Nondestructive molecular sex determination of free-ranging star-nosed moles (*Condylura cristata*)

by

Nadine T. Price

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements for the degree of

MASTER of SCIENCE

Department of Biological Sciences

University of Manitoba

Winnipeg, Manitoba

Abstract:

Molecular techniques, particularly noninvasive genetic sampling (NGS) and nondestructive sampling (NDS), are increasingly being used as tools to study the ecology of free-ranging mammals. A specific application of these methods is the molecular sexing of species for which external sex differentiation is challenging. Star-nosed moles (*Condylura cristata*) are a little-studied species in which females possess a peniform clitoris making them externally indistinguishable from males. To my knowledge, no studies have employed NDS to study any aspect of their ecology. I therefore sequenced fragments of one X-chromosome (*Zfx*) and two Y-chromosome (*Sry* and *Zfy*) genes from known-sex specimens, and designed species-specific primers to co-amplify these loci from hair, claw and fecal samples of 16 star-nosed moles. I found all tissue types were highly (90-100%) reliable for sex determination. I envision that this NDS method will facilitate future capture-and-release studies on the natural history and social structure of this fascinating, semi-aquatic mammal.

Acknowledgements

I found my Master's program to be one of the most challenging yet rewarding experiences of my life. If it were not for the many people that provided assistance and support along the way, my project would not have been completed.

First and foremost, I would like to thank my supervisors, Dr. Kevin Campbell and Dr. Craig Willis. Kevin provided a convincing argument for studying the star-nosed mole and got me hooked on this fascinating species from the beginning. After my original field project took a drastic turn for the worst and I had to completely change gears and do a molecular project instead, Kevin and Craig were an ongoing source of motivation and encouragement. I learned so much from them and appreciate the patience and kindness they showed towards me.

I would also like to thank my advisory committee: Dr. Nicola Koper and Dr. Jim Roth. I am grateful for their advice and support, especially in the matters of statistics and enjoyable weekly lab meetings. I met many new fellow graduate students through Jim's lab meetings and became good friends with them.

If it were not for the assistance and support from my lab mates, Tony Signore, Mike Gaudry and Dr. Kai He, my project would not have been possible. Tony and Kai taught me most of what I know about molecular lab techniques and Mike spent countless hours with me in the field trapping moles and on campus caring for moles.

I am very thankful to Dr. Jane Waterman and Dr. Margaret Docker for lending me space in their labs. This allowed me to complete the remainder of my project hassle-free. I am also thankful to Dr. Jim Hare for allowing me to share office space with his students in my first year here; I really enjoyed getting to know him and his students.

I am also grateful for the help and support I received from the university veterinarians (Dr. Valerie Smid, Dr. Gordon Clark and Dr. Richard Hodges), the Animal Holding Facility staff members, Dr. Stephen Petersen and Lucy Johnson. The veterinarians and Animal Holding Facility staff helped me with the caring of live moles and collection of samples while Stephen and Lucy provided advice on using the Zinc finger gene for molecular sexing of moles.

Financial assistance for this research was provided by a National Sciences and Engineering Research Council of Canada Discovery Grant to Kevin Campbell, a Faculty of Science Graduate Scholarship (University of Manitoba) and the Richard C. Goulden Memorial Award (Manitoba Chapter of The Wildlife Society).

Finally, I would like to thank my family for their constant support and encouragement and my friends for being there no matter what.

Table of Contents

List of Tables

List of Figures

Abbreviations used in text

Amelx	Amelogenin X
Amely	Amelogenin Y
bp	base pairs
CCAC	Canadian Council on Animal Care
Dbx	DEAD Box X
Dby	DEAD Box Y
DTT	dithiothreitol
GEE	generalized estimating equations
HMG	high mobility group
indels	insertions or deletions
L	ladder
NDS	nondestructive sampling
NGS	noninvasive genetic sampling
PCR	polymerase chain reaction
SD	standard deviation
Sry	sex-determining region Y
UTX and *UTY*	ubiquitously transcribed tetratricopeptide repeat protein gene
Zfx	Zinc finger X
Zfy	Zinc finger Y
♀	female
♂	male
(-)	negative PCR control

Chapter 1: Thesis Introduction

Introduction

Definitive sex identification is critical when studying the physiology, biology and ecology of mammals. Sex differentiation can be essential for understanding foraging habits, social relationships and networks, kin-based relationships, space-use and dispersal patterns, as all of these aspects of behavioural ecology can be sex-biased (DeYoung and Honeycutt 2005; Janečka et al. 2006; Walker et al. 2006; Bradley et al. 2008). Knowing the sex of mammals is also important for population ecology studies. Quantifying sex ratios (Bellemain et al. 2005; Rodgers and Janečka 2013) and sex-specific population parameters such as growth, mortality, fecundity and survivorship (Lanyon et al. 2009) is important for understanding population processes and trends. Such studies add to the fundamental knowledge of species, communities and ecosystems and are also critical for conservation and management (Bellemain et al. 2005; Goldberg et al. 2011; Sawaya et al. 2011).

For most mammalian species, sex can be determined based on external morphology. For many species, however, this requires capture, restraint and handling, which can be stressful for the animal (Taberlet et al. 1999). Moreover, for some mammals sex can be difficult if not impossible to assess based on external morphology and can be surprisingly difficult to determine for free-ranging or wild-captured individuals (e.g. North American beavers *Castor canadensis*, Kühn et al. 2002). Similarly, studying the foraging habits, population abundances and behavioural ecology of many mammalian species can involve capture, restraint and handling (Piggott and Taylor 2003). Although indirect methods, such as searching for tracks or browsing signs

have been used to study wild mammals, these approaches commonly produce inaccurate results (e.g. Choate et al. 2006; Hibert et al. 2011). More recently, molecular methods have been developed to sex mammalian species and study their ecology in order to overcome the problems associated with traditional approaches.

The star-nosed mole (*Condylura cristata*) is an example of a species from the Talpidae family that is impossible to sex based on external morphology (Rubenstein et al. 2003) while also being difficult to capture. In addition, very little is known about this mole's ecology. Thus, it is a likely candidate for alternatives to morphological sex differentiation and traditional study approaches. In this chapter, I review some of the methods for sex differentiation and ecological studies of mammals in general and then focus on the family Talpidae and likely best approaches for star-nosed moles.

Traditional methods

Depending on the species, different methods can be used to sex an animal. Some species can be sexed through distant field observations of external genitalia (e.g. crowned sifaka *Propithecus coronatus*, Vanpé et al. 2013) or secondary sexual characteristics and association with young (e.g. bottlenose dolphins *Tursiops* sp., Rowe and Dawson 2009). Species harvested in the fur trade, such as bobcats (*Lynx rufus*) and American martens (*Martes americana*), are generally sexed by visual observation of external genitalia or by size of the lower canine tooth (Williams et al. 2011; Belant et al. 2011). Other mammals must be live-captured and sexed while immobilized either due to the danger they may impose on researchers (e.g. brown bears *Ursus arctos*, Bellemain et al. 2005) or because they must be palpated or carefully examined for subtle sexual differences (e.g. North

American beavers, Kühn et al. 2002; shrews *Sorex* sp., Croin Michielsen 1966). Common to these methods are the increased chance of error when the researcher is inexperienced and the animal being examined is a juvenile or reproductively quiescent adult with less pronounced sexual features (Crawford JC et al. 2008; Carraway 2009; Williams et al. 2011). In the case of wild western gorillas (*Gorilla gorilla*), dung size and presence or absence of white hair observed in or near the nesting site has been used as an indicator of sex of the absent animal, however this method results in at least a 10% error (Bradley et al. 2008) likely due to its subjective nature, similar to many other physical and observational methods.

Traditional methods of studying foraging habits, population abundances and behaviour are also problematic. In the most extreme form, animals must be killed in order to examine the stomach contents for dietary studies (e.g. Bodmer 1991). More indirect approaches include searching for browsing signs in the field and collecting fecal samples to macroscopically identify prey remains, but these methods present challenges when the remains are too damaged or digested to identify properly (Kohn and Wayne 1997; Piggott and Taylor 2003; Hibert et al. 2011). Similarly, population and behavioural studies include methods that can be stressful for animals such as trapping and immobilizing for attachment of a radio-collar or less invasive indirect methods such as searching for tracks or other evidence of the animal's presence (Grigione et al. 1999; Choate et al. 2006; Bradley et al. 2008). Capturing and disturbing mammals is not always the best approach as it can negatively affect the health and behaviour of some species (e.g. northern hairy-nosed wombats *Lasiorhinus krefftii*, Hoyle et al. 1995) while other species may be too elusive and difficult to trap or observe in the wild (e.g. Eurasian badgers *Meles meles*,

Scheppers et al. 2007; felids, Rodgers and Janečka 2013). Although indirect methods avoid these problems, there are many errors associated with these studies such as overestimating population sizes and making inaccurate assumptions about the movement behaviour and composition of social groups (Grigione et al. 1999; Choate et al. 2006; Bradley et al. 2008).

