004) A phylogenetic review of the African leaf chameleons: the role of vicariance and climate change in speciation. *Proeedings of the Royal Society of London: Series* , **271**, 1967-1975.

Mayr E (1963) *Animal Species and Evolution.* Belknap Press, Cambridge, MA.

Myers N, Mittermeier RA, Mittermeier CG, daFonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* **403,** 853-858.

Nielsen R & Wakeley J (2001) Distinguishing Migration From Isolation: A Markov Chain Monte Carlo Approach. *Genetics*, **158**, 885-869.

Pohl W & Horkel A (1980) with the collaboration of Neubauer W, Niedermayr G, Okelo RE, Wachira,JK, Werneck W. Notes on the Geology and Mineral Resources of the Mtito Andei-Taita Area (Southern Kenya). *Mitteilungen der Österreichischen Geologischen Gesellschaft*, **73**, 135-152.

Posada D (2008), jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution*, **25**, 1253-1256.

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

Rundle HD, Nagel L, Boughman JW, Schluter D (2000) Natural selection and parallel speciation in sympatric sticklebacks. *Science*, **287**, 306-308.

Schluter D (2000) *The Ecology of Adaptive Radiation.* Oxford University Press, Oxford.

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America*, **87**, 651–701.

Sjöstedt Y (1909) Wissenschaftliche Ergebnisse der Schwedischen Zoologischen Expedition nach dem Kilimanjaro, dem Meru und den umgebenden Massaisteppen. Deutsch-Ostafrikas 1905-1906. Locustodea: 125-148, Acridoidae: 149-200.

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**, 555–562.

Spooner LJ, Ritchie MG (2006) An unusual phylogeography in the bushcricket *Ephippiger ephippiger* from Southern France. *Heredity*, **97**, 308-408.

Stanley WT, Olson LE (2005) Phylogeny, phylogeography, and geographic variation of Sylvisorex howelli (Soricidae), an Endemic Shrew of the Eastern Arc Mountains, Tanzania. *Zoological Society of London*, **266**, 341–354.

Sætre GP, Sæther SA (2010) Ecology and genetics of speciation in *Fidechula* flycatchers. *Molecular Ecology*, **19**, 1091-1106.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* , **24**, 1596-1599.

Trauth MH, Maslin MA, Deino A, Strecker MR (2005) Late Cenozoic moisture history of East Africa. *Science*, **309**, 2051–2053.

Trewick SA, Morgan-Richards M (2005) After the deluge: mitochondrial DNA indicated Miocene radiation and Pliocene adaptation of tree and giant weta (Orthoptera: Anostostomatidae). *Journal of Biogeograpy*, **32**, 295-309.

Voelker G, Outlaw RK, Bowie RCK (2010) Pliocene forest dynamics as a primary driver of African bird speciation. *Global Ecology and Biogeography*, **19**, 111-121.

Voje KL, Hemp C, Flagstad Ø, Sætre GP, Stenseth NC (2009), Climatic change as an engine for speciation in flightless Orthoptera species inhabiting African mountains. *Molecular Ecology,* **18**, 93-108.

Wilcockson WH (1956), Preliminary notes on the geology of Kilimanjaro. *Geological Magazine*, **93**, no. 03, p. 218.

Zachos J, Pagani M, Sloan L, Thomas E, Billups K, (2001) Trends, Rhythms, and Aberrations in Global Climate 65 Ma to Present. *Science,* **292**, 686-693.

