UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA ANIMAL

Indexing Population Size of Elusive Species From Sign Surveys: An Evaluation Using Non-invasive Genetics of Cabrera Voles

Maria Teresa Courinha de Torres Vaz Freire

Mestrado em Biologia da Conservação

Dissertação orientada por:
Professora Doutora Maria da Luz Mathias
Doutor Ricardo Pita

2017

Agradecimentos

Esta não é a primeira nem será, certamente, a última vez que terei que agradecer às pessoas maravilhosas que tenho na minha vida. Não foi um processo fácil, aliás, foram mais os percalços do que as coisas que efetivamente correram sempre bem, mas é assim que se faz ciência.

Gostaria muito de poder agradecer individualmente a todas as pessoas que expressaram o seu apoio e partilharam do meu #fuisóalifazeratese, mas temo que não haja espaço para todas. Assim, peço que a quem não for diretamente referido que não se sinta menos importante.

Ao Doutor Ricardo Pita tenho que agradecer profundamente toda a disponibilidade e paciência que demonstrou, mesmo quando as coisas pareciam condenadas, e pela oportunidade que me deu de integrar o projeto.

À Professora Maria da Luz Mathias por ter aceite ser minha orientadora.

Ao Professor António Mira por ter sempre um lugar disponível para esta aluna que parece não querer deixar o ponto de partida.

À Joana Paupério por me ter recebido numa casa nova de uma nova cidade, por se ter mostrado sempre disponível e por ser incansável.

Aos meus colegas de campo Dinora Peralta e Tiago Mendes por me deixarem ser eu a pessimista e se assegurarem que, apesar de tudo, haveria amostras, e à minha colega de laboratório Clara Ferreira que teve uma paciência infinita para quem, afinal, nunca foi um "ratinho de laboratório". A todas as outras pessoas da Unidade de Biologia da Conservação da Universidade de Évora ou do CIBIO de Vairão que tenham passado pelo meu caminho.

Ao projeto MATEFRAG - PTDC/BIA-BIC/6582/2014, por todo o apoio logístico na realização do trabalho prático.

À minha companheira de empreendimento, Cláudia Afonso, e aos meus companheiros de cafés vários, porque procrastinar também faz parte. À Marta Escobar, por me ter aturado e me ter feito sentir em casa, no Porto. À Dinora, mais uma vez, porque não foi só colega de campo.

Ao João, por tudo.

Por fim, e porque o fim é o princípio, queria agradecer à minha família porque sem eles nada disto seria possível. À minha irmã, por ser a mais velha e cheia de experiências válidas de mestranda, e aos meus pais pela oportunidade de, mais uma vez, deixar o ninho para outras paragens.

Obrigado a todos!

#fuisóalifazeratese

Nota Prévia

Na escrita desta dissertação optou-se pela língua inglesa por ser a língua universalmente aceite para publicações científicas internacionais. A tese foi organizada em 3 capítulos: um de contextualização geral do problema, um capítulo principal em forma de artigo científico, e um capítulo de síntese dos principais resultados e conclusões.

Na escrita do capítulo principal, a preferência pelo formato de artigo científico deveu-se ao facto de este ser o formato mais comum na disseminação de estudos científicos e por, no futuro, se pretender submeter o trabalho a uma revista da respetiva área científica, indexada na base ISI. Assim, embora o artigo não tenha seguido a formatação específica de nenhuma revista em particular, a estrutura corresponde ao que é normalmente aceite em publicações científicas (e.g. Elsevier, ou Springer).

A autora esclarece ainda que os trabalhos foram feitos em colaboração, e que em todos eles participou ativamente na sua recolha, análise e redação.

Resumo

O conhecimento do tamanho e abundância de populações selvagens é fundamental para a gestão de recursos naturais. Uma das formas de estimar o tamanho de uma população é o de capturar, marcar, libertar e, mais tarde, recapturar indivíduos. Os estudos de captura-marcação-recaptura (CR) permitem ainda obter parâmetros demográficos como a taxas de sobrevivência, migração, e crescimento populacional. Mas, a obtenção de informação relativa à distribuição e tamanho das populações de espécies elusivas e/ou com problemas de conservação, é, muitas vezes, difícil recorrendo aos métodos tradicionais de armadilhagem, quer por poder haver uma resposta negativa dos animais às armadilhas, quer devido ao risco de causar danos aos animais. Para além disso, os métodos tradicionais implicam grandes quantidades de recursos materiais e humanos, sendo muito exigentes em termos de tempo, e estando maioritariamente reservados a pessoal treinado. Estas limitações muitas vezes superam os benefícios da obtenção de informação, o que constitui um desafio em estudos de ecologia e conservação. A alternativa mais comum aos métodos de captura-recaptura é o uso de métodos indiretos baseados em indícios de presença cujos resultados permitem inferir abundâncias relativas. No entanto, os índices de abundância populacional não devem ser utilizados na gestão das populações, sem uma validação prévia que avalie se as variações no índice refletem as variações de tamanho das populações. Contudo, dada a dificuldade de se conhecer o verdadeiro tamanho de uma população selvagem, a validação de índices de abundância torna-se geralmente uma tarefa difícil. Assim, uma tentativa para estimar verdadeiras abundâncias ou densidades a partir de um índice, só será suportada se houver estudos *a priori* onde se comparam os valores desse índice com valores de densidades reais, através de modelação estatística.

Os recentes desenvolvimentos em técnicas de genética não-invasiva que permitem a captura do DNA de um individuo, em vez do individuo propriamente dito, representam uma melhoria significativa para estudos de abundância, uma vez que permitem que seja recolhida uma maior quantidade de informação. Os benefícios da CR genética incluem a diminuição do risco e perturbação para o animal, uma vez que este nunca está fisicamente preso e pode deixar várias amostras contendo material genético em múltiplas localizações durante o tempo de amostragem, baixando também o enviesamento causado pela resposta às armadilhas. A monitorização de tendências populacionais através de métodos de genética não-invasiva pode ser também uma mais valia para identificar declínios e ameaças de extinção e também para avaliar a eficácia das medidas de conservação aplicadas a uma população. O genótipo de cada indivíduo pode ser obtido de amostras não-invasivas de várias fontes. Cada genótipo específico é tratado como sendo uma "marca", enquanto uma "recaptura" acontece de cada vez que o mesmo genótipo é encontrado em duas amostras diferentes. As fezes são o material mais utilizado em estudos de genética não-invasiva, uma vez que para a maioria das espécies, são relativamente fáceis de encontrar, e conseguem providenciar uma grande quantidade de informações relevantes (ex. perfil genético dos indivíduos, dieta, regulação hormonal, presença de DNA de parasitas, etc.).

Em estudos de genética não-invasiva, é possível a utilização de diferentes marcadores para obtenção de diferentes tipos de informação. Os marcadores mitocondriais permitem a identificação da espécie enquanto que marcadores nucleares, como os microssatélites, permitem a construção do perfil genético dos indivíduos, bem como a identificação do sexo.

Neste estudo utilizamos o rato de Cabrera como espécie modelo para testar a definição de índices de tamanho populacional a partir de indícios de presença, validados com estimativas obtidas a partir de amostragem genética não-invasiva. Pretendemos também perceber qual o tipo de indícios mais adequados à construção dos índices, bem como qual o esforço de amostragem mínimo necessário para a obtenção de um índice que reflita a variação no tamanho da população.

O rato de Cabrera (*Microtus cabrerae*, Thomas 1906) é uma espécie de arvicolídeo de tamanho médio, endémico da região Iberoccitana e, atualmente, restrito à Península Ibérica, onde a sua distribuição se encontra largamente fragmentada, situação que se tem vindo a agravar nas últimas décadas.

Uma vez que, para esta espécie, é difícil a obtenção de dados de abundância por armadilhagem, devido à sua resposta negativa às armadilhas, a maioria dos estudos até agora realizados têm sido baseados em regurgitações de rapinas e indícios de presença específicos, os quais são por vezes usados na construção de índices de abundância. Os indícios de presença típicos incluem túneis, fragmentos de erva cortada, e os dejetos verde escuros com 4 a 8 mm de comprimento que são frequentemente encontrados acumulados em latrinas, as quais podem ser utilizadas por um ou vários indivíduos. Estes estudos têm, até agora, permitido a obtenção de informação importante sobre abundâncias locais e dinâmicas populacionais, os quais permitiram delinear propostas de medidas de conservação a nível local e regional.

O trabalho de campo foi realizado no SW de Portugal, integrando parcialmente o Parque Natural da Costa Vicentina e Sudoeste Alentejano. Foi feita uma identificação prévia dos habitats adequados à espécie, e uma procura por indícios de presença pela mesma. Em 8 dos habitats em que a presença da espécie foi confirmada, foram identificadas e contadas latrinas e recolhidas amostras de dejetos frescos usando esforço total de prospeção constante (ca. 60min/0.1ha), ao longo de quatro dias consecutivos, tendo-se recolhido um total de 601 amostras. As amostras foram extraídas e genotipadas segundo o protocolo proposto por Ferreira et al. (submetido) em condições de esterilização para minimizar os riscos de contaminação. Foram selecionadas para análise todas as amostras com pelo menos 3 dejetos, que se encontrassem a uma distância mínima de 2 metros da recolha mais próxima, totalizando 509 amostras. Estas amostras foram genotipadas para um total de 9 microssatelites, que permitem a identificação individual, e sexadas. Com os valores obtidos pelas análises de genética não-invasiva construíram-se duas curvas de rarefação, segundo os pressupostos descritos por Kohn (1999) e Eggert (2003) e com uma randomização de 100 vezes para cada habitat. Adicionalmente, como forma de verificação de concordância entre estimadores, correram-se modelos CR no Capwire (para R).

Do número total de amostras recolhidas e extraídas, obtiveram-se genótipos consensuais para 129, correspondendo a um sucesso de genotipagem de cerca de 25%, semelhante ao verificado para outras espécies de pequenas dimensões ou que habitam em zonas húmidas, em que a quantidade e qualidade de DNA presente nas fezes é reduzida. As restantes amostras foram excluídas por apresentarem baixa qualidade ou estarem contaminadas. Com recurso às curvas de rarefação, foi possível obter estimativas estáveis e com valores de erro relativamente baixas. Optou-se por utilizar os resultados da curva de Eggert como referência do tamanho real das populações, uma vez que este estimador é considerado melhor que o método de Kohn quando os tamanhos populacionais são baixos, e que os modelos do Capwire só produziram resultados para 3 patches, provavelmente devido ao reduzido tamanho das populações presentes nos restantes (<7 indivíduos por *patch*). Das contagens de latrinas geraram-se 9 índices de tamanho populacional (PSI), contruídos segundo três esforços de amostragem distintos: 60min/0.1ha, 30min/0.1ha, 15min/0.1ha, e considerando latrinas grandes (>20 dejetos), latrinas pequenas (< 20 dejetos) e número total de latrinas. Os PSI foram correlacionados com os tamanhos populacionais de referência, através de Modelos Generalizados de Efeitos Mistos (GLMM), incluindo como fatores aleatórios o mês de amostragem e os 4 dias consecutivos de amostragem em cada habitat (pseudo-réplicas), de modo a identificar qual o PSI que melhor se correlaciona com os tamanhos populacionais de referência. Os resultados dos GLMM foram significativos apenas para o PSI que considera as contagens de latrinas pequenas com esforços de amostragem maiores, o qual se correlaciona positivamente com as estimativas de referência de tamanhos populacionais.

Os resultados deste trabalho, mostraram que é possível obter um índice de tamanho populacional de forma expedita e pouco dispendiosa, com um esforço de amostragem razoável. De qualquer forma, será pertinente aumentar o número de habitats estudado de forma a validar este resultado. Com base no modelo estatisticamente suportado, a contagem de latrinas pequenas, com o esforço de amostragem maior, poderá servir para fazer inferências sobre o tamanho das populações. Como latrinas maiores possivelmente são mais usadas em épocas de acasalamento, i.e., quando indivíduos se encontram mais ativos, desaconselha-se a utilização das mesmas como índice de tamanho populacional. De ressalvar que, para além da importância de se confirmar estes resultados incluindo um maior número de habitats, estudos futuros deverão considerar a inclusão

de outros indícios de presença (túneis, restos de erva cortada), assim como de esforços de amostragem maiores, de forma a quantificar mais precisamente o esforço ideal à amostragem de cada habitat. Assim, consideramos que este estudo representa o início de um caminho que poderá conduzir a novas e melhores formas de estudar as flutuações das populações do rato de Cabrera para que, no futuro, possam ser aplicados mais e melhores planos de gestão da espécie.

Palavras-chave: Captura-marcação-recaptura; indícios de presença; amostragem genética-não-invasiva; índices de tamanho populacional.

Abstract

Understanding the factors that affect the abundance of animals in wild populations is crucial in ecological research and species conservation. However, obtaining the necessary data on elusive and endangered species is usually a challenge. Live-trapping has been used for a long time for obtaining population size estimates on several mammal species, particularly small mammals. However, live-trapping is often difficult to implement in the field and often provides insufficient data to infer population size, particularly for species that occur at low densities, or that have a negative response to trapping, or that have very extensive habitats. Recently, genetic non-invasive sampling (gNIS) has been used as an alternative to live-trapping for obtaining more data on several population parameters, including population density and abundance. By genotyping samples from wild animals without capturing or disturbing them (e.g. faeces), gNIS allows the obtainment of unique genetic fingerprint of the individuals in a population. However, when looking for an expedite and cost-effective method to sample populations, most researchers rely on population size indexes (PSI), which are usually based on counts of species presence signs, that are assumed to have a positive relation to true population size. Despite, their utility to infer population abundance, PSI's are seldom tested against known numbers. This may result in important bias, which in turn may affect the effectiveness of conservation or management measures based solely on PSI's index values. This study addresses the issue of obtaining expedite and cost-effective methods of surveying an elusive small mammal by generating PSI's, and validating them against population sizes obtained via capture-recapture (CR) data from gNIS. For this, we used the Cabrera vole (*Microtus cabrerae*) in SW Portugal farmland as a model species. The Cabrera vole is a "Near-Threatened" Iberian endemism, restricted to relatively small habitat patches dominated by wet herbaceous vegetation. These habitats have decreased dramatically during the past, mainly due to agricultural intensification. Between November-16 and March-17, we selected 8 occupied habitat patches, and in each one we searched and counted vole latrines, and collected fresh faeces samples for genotyping, based on 9 microsatellite loci and 1 mitochondrial loci. From the vole latrine counts we generated 9 PSI indexes based on varying sampling efforts and type presence sign considered. We then used GLMM to assess the relationship between each of the 9 PSI's and the estimates of population size, obtained from Eggert's rarefaction curve applied to the gNIS data. Results indicated that, despite the relatively low genotyping success (ca. 25%), population size estimates based on gNIS were robust, and correlated significantly with the PSI based on counts of small latrines (<20 faeces) under higher sampling efforts (≥60 min/0.1ha). Our study suggests that it is possible to use PSIs for monitoring Cabrera vole's populations, which may thus provide an opportunity for monitoring the species over large spatial and temporal scales.

Keywords: Population size; Indexes; Genetic non-invasive sampling; Cabrera vole; Capture-recapture

General Table of Contents

Table of Figures

Table of Tables

1. General Introduction

1.1 The importance of monitoring wild populations

The ability to monitor the status of wild populations is crucial for population ecology and management (Engeman 2005; Witmer 2005; Jareño et al. 2014). Understanding the characteristics of a population (e.g. size, structure, intraspecific relations), its relation to other populations and species, and how it is affected by the environment are key issues for wildlife conservationists and managers.