Noninvasive and nondestructive molecular methods

In order to address the shortcomings and errors associated with traditional methods, noninvasive and nondestructive molecular techniques have been developed. Noninvasive genetic sampling (NGS) involves collection of a biological sample as a DNA source without the animal present (Taberlet et al. 1999). Shed hair or feces are most commonly collected for NGS (Waits and Paetkau 2005) as these types of samples are most often left behind by wild mammals, making it easier to study elusive species as well as animals that pose potential threats to the researcher. Nondestructive sampling (NDS) involves capturing the animal or using a biopsy dart gun (in the case of many marine mammals) to collect a biological sample as a DNA source and can involve varying levels of invasiveness depending on the sample type being collected (Taberlet et al. 1999). NDS can include collection of minimally invasive tissues such as feces or hair to collection of more invasive tissues such as ear biopsy punches (e.g. Lachish et al. 2011) or skin biopsies (e.g. Lanyon et al. 2009). Despite the more invasive manner of NDS, this method is quick and can involve short handling times. Most importantly, analyses based on NGS and NDS tend to be highly accurate when compared to traditional methods as they avoid factors that increase errors such as subjective assessment of sex differences

(e.g. Vanpé et al. 2013) or misinterpreting evidence left by one individual (e.g. tracks, scat) as coming from multiple individuals (e.g. Bradley et al. 2008).

Molecular sexing methods

Once tissue has been collected, through either NGS or NDS methods, DNA is extracted and target genes are amplified via polymerase chain reactions (PCR). Depending on the type of study and species involved, different genes are targeted for amplification and/or sequencing using short synthetic oligonucleotide sequences (~20 base pairs; bp) called primers (DeYoung and Honeycutt 2005). In order to determine sex using molecular techniques, primers are designed to specifically target loci found on the X and Y sex chromosomes. Several different genes or combinations of genes and methods have been used for this purpose. Apart from the Transcaucasian mole vole (*Ellobius lutescens*), Zaisan mole vole (*E. tancrei*), spiny rats (*Tokudaia osimensis osimensis* and *T. osimensis*) and the monotremes, all male mammals have a Y chromosome where the sex-determining region Y (*Sry*) gene is found (Just et al. 1995; Sutou et al. 2001; Wallis et al. 2008; Ferguson-Smith and Rens 2010). PCR amplification of this gene indicates the DNA sample is from a male while no amplification indicates a female (Sánchez et al. 1996). Although many studies have used this method (e.g. Matsubara et al. 2001; Lynch and Brown 2006; Campbell et al. 2010), PCR amplification failure can lead to incorrect sex assignment, i.e. a failed PCR amplification on a sample from a male will result in a female sex assignment (Waits and Paetkau 2005; Robertson and Gemmell 2006). In order to avoid such errors, many studies have incorporated an autosomal or X-linked gene fragment as an internal positive control that is co-amplified

with *Sry* in a duplex (two primer pair) PCR (e.g. Goldberg et al. 2011; Vanpé et al. 2013). This distinguishes the absence of *Sry* from a PCR failure, reducing the chance of misidentifying the sex of an individual (Lachish et al. 2011).

As an alternative to the *Sry* duplex PCR method, many studies have taken advantage of genes with copies on both the X and Y chromosomes, termed "gametologs" (García-Moreno and Mindell 2000). Some examples of these genes include Zinc finger X and Y (*Zfx* and *Zfy*; Shaw et al. 2003), Amelogenin X and Y (*Amelx* and *Amely*; Yamamoto et al. 2002), DEAD Box X and Y (*Dbx* and *Dby*; Vidal et al. 2010), and the ubiquitously transcribed tetratricopeptide repeat protein gene (*UTX* and *UTY*; Villesen and Fredsted 2006). Each pair of genes has similar coding sequences but multiple indels between the X- and Y-linked copies result in a length polymorphism, allowing simultaneous amplification of X and Y for males and amplification of X only for females from a single primer pair (Yamamoto et al. 2002; Han et al. 2010). Occasionally, a three-primer protocol must be used to target smaller fragments of these genes when working with degraded DNA samples such as from feces or hair (e.g. Vidal et al. 2010). An advantage to using these gene pairs is they occur in equimolar amounts in males which should theoretically avoid the problem of PCR competition and allelic dropout (i.e. the amplification failure of one gametolog, allele or locus occurring simultaneously with the amplification success of the other gametolog, allele or locus in a DNA sample from males), which can be experienced when using *Sry* with an autosomal gene of higher (2x) copy number (Shaw et al. 2003; Durnin et al. 2007). Curtis et al. (2007), however, argued that co-amplifying both alleles may lead to PCR competition and allelic dropout and

therefore incorrect sex assignment, and instead used two PCRs and primer pairs to amplify *Zfx* and *Zfy* from each sample separately.

Another molecular sexing method combines co-amplification of *Sry* and an X and Y-linked gene pair into a single multiplex PCR (e.g. Pomp et al. 1995). This method results in simultaneous amplification of two Y and one X-linked loci for males and only the single X-linked locus for females (e.g. Gokulakrishnan et al. 2012). Unlike the other methods, this multiplex method incorporates two Y-specific markers, which act as controls for both verification of amplification and confirmation of sex, thereby increasing the reliability of the sex-specific test (Gokulakrishnan et al. 2012; Bidon et al. 2013).

Although arguments have been made for and against each variation of the molecular sexing method, most studies have demonstrated high reliability for determining sex (e.g. 93 – 100%; Pilgrim et al. 2005; Lanyon et al. 2009; Lachish et al. 2011) regardless of the technique used. This high reliability of the molecular method allows accurate sex determination of mammals that are difficult to differentiate even after capture (e.g. North American beavers, Goldberg et al. 2011). Furthermore, the use of NGS molecular methods to determine sex allows researchers to avoid encounters with dangerous mammals, reduce stress on animals and correctly sex and study elusive species.

Applications of NGS and NDS methods

NGS and NDS molecular methods are not restricted only to sex determination. Another application of these methods is DNA barcoding. This technique can be used in dietary studies to identify the forage or prey species in fecal samples, which are typically

too damaged or digested to identify using standard techniques (e.g. Dodd et al. 2012). DNA barcoding involves obtaining a short nucleotide sequence (~650 bp or less) from the mitochondrial gene cytochrome *c* oxidase I (in animals) from the specimen of interest (Hebert et al. 2003; Hajibabaei et al. 2007). This sequence, or unique 'barcode', differs from comparable sequences from other species because of nucleotide substitutions (Hebert et al. 2003). The 'barcode' is then compared to a library of reference sequences to identify to species level (Hebert et al. 2003; Hajibabaei et al. 2007).

As an alternative to DNA barcoding and molecular techniques, hair samples can be used to investigate foraging habits with stable isotope analysis. Unique carbon and nitrogen stable isotope signatures of different animal tissues are found by determining the ratio of heavy to light isotopes in a sample and measuring this ratio against a standard (Kelly 2000; Crawford K et al. 2008). These signatures are then compared to the stable isotope signatures of potential prey items in order to understand more about the animal's diet (Crawford K et al. 2008). As hair reflects the diet of an individual at the time it was grown (Tieszen et al. 1983; Macko et al. 1999), it is particularly useful for studying seasonal changes in foraging habits.

Any noninvasive or nondestructive tissue can be used as a DNA source for genotyping or identifying individuals with the microsatellite loci technique (e.g. Bellemain et al. 2005). Microsatellite loci are short (< 100 bp) nucleotide sequences consisting of di-, tri- or tetra-nucleotide repeat motifs found interspersed within eukaryotic genomes (Tautz 1989; DeYoung and Honeycutt 2005). Variations between microsatellite alleles are based on length polymorphisms as a result of insertions or deletions (indels) of these repeat motifs (Tautz 1989; DeYoung and Honeycutt 2005).

Microsatellites can be extremely useful compared to traditional methods for estimating regional population densities and social group sizes. For example, Bradley et al. (2008) used this technique to demonstrate that traditional methods were overestimating western gorilla social group sizes by up to 40% due to the mistaken assumption that different individuals constructed every nest. Microsatellites can also be used to investigate kinship relationships and study population genetics (e.g. Janečka et al. 2006; Guerier et al. 2012), topics that cannot be addressed without molecular techniques and are now even more feasible with NGS and NDS methods.