Appendix

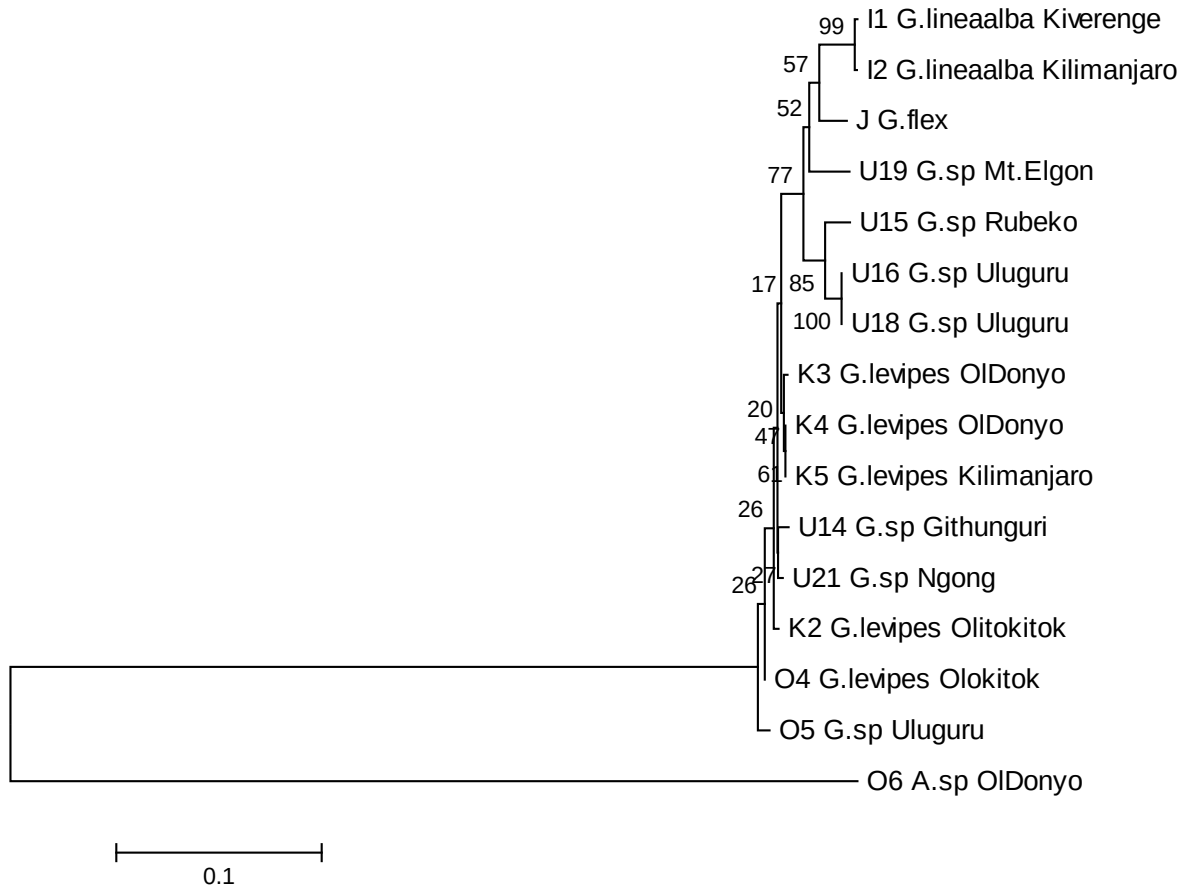

Fig. F1 Neighbor joining tree of the phylogeny of the genus *Gymnobothroides* based on the 16S-gene. The nucleotide substitution model used was the maximum composite likelihood. The branch support values are bootstrap values based on 100 replicates.