Population monitoring may be performed at various levels of resolution. Biologists may either need to know where a particular species occurs (i.e., species distribution) or how many animals are in an area (i.e., species abundance; Gese 2001). Estimating the size or abundance of wild populations plays a vital role in managing populations and conserving rare and endangered species. Animal abundance may be assessed as relative abundance and absolute abundance. Relative abundance is obtained through indexes based on raw counts of captured animals per effort unit, or indirect signs of animals (e.g., footprints or tracks, faeces, food remains, etc; Gese 2001; Bonesi & Macdonald 2004; Witmer 2005) that can be compared over time or between areas, but, in itself does not provide animal numbers. In contrast, absolute abundance involves using methods to count animals and then estimate the number or density of animals in the population (Village & Myhill 1990). With repeated sampling over time, both relative indexes and absolute estimates of animal abundance can be used to monitor population trends (Gese 2001).

Despite being much less informative than absolute measures, abundance indexes, often provide baseline information about the status of animal populations, being considerably easier to obtain in the field than absolute abundances. In particular, assessing the abundance of small mammal species is difficult because small mammals are mainly nocturnal and elusive, and some species are endangered and occur in small numbers.

1.2 Methods for estimating abundances of wild populations

1.2.1. Capture-recapture techniques

If animals can be trapped or observed in a single survey it is possible to generate an abundance index based on the number of captures per sampling effort which may (or may not) correlate with the unknown absolute number of animals present. Obtaining absolute measures of small mammal abundance, requires capture-mark-recapture (CR) methods involving multiple surveys. CR methods consist in tagging animals with individual marks at first capture in order to allow their identification in subsequent surveys, and then estimate the total number of animals within a population using CR models (e.g. Miller et al. 2005; Petit & Valiere 2005). CR models may be implemented based on several sampling techniques, such as live-trapping, camera trapping or, more recently, genetic non-invasive sampling (gNIS) (De Bondi et al. 2010; Torre et al. 2010; Kilpatrick et al. 2013; Sabino-Marques et al. submitted). An important feature of CR models is that they require high detection probabilities and large number of individuals captured and recaptured, which has proven to be a challenge even when using the most recent gNIS techniques (Anderson 2001; Bain et al. 2014).

1.2.2. Capture-recapture based on live-trapping

At present, live-trapping is still the traditional approach for estimating small mammal population size and abundance (Tasker & Dickman 2001). These methods rely on the physical capture of individuals (e.g. with Elliot traps and Sherman traps), and have been successfully used to detect species composition and abundance in several distinct conditions. Live-trapping provides unambiguous observations of species and enable the collection of detailed information, such as, sex, reproductive status, and physical condition (Jareño et al. 2014). Some of the major advantages is that live-trapping is that it allows for simultaneous monitoring of multiple species (possibly including relatively rare or protected species), providing data that can easily be

compared between species over space and time (Flowerdew et al. 2004). In addition, live-trapping can be conducted under more severe weather conditions (e.g. rainy days).

For some species, capturing individuals is frequently the only feasible method of estimating abundance but, as only a proportion of the population is likely to be captured in a trapping session – different capture probabilities among individuals (Burnham & Overion 1979; Slade et al. 2012) -, it becomes very hard and time-consuming to implement this type of survey, as it requires a lot of material and human resources, especially mark–recapture methods based on several days of continuous trapping (Tellería 2004; Witmer 2005). Other sources of bias can arise during a trapping study, such as assuming that a population is closed within each sampling occasion – no births, deaths, immigration or emigration happens -, when it is not (Crosbie & Manly 1985), and eventual losses of the individuals mark, or the capture event negatively influencing the animal's probability of recapture or survival (Crosbie & Manly 1985).

While live-trapping has been the "century-old" method for sampling small mammals (e.g.; (Chitty & Kempson 1949; Stickel 1954; Sibbald et al. 2006; De Bondi et al. 2010; Torre et al. 2010) it may be largely unfeasible for species that occur in such large numbers and are considered a plague, or for species that are endangered, elusive or have a negative response to trapping, as it involves logistics that are too complicated and time-and-money consuming. This justifies the urgent need for cost-effective methods of estimating population sizes and fluctuations, which can be used for conservation management.

1.2.3. Capture-recapture based on genetic non-invasive sampling

gNIS has become an important tool for the estimation of animal abundance, especially in the case of rare and elusive species, as it is based in presence signs (e.g. faeces, shed hair or feathers, sloughed skin; Taberlet et al. 1999; Pearse et al. 2001) and therefore allows the capture of an individual's DNA without having direct contact with the animal (Taberlet et al. 1996; Taberlet et al. 1999; Frantz et al. 2004). Faeces are the most common non-invasive material (Beja-Pereira et al. 2009) because, for most species, faeces are always available for most individuals (Ramón-Laca et al. 2015), being relatively easy to identify in the wild. Besides, feaces allows to assemble a higher variety of information (e.g. diet, stress hormone status, reproductive hormones, parasite infection) than other sampling techniques (Beja-Pereira et al. 2009). The DNA from faeces is isolated from cells shed from the intestinal walls of the animals (Kohn & Wayne 1997).

Non-invasive samples can be collected systematically across the study area and, after the DNA extraction, typed for instance for mitochondrial markers to confirm species' identity (Kohn et al. 1999). These samples may be then typed for nuclear markers that allow for identification and sexing of the individuals (Kohn et al. 1999). Microsatellites are nuclear, single-copy DNA markers composed by short-sequence motifs repeated in tandem (Kohn & Wayne 1997), that are scattered through the genome and are highly polymorphic. By analyzing multiple variable microsatellite *loci*, one can obtain an individual's multilocus genotype, or it's genetic fingerprint (Kohn & Wayne 1997, McKelvey & Schwartz 2004). There are many reasons to be using microsatellite markers for genetic non-invasive sampling (Mills et al. 2000), such as their variability in many species, their selection neutrality and independent evolution and the fact that there are already many markers available for several species.

Extracting and genotyping DNA from faeces might, however, be challenging (Taberlet et al. 1996; Frantz et al. 2004) due to i) contaminants present in the samples that may inhibit the PCR; ii) inconsistencies in the data that arise from fragmented and degraded DNA (Bellemain et al. 2005), and, iii) possibility of contamination by other individuals or even distinct species' (Kohn & Wayne 1997). Moreover, several sources of genotyping errors can arise from the process, such as false alleles, allele dropout, misreading of banding patterns or transcription error from transferring the genetic data to a data file (Taberlet et al. 1996; McKelvey & Schwartz 2004). These genotyping errors usually mislead to the presence of "new" captures, because it's very unlikely that they duplicate the identical random pattern of another genotype, erroneously decreasing the recapture rate (Taberlet et al. 1999; McKelvey & Schwartz 2004). Shadow effect

may also be a source of errors (Mills et al. 2000), since multiple individuals are assigned the same genotype, as a result of using a set of loci with low variability (McKelvey & Schwartz 2004). This limitation may however be overcome by using a large number of markers that are highly variable in the genotyping process. In addition, performing several replicates when genotyping the chosen markers will allow to obtain more reliable consensus genotypes (e.g. Rollins et al. 2012; Hedges et al. 2013). Other preventive steps may be undertaken to minimize genotyping errors and increase genotyping success, through the gNIS process, from the collection and extraction procedures to the genotyping process. These steps include for instance, storing the samples in 96% ethanol and freezing them immediately including an extra step in the DNA extraction, during incubation, for removing inhibitors (Ramón-Laca et al. 2015) or even replicating the purification step of the protocol to maximize DNA yield in samples with poor quality or conditions (Costa et al. 2016).

Therefore, by optimizing the extraction and genotyping protocols it is possible to obtain better quality and quantity of DNA, reduce genotyping errors, and obtain reliable data for CR analyses. Moreover, gNIS has some major advantages in relation to live-trapping. In gNIS the "marks" are never lost, since they are a part of the genetic constitution of the individuals (McDonald et al. 2003; Lukacs & Burnham 2005). For most species, the fieldwork needed to collect non-invasive samples is cheaper and faster than the logistics involved in live-trapping, this being even more noticeable in difficult-to-access sites where the whole data set (i.e. DNA samples) for CR analysis may be collected during a single visit (Cheng et al. 2017), making it possible to survey more locations (Kilpatrick et al. 2013). Indeed, unlike live-trapping, in gNIS estimates can be obtained from a single sampling session as the recaptures happen when the same genotype appears in two different samples, and not when an animal is trapped in two separate occasions in time (Luikart et al. 2010), which makes the method very appealing and helpful for species that are difficult to sample. For this reason, in recent years, gNIS combined with CR models have been used as an alternative to traditional live-trapping for estimating the size of populations (Kohn & Knauer 1998; Lampa et al. 2013), based on the idea that faecal samples with the same *multilocus* genotypes represent recaptures (e.g. forest elephants, Eggert et al. 2003; wild ungulates, Maudet et al. 2004; cottontail rabbits, Kilpatrick et al. 2013; jaguar populations, Roques et al. 2014).

A major shortcoming of CR studies, however, is that, when population sizes are very small, it is difficult to obtain accurate estimates of population size based, particularly under low recapture rates. In this context, gNIS may help to overcome this limitation, as it potentially allows the collection of a larger number of samples. Indeed, genetic CR can produce robust abundance estimates, as well as other population parameters, also important for management and conservation (Marucco et al. 2011) . Population size estimates based on gNIS have been used for small mammal population sizes (Cheng et al. 2017; Ferreira et al. submitted; Sabino-Marques et al. submitted), but contrary to live-trapping, gNIS-based estimates were never used for validating population size indexes developed for small mammals, (e.g. Bonesi & Macdonald 2004; Gervais 2010). Since gNIS presents an array of important advantages over live-trapping, it should provide a useful technique to validate abundance indexes for many species, including small mammals.

1.2.4. Abundance indexes based on presence signs

In most studies, the sample design provides results regarding relative abundance (the roughly estimated number of individuals present during a sampling season; Bonesi & Macdonald 2004; Wheeler 2008), which is then extrapolated for the entire population as an index. There is a large variety of field methods for obtaining an index of abundance (e.g. hunting and fishing harvests, antler counts, aerial surveys, road mortality counts), depending on the species of interest and the landscape where the research is taking place. In the case of mammal species, presence sign searches (e.g. tracks, faeces, runways, burrows…) has been the most common method for generating population size indexes (PSI) (Güthlin et al. 2014).

Whatever the type of index used, indirect surveys must be based on known rates of production of the target indirect signs, and the proportion of the population that actually leaves the 'detectable' signs, so that population abundance can be calculated from the PSI used. However, In the process of indexing population size from presence signs, some errors might appear that influence the

results, either by over or under extrapolating the results (Bonesi & Macdonald 2004). Therefore, an accurate index should possess some desirable properties to be practical and applicable. In particular, PSIs should i) be simple and easy to apply; ii) have a low rate of observer bias; and above all, iii) reflect the population changes (Engeman 2005). Also, it's important that PSIs are cost effective and provide reliable data within relatively short time-periods (Whisson et al. 2005). These properties are of major importance for managers using PSIs as a basis for decision making, either in the context of species control programs, or for assessing the effectiveness of conservation measures (Whisson et al. 2005). A further important aspect when using PSIs is the need to understand the factors that determine the probability of signs being detected by surveyors, and what determines the probability of signs being deposited (Bonesi & Macdonald 2004), (Tellería 2004; Witmer 2005).

While PSIs are generally assumed to be proportional to (and therefore an index of) absolute measures of abundance (Slade & Blair 2000), these indexes may not allow strong inferences of population size. Thus, PSIs based on simple counts of signs that are known to under-represent the numbers of animals in a population (Slade & Blair 2000; Watkins et al. 2010), should first be validated against known standards to confirm that the index indeed reflects the changes in population size (e.g. (Engeman 2003; 2005; Watkins et al. 2010). However, the true population size is very unlikely to be known in wild populations, thereby hampering the validation and calibration of abundance indexes (Anderson 2001). This implies that more informative methods, such as seems to be the case of gNIS, should be increasingly used to assess the "true" population sizes that can be used to validate PSIs.

Although PSIs may be difficult to validate, and may present proportionality problems (i.e. the assumption that a given amount of effort will detect a relatively constant proportion of the population that is related to "true" population size in mostly the same way; Watkins et al. 2010), they are generally more practical than direct observation of elusive animals, making them well suited for long-term and broad-scale monitoring programs requiring knowledge of relative population density but not necessarily absolute numbers (Gervais 2010). In addition, abundance indexes often are easier and more cost-effective to obtain than estimators of absolute population abundance; are much less intrusive, minimize harm and disturbance to individuals (Whisson et al. 2005) and observer bias (Pollock et al. 2002; Whisson et al. 2005; Siddig et al. 2015).

In the case of small mammals, alternative methods to live-trapping are usually based on the presence/absence of vegetation clippings and/or droppings, which have been successfully used for estimating the relative abundance of several species (Giraudoux et al. 1995; Village & Myhill 1990; Quéré et al. 2000; Santos et al. 2006; Wheeler 2008; Jareño et al. 2014). However, there are still few examples testing explicitly how PSIs relate to the true population size. Therefore, In the present study, we intend to build presence sign-based indexes for a small mammal of conservation concern, the Cabrera vole (*Microtus cabrerae*, Thomas 1906 Rodentia, Crecetidae) ; and to evaluate their accuracy by relating them with population size estimates obtained from gNIS of voles' faeces. Therefore, we intend to infer the utility of indexes as a broad-scale monitoring tool for the species.

1.3 The Cabrera vole (*Microtus cabrerae*) as a model species

The Cabrera vole is a medium sized Arvicolinae species (Fig.1.1) endemic to the Iberoccitane region, and at present restricted to the Iberian Peninsula (Fernandez-Salvador, 2008; Fernandes et al. 2008). It is classified as "Near Threatened" by the International Union for Conservation of Nature's (IUCN) Red List (Fernandes et al. 2008) and listed in the Portuguese and Spanish Red Lists (Blanco & Gonzaléz 1992; Cabral et al. 2005) in the European Community Habitats Directive (92/43/EEC), and in the Bern Convention (82/72/CEE) due to loss and fragmentation of the species' habitat.

The species is found in agro–silvo–pastoral systems mainly composed by cork oak and holm oak (e.g. San Miguel Ayanz 1992; Santos et al. 2006), as well as in open farmlands with relatively spaced woodlots (e.g. Pita et al. 2007). Suitable habitat includes patches with dense and tall wet

herbaceous vegetation, including abundant grasses, sedges, rushes and reeds, often with scattered shrubs and trees (Fig.1.2; Pita et al. 2011). These are largely restricted to temporary ponds or field margins with irrigation ditches or other water sources (Pita et al. 2006), which are subject to seasonal fluctuations in environmental conditions, and are frequently disturbed by farming activities (Pita et al. 2007; Pita et al. 2016). In intensive Mediterranean farmland, suitable habitat patches are often very small (ca. 500 m^2; Pita et al. 2007; Santos et al. 2007), and are usually surrounded by an inhospitable matrix. This often leads to a metapopulation-like spatial structure, in which local populations or colonies are more or less isolated consisting of a few individuals, frequently organized as a monogamous breeding pair and its offspring (Pita et al. 2006, Do Rosário & Mathias 2007). Individuals generally show strong site fidelity to their patches, with home-ranges averaging around 300–400 m^2 (Pita et al. 2014). The recent intensification of human activities related to agriculture, cattle overgrazing, road construction and urbanization (e.g. Fernández-Salvador et al. 2005; Pita et al. 2007) has caused a reduction in availability of suitable habitat for the species in many localities of the species range (e.g. Landete-Castillejos et al. 2000; Pita et al. 2014). Also, the recent increasing in drought events negatively affect streams and other wetland areas where the species occurs (Fernandes et al. 2008).