Disadvantages of NGS and NDS molecular methods

While there are many advantages to NGS and NDS methods, there may also be disadvantages, especially as it relates to fecal and hair samples as DNA sources. Unlike tissues with high cell densities, such as muscle or liver, feces and hair have low cell densities meaning they yield much lower quantities of DNA (Taberlet et al. 1999; Bengtsson et al. 2012). The presence of PCR inhibitors in both feces and hair negatively affect the quality of the DNA extract (Taberlet et al. 1999; Eckhart et al. 2000; Oliveira et al. 2010). Additionally, the freshness or condition of the hair or fecal sample when collected and the portion of the hair (i.e. root or shaft) used for DNA extraction may affect the degree of degradation and therefore fragmentation of the DNA (Müller et al. 2007; Bengtsson et al. 2012; Reddy et al. 2012). Furthermore, the relatively low quality and quantity of DNA from hair and fecal samples increases the chance of contamination being introduced during laboratory processing (Taberlet et al. 1999; Waits and Paetkau

2005). All of these factors affect the amplification success of the target DNA and therefore the results of NGS and NDS studies.

A number of methods have been developed to address these challenges. Choosing the optimal DNA extraction method that both removes PCR inhibitors and increases DNA yield from the sample will improve amplification success (Müller et al. 2007; Vallet et al. 2008; Reddy et al. 2012). As DNA extraction quality and PCR results from the same sample can vary, repeating DNA extractions, performing PCR replicates of the same sample and re-amplifying PCR products can also help improve amplification success (Taberlet et al. 1996; Arandjelovic et al. 2009; Vanpé et al. 2013). Designing primers to amplify smaller fragment sizes (ideally <300 bp) generally allows fragmented DNA samples to amplify more readily (Waits and Paetkau 2005; Lindsay and Belant 2008). In addition, designing primers that are species-specific reduces the chance of DNA contamination from exogenous sources (Bidon et al. 2013). These approaches, along with careful field collection, laboratory processing and bookkeeping (Waits and Paetkau 2005; Robertson and Gemmell 2006), can help optimize the effectiveness of NGS or NDS studies.

Small mammal studies using NGS and NDS

Both NGS and NDS molecular methods are well established for many species of large mammals (e.g. western gorillas, Bradley et al. 2008; felids, Rodgers and Janečka 2013; canids, Seddon 2005; ursids, Bidon et al. 2013; sirenians, Lanyon et al. 2009; lowland tapir *Tapirus terrestris*, Hibert et al. 2011) but few studies have been published exploring these methods for small mammals. Due to their small size and generally elusive

nature, small mammals may be difficult to trap and observe in the wild and finding shed hair or feces is a challenge for many species (e.g. lagomorphs, Fontanesi et al. 2008, Henry and Russello 2011; rodents, Barbosa et al. 2013; rodents and shrews, Moran et al. 2008). Earlier studies used invasive ear biopsies (Ray et al. 2002; Meyer et al. 2006) or other tissues such as tails or toes (Bryja and Konečný 2003). Hair trap methods have only recently been tested for remotely collecting hair from some species (e.g. American pika *Ochotona princeps*, Henry and Russello 2011; common hamster *Cricetus cricetus*, Reiners et al. 2011), while fecal samples are typically collected from trapped individuals (e.g. Moran et al. 2008; Alasaad et al. 2011; Dodd et al. 2012) with the exception of fecal collection from beneath clusters of bats (Puechmaille and Petit 2007; Boston et al. 2012). Of these studies, most are focused on species identification from NGS or NDS (e.g. Alasaad et al. 2011; Barbosa et al. 2013) and few are focused on molecular sexing (e.g. Henry and Russello 2011), dietary studies (e.g. Dodd et al. 2012) or the use of microsatellites for population (e.g. Puechmaille and Petit 2007) or behavioural studies (e.g. Streatfeild et al. 2011).

Talpidae

Even fewer studies have focused on developing NGS or NDS methods for true moles and little is known about the ecology of this family (Talpidae). Similar to many other small mammals, most talpid species spend most of their lives underground and are difficult to study and observe in the wild (Stone and Gorman 1985). Although talpids are found throughout North America, Europe and Asia (Gorman and Stone 1990), there are few population studies (e.g. Waldron et al. 2000; García-López de Hierro et al. 2013) and

many species are consequently listed as having unknown population trends (IUCN 2013). Of the few population studies conducted, most have relied on traditional methods such as presence or absence of molehills and tunnels (e.g. Zurawska-Seta and Barczak 2012; García-López de Hierro et al. 2013), or in the case of the endangered semi-aquatic Pyrenean desman (*Galemys pyrenaicus*), scat sampling or live-trapping (González-Esteban et al. 2003a). I am aware of only one study that has made use of NGS to investigate genetic diversity of Pyrenean desman populations via fecal samples (Igea et al. 2013). Similarly, behavioural and space-use patterns have only been examined for a few talpid species. For example, using radiotelemetry, Stone and Gorman (1985) concluded that both European moles (*Talpa europaea*) and Pyrenean desmans purposely avoid conflict with neighbours by using different parts of their home range during active times of the day. Conversely, Macdonald et al. (1997) showed that the movement of European moles is actually dependent on the location of food resources, while Melero et al. (2012) suggested that Pyrenean desmans share space with other individuals. No published studies thus far have used NGS or NDS molecular techniques to investigate behaviour in talpid moles. Furthermore, although it is known that talpids predominantly eat invertebrates (Gorman and Stone 1990), little research has been conducted on their foraging ecology and seasonal dietary patterns. Of the few studies investigating their foraging ecology, many have used the invasive stomach content analysis approach (e.g. Hartman et al. 2000; Beolchini and Loy 2004; Kashimura et al. 2010) and none have made use of NDS or NGS molecular approaches.

Star-nosed mole

The North American moles have been even less of a focus for ecological studies and no published studies have used NGS or NDS molecular methods to study any of the seven recognized species. A particularly interesting North American talpid is the star-nosed mole. This mole is the only North American talpid that routinely exploits both aquatic and subterranean habitats for its invertebrate prey and can often be found in wetland areas with a stream nearby (Hamilton 1931; Schmidt 1931; Rust 1966; Catania 2012). The star-nosed mole is highly adapted to an amphibious lifestyle, using its extraordinary nose – consisting of 22 mobile, fleshy and highly tactile sensory appendages – to explore its environment and forage for terrestrial prey (Catania 1999; Catania and Remple 2005). This species also performs underwater olfaction, in which air bubbles are rapidly exhaled onto objects and then re-inspired, allowing potential aquatic prey items to be more easily identified while submerged (Catania 2006). So far it is unclear however whether there is a preference for either environment or if foraging habits change seasonally, as dietary habits have only been studied via terminal stomach content analysis (Hamilton 1931; Schmidt 1931; Rust 1966).

Star-nosed moles range farther north than any other North American mole (Petersen and Yates 1980) and must survive subzero temperatures and scarce terrestrial invertebrate food resources during winter (Aitchison 1979; Danks 2007). The small body size (~50 g; Hamilton 1931) and elevated metabolic rate and body temperature relative to other fossorial mole species (Campbell et al. 1999) suggest exceptional energetic adaptations for a small mammal in a cold northern climate. Other potential energetic adaptations have yet to be investigated in this species. For example, it is possible that

star-nosed moles alter their behaviour by huddling with conspecifics to survive harsh winter conditions but only anecdotal observations exist to support this hypothesis (Hamilton 1931). Virtually nothing is known about any other aspects of their behaviour, home range, activity patterns or population status in the wild. To my knowledge, only a few studies exist that included star-nosed moles as part of their small mammal live-trapping surveys, however all failed to capture and observe more than a few moles in the study area while hundreds of individuals of other species were captured (e.g. Stewart et al. 2008). Although star-nosed moles are listed as "least concern" with a stable population trend (IUCN 2013), they inhabit wetlands, which are rapidly diminishing in North America (Dahl 1990; Ducks Unlimited Canada 2010), suggesting there may be cause for concern but no data exist with which to test this hypothesis. The lack of knowledge about the ecology and conservation status of star-nosed moles, and the challenges associated with studying them in the wild, suggest this species would be an ideal candidate for development of NDS methods.

Traditional and molecular methods for sexing talpids

In all talpid species examined so far, including star-nosed moles, females have a peniform clitoris (Rubenstein et al. 2003) making external genitalia indistinguishable for most of the year, except during the reproductively active season (approximately February to June) when the vaginal seal opens and male testes increase in size (Hamilton 1931; Eadie and Hamilton 1956; Gorman and Stone 1990). Although some non-molecular sexing techniques have been used for reproductively quiescent species, such as measuring the urogenital distance in European moles (Gorman and Stone 1990) or extruding the

glans in male Pyrenean desmans (González-Esteban et al. 2003b), these methods may not work in all talpid species (e.g. broad-footed mole *Scapanus latimanus*, Rubenstein et al. 2003). These methods may also require potentially stressful physical restraint and manipulation of the animal and may be prone to error (e.g. Vidal et al. 2010).