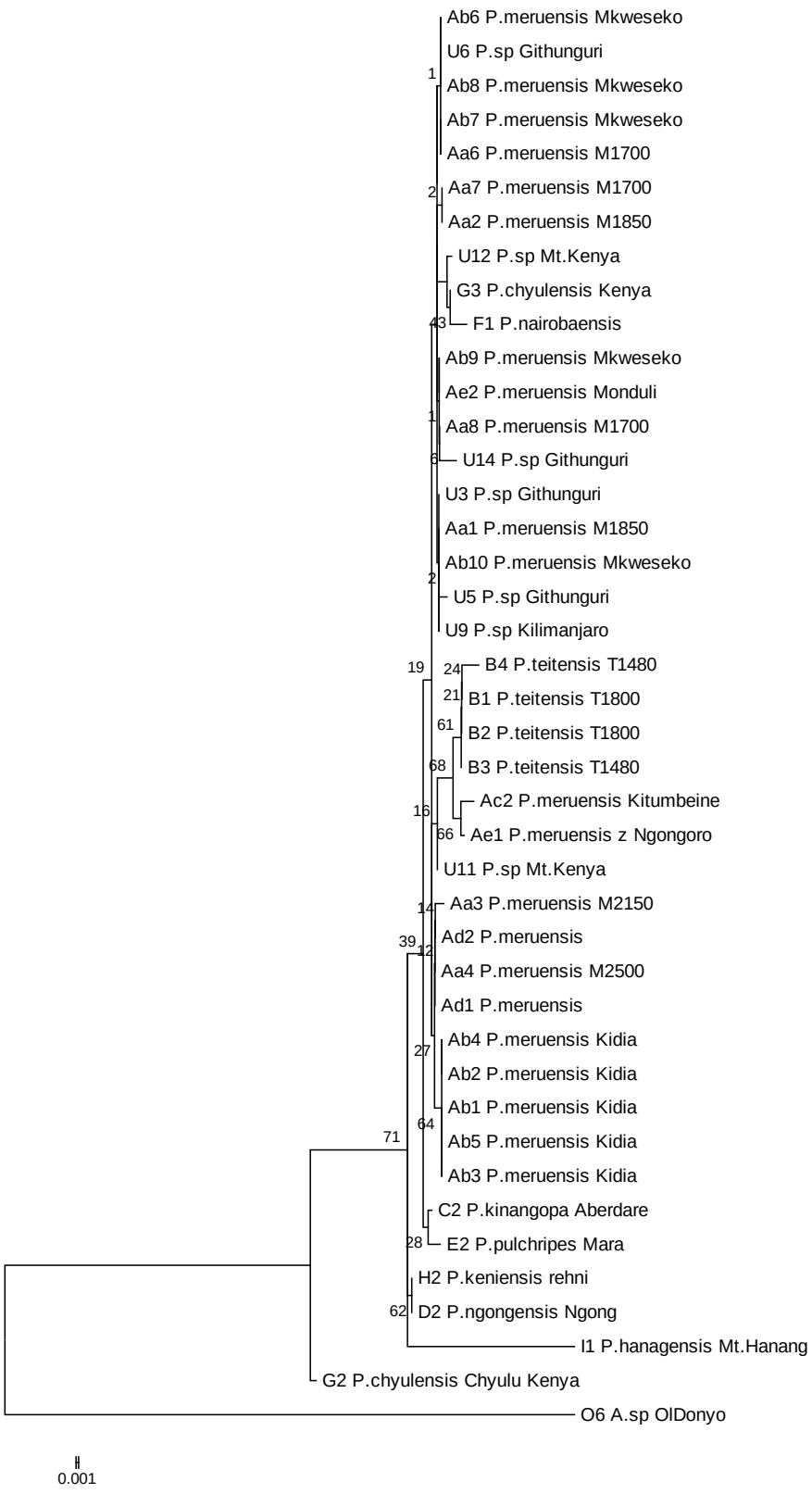

Fig. F2 Neighbor joining tree of the phylogeny of the genus *Parasphena* based on the 16S-gene. The nucleotide substitution model used was the maximum composite likelihood. The branch support values are bootstrap values based on 100 replicates.