The Cabrera vole is not frequently trapped in wildlife surveys (Alasaad et al. 2011) and has been considered trap-shy (Fernández-Salvador 1998), which makes the estimation of population sizes and dynamics particularly hard. For this reason, first studies aiming to assess the population ecology of the species are mostly based on raptor pellets (e.g. *Tyto alba*, Mira et al. 2008) and species-specific presence signs (Pita et al. 2007; Santos et al. 2006). Typical presence signs for the species include grass clippings, runaways and latrines or isolated dark-green faeces (Fig.1.3). The Cabrera's vole is a good model species for the purposes of this study, as its typical presence signs are relatively easy to identify in the field, and apparently it shows negative responses to trapping (Fernández-Salvador 1998). Furthermore, because the species faces at present major threats related to habitat loss, it would be desirable to assess its population status through its range, by using a simple abundance index based on the species presence signs. Thus, finding innovative techniques for monitoring Cabrera vole populations seems a timely and opportune research topic, particularly if such techniques are less expensive, and have the potential to yield better results than the traditional live-trapping methods, as it seems to be the case of gNIS. Indeed, gNIS of Cabrera vole faeces have been already successfully tested and used to infer local population size of the species in patchy habitats, based on CR models (Ferreira et al. submitted; Sabino-Marques et al. submitted), thereby providing a potentially useful approach to validate putative abundance indexes for the species. However, since gNIS has still considerable costs associated, and requires time-consuming laboratory work, it would be important to have a sampling method based on presence-sign searches for generating abundance indexes that, despite only providing relative abundances, would allow for the species long term and broad-scale monitoring.

Fig.1.1 – Female Cabrera Vole (*Microtus cabrerae*, Thomas 1906) captured near Cabo Sardão in December 2006, (photographed by R. Brito).

Fig.1.2 – Typical suitable habitat for the Cabrera's vole in Southwest Portugal. (photographed by R. Pita).

Fig.1.3 – Example of a medium latrine with typical 4-8 mm long green droppings (Pita et al. 2014). Photographed in Southwest Alentejo in 2017 by D. Peralta.

1.4 Main objectives and expected results of this study

Given the need for cheap and easy implementable sampling schemes for the Cabrera vole, presence sign-based indexes would be a useful alternative to more demanding methods for assessing local abundances. Cabrera vole presence signs have already been used to build abundance indexes (e.g. San Miguel Ayanz 1992; Santos et al. 2006), which provided relevant insights on local abundances and population dynamics. However, these indexes were never validated and their relation with to "true" population size remained unknown.
The present study aims to;

1. Validate the use of presence sign counts under several sampling efforts as indexes of Cabrera vole abundance, by correlating them with estimates obtained from CR methods based on gNIS data;
2. Understand which latrine size works best to generate presence sign indexes for the species;
3. Test the most informative sampling effort to build a PSI.

Overall, results from the study are expected to provide further insights regarding the use of presence sign-based abundance indexes of the Cabrera vole in ecological studies; and the potential utility of such indexes to inform conservation planning targeting the species.

1.5 Literature cited

Alasaad S, Soriguer RC, Jowers MJ, Marchal JA, Romero I, Sánchez A. 2011. Applicability of mitochondrial DNA for the identification of Arvicolid species from faecal samples: A case study from the threatened Cabrera's vole. Molecular Ecology Resources **11**:409–414.

Anderson DR. 2001. The need to get the basics right in wildlife field studies. Wildlife Society Bulletin **29**:1294–1297.

Bain K, Wayne A, Bencini R. 2014. Overcoming the Challenges of Measuring the Abundance of a Cryptic Macropod: Is a Qualitative Approach Good Enough? Wildlife Research **41**:84–93.

Beja-Pereira A, Oliveira R, Alves PC, Schwartz MK, Luikart G. 2009. Advancing ecological

understandings through technological transformations in noninvasive genetics. Molecular Ecology Resources **9**:1279–1301.

Bellemain E, Swenson J, Tallmon D, Brunberg S, Taberlet P. 2005. Estimating Population Size of Elusive Animals With DNA From Hunter Collected Feces: Four Methods for Brown Bears. Conservation Biology **19**:150–161.

Blanco JC, Gonzaléz JL. 1992. Fichas descriptivas de las especies y subespecies amenazadas: mamiferos. Pages 515–681in J. C. Blanco and J. L. Gonzaléz, editors.Libro Rojo de los Vertebrados de España. Instituto para la Conservación de la Naturaleza, Ministerio de Agricultura, Pesca y Alimentación, Madrid, Spain.

Bonesi L, Macdonald DW. 2004. Evaluation of Sign Surveys as a Way to Estimate the Relative Abundance of American Mink (*Mustela vison*). Journal of Zoology **262**:65–72. Available from http://doi.wiley.com/10.1017/S0952836903004448.

Burnham KP, Overion WS. 1979. Robust Estimation of Population Size When Capture Probabilities Vary Among Animals. Ecology **60**:927–936.

Cabral MJMJ et al. 2005. Livro Vermelho dos Vertebrados de Portugal. Page (Cabral MJ, Almeida J, Almeida PR, Dellinger T, Ferrand de Almeida N, Oliveira ME, Palmeirim J, Queiroz AI, Rogado L, Santos-Reis M, editors). Instituto da Conservação da Natureza, Lisboa.

Cheng E, Hodges KE, Sollmann R, Mills LS. 2017. Genetic Sampling for Estimating Density of Common Species. Ecology and Evolution:1–11.

Chitty D, Kempson DA. 1949. Prebaiting Small Mammals and a New Design of Live Trap. Ecological Society of America **30**:536–542.

Costa V, Rosenbom S, Monteiro R, O'Rourke SM, Beja-Pereira A. 2016. Improving DNA quality extracted from fecal samples—a method to improve DNA yield. European Journal of Wildlife Research **63**:3. European Journal of Wildlife Research. Available from http://link.springer.com/10.1007/s10344-016-1058-1.

Crosbie SF, Manly BFJ. 1985. Parsimonious Modelling of Capture-Mark-Recapture Studies. J. Manly Source: Biometrics BIOMETRICS **41**:385–398.

De Bondi N, White JG, Stevens M, Cooke R. 2010. A Comparison of the Effectiveness of Camera Trapping and Live Trapping for Sampling Terrestrial Small-mammal Communities. Wildlife Research **37**:456–465.

Do Rosário IT, Mathias MDL. 2007. Post-fire Recolonisation of a Montado Area by the Endangered Cabrera Vole (*Microtus cabrerae*). International Journal of Wildland Fire **16**:450–457.

Eggert LS, Eggert JA, Woodruff DS. 2003. Estimating population sizes for elusive animals: The forest elephants of Kakum National Park, Ghana. Molecular Ecology **12**:1389–1402.

Engeman RM. 2003. More on the Need to Get the Basics Right: Population Indexes. Wildlife Society Bulletin **31**:286–287.

Engeman RM. 2005. Indexing principles and a widely applicable paradigm for indexing animal populations. Wildlife Research **32**:203–210.

Fernandes M, Pita R, Mira A. 2008. *Microtus cabrerae*. Page The IUCN Red List of Threatened Species. The IUCN Red List of Threatened Species.

Fernández-Salvador R. 1998. Topillo de Cabrera, *Microtus cabrerae* THOMAS, 1906. Galemys:5–18. Available from http://www.secem.es/wp-content/uploads/2013/03/G-10-2-01-Fdez-Salvador-5-18.pdf (accessed March 3, 2017).

Fernández-Salvador R, Ventura J, García-Perea R. 2005. Breeding patterns and demography of a

population of the Cabrera vole, <i> Microtus cabrerae <\i>. Animal Biology **55**:147–161.

Ferreira CM et al. (submitted). High Cost-effectiveness of Genetic Non-invasive Sampling (gNIS) for Monitoring Elusive Small Mammals.

Flowerdew JRF, Shore RF, Poulton SMC, Ks THS. 2004. Live Trappin to Monitor Small Mammals in Britain. Mammal Review **34**:31–50.

Frantz AC, Schaul M, Pope LC, Fack F, Schley L, Muller CP, Roper T. 2004. Estimating population size by genotyping remotely plucked hiar: the European badger. Journal of Applied Ecology **41**:985–995.

Gervais J a. 2010. Testing sign indexes to monitor voles in grasslands and agriculture. Northwest Science **84**:281–288.

Gese EM. 2001. Monitoring of Terrestrial Carnivore Populations. Carnivore Conservation:372–396..

Giraudoux P, Pradier B, Delattre P, Deblay S, Salvi D, Defaut R. 1995. Estimation of water vole abundance by using surface indexes. Acta Theriologica **40**:77–96.

Güthlin D, Storch I, Küchenhoff H. 2014. Toward Reliable Estimates of Abundance: Comparing Index Methods to Assess the Abundance of a Mammalian Predator. PLoS ONE **9**.

Hedges S, Johnson A, Ahlering M, Tyson M, Eggert LS. 2013. Accuracy, Precision, and Cost-Effectiveness of Conventional Dung Density and Fecal DNA Based Survey Methods to Estimate Asian Elephant (*Elephas maximus*) Population Size and Structure. Biological Conservation **159**:101–108. Elsevier Ltd.

Jareño D, Viñuela J, Luque-Larena JJ, Arroyo L, Arroyo B, Mougeot F. 2014. A comparison of methods for estimating common vole (*Microtus arvalis*) abundance in agricultural habitats. Ecological Indicators **36**:111–119.

Kilpatrick HJ, Goodie TJ, Kovach AI. 2013. Comparison of Live-trapping and Noninvasive Genetic Sampling to Assess Patch Occupancy by New England Cottontail (Sylvilagus transitionalis) Rabbits. Wildlife Society Bulletin **37**:901–905.

Kohn MH, Knauer F. 1998. Philogeography of Brown Bears in Europe and Excremental PCR-The New Tool in the Genetic Analysis of Animals in the wild. Ursus **10**:315–321.

Kohn MH, Wayne RK. 1997. Facts from feces revisited. Trends in Ecology & Evolution **12**:223–227.

Kohn MH, York EC, Kamradt D a, Haught G, Sauvajot RM, Wayne RK. 1999. Estimating population size by genotyping faeces. The Royal Society **266**:657–663.

Lampa S, Henle K, Klenke R, Hoehn M, Gruber B. 2013. How to overcome genotyping errors in non-invasive genetic mark-recapture population size estimation - A review of available methods illustrated by a case study. Journal of Wildlife Management **77**:1490–1511.

Landete-Castillejos T, Andrés-Abellán M, Argandoña JJ, Garde J. 2000. Distribution of the Cabrera vole (*Microtus cabrerae*) in its first reported areas reassessed by live trapping. Biological Conservation **94**:127–130.

Luikart G, Ryman N, Tallmon DA, Schwartz MK, Allendorf FW. 2010. Estimation of census and effective population sizes: The increasing usefulness of DNA-based approaches. Conservation Genetics **11**:355–373.

Lukacs PM, Burnham KP. 2005. Review of Capture-recapture Methods Applicable to Noninvasive Genetic Sampling. Molecular Ecology **14**:3909–3919.

Marucco F, Boitani L, Pletscher DH, Schwartz MK. 2011. Bridging the Gaps Between Non-invasive Genetic Sampling and Population Parameter Estimation. European Journal of Wildlife

Research **57**:1–13.

Maudet C, Luikart G, Dubray D, Von Hardenberg A, Taberlet P. 2004. Low genotyping error rates in wild ungulate faeces sampled in winter. Molecular Ecology Notes **4**:772–775.

McDonald TLT, Amstrup SSC, Manly BFJB. 2003. Tag loss can bias Jolly-Seber capture-recapture estimates. Wildlife Society Bulletin **31**:814–822.

McKelvey KS, Schwartz MK. 2004. Providing Reliable and Accurate Genetic Capture-Mark-Recapture Estimates in a Cost- Effective Way. Journal of Wildlife Management **68**:453–456.

Miller CR, Joyce P, Waits LP. 2005. A new method for estimating the size of small populations from genetic mark – recapture data. Molecular Ecology **14**:1991–2005.

Mills LS, Citta JJ, Lair KP, Schwartz MK, Tallmon DA. 2000. Estimating animal abundance using noninvasive DNA sampling: Promise and pitfalls. Ecological Applications **10**:283–294.

Mira A, Marques CC, Santos SM, Rosário IT, Mathias ML. 2008. Environmental Determinants of the Distribution of the Cabrera Vole (*Microtus cabrerae*) in Portugal: Implications for Conservation. Mammalian Biology - Zeitschrift für Säugetierkunde **73**:102–110.

Pearse DE, Eckerman CM, Janzen FJ, Avise C. 2001. A genetic analogue of "mark–recapture" methods for estimating population size: an approach based on molecular parentage assessments. Molecular Ecology **10**:2711–2718.

Petit E, Valiere N. 2005. Estimating Population Size with Noninvasive Capture-Mark-Recapture Data. Conservation Biology **20**:1062–1073.

Pita R, Beja P, Mira A. 2007. Spatial population structure of the Cabrera vole in Mediterranean farmland: The relative role of patch and matrix effects. Biological Conservation **134**:383–392.

Pita R, Lambin X, Mira A, Beja P. 2016. Hierarchical Spatial Segregation of Two Mediterranean Vole Species: the Role of Patch-Network Structure and Matrix Composition. Oecologia **182**:253–263. Springer Berlin Heidelberg.

Pita R, Mira A, Beja P. 2006. Conserving the Cabrera vole, Microtus cabrerae , in intensively used Mediterranean landscapes. Agriculture, Ecosystems & Environment **115**:1–5.

Pita R, Mira A, Beja P. 2011. Assessing habitat differentiation between coexisting species: The role of spatial scale. Acta Oecologica **37**:124–132.

Pita R, Mira A, Beja P. 2014. *Microtus cabrerae* (Rodentia: Cricetidae). Mammalian Species **912**:48–70.

Pollock KH, Nichols JD, Simons TR, Farnsworth GL, Bailey LL, Sauer JR. 2002. Large scale wildlife monitoring studies: Statistical methods for design and analysis. Environmetrics **13**:105–119.

Quéré JP, Raoul F, Giraudoux P, Delattre P. 2000. An Index Method of Estimating Relative Population Densities of the Common Vole (Microtus Arvalis) at Landscape Scale. Revue d'écologie (La terre et la Vie) **55**:25–32.

Ramón-Laca A, Soriano L, Gleeson D, Godoy JA. 2015. A simple and effective method for obtaining mammal DNA from faeces. Wildlife Biology **21**:195–203.

Rollins LA, Browning LE, Holleley CE, Savage JL, Russell AF, Griffith SC. 2012. Building genetic networks using relatedness information: A novel approach for the estimation of dispersal and characterization of group structure in social animals. Molecular Ecology **21**:1727–1740.

Roques S, Furtado M, Jácomo a. T a., Silveira L, Sollmann R, Tôrres NM, Godoy J a., Palomares F. 2014. Monitoring jaguar populations *Panthera onca* with non-invasive genetics: a pilot study in Brazilian ecosystems. Oryx **48**:1–9.

Sabino-Marques H et al. (submitted). Combining noninvasive genetics with spatially explicit capture-recapture models for density estimation of a patchily distributed small mammal.

San Miguel Ayanz A. 1992. Inventario de la Poblacion Española del Topillo de Cabrera (Microtus cabrerae Thomas, 1906). Madrid.

Santos SM, Mathias MDL, Mira A, Simões MP. 2007. Vegetation structure and composition of road verge and meadow sites colonized by Cabrera vole (*Microtus cabrerae* Thomas). Polish Journal of Ecology **55**:481–493.

Santos SM, Simões MP, Mathias MDL, Mira A. 2006. Vegetation analysis in colonies of an endangered rodent, the Cabrera vole (*Microtus cabrerae*), in southern Portugal. Ecological Research **21**:197–207.