To my knowledge, only two previous studies have designed molecular sexing methods for talpids. Sánchez et al. (1996) used a fragment of *Sry* to determine sex for three mole species (*Talpa occidentalis*, *T. romana*, and *T. europaea*). However, they failed to incorporate an internal positive control in the PCR, an important aspect of molecular sex determination tests (Lachish et al. 2011). Additionally, Sánchez et al. (1996) used primers that also amplify samples from humans and other mammal species, which could potentially introduce contamination (Waits and Paetkau 2005). Furthermore, they failed to mention what tissue was used as the DNA source in their study (Sánchez et al. 1996). Vidal et al. (2010) used gametologous sex-linked genes (*Dbx* and *Dby*) to determine sex of Pyrenean desmans. They successfully amplified and identified sex for all nondestructively collected hair samples, but only 11% of noninvasively collected fecal samples amplified, suggesting a problem related to DNA quality of these samples or protocol optimization (Vidal et al. 2010). Results of these two studies suggest a need to develop and test additional molecular sexing methods using NGS or NDS on talpid moles.

Purpose of this study

Currently, no species-specific molecular sexing test exists for the star-nosed mole. Developing a species-specific molecular sexing test using NDS techniques would

therefore facilitate further study of this virtually unstudied insectivore. My main study objectives were therefore to: 1) develop a multiplex sex determination PCR using star-nosed mole specific primers designed to amplify *Sry*, *Zfx* and *Zfy*; and 2) validate this test using nondestructive samples of feces, hair and claws.

Chapter 2: Nondestructive molecular sex determination of free-ranging star-nosed moles (Condylura cristata)

Introduction:

Molecular techniques can be used to investigate research questions related to behaviour, dietary habits and population status of a species. DNA collected from individuals can provide information such as individual identity via genotyping, relatedness among individuals, and genetic diversity between populations (Janečka et al. 2006; Walker et al. 2006; Reed et al. 2007). The sex of individuals can also be determined from molecular markers, which is critical for addressing questions about behaviour and population ecology of a species and for informing conservation and management (Bradley et al. 2008; Rodgers and Janečka 2013).

Molecular methods require collection of a tissue sample from an individual, from which DNA is then extracted and amplified in a polymerase chain reaction (PCR) for specific genes of interest. Blood and other invasive tissue samples (e.g. liver, spleen; Bryja and Konečný 2003; Shaw et al. 2003) are commonly collected from mammals as DNA sources due to their yield of high molecular weight DNA (Taberlet et al. 1999; Bengtsson et al. 2012). Collection of these types of samples often involves stressful handling and may require terminal sampling, options that are not feasible for behavioural ecology studies or studies on species of conservation concern. Additionally, many mammals are difficult to capture in the wild making it a challenge to obtain tissue samples (e.g. southern hairy-nosed wombats *Lasiorhinus latifrons*, Walker et al. 2006; Eurasian badgers *Meles meles*, Scheppers et al. 2007; felids, Rodgers and Janečka 2013).

To address these limitations, noninvasive genetic sampling (NGS) methods have been developed over the past two decades (Waits and Paetkau 2005). NGS involves collection of tissue without the animal present, allowing researchers to collect genetic information from wildlife without capture or disturbance (Taberlet et al. 1999). Shed hair and fecal samples are most commonly used as DNA sources for NGS (Waits and Paetkau 2005), as they are often left behind by an animal and can easily be collected in the field. Unlike NGS, nondestructive sampling (NDS) often involves capture and handling of the animal in order to collect a tissue sample and may involve varying levels of invasiveness (Taberlet et al. 1999). Some examples of tissues collected for NDS include feces, clipped or plucked hair, claw clippings, blood, and ear biopsy punches (Hedmark and Ellegren 2005; Durnin et al. 2007; Lachish et al. 2011). Although NDS involves more disturbance of the animal than NGS, NDS methods can be minimally stressful if tissues such as feces, hair or claws only are collected.

DNA extracted from fecal and hair samples tend to be of lower quality and quantity due to the presence of PCR inhibitors, cell degradation and low cell density (Eckhart et al. 2000; Oliveira et al. 2010; Bengtsson et al. 2012). Consequently, much effort has focused on improving amplification success rates (e.g. Broquet et al. 2007; Vallet et al. 2008; Arandjelovic et al. 2009), with optimization of the DNA extraction method, repeating DNA extractions, replicating DNA samples in PCRs and re-amplifying PCR products being the most effective (Taberlet et al. 1996; Broquet et al. 2007; Vallet et al. 2008; Arandjelovic et al. 2009; Reddy et al. 2012; Vanpé et al. 2013). Additionally, using shorter gene fragment sizes (<300 base pairs, bp; Waits and Paetkau 2005) for amplification of DNA from NGS or NDS generally improves amplification success rates

and decreases allelic dropout rates (i.e. the failure to amplify one allele or gene fragment when two or more are being amplified simultaneously) (Broquet et al. 2007; Amory et al. 2007). In order to reduce the chance of contamination, it is important to design species-specific primers to avoid erroneous results that may happen if species-specificity had not been considered (Bidon et al. 2013). This is especially important when working with fecal DNA samples from carnivores as prey DNA is also extracted in the same sample (Seddon 2005).

Along with reducing disturbance of wild mammals, NGS and NDS methods can be used to determine the sex of individuals, especially in species for which sex differentiation based on external morphology is difficult even after capture (e.g. North American beavers *Castor canadensis*, Kühn et al. 2002; shrews *Sorex* sp., Croin Michielsen 1966; talpid moles, Rubenstein et al. 2003). Some genes that have commonly been used for this purpose are Zinc finger X and Y (*Zfx* and *Zfy*) and sex-determining region Y (*Sry*). The *Zfx* allele or "gametolog" is found on the X chromosome and evolves independently of the Y chromosomal *Zfy* allele or "gametolog" (García-Moreno and Mindell 2000). The two alleles can be distinguished by length polymorphisms in the final intron of many mammalian species (Shaw et al. 2003). Primers designed to target this region of the gene allow a single PCR to simultaneously amplify both alleles from male DNA samples while only an amplification product of the X allele arises from female samples (Shaw et al. 2003). As the *Sry* gene is endemic to the Y chromosome, its presence following PCR amplification indicates the sample is from a male, while its absence is indicative of a female (Matsubara et al. 2001). A shortcoming of this method, however, is its inability to differentiate between a failed PCR amplification and a female

DNA sample. Thus, many studies have incorporated an internal positive control by co-amplifying an autosomal gene with the *Sry* fragment (e.g. Crawford JC et al. 2008; Goldberg et al. 2011; Lachish et al. 2011). Other studies have co-amplified the *Sry* locus with *Zfx* and *Zfy* (e.g. Lindsay and Belant 2008; Lanyon et al. 2009) to increase the sensitivity of the molecular sex determination test.

Although NGS and NDS methods are well established for ecological studies of large mammals (e.g. brown bears *Ursus arctos*, Proctor et al. 2010; western gorillas *Gorilla gorilla*, Bradley et al. 2008; felids, Rodgers and Janečka 2013), these techniques have seldom been used for small mammals. Indeed, NGS methods have only recently been developed for collection of hair from some small mammals (e.g. American pika *Ochotona princeps*, Henry and Russello 2011; common hamster *Cricetus cricetus*, Reiners et al. 2011), while NDS fecal sample collection generally requires trapping (e.g. rodents and shrews, Moran et al. 2008; but see Puechmaille and Petit 2007; Boston et al. 2012). Most species of the Family Talpidae exhibit fossorial, semi-fossorial or semi-aquatic habits (Shinohara et al. 2003) and are extremely challenging to observe or capture in the wild. Consequently, only a few published studies have employed NGS or NDS methods to study their biology, and these have focused exclusively on Old World talpids (Vidal et al. 2010; Igea et al. 2013).

Among the North American talpids, the star-nosed mole (*Condylura cristata*) is one of the least studied but potentially most unusual in terms of its natural history. Its range extends much farther north than any other North American mole (Petersen and Yates 1980). It appears to remain active throughout the winter, as it has been observed tunnelling under the snow and even swimming under ice in winter (Merriam 1884;

Hamilton 1931). Indeed, the star-nosed mole is an accomplished diver (McIntyre et al. 2002) that is able to exploit both subterranean and aquatic environments for its diet of invertebrates (Petersen and Yates 1980). Given the low ambient temperatures and scarcity of terrestrial invertebrate food resources during winter (Aitchison 1979; Danks 2007), the success of this small mammal (~50 g; Hamilton 1931) in harsh northern climates suggests that it must have specialized thermoregulatory and energetic adaptations. Indeed, it is known from laboratory studies that the star-nosed mole has a higher metabolic rate and body temperature compared to other mole species (Campbell et al. 1999) but little is known about any other aspect of the star-nosed mole's biology.