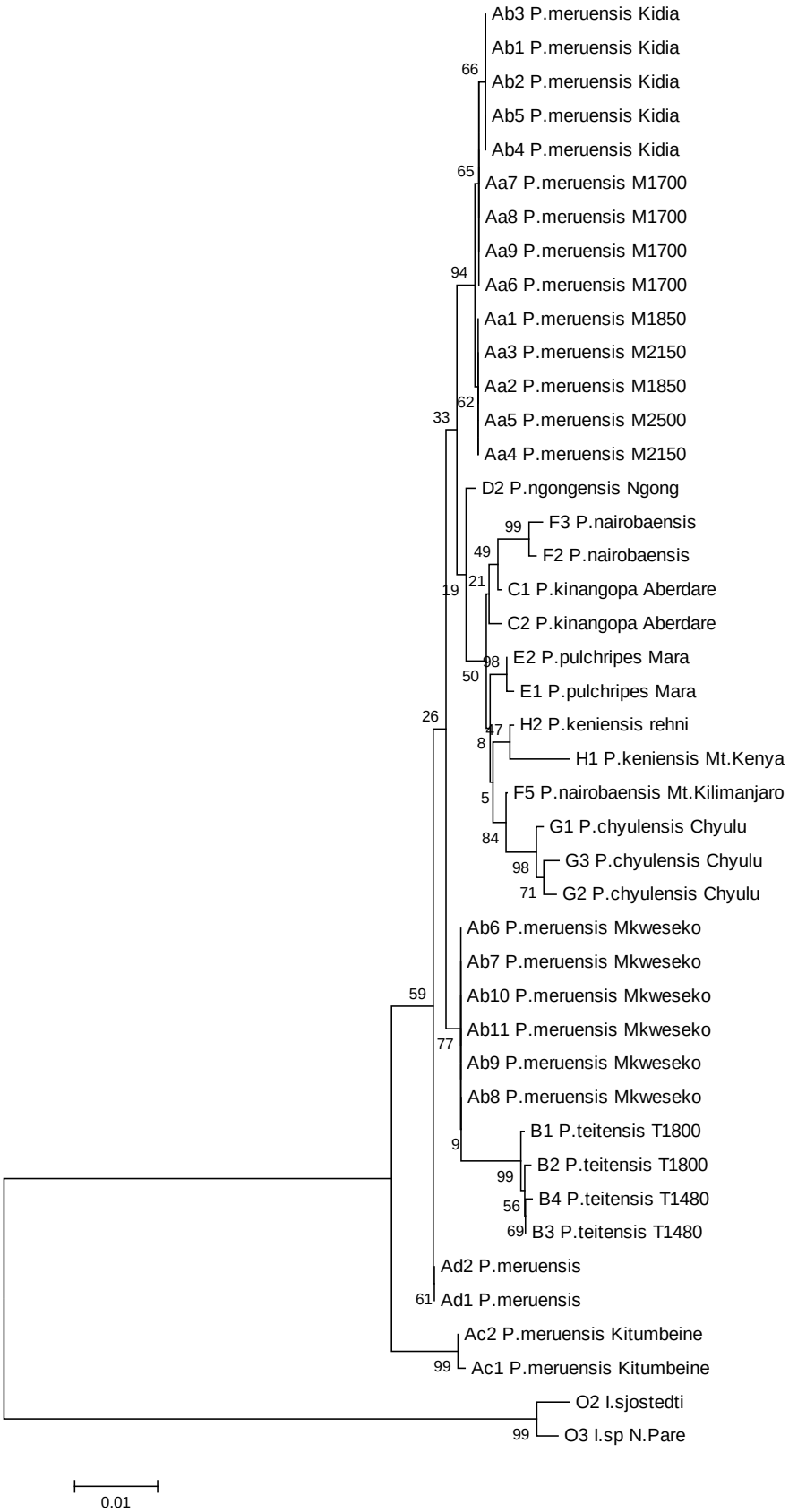

Fig. F3 Neighbor joining tree of the phylogeny of the genus *Parasphena* based on the COI-gene. The nucleotide substitution model used was the maximum composite likelihood. The branch support values are bootstrap values based on 100 replicates.