Sibbald S, Carter P, Poulton S. 2006. Proposal for a National Monitoring Scheme for Small Mammals in the United Kingdom and the Republic of Eire. Page The Mammal Society Research report No. 6.

Siddig A a., Ellison AM, Jackson S. 2015. Calibrating abundance indexes with population size estimators of red back salamanders (*Plethodon cinereus*) in a New England forest. PeerJ **3**:e952. Available from https://peerj.com/articles/952.

Slade NA, Blair SM. 2000. An Empirical Test of Using Counts of Individuals Captured as Indexes of Population Size. Journal of Mammalogy **81**:1035–1045.

Stickel LF. 1954. A Comparison of Certain Methods of Measuring Ranges of Small Mammals. Journal of Mammalogy **35**:1–15.

Taberlet P, Griffin S, Goossens B, Questiau S, Manceau V, Escaravage N, Waits LP, Bouvet J. 1996. Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic acids research **24**:3189–3194.

Taberlet P, Luikart G, Waits LP, Luikart G. 1999. Non-invasive genetic sampling and individual identification. Biological Journal of the Linnean Society **68**:41–55. Oxford University Press.

Tasker EM, Dickman CR. 2001. A Review of Elliott Trapping Methods for Small Mammals in Australia. Australian Mammalogy **23**:77–87.

Tellería JL. 2004. Métodos de Censo en Vertebrados Terrestres. Dep. Biología Animal I. Fac. Biología, Univ. Camplutense, Madrid.

Thomas O. 1906. A new Vole from Spain. Annals and Magazine of Natural History **7**:576–577.

Torre I, Guixé D, Sort F. 2010. Comparing Three Live Trapping Methods for Estimating Small Mammal Sampling in Cultivated Areas of NE Spain. Italian Journal of Mammalogy **21**:147–155.

Village A, Myhill D. 1990. Estimating Small Mammal Abundance for Predator Studies - Snap-Trapping Versus Sign Indexes. Journal of Zoology **222**:681–689.

Watkins AF, McWhirter JL, King CM. 2010. Variable Detectability in Long-term Population Surveys of Small Mammals. European Journal of Wildlife Research **56**:261–274.

Wheeler P. 2008. Effects of sheep grazing on abundance and predators of field vole (*Microtus agrestis*) in upland Britain. Agriculture, Ecosystems and Environment **123**:49–55.

Whisson DA, Engeman RM, Collins K. 2005. Developing relative abundance techniques (RATs) for monitorin rodent population. Wildlife Research **32**:239–244.

Witmer GW. 2005. Wildlife population monitoring: Some practical considerations. Wildlife Research **32**:259–263.

2. Paper: Indexing population size of elusive species from sign surveys: an evaluation using genetic non-invasive sampling of Cabrera voles

T. Vaz-Freire [a], D. Peralta [b], T. Mendes [b], C. Ferreira [c], J. Paupério [c], R. Pita [b].

[a] – Departamento de Biologia Animal, Faculdade de Ciências, Universidade de Lisboa, R. Ernesto de Vasconcelos, Ed. C2, 2º Piso, Campo Grande, 1749-016 Lisboa

[b] – InBIO/CIBIO-UE, Centro de Investigação em Biodiversidade e Recursos Genéticos, Pólo de Évora, Universidade de Évora, Núcleo da Mitra, apartado 94, 7002-554, Évora, Portugal.

[c] – InBIO/CIBIO -UP, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal.

Abstract

Knowledge of the size of the population and its status is fundamental for the implementation of appropriate conservation measures for rare, elusive and threatened species. Capture–recapture (CR) studies often provide such information, though these are largely impractical and cost-prohibitive to conduct at large spatial scales. Therefore, rapid survey techniques based on presence signs are attractive options to provide an index of population size. The aim of our study was to assess the utility of latrine counts as indexes of population size for the Cabrera vole, analyzing how such indexes relate to the "true" population sizes. Particularly, we intended to identify the type of latrine and the ideal sampling efforts that yield more informative data to monitor the species. For this, we collected fresh faeces from latrines in 8 habitat patches in SW Portugal, for microsatellite genotyping. Rarefaction curves and CR models were used to obtain "true" population size from the gNIS data. These estimates were then related via GLMM with 9 populations size indexes (PSI), which were built from counts of latrines of different sizes under varying sampling efforts. Although genotyping success was relatively low (ca. 25%), genotyping errors were quite low (very close to 0), thus suggesting low potential for bias in the results**.** We found that, at least one PSI seems to provide a promising tool to assess local population sizes of Cabrera voles. Specifically, counts of small latrines (<20 faeces) under a minimum sampling effort of 60min/0.1ha, provided a reasonable PSI for Cabrera voles (R^2= 0,81). Although further research is needed to confirm these results, we suggest that Cabrera vole's population sizes may be effectively inferred from expedite and inexpensive methods based on presence sign counts, which may be used to monitor and guide conservation planning for the species over large spatial and temporal scales.

Keywords: population size; abundance indeces; gNIS; faeces; *Microtus cabrerae*

2.1. Introduction

Understanding the factors that affect the spatial and temporal variation of species population size provides baseline guidance for adequate conservation management, either targeting the protection of endangered species, or the control of pest outbreaks (Tellería 2004; Witmer 2005). Because counting every member in a population is mostly impractical or impossible, wildlife managers usually employ sampling and modelling techniques to make inferences on population size and abundance (Village & Myhill 1990; Engeman 2005). Despite the great efforts and developments in such techniques over time, knowledge of the size and abundance of wild populations still remains a main challenge in conservation ecology studies, particularly for elusive species of conservation concern (Miller et al. 2005).

Indirect methods based on counts of species presence signs (e.g. nests, faeces, tracks, food remains) provide the simplest method for assessing variations in animal population size and abundance (Ferron et al. 2008). Presence sign surveys are often more efficient than direct observation of elusive animals, making them well suited for long-term and broad-scale

monitoring programs requiring knowledge of relative population density but not necessarily absolute numbers (Gervais 2010; Watkins et al. 2010; Bain et al. 2014). A major assumption of indirect methods for indexing population size is that the relation between indexes and the true population size is linear, monotonic, and stable (Engeman 2005; Gervais 2010). However, this relation is often complex, and despite the widespread use of presence signs surveys for inferring animal abundance, this assumption is seldom tested, which may lead to biased estimates (McKelvey & Pearson 2001; Tellería 2004; Witmer 2005; Güthlin et al. 2014). Because presence signs left by animals may vary with individual activity, and environmental conditions may affect their detection (Piggott 2004; Gervais 2010), wildlife managers and conservation practitioners should benefit from a better understanding of the link between indirect population indexes and the actual population size (Tellería 2004; Witmer 2005; Jareño et al. 2014). Before being used in wildlife management decisions, abundance or population size indexes need therefore to be validated against known standards, in order to find out whether variation in those indexes really match variation in population size (Anderson 2001; Engeman 2003, 2005; Gervais 2010; Garel et al. 2010). However, knowledge on the true size of a population is cost-prohibitive for many species, making the validation of the abundance indexes problematic, particularly for species that are difficult to sample with traditional techniques involving direct visualisation and individual identification of the animals (Watkins et al. 2010).

The most popular method for approaching the true size of animal populations involves capture-recapture (CR) sampling techniques, which require tagging or marking a number of individuals in a population, and subsequently recapture some of them (e.g. Crosbie & Manly 1985; Village & Myhill 1990). Using statistical analyses such as rarefaction curves (e.g. Kohn et al. 1999; Eggert et al. 2003) or CR models (e.g. Mills et al. 2000; Pearse et al. 2001), it is possible to obtain credible estimates of population size. In the first case, the estimated population size corresponds to the projected asymptote of a function (rarefaction curve) describing the total number of captures versus the number of individuals captured (Kohn et al. 1999; Eggert et al. 2003). In the second case, for the simplest CR model with two sampling occasions, the number of individuals captured in both occasions (recaptures) and the number captured in just one sample, are used to estimate the number of individuals that weren't captured in either sample, thus providing an estimate of the total population size. During the past decades, there has been an extensive effort in the development of more complex CR models (e.g. spatially explicit open CR models, e.g. Royle & Young 2008; Borchers & Fewster 2016). This has greatly contributed to increase the accuracy and precision of population size estimates for many species and circumstances (e.g. Asian elephants, Hedges et al. 2013; felids, Rodgers & Janečka 2013), allowing reliable hypothesis testing on the factors affecting variations in animal population size, as well as in other population parameters such as survival, migration and fecundity (Burnham & Overion 1979; Petit & Valiere 2005). Despite their utility, traditional CR-based techniques are particularly costly to implement over large areas and over long sampling periods, making their use cost-prohibitive either for population monitoring or even to validate simpler abundance indexes (Tellería 2004; Witmer 2005). Moreover, accurate parameter estimation based on CR modelling is often difficult to achieve, as it requires large sample sizes which are particularly difficult to obtain for many elusive species (Anderson 2001; Watkins et al. 2010; Bain et al. 2014). Recently, this shortcoming has been greatly reduced with the new advances of DNA extraction from presence signs (genetic non-invasive sampling, gNIS), which allows the capture of individual's DNA instead of the individual itself (Taberlet et al. 1997; Miller et al. 2005). Indeed, gNIS represents an excellent alternative for obtaining data on elusive or rare species that can be used in classical CR studies, as it allows identifying individuals from indirect presence signs (faeces, urine, shed hair, shed skin, etc; without the risk of harming the animals (Taberlet et al. 1999; Pearse et al. 2001; Bonesi & Macdonald 2004; Beja-Pereira et al. 2009). The benefits of DNA-based CR relative to classic trap-based studies include the sampling being performed by replacement, since an individual is not physically confined at any time and may leave multiple presence signs at multiple locations during a sampling session (Taberlet et al. 1999; Miller et al. 2005). Moreover, this method allows for increasing the number of observations *per* sampling season, while shortening the sampling period to better approximate the closure assumption, and reducing animal stress, mortality, and

capture bias caused by individual heterogeneity in trap response (Miller et al. 2005). However, despite these advantages, few studies have used CR techniques based on gNIS as a tool to validate abundance indexes based on presence sign counts of elusive species (e.g. Miller et al. 2005; Marucco et al. 2011; Rodgers & Janečka 2013).

This study addresses these issues using the Cabrera vole (*Microtus cabrerae*) as model system. The Cabrera vole is a medium sized arvicolinae, endemic to the Iberian Peninsula, where its distribution has decreased considerably in the past few years. It is considered Near-Threatened by the International Union for Conservation of Nature (IUCN, Fernandes et al. 2008), implying that specific and adequate conservation measures are required for the species. The Cabrera vole lives hidden among wet herbs and bushes, and is infrequently trapped in wildlife surveys (Alasaad et al. 2011), being considered as a trap-shy species (Fernández-Salvador 1998). Many of the studies involving Cabrera vole population sampling have been largely based on raptor (e.g., barn owl, *Tyto alba*) pellet analysis (e.g., (Mira et al. 2008), and mostly on surveys of species-specific presence signs (e.g. Santos et al. 2006; Pita et al. 2007). Typical presence signs of the species include surface runways, grass clippings, and the dark-green faeces about 4–8 mm long, which are frequently found in latrines (Pita et al. 2014). While these presence signs have been used to build abundance indexes (e.g. San Miguel Ayanz 1992; Santos et al. 2006), and may be useful to provide an idea of local abundance and population dynamics, no study has ever evaluated how such indexes correlate with the 'true' population size.

Given the need for expedite, low-cost, and easily implementable sampling schemes to monitor Cabrera vole populations over large spatial and temporal scales, it seems therefore crucial that presence sign-based indexes describing Cabrera voles' local population sizes are validated against robust estimates provided for instance from CR techniques (Garel et al. 2010). In this context, the general goal of this study is to validate the use of presence sign counts of the Cabrera vole under varying sampling efforts, by correlating them with estimates obtained from CR techniques. Because trapping Cabrera voles is logistically difficult, 'true' population sizes will be estimated based on gNIS, which has been recently optimized for this species and already proved to be suited for individual identification and CR modelling (Ferreira et al. submitted; Sabino-Marques et al. submitted). Specifically, the study addresses the following questions:

1. Are presence-signs counts suitable for indexing local population size of Cabrera voles?
2. Which latrine size provides the best index for inferring population sizes?
3. What is the ideal sampling effort in presence sign counts for obtaining a good index?

2.2. Material and Methods

2.2.1. Study area and species

The study was carried out on the coastal plateau of southwestern Portugal in an area roughly within the Natural Park of Southwest Alentejo and Vicentina Coast with an approximate area of 65700 ha (See Fig. 2.1). Climate in this region is Mediterranean with oceanic influence. Mean monthly temperatures range between 6 and 29 ◦C, and average annual rainfall is around 650 mm, of which >80% falls in October to March (AEmet-IM 2011). The landscape is manly agricultural with almost half the land dominated by irrigated annual crops and greenhouses. The production of beef cattle is also important, resulting in large areas occupied by pastures, while wood cover is restricted to arboreal windbreaks and a few woodlots (Pita et al. 2009).

Cabrera voles, in the study area, occupy patches of suitable habitat, surrounded by a matrix of unsuitable agricultural habitats for the species (Pita et al. 2007). Within occupied patches, voles are typically grouped in subpopulations or colonies consisting of a few individuals, often organized as a monogamous breeding pair and its offspring (Pita et al. 2006, Rosário & Mathias 2007). Individuals generally show strong site fidelity, with home-ranges averaging around 300–400 m^2 (Pita et al. 2014). Home-ranges are typically scent-marked by deposition of faeces in latrines of variable sizes (up to several dozens of faeces), which are thought to be related to individual communication for territory defence and mate advertisement (Gomes et al. 2013).

2.2.2. Study design and surveys

Within the study area, 8 habitat patches (see Fig. 2.1), occupied by Cabrera voles and ranging between 600 and 3200 m^2, were surveyed for fresh faeces of the species. Among the presence signs typically left by the species, fresh faeces are considered the most reliable for assessing its presence, as other signs such as runways on grasses, borrows, and grass clippings may be equivocal when other similar species are also present (e.g. *Arvicola sapidus*) (e.g. Garrido-García et al. 2013; Pita et al. 2016). In our study area, no species produces similar faeces to those of Cabrera voles, as faeces made by *Arvicola sapidus* are much larger . Surveys were performed between December 2016 and March 2017, thus including different climatic conditions to which habitats, voles and their faeces were subjected. Survey consisted in systematically searching for fresh faeces at a constant effort of 60 min per 0.1 ha, as recommended in other studies (e.g. Sabino-Marques et al. submitted). Because surveys were conducted by 3 observers, the sampling effort corresponded to 20 min per 0.1 ha per person, which was measured using handling chronometers. All Cabrera vole faeces identified within each sampling period were recorded and latrines were classified as either large (LL, ≥ 20 droppings) or small (SL, <20 droppings) (see Appendix 2). At each record, the time in seconds was also registered using handling chronometers. This procedure was repeated for each habitat in four consecutive days, and the data were used to generate abundance indexes based on presence sign counts under different sampling efforts (see data analyses).

During surveys, we also conducted gNIS of vole faeces, which consisted in collecting a sample of up to 12 droppings (mean ± SD = 5,51 ± 1,87) from each fresh latrine that contained at least 3 droppings, and that was distanced by at least 2m from the nearest collected sample. These samples were used for DNA extraction and individual genotyping, and the data was used to obtain estimates of population size within a CR framework (see data analyses). Samples were collected using sterilized tweezers and were stored in individual 2 mL micro tubes containing 96% alcohol, at -20 ◦C until extraction. At each sample collection, chronometers used for counting fresh faeces were paused and then switched again when the searches restarted. In order to collect a representative sample size at each survey, we often extended feacal sampling times beyond the searching period predefined for fresh faeces counts, which resulted in a mean sampling effort of 73 min/0.1ha (SD= 18) for fecal sampling.