As in all talpids examined, female star-nosed moles possess a peniform clitoris, making them indistinguishable from males based on external morphology (Rubenstein et al. 2003). Female spotted hyenas (*Crocuta crocuta*) are better known for this masculinized genital, which is associated with a well-studied, matriarchal dominance social system with females dominant to males (Frank 1986). Conversely, virtually nothing is known about the social behaviour or mating system of star-nosed moles but this morphological similarity could indicate a similar social system in this semi-aquatic talpid. To explore these types of questions, the ability to accurately differentiate between the sexes is critical. Therefore, developing a molecular sex determination test based on NDS methods for this species would facilitate further research into the ecology of this amphibious mole.

Only two published studies employing molecular sexing methods are available for talpid mole species (*Talpa occidentalis*, *T. romana*, *T. europaea*, Sánchez et al. 1996; *Galemys pyrenaicus*, Vidal et al. 2010), with both studies exclusively using European

(Old World) species. One of these studies had limited success with noninvasively-collected fecal samples (11% success rate), while nondestructively-collected hair and invasively-collected muscle tissue samples were more successful (100%; Vidal et al. 2010). Sánchez et al. (1996) did not specify what types of tissues were used, the primers were not species-specific and no internal positive control was used to co-amplify with *Sry*. As some moles, such as the star-nosed mole, share their environment with closely related species in many parts of their distribution range (e.g. American water shrew *Sorex palustris*, northern short-tailed shrew *Blarina brevicauda*, masked shrew *Sorex cinereus*, eastern mole *Scalopus aquaticus* and hairy-tailed mole *Parascalops breweri*; Stewart et al. 2008), designing species-specific primers is important to reduce the risk of contamination in molecular sex determination tests (Bidon et al. 2013). Thus, more work is needed to develop molecular sexing tests for talpid moles, generally, and North American species in particular.

Given the absence of species-specific NGS and NDS methods for sexing star-nosed moles, my objectives were to: (1) sequence fragments of the *Sry*, *Zfx* and *Zfy* genes and develop star-nosed mole-specific primer pairs from these fragments, (2) optimize the PCR conditions of the mole-specific primer pairs for co-amplifying *Sry*, *Zfx* and *Zfy* in a multiplex PCR from star-nosed mole genomic DNA extracted from liver or kidney tissue; (3) use hair, claw clippings and feces (collected from live-trapped star-nosed moles and preserved specimens) as DNA sources to validate the multiplex PCR as a nondestructive sex determination method; and (4) identify nondestructive tissues that are most reliable for sex assignment using this approach.

To address my final objective, I compared amplification success rates, accuracy rates, number of PCR methods needed and allelic dropout rates among tissue types. Hair and feces contain PCR inhibitors (Eckhart et al. 2000; Oliveira et al. 2010), which can affect the quality of DNA extracted from these tissues. Therefore I predicted that DNA extracted from claws would result in higher amplification success rates and higher accuracy rates; require fewer PCR methods to amplify and have lower allelic dropout rates than DNA extracted from hair or fecal samples. In addition to containing melanin, a PCR inhibitor (Eckhart et al. 2000), hair shafts generally contain lower quality DNA as a result of the keratinization process when compared to hair roots (Bengtsson et al. 2012). Because I was only able to use hair shafts in my study, I also predicted that DNA from hair samples would result in lower amplification success rates and lower accuracy rates, require more PCR methods to amplify and have higher allelic dropout rates than DNA from claw or fecal samples.

Methods:

Samples and visual sex determination

I sampled tissues from 16 star-nosed moles (11 frozen carcasses and five live animals) collected from a field site near the town of Piney (N49°6.183' W095°58.679'), Whiteshell Provincial Park (N49°47.139' W095°13.383') and Nopiming Provincial Park (N50°28.209' W095°16.508'), Manitoba, and one found dead just north of Kenora, Ontario (N49°837' W094°596') by R.A. MacArthur. Five of these were frozen carcasses from a previous study (Campbell KL, pers. comm.), while five others were individuals captured in the summer of 2011 as part of this project. I thawed each carcass (stored at

either -20 or -80°C), dissected the reproductive tracts to determine sex (7 ♂, 3 ♀) and then harvested approximately 20 mg of liver or kidney tissue for DNA extraction (see below). I also collected hair, claw clippings and fecal samples from all ten frozen specimens and stored these samples at -20°C for later DNA extraction. From the carcass found in Ontario in the fall of 2011, I collected hair and claw clippings and determined the sex by PCR before confirming its sex through examination of the reproductive tract. Finally, I collected five hair and two claw clipping samples from five anaesthetized star-nosed moles of unknown sex captured with Sherman traps in 2011 and later released at their site of capture in Whiteshell and Nopiming Provincial Parks. I collected fecal samples by dissecting them from the large intestinal tract of mole carcasses, or in one case excreted from a live mole. I collected claw samples by clipping them from the tips of either the fore or hind claws, and collected hair shaft samples by clipping hair from the abdominal area of all moles. All mole specimens were collected under wildlife scientific permits (WSP 02012, WSP 03002, and WB12563) and handled in accordance with the standards of the Canadian Council on Animal Care (CCAC) following a university approved protocol (#F11-007).

Testing primers and obtaining sequence fragments

As a positive control for *Sry*, I initially extracted genomic DNA from liver or kidney tissue of two male specimens using the Qiagen DNeasy Blood and Tissue Kit (Qiagen, Inc., Toronto, ON, Canada). I then quantified the amount of DNA extracted using the NanoVue Spectrophotometer (GE Life Sciences, Piscataway, NJ, USA), and prepared 20 ng/µL working solutions for use in PCR.

To test optimal amplification conditions for genomic DNA, I initially used primers (Sry-HMG-Box-F and Sry-HMG-Box-R; Table 1) known to amplify a 202 bp fragment of the *Sry* HMG (high mobility group) box of human, rodents, bats, Algerian hedgehog (*Atelerix algirus*), soricid shrews and European mole species of the genus *Talpa* (Sánchez et al. 1996; Bullejos et al. 2000; Bryja and Konečný 2003). I carried out PCR amplifications in 20 µL reaction mixtures consisting of 200 µM (each) dNTP, 10mM Tris-HCl, 50 mM KCl, 1.5 mM $MgCl_2$, 0.5 µM SRY-HMG-Box-F, 0.5 µM SRY-HMG-Box-R, 1 unit *Taq* DNA polymerase (New England Biolabs, Ipswich, MA, USA), 20 ng template DNA and nuclease-free water to 20 µL. I used an MJ Mini™ Gradient Thermal Cycler (Bio-Rad Laboratories Ltd., Mississauga, ON, Canada) with the following thermal cycling profile: initial denaturation at 95°C for 2.5 min, 35 cycles of 95°C for 30 s, 50°C for 20 s and 68°C for 20 s, and ending with a final extension at 68°C for 5 min. These PCR conditions produced a band of expected size for the *Sry* HMG box fragment (~200 bp) along with bands of higher molecular weights after electrophoresis on a 2.5% agarose gel. I excised the target band from both male samples, purified this band using the Illustra™ GFX™ PCR DNA and Gel Band Purification Kit (GE Life Sciences, Piscataway, NJ, USA) and sequenced it (see below for details). Unfortunately, sequence reads were of low quality and contained human contamination (data not shown).

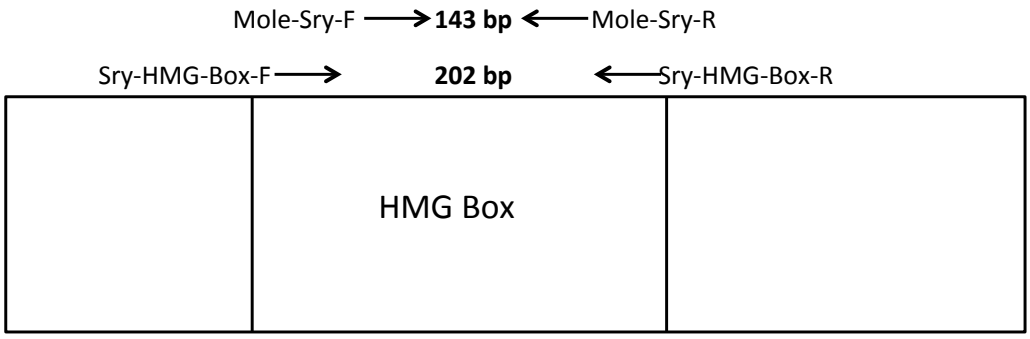
Table 1. Oligonucleotide primers used to amplify and sequence *Sry* or *Zfx* and *Zfy* genes of star-nosed moles.