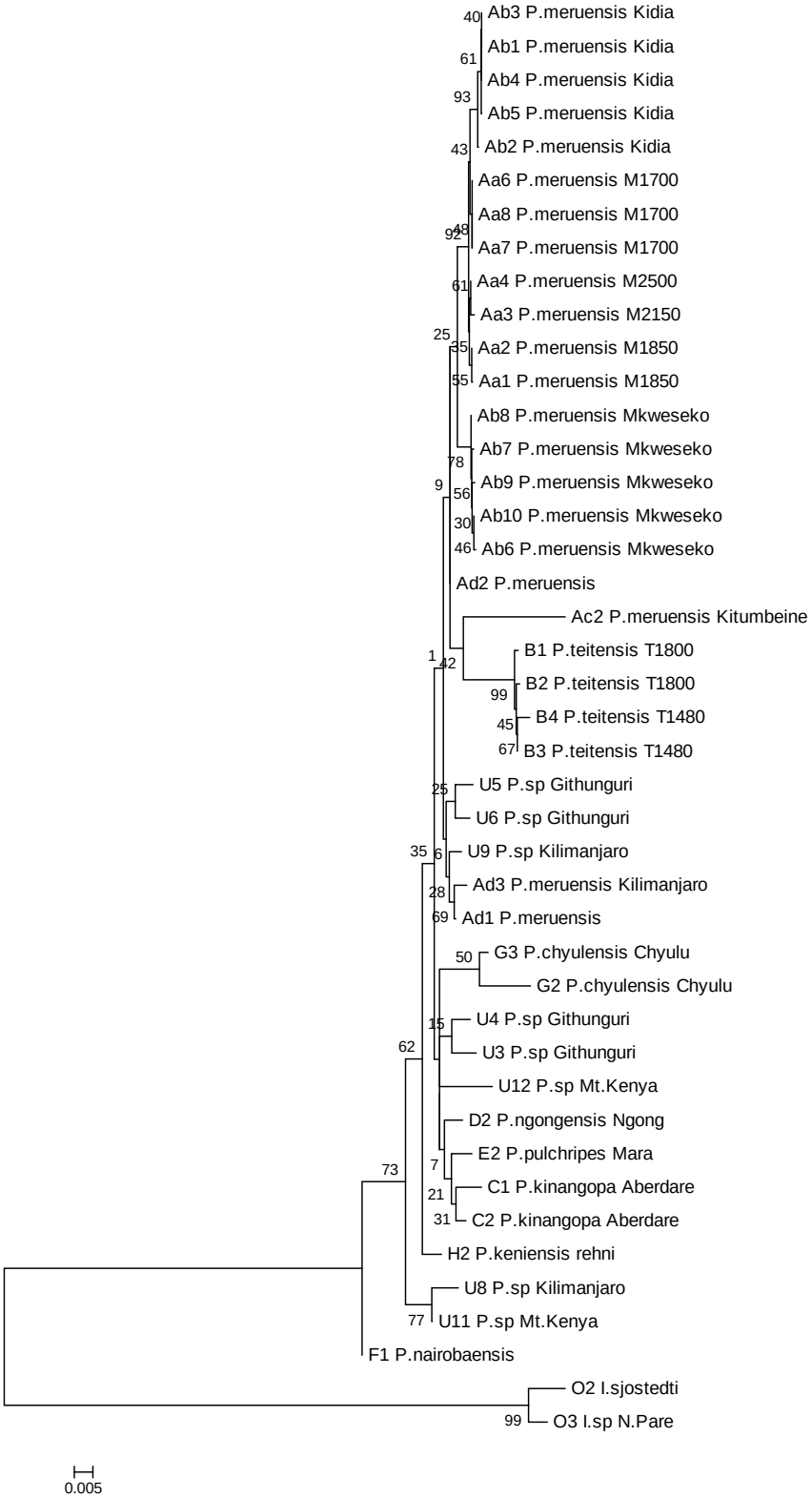

Fig. F4 Neighbor joining tree of the phylogeny of the genus *Gymnobothroides* based on both the 16S and COI-genes. The nucleotide substitution model used was the maximum composite likelihood. The branch support values are bootstrap values based on 100 replicates.