2.2.3 DNA Extraction and Genotyping

From the overall feacal samples collected, ca. 85% were selected based on their freshness for DNA extraction (see 'Results'). For this, we used the E.Z.N.A.® Tissue DNA Kit (OMEGA bio-tek) following the manufactures instructions, with an initial digestion step using a lysis washing buffer (Maudet et al. 2004) for 15 minutes at 56ºC.

The faecal samples were genotyped for a set of nine microsatellites following the protocol described by Ferreira et al. (submitted). This protocol includes an initial screening of the DNA quality using three species-specific microsatellites (Mix 1, see Table A1.3, Appendix 1). The samples that amplified for the three loci, were then amplified for an additional 6 microsatellite loci. For these samples, the species ID was confirmed using a small fragment of the Dloop gene (Alasaad et al. 2011). The samples were also sexed using two small-sized sex chromosomes introns (DBX5-S and DBY7-S, Ferreira et al. submitted). To account for genotyping errors (e.g. allele dropout and false alleles) and obtain a consensus genotype, each multiplex reaction was replicated four times (three times for the sex chromosome introns amplification). PCR reactions were performed in a final volume of 10 µL, consisting of 4 µL of Qiagen© Multiplex PCR Kit Master Mix, 1µL of DNA, and primer concentrations and thermal profiles according to Ferreira et al. (submitted). All products were sequenced on a ABI3130 Capillary Sequencer (Applied Biosystems).

The extractions and PCR reactions of the non-invasive samples were performed in physically isolated rooms, and all the equipment used was sterilized with bleach and ethanol, and exposed to UV light before and after usage. Aerosol-resistant pipette tips were used, and negative controls

were included in each manipulation, maintaining conditions to monitor and reduce risk of DNA contamination (Beja-Pereira et al. 2009; Barbosa et al. 2013; Costa et al. 2017).

Allele calling of the microsatellite loci and sex chromosome introns was performed using GeneMapper (v.4.0; Applied Biosystems) while Dloop sequences were analysed with Geneious (v.8.0; Kearse et al. 2012). Consensus genotypes for each sample were obtained by analysing all replicate genotypes with the software Gimlet (v.1.3.3, Valière 2002). For genotypes differing by up to two loci or with up to two missing data, additional PCR replicas were performed, to try to complete the genotypes, and check for the presence of genotyping errors. Sample consensus genotypes were then compared with each other to identify individuals. The criteria used to assign samples to individuals was very strict since only individuals that differed in more than two alleles were assigned as new captures. Genotyping error rates were estimated by the software Pedant (Johnson & Haydon 2007), using 10,000 search steps. Since the software only allows the comparison of two replicates, we randomly selected four of them, with a random function in Microsoft Excel 2016, and averaged all possible pairwise comparisons.

2.2.4. Data Analysis

Population size indexes

9 population size indexes (PSI) were generated by counting latrines, considering i) each latrine size category alone (LL versus SL) and together (TL), and ii) three sampling efforts: the baseline effort of 60min/0.1ha (E3), and the efforts of 30min/0.1ha (E2) and 15min/01ha (E1). These indexes will be hereafter referred to as E1-LL, E1-SL, E1-TL, E2-LL, E2-SL, E2-TL, E3-LL, E3-SL, and E3-TL, according to the effort and type of presence sign used (see Table 2.1). Each index was estimated for each habitat and sampling day, thus providing a total of 32 measurements per PSI.

Table 1.1 - Abundance indexes estimated for the Cabrera vole based on fresh faeces counts under different sampling efforts and considering different latrine sizes.

Abundance index (code)	Sampling effort	Latrine type
E1-LL	15min/0.1ha	Large latrines
E1-SL	15min/0.1ha	Small latrines
E1-TL	15min/0.1ha	Overall latrines
E2-LL	30min/0.1ha	Large latrines
E2-SL	30min/0.1ha	Small latrines
E2-TL	30min/0.1ha	Overall latrines
E3-LL	60min/0.1ha	Large latrines
E3-SL	60min/0.1ha	Small latrines
E3-TL	60min/0.1ha	Overall latrines

Population size estimates

For estimating population size, we analysed the gNIS pooled data of the 4 sampling days for each habitat, using the two most common accumulation curve techniques for CR data. The first corresponded to the hyperbolic curve proposed by Kohn et al. (1999), which is given as:

$$E(x) = \frac{ax}{(b+x)}$$

where *x* represents the number of genotyped samples, *E(x)* is the cumulative number of unique genotypes found in *x* genotyped samples, *a* is the asymptote of the function, and therefore the estimated population size, and *b* is the non-linear slope of the function.

The second curve was based on the exponential function of Eggert et al. (2003), given as:

$$E(x) = a(1 - e^{(bx)})$$

Both accumulation techniques assume that populations were closed (i.e. no immigration, emigration, birth or death over the sampling period), that the area over which the population size was to be estimated had been sampled extensively, and that all individuals had an equal probability of capture (Kohn et al. 1999; Eggert et al. 2003). In Microsoft Excel 2016, we used the 'Solver' add-in for optimization problems to run an iterative least mean squares regression fit

of these equations to the data from each of the 8 habitat patches, in order to estimate the respective values of *a* and *b*. As restriction rules for minimizing residual sum of squares, the number of genotypes for the first sample was set to be 1, while the value of *a* was assumed to be equal or greater than the minimum number of genotypes identified in each patch. Since the order in which the samples are added may influence the shape of the curve (Colwell & Coddington 1994; Eggert et al. 2003), each dataset was randomized 100 times for each of the 8 habitat patches, and the value of *a* and *b* was estimated for each replicate. The estimate of *a* for the dataset was the average of all replicates, and the precision was assessed by calculating the respective SD (Colwell & Coddington 1994).

In addition, when possible, we also attempted to build simple urn CR models based on sampling with replacement (Miller et al. 2005), in order to corroborate rarefaction estimates (e.g. Perez et al. 2006). These models were implemented in version 3.3.2 of the R program (R Development Core Team 2016), using the package 'Capwire' (Pennell et al. 2013), which uses maximum likelihood estimation. For simplicity and comparative purposes, we used the Equal Capture Model (ECM, Miller et al. 2005), as it relies on the same assumptions as rarefaction curves. Only the estimator that yielded the most accurate results was to be applied in further analysis.

Relating PSI with estimates of population size

The last step of the analysis consisted in assessing the relationship between each PSI and the CR-based estimator that yielded more precise population size estimates for all habitat patches surveyed. This consisted in building Generalized Linear Mixed-Effects Models (GLMM) with Poisson distribution for count data, testing whether abundance indexes estimated based on fresh faeces counts at each patch and day (n=32) could be explained by the actual local population size as assessed from the gNIS data, while including as random effects the patch and the month of the surveys. Using the 'lme4' package (Bates et al. 2015), a model was built for each PSI, and then compared with the respective null model (including only the random effects) using the Aikaike Information Criterium (AIC) corrected for small sample sizes (AICc; Burnham & Anderson 2002). Models with AICc values of >2 units lower than the corresponding null model were assumed to provide support for the relationship between PSI and actual population sizes. Models' support was also assessed by estimating the respective AICc-weighs, which provide the probability of the model being correct (Burnham & Anderson 2002). The variance explained by supported models was assessed by estimating the adjusted pseudo-R^2 for GLMM (Nakagawa & Schielzeth 2013), using the MuMIn package (Kamil 2016).

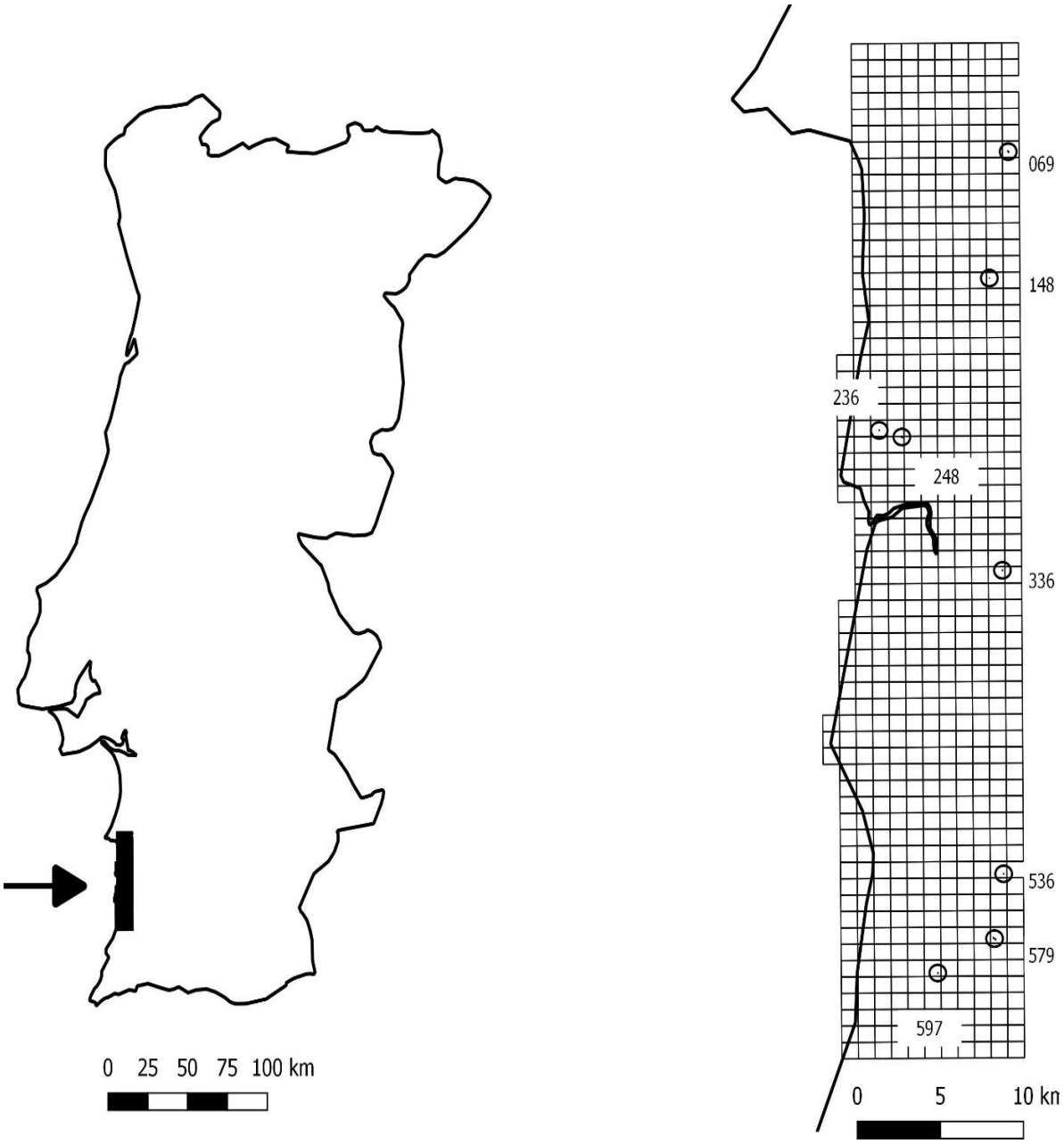

Figure 2.1- Location of the study area in SW Portugal (on the left) and location of the 8 habitat patches sampled, (on the right)

2.3. Results

In the 8 surveyed habitat patches, we recorded a total of 544 fresh latrines (mean ± SE = 68±20,69) (see Table A1.1 in Appendix 1), of which about 70% were classified as small.

A total of 601 faecal samples were collected during gNIS in the 8 patches, with a mean ± SE of 75,13 ± 14,23 (range: 59-100) samples per patch (Table 2.2). Between 48 and 77 samples were extracted in each patch (mean ± SE = 63,625 ± 9,43), which corresponds to around 85% (n=509) of the total samples collected (Table 2.2). Of the extracted samples, 129 were successfully genotyped for 8 or more loci, resulting in an overall genotyping success of approximately 25%.

Table 2.2 - Details on the surveys of the 8 habitat patches, including patch area, total number of available samples, and gNIS results: extracted and genotyped samples, number of individuals identified and recaptured.

Habitat	Area(m^2)	Total samples	Extracted	Genotyped	Success (%)	Individuals	Recaptures
069	2696	59	48	11	22,92	5	6
148	1313	64	62	19	30,65	8	11
236	627	83	68	13	19,12	4	9
248	2968	81	59	17	28,81	2	15
336	1735	80	70	12	17,14	5	7
536	597	58	55	16	29,09	3	13
579	3184	100	77	20	25,97	12	8
597	984	76	70	21	30,00	7	14
Total	**-**	**601**	**509**	**129**	**-**	**46**	**83**

The amplification of the Dloop fragment in the 129 genotyped samples, confirmed the species identification as Cabrera vole for all of them. Moreover, the total allelic dropout rate was 0,01 (range = [0,000339-0,036599]) and the false allele rate was close to 0. The error rates, per locus, are presented in supplementary material (see Table A1.2 in Appendix 1).

The gNIS allowed the identification of a total of 46 individuals (mean ± SE [range] = 5,75 ± 3,2 [2-12] individuals per patch), and 83 recaptures (mean ± SE [range] = 11 ± 3,34 [6-14]).

The results from CR-based population size estimates (Table 2.3) indicated that, in general, the number of animals captured within each habitat patch was close to the estimates produced by accumulation curves, suggesting that our gNIS protocol allowed the identification of a great proportion of the individuals occupying the surveyed patches. However, the Kohn's estimator provided higher (up to 25%) and less precise estimates then those obtained by the Eggert's curve (Table 2.3). Capwire results were only available for 3 of the 8 habitat patches as the model was unable to produce a stable result for habitats with fewer than 7 individuals. Therefore, we considered that the Eggert estimator provided the best population size estimates, and used these estimates as reference of the "true" population sizes in GLMM analyses.

Table 2.2- Summary results of the population size estimates using gNIS for the 8 habitat patches. The results of the rarefaction curves using two different estimators (Kohn and Eggert) and of the Capwire package, are shown.

Habitat	Population size estimates (mean ± SE)		
	Rarefaction curves		C-R models
	Kohn estimator	Eggert estimator	Capwire estimator
069	9,00 ± 2,51	6,00 ± 1,07	-
148	19,00 ± 12,47	12,00 ± 5,82	8.00 ± 0.00
236	6,00 ± 1,15	5,00 ± 0,28	-
248	2,00 ± 0,10	2,00 ± 0,00	-
336	7,00 ± 2,45	6,00 ± 0,91	-
536	3,00 ± 0,56	3,00 ± 0,07	-
579	26,00 ± 13,80	16,00 ± 6,06	17.00 ± 0.00
597	11,00 ± 4,50	8,00 ±1,63	7.00 ± 0.73

We found that only the model testing the relationship between the E3-SL index and the Eggert's population size estimates, provided more support than the respective null model, with an AICc-weigh of 0.86 (see Table 2.4). It is interesting to note that, in general, there was a tendency for decreasing AICc weights as the sampling effort decreases, and when large latrines are considered. The model testing the

E3-SL index indicated a positive relationship between this variable and the Eggert's estimator (see Figure 2.2), with a coefficient ± SE of 0,04 ± 0,01 ($p < 0,001$), and an adjusted pseudo-R^2 of 0,81.

Table 2.3- Results for the GLMM models regarding the three sampling efforts and the three latrine size categories. Model E3_SL (in bold) is the only that presented a significant result. Models are considered significant if the results from the AICc null are 2+ units higher than the AICc value and are considered robust if the AICc's weight is >0.50.