Primer	Sequence (5'➔3')	Reference
Sry-HMG-Box-F	GTC AAG CGC CCC ATG AAT GCA T	Sánchez et al. 1996
Sry-HMG-Box-R	AGT TTG GGT ATT TCT CTC TGT G	Sánchez et al. 1996
Mole-Sry-F	GAG ATC AAA GGC GCA AGG	This study
Mole-Sry-R	CGC TGC CTG TAG TCT CTG TG	This study
Cc-SRY-F1	GAG ATC AAA GGC GCA AGT T	This study
Cc-SRY-R1	TTC CGA TTC CGT CAA CAT T	This study
LGL335	AGA CCT GAT TCC AGA CAG TAC CA	Cathey et al. 1998
LGL331	CAA ATC ATG CAA GGA TAG AC	Cathey et al. 1998
Cc-ZFY-F1	TTT GCT ACA ATT ACT AAC CAT	This study
Cc-ZFY-R1	ATT TCT ACT AAG ACC ACC TG	This study
Cc-ZFX-F1	CGG TAA CCA AAG CAG AAG TGT	This study
Cc-ZFY-Inter-F1	GGG TAA GTT TGC TAC AAT TAC TAA CC	This study
Cc-ZFX-R1	GGG CCA ATG ATT ATT GCT AAG	This study

I thus extracted a new DNA sample from liver tissue of a third male mole and amplified it following the same procedure as above. I purified the ~200 bp product and used it as template for two separate hemi-nested PCR amplifications using the primers from Sánchez et al. (1996) and nested primers, Mole-Sry-F and Mole-Sry-R (Table 1), designed from known talpid *Sry* HMG box sequence (Figure 1). The hemi-nested PCR consisted of 200 µM (each) dNTP, 1 X Standard One*Taq* reaction buffer, 0.2 µM SRY-HMG-Box-F (or Mole-Sry-F), 0.2 µM SRY-HMG-Box-R (or Mole-Sry-R), 1 unit One*Taq* DNA polymerase (New England Biolabs, Ipswich, MA, USA), 1 µL template (10% working solution of the purified PCR product) and nuclease-free water to 20 µL. I carried out amplification with an initial denaturation at 94°C for 30 seconds, followed by 35 cycles of 94°C for 30 s, 57°C for 15 s, and 68°C for 20 s, and ending with a final extension at 68°C for 5 min. I then electrophoresed both hemi-nested amplified products (or "amplicons") on a 2.5% agarose gel, excised these products from the gel and purified them.

I used the purified hemi-nested amplicons in four separate sequencing reactions with the BigDye® Terminator v3.1 Sequencing Kit (Applied Biosystems, Carlsbad, CA, USA), 3.0 ng DNA and the primers Sry-HMG-Box-F, Sry-HMG-Box-R, Mole-Sry-F and Mole-Sry-R. These reaction mixtures were sequenced in both directions using a four-capillary Applied Biosystems 3130 Genetic Analyzer. I used Sequencher™ software (Version 4.6, Gene Codes Corp., Ann Arbor, MI, USA) to align the sequences along with known mole *Sry* sequences (GenBank accession #: X90843, X90863 and X95595) to obtain a consensus sequence of the (partial) star-nosed mole *Sry* HMG box. I used this sequence information to design new primers, Cc-SRY-F1 and Cc-SRY-R1

Figure 1. Schematic diagram showing the intronless *Sry* gene structure, including relative position of all primers (represented by arrows) used to sequence the star-nosed mole HMG box fragment. Numbers shown in bold give the size (in bp) of the product amplified from each pair of primers.


Mole-Sry-F
143 bp
Mole-Sry-R
Sry-HMG-Box-F
202 bp
Sry-HMG-Box-R
HMG Box

(Table 1), with Primer Premier 5.0 software (PREMIER Biosoft International, Palo Alta, CA, USA), with an expected amplicon size of 101 bp (including primers).

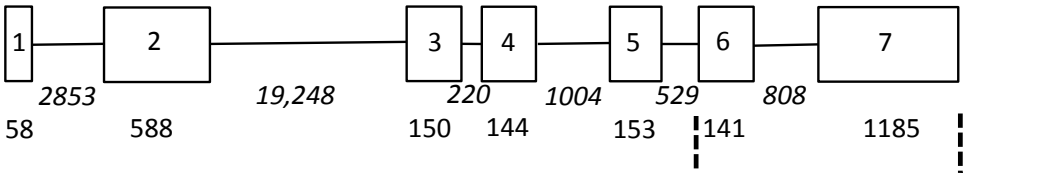
To amplify star-nosed mole *Zfx* and *Zfy* gene fragments, I used the universal LGL335 and LGL331 primers (Table 1) designed by Cathey et al. (1998), and tested on many mammalian species by Shaw et al. (2003). I first tested optimal conditions of these primers by amplifying star-nosed mole genomic DNA extracted from liver or kidney of five moles of known sex (3 males, 2 females). I performed PCR amplifications using 20 µL reaction mixtures consisting of: 200 µM (each) dNTP, 1 X Standard One*Taq* reaction buffer, 0.2 µM LGL335, 0.2 µM LGL331, 2 µg bovine serum albumin (BSA, New England Biolabs, Ipswich, MA, USA), 1 unit One*Taq* DNA polymerase, 20 ng template DNA and nuclease-free water to 20 µL. I used the following thermal cycling profile: initial denaturation at 94°C for 5 min, 55°C for 30 s, 72°C for 30 s, 30 cycles of 94°C for 30 s, 55°C for 30s and 72°C for 20 s, and ending with a final extension at 72°C for 2 min. After electrophoresis on a 1.5% agarose gel, I observed bands of the expected size for *Zfx* (892 bp) and *Zfy* (830 bp) from DNA extracted from males and the expected sized *Zfx* band only for DNA extracted from females, however many of these bands were weak.

To prepare the *Zfx* and *Zfy* amplicons for sequencing, I chose one of the above samples, extracted from kidney tissue of a male specimen, and replicated the above PCR amplification ten times on this sample. In order to ensure I would obtain enough DNA for clear sequence results, I combined and then purified the ten PCR product replicates using a GeneJET^{TM} PCR Purification Kit protocol (Thermo Scientific Molecular Biology, Ottawa, ON, Canada). Next, I electrophoresed this purified PCR product on a 1.5% agarose gel, excised both *Zfx* and *Zfy* bands from the gel and purified as above.

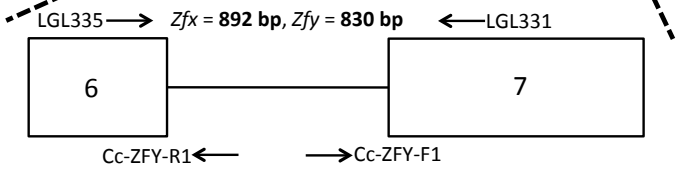
To obtain sequence information for both the purified *Zfx* and *Zfy* amplicons, I first used each in two separate sequencing reactions (for a total of four reactions) with the BigDye® Terminator v3.1 Sequencing Kit and the LGL335 and LGL331 primer pair (which had worked to amplify both *Zfx* and *Zfy* simultaneously), and sequenced each purified amplicon in both directions. To obtain cleaner sequence data for the first and last ca. 200 bp of the *Zfy* fragment, I designed new sequencing primers (Cc-ZFY-F1 and Cc-ZFY-R1; Table 1; Figure 2a and 2b) and re-sequenced the above *Zfy* amplicon. I then used Sequencher™ software (Version 4.6) to align the obtained *Zfy* sequences and determine a consensus sequence of the star-nosed mole *Zfy* gene fragment. I also used this software package to align the consensus *Zfx* sequence with a publically available genome of a female star-nosed mole (GenBank accession #: AJFV00000000.1). I first annotated the *Condylura Zfx* locus by aligning known human *ZFX* coding sequence (GenBank accession #: NM_001178085.1) to the contig (Appendix 1), revealing seven exons and six introns (Figure 2a). I then used MUSCLE software (Edgar 2004) to align the resulting star-nosed mole *Zfx* and *Zfy* sequences. Finally, I used regions of non-homology between these loci to design new primers, Cc-ZFX-F1, Cc-ZFY-Inter-F1 and Cc-ZFX-R1 (Table 1), to amplify shorter fragments (349 bp for *Zfx* and 236 bp for *Zfy*) within intron 6 of this locus.