Tab. T1 P-distances between all *Gymnobothroides* samples on the 16S-gene.

	1	2	3	4	5	6	7	8	9	10	11	12
1)#I1_G.lineaalba_Kiverengi												
2)#I2_G.lineaalba_Kilimanjaro	0.001											
3)#U14_G.sp_Githunguri	0.021	0.019										
4)#U15_G.sp_Rubeko	0.025	0.024	0.023									
5)#U16_G.sp_Uluguru	0.021	0.02	0.017	0.01								
6)#O5_G.sp_Uluguru	0.024	0.025	0.012	0.021	0.016							
7)#U19_G.sp_Mt.Elgon	0.018	0.017	0.017	0.027	0.018	0.024						
8)#U21_G.sp_Ngong	0.021	0.019	0.003	0.02	0.014	0.009	0.017					
9)#J_G.flex	0.016	0.017	0.021	0.021	0.021	0.022	0.024	0.019				
10)#O4_G.levipes_Olitokitok	0.019	0.021	0.006	0.021	0.016	0.006	0.018	0.004	0.018			
11)#K2_G.levipes_Olitokitok	0.018	0.019	0.005	0.021	0.017	0.009	0.02	0.002	0.017	0.004		
12)#K3_G.levipes_OlDonyo	0.019	0.021	0.006	0.02	0.016	0.008	0.018	0.004	0.018	0.002	0.004	
13)#K4_G.levipes_OlDonyo	0.021	0.019	0.005	0.018	0.014	0.009	0.017	0.002	0.019	0.004	0.005	0.001

Tab. T2 P-distances between both *Parasphena* and *Gymnobothroides* for the 16S-gene.