	Model evaluation			
Code	**AICc**	**AICc_null**	**deltaAICc**	**AICc-weight**
E3_LL	183.7648	181.2969	-2.4679	0.23
E3_SL	**201.199**	**238.5523**	**37.3533**	**0.86**
E3_TL	269.0834	267.6816	-1.4018	0.33
E2_LL	149.5693	147.1398	-2.4295	0.23
E2_SL	201.199	199.6711	-1.5279	0.32
E2_TL	221.0663	219.186	-1.8803	0.28
E1_LL	194.1165	122.9334	-71.1831	0.21
E1_SL	176.8198	174.7479	-2.0719	0.26
E1_TL	194.1165	192.0737	-2.0428	0.26

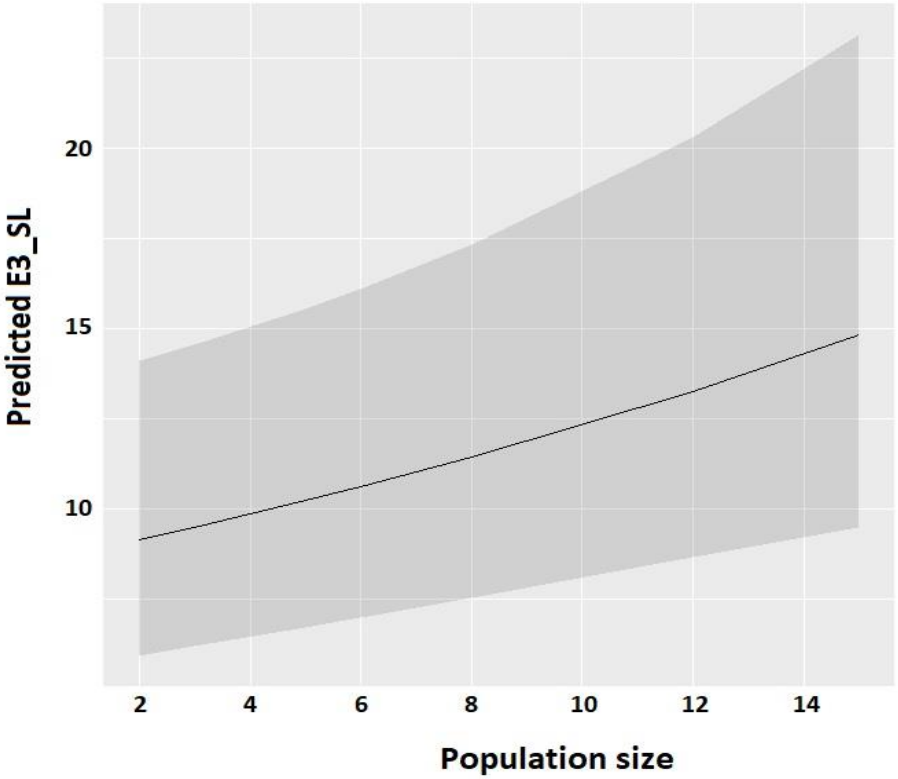

Figure 2.2 *– Variation in the predicted number of small fresh latrines recorded under the sampling effort of 60min/0.1ha (E3-SL) relative to the population size, as estimated from the Eggert's rarefaction curve using gNIS data. The graphical representation was made using the packages 'sjPlot' and 'sjmisc'(Lüdecke 2017).*

2.4. Discussion

When proposing this work, we asked three main questions related to the study of population size indexes for Cabrera voles. Our results provided answers to those questions as we were able to generate one population size index based on the counts of fresh small latrines under a minimum sampling effort of 60 min/ 0.1 ha that calibrated well for the habitat patches and conditions to test.

Knowledge about the variations in animals' abundance is at the core of population ecology and conservation biology (Witmer 2005; Marucco et al. 2011). However, despite the importance of accurate estimates of population size, ecologists and conservation practitioners often must rely on only crude indexes (usually based on raw counts of presence signs) which are rarely calibrated with the 'true' values (Engeman 2003; Witmer 2005; Siddig et al. 2015). This is particularly notorious for small and elusive species of conservation concern, such as the Cabrera vole, for which detailed studies based on invasive sampling (e.g. live-trapping) typically involve costly logistics, and often yield insufficient data for accurate estimation of population size (Tellería 2004; Watkins et al. 2010; Kilpatrick et al. 2013). Here, we investigated for the first time whether presence-sign counts at sites occupied by Cabrera voles may provide a useful surrogate of local population size estimates based on gNIS of their faeces. Our results showed that from the 9 abundance indexes estimated based on counts of fresh latrines under varying sampling efforts, at least one seems to calibrate reasonably well, potentially allowing for credible inferences regarding Cabrera voles' population size. Although further data are needed to confirm these results, this study provides evidence that relatively rapid surveys of Cabrera voles latrines may provide a promise, expedite method to inform about population size variation of the species, which may be of crucial utility for cost-effective assessment of population status (Bonesi & Macdonald 2004; Slade et al. 2012; Pita et al. 2014; Sabino-Marques et al. submitted). Such information may hence be used for species conservation planning over large spatial and temporal scales, without the need of cost-prohibitive logistics and man power.

Despite the potential of abundance indexes based on presence sign counts for inferring Cabrera vole local population size, our results further suggest that they may also be a function of other variables, associated for instance to sampling effort and type of presence signs being surveyed. Indeed, a major outcome of this study is that both the sampling effort and the type of presence sign considered for generating abundance indexes critically affected the strength of inferences regarding Cabrera vole population sizes. In particular, our results indicated that reduced sampling efforts (<60min/0.1ha searches) should provide less accurate abundance indexes, suggesting that, even considering the expeditious nature inherent to presence sign count methods relative to traditional live-trapping techniques, insufficient or inappropriate sampling effort may fail to provide accurate indexes of Cabrera voles' abundance. This is in accordance to the general prediction that the accuracy and effectiveness of abundance indexes should increase with increasing sampling effort (Engeman 2003). We thus stress that careful planning of the appropriate sampling effort is needed to generate an acceptable abundance index based on presence-signs. This would require assessing other potential sources of variation affecting latrine detectability such as local habitat characteristics, observer experience, or seasonal variations in defecation rates and DNA preservation (Anderson 2001; Barnes 2001; Murray et al. 2002; Marucco et al. 2011; Kilpatrick et al. 2013; Bain et al. 2014). Although seasonal effects were included in the random part of the GLMMs relating abundance indexes and population size estimates, further studies including more spatial and temporal replicates should explicitly account for such differences (Piggott 2004; Marucco et al. 2011).

On the other hand, it seems that not all types of presence signs may provide the level of precision required to obtain unbiased estimates. Our results suggest that counts of small fresh latrines provided the most suitable presence sign for generating abundance indexes, while large latrines seem to yield relatively poor surrogates of population size. This result can be related to the fact that latrines should play different roles, with small latrines being presumably more involved in individual familiarization with the environment and territoriality, and large latrines more related to inter-individual transmission of social signs related with reproductive advertisement and mating, as observed in other vole species (Rozenfeld et al. 1987; Woodroffe et al. 1990). Therefore, we suggest that, while large accumulations of faeces from repeated defecation by multiple voles could possibly be indicative of mating behavior intensity, they should be inadequate to infer local population abundance of Cabrera voles, which otherwise should be better predicted based on small accumulations of faeces made by animals along their routine activities.

While our results seem to provide evidence for the utility of latrine counts in indexing local population size of Cabrera voles, it is noteworthy that our reference population sizes were not free from bias, as these were derived from asymptotic estimators based on genetic non-invasive sampling. Asymptotic approaches provide the simplest population size estimators from CR data, relying on overly naive

assumptions (e.g. random spatial distribution, homogeneous capture probabilities in space and time) that should be only barely accomplished under natural systems (Waits & Leberg 2000; Miller et al. 2005; Luikart et al. 2010). However, since the total number of individuals identified in each patch were very close to the estimates produced by asymptotic estimators and also, at least in part, with those generated by *capwire* models, we assumed that our sampling procedures and rarefaction-based methods provided a reliable basis to assess local population size of the studied populations. In particular the Eggert's curve (the one selected in GLMM analyses), seems to perform much better than the Kohn's estimator, which tends to overestimate population size (Eggert et al. 2003), particularly when sample size is low (Frantz & Roper 2006), such as in the present study.

A further potential limitation is that, as in other studies using gNIS (e.g. Ernest et al. 2000; Reddy et al. 2011; Karmacharya et al. 2011), a significant number of samples from the total number that was extracted didn't produce results for one or more microsatellites, or was contaminated, which resulted in relatively low genotyping success, with potential impacts on population size estimates (Waits & Leberg 2000). This general drawback of low DNA quality in gNIS may be particularly notorious when dealing with the small size of feacal material of small mammals, making it more difficult to identify genotypes (Lampa et al. 2013; Barbosa et al. 2013). Genotyping rates achieved here (ca. 25%) were indeed much reduced compared to those obtained from feacal samples of larger mammals (e.g. *Lepus americanus*: 54-69%, Schwartz et al. 2007; North African ungulates: 80%, Silva et al. 2015). However, they were quite comparable with those obtained for aquatic mammals (*Lutra lutra*: 41-46%, Prigioni et al. 2006, Lampa et al. 2013; *Lontra canadensis*: 24%, Mowry et al. 2011), which can be related to the fact that Cabrera voles occupy mostly wet habitats, where DNA degradation is faster, due to moist local environmental conditions (Ferreira et al. submitted). Apart from that, we obtained relatively low genotyping error rates, both the allelic dropout and false alleles rates close to 0, suggesting that the use of the careful stepwise protocol followed here, was effective in providing accurate results (Ferreira et al. submitted). In particular, the multiplex amplification process likely decreased genotyping errors and increased amplification success for samples with low quality and quantity of template DNA (Lampa et al. 2013, Ferreira et al. submitted). However, when using the multiplex process, allelic dropout can occur more frequently, tan with conventional PCR, thus increasing allelic dropout rates. The results from genotyping errors might have also been influenced by the problems in amplification of the markers Ma25 and Mc02, that had not been a problem in previous studies (Ferreira et al. submitted; Sabino-Marques et al. submitted), while here, they ended up representing the highest dropout rates from all the markers (around 0,03). Much efforts have been made to reduce genotyping errors in gNIS, starting from field data collection procedures, to laboratory procedures and data analysis (see e.g. Lampa et al. 2013, for a review). Despite the efforts to prevent and account for genotyping errors, we stress that if genotyping errors become a source of severe bias, further preventive measures should be adopted. For instance, eliminating flawed markers (Waits & Leberg 2000; Cheng et al. 2017) may be of help to reach a solution where genotyping errors do not affect the final result (Lukacs & Burnham 2005), as well as choosing a model that can account for genotyping errors. It is worth mentioning that some preventive steps such as the removal of all low quality samples can however reflect negatively on the probability of capture, because better quality samples tend to be left by individuals that shed more cells than others, which increases the heterogeneity of capture between individuals (Lukacs & Burnham 2005), that not being advisable when using models assuming equal capture probabilities, as in most genetic CR studies (e.g. Miller et al. 2005).

Finally, it should be mentioned that our results reinforce the idea that local population sizes of Cabrera voles are in general very low, potentially comprising only a few breeding pairs and their offspring (Pita et al. 2007). This highlights the need for conservation measures targeting the species and its habitats, particularly in intensively used Mediterranean farmland, where local habitats are often destroyed for agricultural activities (Pita et al. 2014). In this context, the use of small-latrine counts for indexing the size of local population of the Cabrera vole, should provide a very useful tool to monitor its populations over large spatial scales and along long-time periods. Although we should reinforce the idea that more research is needed to on this, our study provides evidence that new cost-effective tools based on non-invasive sampling may be developed and possibly used in decision-making involving the conservation of Cabrera vole populations.

2.5. Literature cited

AEmet-IM. 2011. Iberian Climate Atlas.

Alasaad S, Soriguer RC, Jowers MJ, Marchal JA, Romero I, Sánchez A. 2011. Applicability of mitochondrial DNA for the identification of Arvicolid species from faecal samples: A case study from the threatened Cabrera's vole. Molecular Ecology Resources **11**:409–414.

Anderson DR. 2001. The need to get the basics right in wildlife field studies. Wildlife Society Bulletin **29**:1294–1297.

Bain K, Wayne A, Bencini R. 2014. Overcoming the Challenges of Measuring the Abundance of a Cryptic Macropod: Is a Qualitative Approach Good Enough? Wildlife Research **41**:84–93.

Barbosa S, Pauperio J, Searle JB, Alves PC. 2013. Genetic identification of Iberian rodent species using both mitochondrial and nuclear loci: Application to noninvasive sampling. Molecular Ecology Resources **13**:43–56.

Barnes RFW. 2001. How reliable are dung counts for estimating elephant numbers? African Journal of Ecology **39**:1–9.

Bates DM, Mächler M, Bolker B, Walker S. 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software **67**:1–48.

Beja-Pereira A, Oliveira R, Alves PC, Schwartz MK, Luikart G. 2009. Advancing ecological understandings through technological transformations in noninvasive genetics. Molecular Ecology Resources **9**:1279–1301.

Bonesi L, Macdonald D. 2004. Evaluation of sign surveys as a way to estimate the relative abundance of American mink (*Mustela vison*). Journal of Zoology **262**:65–72.

Borchers D, Fewster R. 2016. Spatial Capture–Recapture Models. Statistical Science **31**:219–232. Available from http://projecteuclid.org/euclid.ss/1464105039.

Burnham KP, Anderson DR. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach (2nd Edition). Page Ecological Modelling. Available from http://linkinghub.elsevier.com/retrieve/pii/S0304380003004526.

Burnham KP, Overion WS. 1979. Robust Estimation of Population Size When Capture Probabilities Vary Among Animals. Ecology **60**:927–936.

Cheng E, Hodges KE, Sollmann R, Mills LS. 2017. Genetic Sampling for Estimating Density of Common Species. Ecology and Evolution:1–11.

Colwell RK, Coddington J a. 1994. Estimating Terrestrial Biodiversity through Extrapolation. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **345**:101–118.

Costa V, Rosenbom S, Monteiro R, O'Rourke SM, Beja-Pereira A. 2016. Improving DNA quality extracted from fecal samples—a method to improve DNA yield. European Journal of Wildlife Research **63**:3. European Journal of Wildlife Research. Available from http://link.springer.com/10.1007/s10344-016-1058-1.

Crosbie SF, Manly BFJ. 1985. Parsimonious Modelling of Capture-Mark-Recapture Studies. J. Manly Source: Biometrics BIOMETRICS **41**:385–398.

Do Rosário IT, Mathias MDL. 2007. Post-fire Recolonisation of a Montado Area by the Endangered Cabrera Vole (*Microtus cabrerae*). International Journal of Wildland Fire **16**:450–457.

Eggert LS, Eggert JA, Woodruff DS. 2003. Estimating population sizes for elusive animals: The forest elephants of Kakum National Park, Ghana. Molecular Ecology **12**:1389–1402.

Engeman RM. 2003. More on the Need to Get the Basics Right: Population Indexes. Wildlife Society

Bulletin **31**:286–287.

Engeman RM. 2005. Indexing principles and a widely applicable paradigm for indexing animal populations. Wildlife Research **32**:203–210.

Ernest HB, Penedob MCT, May P, Syvanen M, Boyce WM. 2000. Molecular Tracking of Mountain Lions in the Yosemite Valley Region in California: Genetic Analysis Using Microsatellites and Faecal DNA. Molecular Ecology **9**:433–441.

Fernandes M, Pita R, Mira A. 2008. *Microtus cabrerae*. Page The IUCN Red List of Threatened Species. The IUCN Red List of Threatened Species.