Figure 2. a) Schematic of *Condylura cristata Zfx* gene structure. Exons are denoted by boxes and introns represented by horizontal lines. The number of each exon is shown inside the box, while the numbers below boxes and lines represent the size (in bp) of each exon and intron, respectively. b) Close-up of exons 6 and 7 illustrating relative position of primers (represented by arrows) used for sequencing. Numbers shown in bold give the size (in bp) of the *Zfx* and *Zfy* products amplified from the universal LGL335 and LGL331 primers.

a)


1
2
3
4
5
6
7
2853
19,248
220
1004
529
808
58
588
150
144
153
141
1185

b)


LGL335
Zfx = 892 bp, Zfy = 830 bp
LGL331
6
7
Cc-ZFY-R1
Cc-ZFY-F1

Testing species-specificity of the new primers and combining in a multiplex PCR

To find the optimal annealing temperature and PCR conditions, I first tested the newly designed *Sry*, *Zfx* and *Zfy* primers in separate PCRs using 10 ng/μL each of template DNA extracted from liver or kidney tissue of several known male and female star-nosed moles. To check species-specificity of these new primers, I then tested the primers in separate PCRs using one DNA sample extracted from liver or muscle tissue of each of 11 different talpid species (*Scaptonyx fusicaudus*, *Parascaptor leucura*, *Uropsilus gracilis*, *Euroscaptor longirostris*, *Scaptochirus moschatus*, *Scalopus aquaticus*, *Scapanus townsendii*, *Neurotrichus gibbsii*, *Mogera wogura*, *Scapanulus oweni*, and *Desmana moschata*) of both known (*N. gibbsii* is male; *S. fusicaudus*, *P. leucura*, and *S. townsendii* are female) and unknown sex (remainder of the talpid species listed) along with a DNA sample from kidney tissue of a known male star-nosed mole. For the *Zfx* and *Zfy* PCR, I used single 10 uL reaction mixtures consisting of: 200 μM (each) dNTP, 1 X Standard One*Taq* reaction buffer, 0.05 μM Cc_ZFX_F1, 0.075 μM Cc_ZFY_Inter-F1, 0.125 μM Cc_ZFX_R1, 0.4 μg BSA, 0.5 unit One*Taq* DNA polymerase, approximately 15 ng template DNA and 6.75 μL of nuclease-free water. I used the following thermal cycling conditions for this PCR: initial denaturation at 94°C for 3 min, 10 cycles of 94°C for 30 s, 62°C for 30 s (decreasing by 0.5°C every cycle down to 58°C), 68°C for 30 s, 30 cycles of 94°C for 30 s, 58°C for 30 s and 68°C for 30 s, and ending with a final extension at 68°C for 5 min. For the *Sry* PCR, I used single 10 uL reaction mixtures consisting of: 200 μM (each) dNTP, 1 X Standard One*Taq* reaction buffer, 0.2 μM Cc_Sry_F1, 0.2 μM Cc_Sry_R1, 1 μg BSA, 0.5 unit One*Taq* DNA polymerase, approximately 15 ng template DNA and 6.3 μL of nuclease-free water. I used the

following thermal cycling conditions for this PCR: initial denaturation at 94°C for 30 s, 35 cycles of 94°C for 30 s, 57°C for 15 s and 68°C for 20 s, and ending with a final extension at 68°C for 5 min. The lack of amplified bands from non-*Condylura* samples (and the amplification success from the DNA sample of the star-nosed mole) for both the *Sry* and the *Zfx* and *Zfy* PCR tests is consistent with the hypothesis that the primers are star-nosed mole specific.

Finally, I combined these primers into a multiplex PCR, amplifying all loci together in single 10 uL reaction mixtures consisting of: 200 µM (each) dNTP, 1 X Standard One*Taq* reaction buffer, 1.0 mM $MgCl_2$, 0.05 µM Cc_ZFX_F1, 0.075 µM Cc_ZFY_Inter-F1, 0.125 µM Cc_ZFX_R1, 0.1 µM Cc_Sry_F1, 0.1 µM Cc_Sry_R1, 0.4 µg BSA, 0.3 unit One*Taq* DNA polymerase, 10 ng template DNA (extracted from liver or kidney tissue of star-nosed moles) and 6.19 µL of nuclease-free water. For this multiplex PCR, I used the following thermal cycling conditions: initial denaturation at 94°C for 4 min, 10 cycles of 94°C for 35 s, 60°C for 35 s (decreasing by 0.5°C every cycle down to 55°C), 68°C for 35 s, 30 cycles of 94°C for 35 s, 55°C for 35 s and 68°C for 35 s, and ending with a final extension at 68°C for 5 min.

Collection and DNA extraction of nondestructive samples

I performed DNA extractions on 2 to 124 mg of fecal sample collected from each of the 10 known-sex moles using the QIAamp DNA Stool kit (Qiagen, Inc., Toronto, ON, Canada), following the manufacturer's protocol (modified to include a first and second elution at the end of the procedure). Immediately following the extraction procedure, I quantified the DNA samples with the NanoVue Spectrophotometer.

Prior to DNA extractions, I washed off any excess debris observed on the claw clippings (n = 13) or hair shaft samples (n = 16) using a 70% ethanol solution followed by a nuclease-free water rinse. I then dried these samples at 50°C and stored them at -20°C. I performed DNA extractions on ~2 mg of claw clippings and ~1.5 mg of hair shaft samples using the QIAamp DNA Investigator kit following the manufacturer's protocol for nail clippings and hair with the following modifications: After weighing the sample in a sterile 1.5 mL tube, I added 300 µL Buffer ATL to the tube and cut up the sample into the smallest pieces possible using autoclaved and ultraviolet-irradiated scissors. Next, I added 20 µL of proteinase K and 20 µL of 1 M dithiothreitol (DTT) to the tube and mixed by pipetting. I then incubated the sample overnight at 56°C and 300 rpm on a ThermoMixer® (Eppendorf Canada, Mississauga, ON, Canada). Once lysis was complete, I centrifuged the sample at 10,000 g for one minute and transferred the supernatant to a fresh 1.5 mL sterile tube. I added 300 µL Buffer AL and 1 µL of dissolved carrier RNA to the sample to increase DNA yield (Kishore et al. 2006). I then incubated the sample at 70°C and 300 rpm for approximately 15 minutes and immediately placed it in a refrigerator at 4°C for 10 minutes. Instead of adding 150 µL ethanol as the manufacturer's protocol suggests, I then added 600 µL of -20°C isopropanol to the sample to improve DNA precipitation (Hänni et al. 1995; Ojeda et al. 2012). I gently mixed the sample tube by inverting several times until the solution appeared homogeneous at which point I placed it in a -20°C freezer for a minimum of 30 minutes. Following the manufacturer's protocol, I subsequently transferred the sample solution from the tube to a QIAamp MinElute column and completed the washing steps. With some modifications from the manufacturer's protocol, I eluted DNA by adding 35

μL of nuclease-free water pre-incubated at 56°C to the QIAamp MinElute column (placed in a sterile 1.5 mL tube), incubated the column with the tube at 56°C for 2 minutes, and then centrifuged them at 14,100 g for one minute. Finally, I transferred the eluted DNA sample from the tube back onto the column membrane, placed the column back in the tube and incubated and centrifuged them a second time to increase the DNA yield of the final extracted sample (Qiagen, pers. comm.).

In order to confirm cross-contamination did not occur during the DNA extraction protocols, I completed one blank extraction (i.e. a DNA extraction with no tissue sample) simultaneously with the unknown-sex claw sample DNA extractions and one blank extraction along with the unknown-sex hair sample DNA extractions. Except for the incubation and centrifugation steps, I completed all hair and claw DNA extractions under a PCR workstation hood which had been irradiated with UV light for 30 minutes immediately prior to beginning the protocol. Immediately following completion of the DNA extractions, I quantified all samples using the NanoVue Spectrophotometer.

Testing molecular sex determination using nondestructive tissues

As with the DNA extractions, I performed set up for all PCRs on nondestructive samples under a UV-light-irradiated PCR workstation hood. I ran each of these samples in the multiplex PCR described earlier along with 1) a positive male control (kidney DNA) to verify the PCR worked and 2) a negative PCR control (no DNA) to confirm the PCR reagents were contamination-free. I also ran blank extraction samples in PCRs alongside hair or claw DNA samples as an additional negative control to check for contamination. Regardless of the amount of DNA present in a sample, I ran all samples in