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]
[1] #B1_P.teitensis_T1800															
[2] #Ab3_P.meruensis_Kidia	0.053														
[3] #Ab2_P.meruensis_Kidia	0.039	0.013													
[4] #Aa6_P.meruensis_M1700	0.053	0.025	0.013												
[5] #Ac2_P.meruensis_Kitumbeine	0.013	0.039	0.026	0.039											
[6] #Ae1_P.meruensis_z_Ngongoro	0.026	0.052	0.039	0.026	0.013										
[7] #O6_A.sp_OlDonyo	0.342	0.352	0.377	0.352	0.377	0.352									
[8] #G3_P.chyulensis_Kenya	0.053	0.025	0.013	0.025	0.039	0.052	0.377								
[9] #G2_P.chyulensis_Chyulu_Keny	0.08	0.081	0.067	0.081	0.095	0.11	0.341	0.053							
[10] #C1_P.kinangopa_Aberdare	0.098	0.067	0.053	0.04	0.082	0.068	0.401	0.067	0.128						
[11] #C2_P.kinangopa_Aberdare	0.067	0.038	0.025	0.013	0.052	0.039	0.328	0.038	0.095	0.026					
[12] #E2_P.pulchripes_Mara	0.053	0.052	0.039	0.025	0.039	0.025	0.303	0.052	0.11	0.04	0.013				
[13] #H3_P.keniensis_keniensis	0.177	0.142	0.158	0.174	0.158	0.174	0.35	0.158	0.128	0.197	0.158	0.142			
[14] #D2_P.ngongensis_Ngong	0.053	0.052	0.039	0.052	0.039	0.052	0.302	0.052	0.081	0.068	0.039	0.025	0.111		
[15] #U4_P.sp_Githunguru	0.133	0.099	0.084	0.069	0.115	0.1	0.268	0.099	0.165	0.087	0.055	0.07	0.241	0.1	
[16] #U8_P.sp_Kilimanjaro	0.068	0.039	0.026	0.013	0.053	0.039	0.361	0.039	0.082	0.054	0.026	0.039	0.176	0.053	0.085
[17] #U11_P.sp_Mt.Kenya	0.026	0.025	0.013	0.025	0.013	0.025	0.35	0.025	0.081	0.067	0.039	0.025	0.142	0.025	0.099
[18]#U12_P.sp_Mt.Kenya	0.039	0.039	0.025	0.039	0.025	0.039	0.35	0.013	0.067	0.082	0.052	0.038	0.142	0.039	0.114
[19] #L1_P.hanaagensis_Mt.Hanang	0.068	0.067	0.053	0.067	0.053	0.067	0.272	0.067	0.096	0.083	0.053	0.039	0.128	0.013	0.117
[20] #I1_G.lineaalba_Kiverenge	0.23	0.236	0.254	0.236	0.235	0.217	0.361	0.254	0.195	0.265	0.218	0.2	0.068	0.199	0.32
[21] #U14_G.sp_Githunguri	0.174	0.177	0.193	0.21	0.193	0.21	0.361	0.21	0.141	0.236	0.193	0.176	0.054	0.144	0.285
[22] #U15_G.sp_Rubeko	0.218	0.223	0.241	0.26	0.241	0.26	0.327	0.241	0.165	0.291	0.241	0.222	0.055	0.186	0.291
[23] #U18_G.sp_Ulugurú	0.196	0.218	0.236	0.254	0.217	0.235	0.361	0.236	0.162	0.285	0.236	0.217	0.054	0.182	0.341
[24] #O5_G.sp_Ulugurú	0.193	0.197	0.214	0.231	0.214	0.231	0.327	0.214	0.144	0.26	0.214	0.196	0.04	0.162	0.312
[25] #U19_G.sp_Mt.Elgon	0.235	0.214	0.231	0.249	0.213	0.23	0.426	0.231	0.199	0.278	0.231	0.213	0.053	0.179	0.332
[26] #J_G.flex	0.21	0.214	0.231	0.214	0.231	0.214	0.303	0.231	0.159	0.241	0.197	0.179	0.053	0.179	0.291
[27] #K3_G.levipes_OlDonyo	0.21	0.214	0.231	0.249	0.231	0.249	0.327	0.231	0.159	0.278	0.231	0.213	0.053	0.179	0.312

[16]	[17]	[18]	[19]	[20]	[21]	[22]	[23]	[24]	[25]	[26]
0.039										
0.053	0.013									
0.067	0.039	0.053								
0.24	0.235	0.235	0.221							
0.213	0.176	0.192	0.162	0.053						
0.264	0.222	0.222	0.207	0.068	0.04					
0.259	0.217	0.217	0.203	0.025	0.026	0.04				
0.235	0.196	0.196	0.182	0.039	0.013	0.026	0.013			
0.253	0.213	0.213	0.199	0.039	0.054	0.069	0.026	0.04		
0.217	0.213	0.213	0.199	0.026	0.026	0.04	0.026	0.013	0.053	
0.253	0.213	0.213	0.199	0.052	0.026	0.026	0.026	0.013	0.053	0.025

Tab. T3 P-distances between all *Parasphena* species on the COI-gene.