Fernández-Salvador R. 1998. Topillo de Cabrera, *Microtus cabrerae* THOMAS, 1906. Galemys:5–18. Available from http://www.secem.es/wp-content/uploads/2013/03/G-10-2-01-Fdez-Salvador-5-18.pdf (accessed March 3, 2017).

Ferreira CM et al. (submitted). High Cost-effectiveness of Genetic Non-invasive Sampling (gNIS) for Monitoring Elusive Small Mammals.

Ferron J, Ferron J, St-Laurent M-H, St-Laurent M-H. 2008. Testing the Reliability of Pellet Counts as an Estimator of Small Rodent Relative Abundance in Mature Boreal Forest. Journal of Negative Results, Ecology & Evolutionary Biology **5**:14–22.

Frantz AC, Roper TJ. 2006. Simulations to assess the performance of different rarefaction methods in estimating population size using small datasets. Conservation Genetics **7**:315–318.

Garel M, Bonenfant C, Hamann J-L, Klein F, Gaillard J-M. 2010. Are abundance indexes derived from spotlight counts reliable to monitor red deer *Cervus elaphus* populations? Wildlife Biology **16**:77–84.

Garrido-García JA et al. 2013. Revisión a nivel ibérico de la distribución del topillo de Cabrera o iberón, *Iberomys cabrerae* (Thomas, 1906). Galemys **25**:35–49. Available from http://secem.es/galemys/index.php/Galemys/article/view/42 (accessed February 22, 2017).

Gervais J a. 2010. Testing sign indexes to monitor voles in grasslands and agriculture. Northwest Science **84**:281–288.

Gomes LAP, Mira APP, Barata EN. 2013. The Role of Scent-marking in Patchy and Highly Fragmented Populations of the Cabrera Vole (*Microtus cabrerae* Thomas, 1906). Zoological science **30**:248–54.

Güthlin D, Storch I, Küchenhoff H. 2014. Toward Reliable Estimates of Abundance: Comparing Index Methods to Assess the Abundance of a Mammalian Predator. PLoS ONE **9**.

Hedges S, Johnson A, Ahlering M, Tyson M, Eggert LS. 2013. Accuracy, Precision, and Cost-Effectiveness of Conventional Dung Density and Fecal DNA Based Survey Methods to Estimate Asian Elephant (*Elephas maximus*) Population Size and Structure. Biological Conservation **159**:101–108. Elsevier Ltd.

Jareño D, Viñuela J, Luque-Larena JJ, Arroyo L, Arroyo B, Mougeot F. 2014. A comparison of methods for estimating common vole (*Microtus arvalis*) abundance in agricultural habitats. Ecological Indicators **36**:111–119.

Johnson PCD, Haydon DT. 2007. Maximum-Likelihood Estimation of Allelic Dropout and False Allele Error Rates From Microsatellite Genotypes in the Absence of Reference Data. Genetics **175**:827–842.

Kamil B. 2016. Package "MuMIn"- Multi-Model Inference. Available from https://cran.r-project.org/web/packages/MuMIn/MuMIn.pdf (accessed August 29, 2017).

Karmacharya DB, Thapa K, Shrestha R, Dhakal M, Janecka JE. 2011. Noninvasive Genetic Population Survey of Snow Leopards (*Panthera uncia*) in Kangchenjunga Conservation Area, Shey

Phoksundo National Park and Surrounding Buffer Zones of Nepal. BMC Research Notes **4**:516. Available from http://www.biomedcentral.com/1756-0500/4/516.

Kearse M et al. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics **28**:1647–1649.

Kilpatrick HJ, Goodie TJ, Kovach AI. 2013. Comparison of Live-trapping and Noninvasive Genetic Sampling to Assess Patch Occupancy by New England Cottontail (Sylvilagus transitionalis) Rabbits. Wildlife Society Bulletin **37**:901–905.

Kohn MH, York EC, Kamradt D a, Haught G, Sauvajot RM, Wayne RK. 1999. Estimating population size by genotyping faeces. The Royal Society **266**:657–663.

Lampa S, Henle K, Klenke R, Hoehn M, Gruber B. 2013. How to overcome genotyping errors in non-invasive genetic mark-recapture population size estimation - A review of available methods illustrated by a case study. Journal of Wildlife Management **77**:1490–1511.

Lüdecke D. 2017. Package "sjPlot"- Data Visualization for Statistics in Social Science. Available from https://cran.r-project.org/web/packages/sjPlot/sjPlot.pdf (accessed August 29, 2017).

Luikart G, Ryman N, Tallmon DA, Schwartz MK, Allendorf FW. 2010. Estimation of census and effective population sizes: The increasing usefulness of DNA-based approaches. Conservation Genetics **11**:355–373.

Lukacs PM, Burnham KP. 2005. Review of Capture-recapture Methods Applicable to Noninvasive Genetic Sampling. Molecular Ecology **14**:3909–3919.

Marucco F, Boitani L, Pletscher DH, Schwartz MK. 2011. Bridging the Gaps Between Non-invasive Genetic Sampling and Population Parameter Estimation. European Journal of Wildlife Research **57**:1–13.

Maudet C, Luikart G, Dubray D, Von Hardenberg A, Taberlet P. 2004. Low genotyping error rates in wild ungulate faeces sampled in winter. Molecular Ecology Notes **4**:772–775.

McKelvey KS, Pearson DE. 2001. Population estimation with sparse data: the role of estimators versus indexes revisited. Canadian Journal of Zoology **79**:1754–1765.

Miller CR, Joyce P, Waits LP. 2005. A new method for estimating the size of small populations from genetic mark – recapture data. Molecular Ecology **14**:1991–2005.

Mills LS, Citta JJ, Lair KP, Schwartz MK, Tallmon DA. 2000. Estimating animal abundance using noninvasive DNA sampling: Promise and pitfalls. Ecological Applications **10**:283–294.

Mira A, Marques CC, Santos SM, Rosário IT, Mathias ML. 2008. Environmental Determinants of the Distribution of the Cabrera Vole (*Microtus cabrerae*) in Portugal: Implications for Conservation. Mammalian Biology - Zeitschrift für Säugetierkunde **73**:102–110.

Mowry RA, Gompper ME, Beringer J, Eggert LS. 2011. River Otter Population Size Estimation Using Noninvasive Latrine Surveys. Journal of Wildlife Management **75**:1625–1636.

Murray DL, Roth JD, Ellsworth E, Wirsing AJ, Steury TD. 2002. Estimating low-density snowshoe hare populations using fecal pellet counts. Canadian Journal of Zoology **80**:771–781.

Nakagawa S, Schielzeth H. 2013. A General and Simple Method for Obtaining R2 from Generalized Linear Mixed-effects Models. Methods in Ecology and Evolution **4**:133–142.

Pearse DE, Eckerman CM, Janzen FJ, Avise C. 2001. A genetic analogue of "mark–recapture" methods for estimating population size: an approach based on molecular parentage assessments. Molecular Ecology **10**:2711–2718.

Pennell MW, Stansbury CR, Waits LP, Miller CR. 2013. Capwire: a R Package for Estimating Population Census Size From Non-invasive Genetic Sampling. Molecular Ecology Resources **13**:154–157. Available from http://doi.wiley.com/10.1111/1755-0998.12019 (accessed May 24,

2017).

Perez I, Geffen E, Mokady O. 2006. Critically Endangered Arabian Leopards *Panthera pardus nimr* in Israel: Estimating Population Parameters Using Molecular Scatology. Oryx **40**:295. Available from http://www.journals.cambridge.org/abstract_S0030605306000846.

Petit E, Valiere N. 2005. Estimating Population Size with Noninvasive Capture-Mark-Recapture Data. Conservation Biology **20**:1062–1073.

Piggott MP. 2004. Effect of Sample Age and Season of Collection on the Reliability of Microsatellite Genotyping of Faecal DNA. Wildlife Research **31**:485–493.

Pita R, Beja P, Mira A. 2007. Spatial population structure of the Cabrera vole in Mediterranean farmland: The relative role of patch and matrix effects. Biological Conservation **134**:383–392.

Pita R, Lambin X, Mira A, Beja P. 2016. Hierarchical Spatial Segregation of Two Mediterranean Vole Species: the Role of Patch-Network Structure and Matrix Composition. Oecologia **182**:253–263. Springer Berlin Heidelberg.

Pita R, Mira A, Beja P. 2006. Conserving the Cabrera vole, Microtus cabrerae , in intensively used Mediterranean landscapes. Agriculture, Ecosystems & Environment **115**:1–5.

Pita R, Mira A, Beja P. 2014. *Microtus cabrerae* (Rodentia: Cricetidae). Mammalian Species **912**:48–70. Available from https://academic.oup.com/mspecies/article-lookup/doi/10.1644/912.1.

Pita R, Mira A, Moreira F, Morgado R, Beja P. 2009. Influence of Landscape Characteristics on Carnivore Diversity and Abundance in Mediterranean Farmland. Agriculture, Ecosystems and Environment **132**:57–65.

Prigioni C, Remonti L, Balestrieri A, Sgrosso S, Priore G, Mucci N, Randi E. 2006. Estimation of European Otter (*Lutra Lutra*) Population Size By Fecal Dna Typing in Southern Italy. Journal of Mammalogy **87**:855–858.

R Development Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reddy PA, Kumaraguru A, Yadav PR, Ramyashree A, Bhagavatula J, Shivaji S. 2011. Studies to Determine Presence or Absence of the Indian Tiger (*Panthera tigris tigris*) in Kawal Wildlife Sanctuary, India. European Journal of Wildlife Research **57**:517–522.

Rodgers TW, Janečka JE. 2013. Applications and techniques for non-invasive faecal genetics research in felid conservation. European Journal of Wildlife Research **59**:1–16. Springer-Verlag. Available from http://link.springer.com/10.1007/s10344-012-0675-6 (accessed February 7, 2017).

Royle J a, Young K V. 2008. A Hierarchical Model for Spatial Capture-recapture Data. Ecology **89**:2281–2289.

Rozenfeld FM, Le Boulangé E, Rasmont R. 1987. Urine Marking by Male Bank Voles (*Clethrionomys glareolus* Schreber, 1780; *Mircrotidae*, Rodentia) in Relation to Their Social rank. Canadian Journal of Zoology **65**:2594–2601.

Sabino-Marques H et al. (submitted). Combining noninvasive genetics with spatially explicit capture-recapture models for density estimation of a patchily distributed small mammal.

San Miguel Ayanz A. 1992. Inventario de la Poblacion Española del Topillo de Cabrera (Microtus cabrerae Thomas, 1906). Madrid.

Santos SM, Simões MP, Mathias MDL, Mira A. 2006. Vegetation analysis in colonies of an endangered rodent, the Cabrera vole (*Microtus cabrerae*), in southern Portugal. Ecological Research **21**:197–207.

Schwartz MK, Pilgrim KL, McKelvey KS, Rivera PT, Ruggiero LF. 2007. DNA Markers for Identifying Individual Snowshoe Hares Using Field-collected Pellets. Northwest Science **81**:316–322.

Siddig A a., Ellison AM, Jackson S. 2015. Calibrating abundance indexes with population size estimators of red back salamanders (*Plethodon cinereus*) in a New England forest. PeerJ **3**:e952. Available from https://peerj.com/articles/952.

Silva TL, Godinho R, Castro D, Abáigar T, Brito JC, Alves PC. 2015. Genetic Identification of Endangered North African Ungulates Using Noninvasive Sampling. Molecular Ecology Resources **15**:652–661.

Slade NA, Blair SM. 2000. An Empirical Test of Using Counts of Individuals Captured as Indexes of Population Size. Journal of Mammalogy **81**:1035–1045.

Taberlet P, Camarra JJ, Griffin S, Uhres E, Hanotte O, Waits LP, DuboisPaganon C, Burke T, Bouvet J. 1997. Noninvasive genetic tracking of the endangered Pyrenean brown bear population. Molecular Ecology **6**:869–876.

Taberlet P, Luikart G, Waits LP, Luikart G. 1999. Non-invasive genetic sampling and individual identification. Biological Journal of the Linnean Society **68**:41–55. Oxford University Press. Available from https://academic.oup.com/biolinnean/article-lookup/doi/10.1111/j.1095-8312.1999.tb01157.x (accessed February 7, 2017).

Tellería JL. 2004. Métodos de Censo en Vertebrados Terrestres. Dep. Biología Animal I. Fac. Biología, Univ. Camplutense, Madrid.

Valière N. 2002. GIMLET : a computer program for analysing genetic individual identification data. Molecular Ecology Notes **1000228**:1471–8278. Available from ftp://pbil.univ-lyon1.fr/pub/logiciel/gimlet/Gimlet paper.pdf (accessed July 12, 2017).

Village A, Myhill D. 1990. Estimating Small Mammal Abundance for Predator Studies - Snap-Trapping Versus Sign Indexes. Journal of Zoology **222**:681–689.

Waits JL, Leberg PL. 2000. Biases Associated with Population Estimation Using Molecular Tagging. Animal Conservation **3**:191–199.

Watkins AF, McWhirter JL, King CM. 2010. Variable Detectability in Long-term Population Surveys of Small Mammals. European Journal of Wildlife Research **56**:261–274.

Witmer GW. 2005. Wildlife population monitoring: Some practical considerations. Wildlife Research **32**:259–263.

Woodroffe GL, Lawton JH, Davidson WL. 1990. Patterns in the Production of Latrines by Water Voles (*Arvicola terrestris*) and Their Use as Indexes of Abundance in Population Surveys. Journal of Zoology **220**:439–445.

3. Synthesis of Main Findings and Implications

Understanding and monitoring the fluctuations in wild populations is key for population ecology and management (e.g. Jareño et al. 2014), though having sufficient data on species population size, structure, distribution, and dynamics is often a challenging task, especially for elusive, cryptic and endangered species (Frantz et al. 2004; Bonesi & Macdonald 2004; Waits & Paetkau 2005; Lampa et al. 2013). The collection of informative data for such species often implies a combination of costly and time-consuming methods (e.g. Frantz et al. 2004; Hedges et al. 2013). For instance, live-trapping capture-recapture (CR) studies have been for long the most popular method for obtaining demographic information for mammal species (Stickel 1954; Village & Myhill 1990; De Bondi et al. 2010). However, this method requires very demanding efforts both in term of logistics and men power, and often provides insufficient data for obtaining accurate estimates of the population parameters of interest (Tellería 2004). In this context, genetic non-invasive sampling (gNIS) is considered at present a promising technique to overcome the main limitations of live-trapping CR studies (Frantz et al. 2004; Miller et al. 2005; Torre et al. 2010; Kilpatrick et al. 2013; Cheng et al. 2017), though there are still very few examples in the literature applying this technique in small mammal species (Sibbald et al. 2006; Jareño et al. 2014). In the case of elusive species, inferences on population size are often made from abundance indexes , as these provide a relatively easier and more expedite alternative to methods involving live-trapping (Bonesi & Macdonald 2004). Most population size indexes (PSI) are based on presence-sign counts, and are assumed to have a linear relation to the true population size. However, in many cases PSIs are seldom validated against "true" population sizes , making their use in conservation management questionable. (Bonesi & Macdonald 2004).

In this thesis, we addressed the difficulties of obtaining reliable information on population size of an elusive small mammal species, the Cabrera vole (*Microtus cabrerae*, Thomas 1906). The study was considered timely and of importance, as the Cabrera vole is "Near-Threatened", and several studies aiming to assess local population status have relied on abundance indexes built from presence-sign surveys (e.g. latrines) (Pita et al. 2007; Barbosa et al. 2013; Mestre et al. 2015; Ferreira et al. submitted; Sabino-Marques et al. submitted), which, however, were never validated with appropriate estimates of the 'true' population sizes. While addressing these problem, we found that PSI based on latrine counts have the potential to provide a good approximation of the relative abundance, though care must be taken regarding the size of latrines considered, and the sampling effort used. In particular, it seems that counts of latrines with less than 20 faeces under sampling efforts of at least 60 min/0.1ha, correlate positively with population size estimates based on gNIS capture-recapture. To our knowledge, this was the first study explicitly testing the utility of PSI built for the Cabrera vole to infer local population sizes of the species.

Although we acknowledge that much is still to be done to improve gNIS-based population estimates (e.g. improved techniques and protocols to minimize genotyping errors and maximize genotyping success), gNIS of Cabrera vole faeces has already been proven to provide accurate population size estimates (Ferreira et al. submitted; Sabino-Marques et al. submitted), and we believe that our reference score based on the Eggert's accumulation curve provided a good approximation of the 'true' population sizes. Outside the study area, the index may need to be adjusted since the conditions where the species inhabits may differ according to the geographic area . In areas where other Arvicolidae species producing similar presence signs are present (e.g. *Microtus agrestis* in Northern Iberia), it may not be so easy to assess the species presence based on sign surveys. In addition, some less fragmented landscapes may contain habitat patches much larger then those considered in this study, thus requiring some readjustments in the sampling design.

Therefore, while further studies including more sampling patches, and considering other additional factors not explicitly addressed here (e.g. seasonal effects), we believe that PSI for the Cabrera vole based on presence-sign surveys may be used to monitor the species, and guide conservation management of its populations and habitats.

3.1. Literature cited

Barbosa S, Pauperio J, Searle JB, Alves PC. 2013. Genetic identification of Iberian rodent species using

both mitochondrial and nuclear loci: Application to noninvasive sampling. Molecular Ecology Resources **13**:43–56.

Bonesi L, Macdonald D. 2004. Evaluation of sign surveys as a way to estimate the relative abundance of American mink (*Mustela vison*). Journal of Zoology **262**:65–72.

Cheng E, Hodges KE, Sollmann R, Mills LS. 2017. Genetic Sampling for Estimating Density of Common Species. Ecology and Evolution:1–11.

De Bondi N, White JG, Stevens M, Cooke R. 2010. A Comparison of the Effectiveness of Camera Trapping and Live Trapping for Sampling Terrestrial Small-mammal Communities. Wildlife Research **37**:456–465.

Ferreira CM et al. (submitted). High Cost-effectiveness of Genetic Non-invasive Sampling (gNIS) for Monitoring Elusive Small Mammals.

Frantz AC, Schaul M, Pope LC, Fack F, Schley L, Muller CP, Roper T. 2004. Estimating population size by genotyping remotely plucked hiar: the European badger. Journal of Applied Ecology **41**:985–995.

Hedges S, Johnson A, Ahlering M, Tyson M, Eggert LS. 2013. Accuracy, Precision, and Cost-Effectiveness of Conventional Dung Density and Fecal DNA Based Survey Methods to Estimate Asian Elephant (*Elephas maximus*) Population Size and Structure. Biological Conservation **159**:101–108. Elsevier Ltd. Available from http://dx.doi.org/10.1016/j.biocon.2012.12.010.

Jareño D, Viñuela J, Luque-Larena JJ, Arroyo L, Arroyo B, Mougeot F. 2014. A comparison of methods for estimating common vole (*Microtus arvalis*) abundance in agricultural habitats. Ecological Indicators **36**:111–119.

Kilpatrick HJ, Goodie TJ, Kovach AI. 2013. Comparison of Live-trapping and Noninvasive Genetic Sampling to Assess Patch Occupancy by New England Cottontail (Sylvilagus transitionalis) Rabbits. Wildlife Society Bulletin **37**:901–905. Available from http://doi.wiley.com/10.1002/wsb.330.

Lampa S, Henle K, Klenke R, Hoehn M, Gruber B. 2013. How to overcome genotyping errors in non-invasive genetic mark-recapture population size estimation - A review of available methods illustrated by a case study. Journal of Wildlife Management **77**:1490–1511.

Mestre F, Pita R, Paupério J, Martins FMS, Alves PC, Mira A, Beja P. 2015. Combining distribution modelling and non-invasive genetics to improve range shift forecasting. Ecological Modelling **297**:171–179. Elsevier B.V. Available from http://dx.doi.org/10.1016/j.ecolmodel.2014.11.018.

Miller CR, Joyce P, Waits LP. 2005. A new method for estimating the size of small populations from genetic mark – recapture data. Molecular Ecology **14**:1991–2005.

Pita R, Beja P, Mira A. 2007. Spatial population structure of the Cabrera vole in Mediterranean farmland: The relative role of patch and matrix effects. Biological Conservation **134**:383–392.

Sabino-Marques H et al. (submitted). Combining noninvasive genetics with spatially explicit capture-recapture models for density estimation of a patchily distributed small mammal.

Sibbald S, Carter P, Poulton S. 2006. Proposal for a National Monitoring Scheme for Small Mammals in the United Kingdom and the Republic of Eire. Page The Mammal Society Research report No. 6.

Stickel LF. 1954. A Comparison of Certain Methods of Measuring Ranges of Small Mammals. Journal of Mammalogy **35**:1–15.

Tellería JL. 2004. Métodos de Censo en Vertebrados Terrestres. Dep. Biología Animal I. Fac. Biología, Univ. Camplutense, Madrid.

Thomas O. 1906. A new Vole from Spain. Annals and Magazine of Natural History **7**:576–577.

Torre I, Guixé D, Sort F. 2010. Comparing Three Live Trapping Methods for Estimating Small Mammal Sampling in Cultivated Areas of NE Spain. Italian Journal of Mammalogy **21**:147–155.

Village A, Myhill D. 1990. Estimating Small Mammal Abundance for Predator Studies - Snap-Trapping Versus Sign Indexes. Journal of Zoology **222**:681–689.

Waits LP, Paetkau D. 2005. Noninvasive Genetic Sampling Tools for Wildlife Boiologists: A Review

of Applications and Recommendations for Accurate Data Collection. Journal of Wildlife Management **69**:1419–1433.

Appendix 1

Table A1.1 - PSI built for each one of the 8 sampled patches, by counting latrines in three different sampling efforts. Latrines were sized as large (LL) and small (SL) and the total number of latrines was also recorded for each patch (TL).

	60 min/ 0.1ha			30 min/ 0.1ha			15 min/ 0.1ha		
Habitat/ Size	**LL**	**SL**	**TL**	**LL**	**SL**	**TL**	**LL**	**SL**	**TL**
69	5	54	59	4	41	45	4	25	29
148	26	37	63	16	25	41	9	10	19
236	19	57	76	3	18	21	1	7	8
248	23	58	81	22	54	76	18	43	61
336	45	35	80	32	28	60	13	17	30
536	15	21	36	6	11	17	3	5	8
579	11	90	101	8	73	81	8	51	59
597	18	30	48	9	12	21	4	6	10
Totals	**162**	**382**	**544**	**100**	**262**	**362**	**60**	**164**	**224**

Table A1.2 - Genotyping error rates calculations, for each one of the 9 tested loci *(Mc18, Mc24, Mc30, Ma25, Mar76, Mc02, MSMM-3, Mar03, Mar16) in combinations of 4 different replicas (rows 3-8). E1 represents Allelic dropout rates and E2 the false alleles rate. For each* locus *the medium value of E1 and E2 was calculated (final row).*

	Mc18		**Mc24**		**Mc30**		**Ma25**		**Mar76**		**Mc02**		**MSMM-3**		**Mar03**		**Mar16**	
	E1	**E2**	**E1**	**E2**	**E1**	**E2**	**E1**	**E2**	**E1**	**E2**	**E1**	**E2**	**E1**	**E2**	**E1**	**E2**	**E1**	**E2**
1-2	0,00000	0,004464	0,011589	0,002776	0,028748	0,004927	0,023	0,012605	0,023127	0,002959	0,022202	0,00000	0,012745	0,002444	0,019187	0,006739	0,00000	0,00000
1-3	0,00000	0,00000	0,014123	0,002834	0,025554	0,002426	0,029335	0,013216	0,021201	0,00000	0,018505	0,00000	0,021204	0,002557	0,014007	0,00000	0,003519	0,00000
1-4	0,000001	0,00000	0,017026	0,002909	0,021988	0,002427	0,036986	0,006853	0,024016	0,00000	0,04387	0,00000	0,023604	0,002478	0,011228	0,00000	0,00000	0,00000
2-3	0,000001	0,00000	0,005461	0,005549	0,018664	0,00000	0,059813	0,012501	0,028992	0,00000	0,040371	0,00000	0,015996	0,00252	0,014151	0,00000	0,003469	0,00000
2-4	0,000001	0,00000	0,006545	0,002821	0,02199	0,00000	0,040158	0,000001	0,02292	0,00000	0,035446	0,00000	0,021035	0,002485	0,01434	0,00000	0,00000	0,002036
3-4	0,003149	0,00000	0,009184	0,002858	0,018423	0,00000	0,030301	0,007141	0,022894	0,00000	0,036882	0,00000	0,021477	0,002553	0,008405	0,00000	0,003517	0,00000
Med	0,000525	0,000744	0,010655	0,003291	0,022561	0,00163	0,036599	0,00872	0,023858	0,000493	0,032879	0	0,019344	0,002506	0,013553	0,001123	0,001751	0,000339

Table A1.3 - Markers used in the genotyping protocol. Mix 1 represents the quality screening and Mix 2 and 3are used a posteriori *for the good quality samples*

Mix 1		**Mix 2**		**Mix 3**	
Marker ID	**Reference**	**Marker ID**	**Reference**	**Marker ID**	**Reference**
Mc18	Ferreira et al. (subm)	*Ma25*	Gauffre et al. 2007	*MSMM-3*	Ishibashi et al. 1999
Mc24	Ferreira et al. (subm)	*Mar76*	Walser & Heckel 2008	*Mar03*	Walser and Heckel 2007
Mc30	Ferreira et al. (subm)	*Mc02*	Ferreira et al. (subm)	*Mar16*	Walser and Heckel 2007

Literature cited

Ferreira CM et al. (submitted). High Cost-effectiveness of Genetic Non-invasive Sampling (gNIS) for Monitoring Elusive Small Mammals.

Gauffre B, Galan M, Bretagnolle V, Cosson JF. 2007. Polymorphic microsatellite loci and PCR multiplexing in the common vole, Microtus arvalis. Molecular Ecology Notes **7**:830–832.

Ishibashi Y, Yoshinaga Y, Saitoh T, Abe S, Iida H, Yoshida M. 1999. Polymorphic Microsatellite DNA Markers in the Field Vole Microtus montebelli. Molecular Ecology **8**:163–164.

Walser B, Heckel G. 2008. Microsatellite markers for the common vole (Microtus arvalis) and their cross-species utility. Conservation Genetics **9**:479–481.

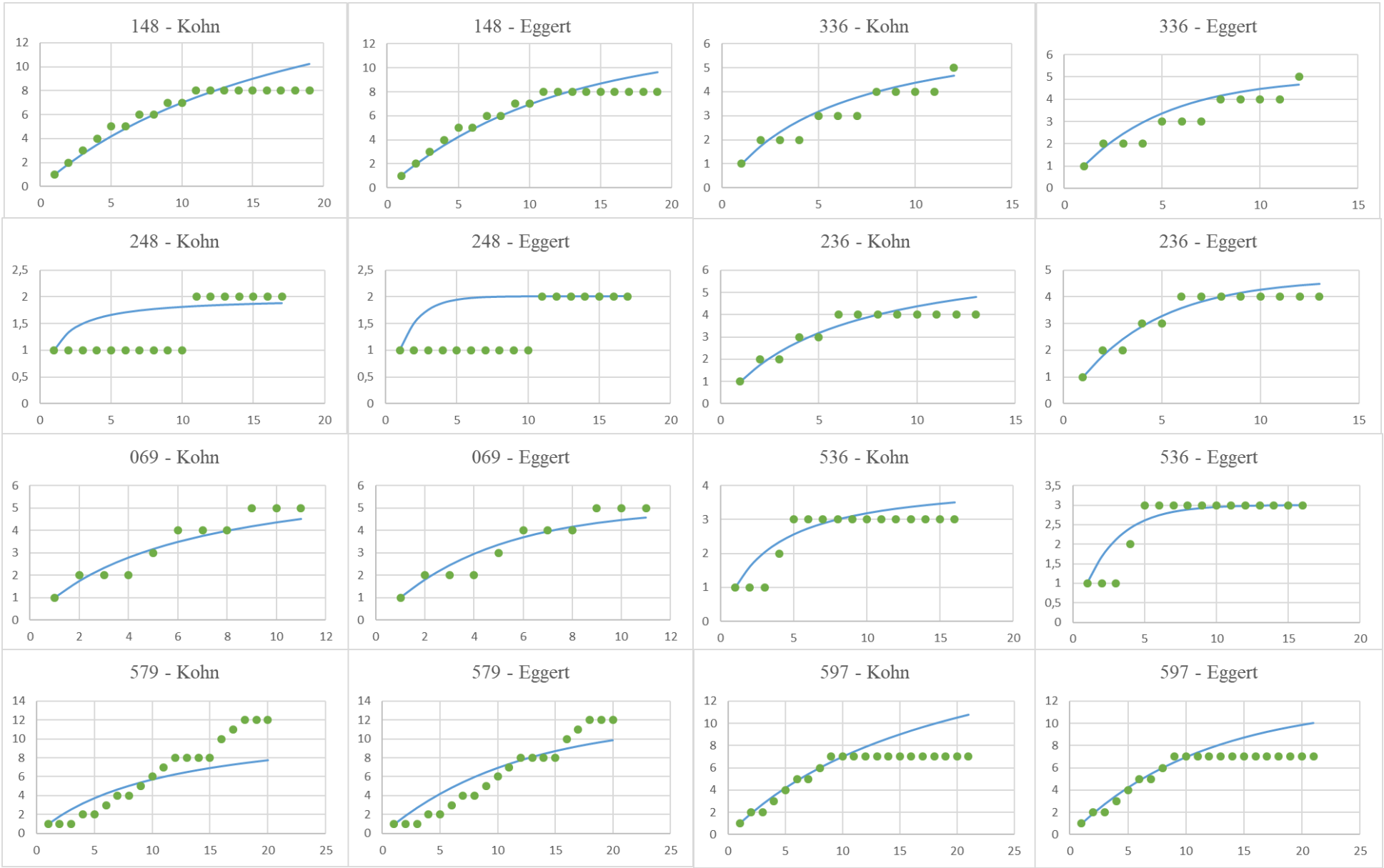

Figure A1.1- Example of rarefaction curve fitting for each of the 8 habitat patches considered in the study. In each case 100 replicas were fitted to obtain a final estimate (mean±SD) (see main text for details).

Appendix 2

– MateFrag –

Pág.: _____/_____

"CAPTURA-RECAPTURA" DE DEJECTOS

Habitat: _____ Área (m^2): _____ Tempo de amostragem (minutos): _____

Data: _____/_____/_____ Observador: _____ (+ _____) equipa

Hora de início: _____ Hora de fim: _____ Notas: _____

Ocupação		Tempo de detecção inicial	GPS	ID Pontos
ARVICOLA				
CABRERA				

Código	GPS		Tempo de detecção	Dejectos	Latrina	Frescura	CAB	Observações
	Nome	Ponto						

Habitat: _____ Data: _____/_____/_____ Observador: _____ Pág.: _____/_____

Figure A2.1 - Field data sheets from the "MATEFRAG" project, used to record Cabrera voles detection/non detection, and habitat-specific variables.

.