	1	2	3	4	5	6	7	8	9	10	11
1) #B4_P.teitensis_T1480											
2) #B3_P.teitensis_T1480	0.002										
3) #B1_P.teitensis_T1800	0.005	0.002									
4) #Ab6_P.meruensis_Mkweseko	0.017	0.014	0.017								
5) #Ab4_P.meruensis_Kidia	0.024	0.021	0.024	0.007							
6)#Aa7_P.meruensis_M1700	0.021	0.019	0.021	0.005	0.002						
7)#Ac2_P.meruensis_Kitumbeine	0.059	0.057	0.059	0.042	0.039	0.037					
8)#Ad2_P.meruensis	0.026	0.024	0.026	0.009	0.007	0.005	0.032				
9)#Ad3_P.meruensis_Kilimanjaro	0.031	0.029	0.026	0.014	0.012	0.009	0.037	0.005			
10)#G3_P.chyulensis_Chyulu	0.051	0.048	0.051	0.034	0.031	0.029	0.057	0.024	0.024		
11)#G2_P.chyulensis_Chyulu	0.056	0.054	0.056	0.039	0.036	0.034	0.062	0.029	0.029	0.005	
12)#F1_P.nairobaensis	0.056	0.054	0.056	0.041	0.039	0.036	0.057	0.031	0.031	0.031	0.031
13)#C1_P.kinangopa_Aberdare	0.031	0.029	0.026	0.019	0.017	0.014	0.042	0.009	0.005	0.024	0.029
14)#E2_P.pulchripes_Mara	0.039	0.036	0.039	0.022	0.019	0.017	0.044	0.012	0.012	0.021	0.021
15)#H2_P.keniensis_rehni	0.036	0.034	0.036	0.024	0.021	0.019	0.042	0.014	0.014	0.024	0.029
16)#D2_P.ngongensis_Ngong	0.029	0.026	0.029	0.012	0.009	0.007	0.039	0.007	0.007	0.021	0.026
17)#U5_P.sp_Githunguri	0.034	0.031	0.029	0.021	0.021	0.019	0.047	0.014	0.009	0.024	0.029
18)#U6_P.sp_Githunguri	0.034	0.031	0.029	0.021	0.021	0.019	0.047	0.014	0.012	0.031	0.036
19)#U4_P.sp_Githunguri	0.034	0.031	0.034	0.017	0.014	0.012	0.039	0.007	0.007	0.017	0.021
20)#U3_P.sp_Githunguri	0.034	0.031	0.034	0.017	0.014	0.012	0.037	0.007	0.007	0.021	0.026
21)#U9_P.sp_Kilimanjaro	0.021	0.019	0.021	0.014	0.012	0.009	0.037	0.005	0.009	0.029	0.034
22)#U11_P.sp_Mt.Kenya	0.059	0.057	0.059	0.044	0.041	0.039	0.054	0.034	0.034	0.039	0.039
23)#U12_P.sp_Mt.Kenya	0.054	0.051	0.054	0.036	0.034	0.031	0.06	0.026	0.026	0.007	0.002

0.029	0.007	0.034	0.024	0.024	0.039	0.031	0.029	0.024	0.024	0.031													12
0.026	0.034	0.009	0.007	0.007	0.007	0.005	0.012	0.014	0.012														13
0.019	0.026	0.017	0.009	0.005	0.019	0.012	0.009	0.012															14
0.026	0.022	0.014	0.012	0.007	0.021	0.014	0.012																15
0.024	0.031	0.012	0.009	0.005	0.019	0.012																	16
0.026	0.034	0.014	0.012	0.007	0.007																		17
0.034	0.041	0.014	0.014	0.014																			18
0.019	0.026	0.012	0.005																				19
0.024	0.026	0.012																					20
0.031	0.036																						21
0.036																							22

Tab. T4 List of species included in this thesis, the location they were found (when known), and the number of samples from each species.

Species name	**Location (when known)**	**No. Samples**
Parasphena meruensis	Kilimanjaro, Mt. Meru, Monduli, Kitumbeine, Ngorongoro	26
Parasphena chyulensis	Chyulu	3
Parasphena teitensis	Taita Hills	4
Parasphena nairobaensis	Kilimanjaro	4
Parasphena kinangopa	Aberdare	2
Parasphena pulchripes	Mara	2
Parasphena keniensis	Mt. Kenya	2
Parasphena ngongensis	Ngong	1
Parasphena hanangensis	Hanang	1
Parasphena sp		7
Gymnobothroides levipes	Oloitokitok, OlDonyo, Kilimanjaro	5
Gymnobothrus lineaalba	Kiverengi, Kilimanjaro	2
Gymnobothroides flex		1
Gymnobothroides sp		7