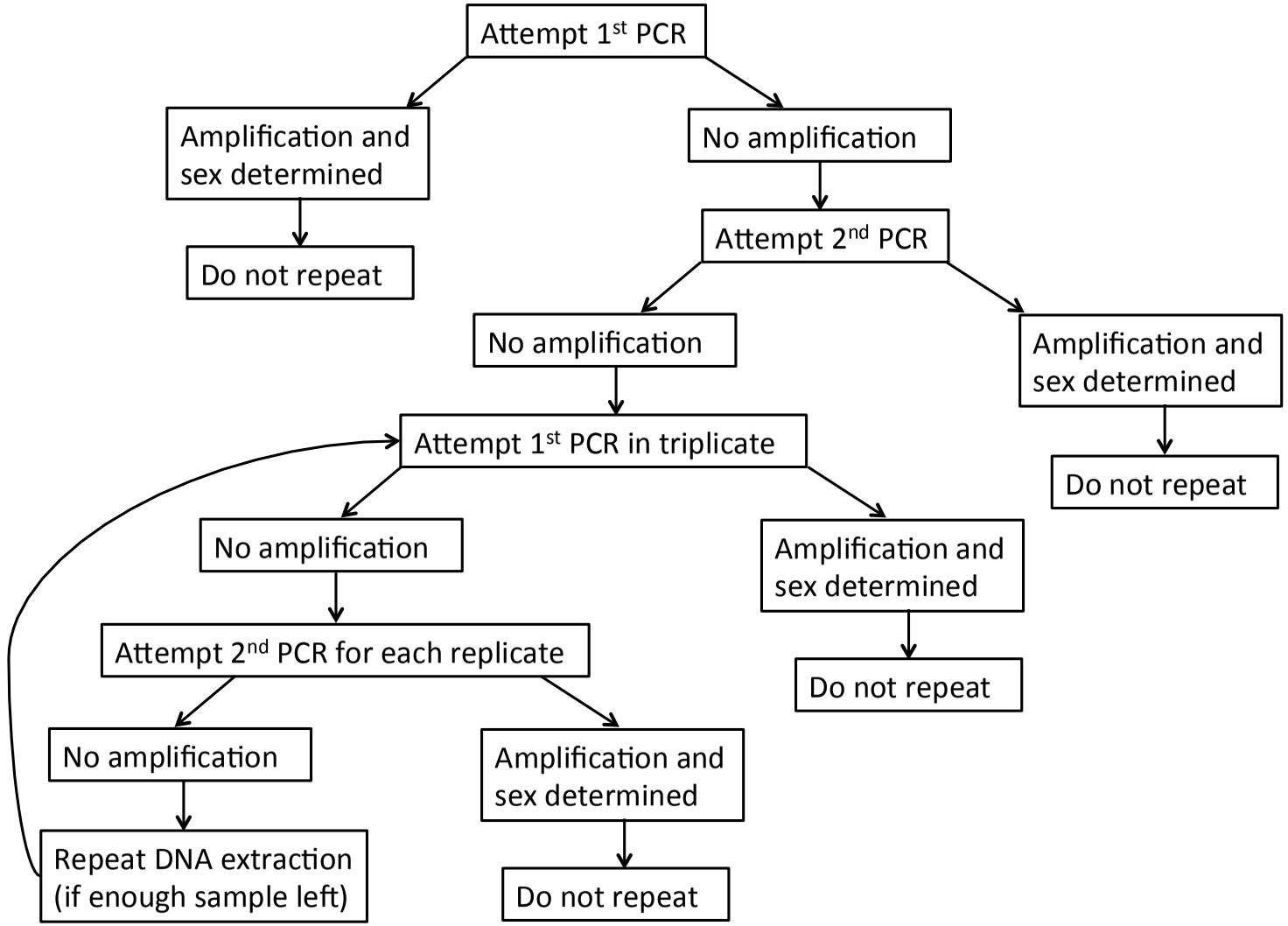
the multiplex PCR using 1 μL of sample, except for two fecal DNA samples for which I used 2 μL of sample due to their low DNA concentrations (<2.5 ng/μL). I added approximately 15 μL of mineral oil to each PCR tube to prevent evaporation of the contents during thermal cycling and kept all other conditions the same. For re-amplifications, or "second PCRs", I took 0.5 μL of the first PCR product and used it as template in the second reaction. These PCRs used the same thermal cycling conditions as outlined above except that the second set of cycles was reduced from 30 to 15 for a total of 25 cycles overall (as opposed to the original total of 40 cycles) to minimize the risk of unspecific amplification in the re-amplified product. I electrophoresed all PCR products on 2.5% agarose gels and photographed these gels with the VersaDoc Imaging System (Bio-Rad Laboratories Ltd., Mississauga, ON, Canada).

To optimize the amplification success and help ensure the correct sex assignment for each sample, I tested a series of PCR methods (Taberlet et al. 1996; Arandjelovic et al. 2009; Vanpé et al. 2013; Figure 3). Initially, I amplified each sample once in the multiplex PCR. If amplification occurred and I could determine the sex (as defined by at least one of the Y fragments amplifying for males and only the *Zfx* fragment amplifying for females) then I did not conduct any further PCRs on that sample (Figure 3). If no amplification products were observed, I attempted a re-amplification using the first PCR product of that sample (as described above). If the amplification again was unsuccessful, then I repeated the initial multiplex PCR for that sample in triplicate (Figure 3). If no amplification occurred for any of the three replicates, then I attempted the second PCR for each of the three replicates (Figure 3). If there were still no successful amplifications after completing all of these steps, if possible (i.e., if enough sample was left) I repeated

the DNA extraction. After completing the second DNA extraction, I would test this sample beginning with the PCR amplification in triplicate as above (Figure 3). Except for a couple of samples for which I only had enough tissue collected to complete one DNA extraction, I was able to repeat two to three DNA extractions (when necessary), which was enough to achieve amplification success.

First, I tested these methods on the nondestructive samples extracted from the 10 moles of known sex and samples obtained from the frozen Ontario mole carcass. After molecular sex determination of the Ontario mole (male), I dissected this mole's reproductive tract to confirm the PCR results. Finally, I applied these multiplex PCR amplification methods to the hair and claw DNA samples collected and extracted from the five live-captured moles of unknown sex to determine sex.

Figure 3. Flow chart showing the series of PCR methods used on nondestructive fecal, hair and claw samples collected from star-nosed moles to achieve amplification success.


Attempt 1st PCR
Amplification and sex determined
Do not repeat
No amplification
Attempt 2nd PCR
No amplification
Amplification and sex determined
Do not repeat
Attempt 1st PCR in triplicate
No amplification
Amplification and sex determined
Do not repeat
Attempt 2nd PCR for each replicate
No amplification
Repeat DNA extraction (if enough sample left)
Amplification and sex determined
Do not repeat

Analysis of sex determination results

I used generalized estimating equations (GEE) within generalized linear models, with a binomial distribution (logit link function), to perform separate statistical tests investigating the effect of the NDS tissue source (pairwise comparisons of hair vs. claw vs. feces) of DNA on amplification success, correct sex assignment (i.e. accuracy), number of PCR methods needed and allelic dropout rate of *Zfx* using the SAS 9.3 software program (SAS Institute Inc., Cary, NC, USA). Use of the GEE analysis allowed me to control for lack of independence between samples from different tissue sources collected from the same individual (Koper and Manseau 2012). I completed these analyses combining results from samples collected from both known- and unknown-sex moles.

I defined amplification success as an observation of at least one of the *Zfx*, *Zfy* or *Sry* products as a band in the gel photograph regardless of the number of PCR methods needed to obtain successful amplification for a sample (n = 39). Using only successfully amplified sample results (n = 37), I concluded that I had assigned sex correctly when I observed at least one of the Y fragments for samples from known males (or males determined as such by the multiplex PCR) and only the *Zfx* fragment from known females (or females determined as such by the multiplex PCR). Sex was assigned based on the first successful PCR amplification of a sample. If I was assessing the sex from a sample amplified in triplicate (n = 7 out of 37 samples), two out of the three results had to match for the same sex and be consistent with the sex of the carcass or the sex assigned using other tissues sources from the same mole if available in order for me to conclude I had assigned sex correctly.

I assessed the quality of different tissues for sex assignment analysis by first comparing the effort required (i.e., the number of PCRs needed) to amplify the necessary markers relative to which tissue was used and, second, by comparing the amplification success rate of the largest marker (*Zfx*) relative to which tissue was used. From the subset of samples I was able to assign correctly (n = 34), I categorized samples that amplified after only the first PCR as "high quality". I defined samples that amplified after two or more PCRs as "low quality". Using only samples from males that amplified successfully (n = 26), I categorized samples that amplified the *Zfx* band in addition to the smaller size Y bands as "high quality". I assigned samples that amplified at least one of the Y bands and did not amplify the *Zfx* band (i.e. allelic dropout of *Zfx*) as "low quality"; likely indicating degraded DNA resulting in shorter fragment sizes (Lindsay and Belant 2008).

I used an alpha level of 0.05 to test for significant effects between tissues on amplification success, correct sex assignment, number of PCR methods needed and allelic dropout rate of *Zfx*. Data are presented as the log odds ratio estimates with 95% confidence intervals for each pairwise comparison.

Results:

Sequencing and new primers

I obtained a 158 bp sequence fragment of the male star-nosed mole *Sry* HMG box (Figure 4). This sequence had 17 nucleotide polymorphisms (substitutions) when compared to *Talpa Sry* and Soricidae *Sry* sequences (Figure 5). The primers I designed from the star-nosed mole *Sry* sequence (Cc-SRY-F1 and Cc-SRY-R1) amplified a 101 bp fragment for males while no amplification product occurred for females. After testing

Figure 4. Sequenced fragment (158 bp) of the *Condylura cristata Sry* gene with star-nosed mole specific primers, Cc-SRY-F1 and Cc-SRY-R1, indicated in bold, italics and underlining.

Cc-SRY-F1 ***G A G A I C A A A G G C G C A A G I I***

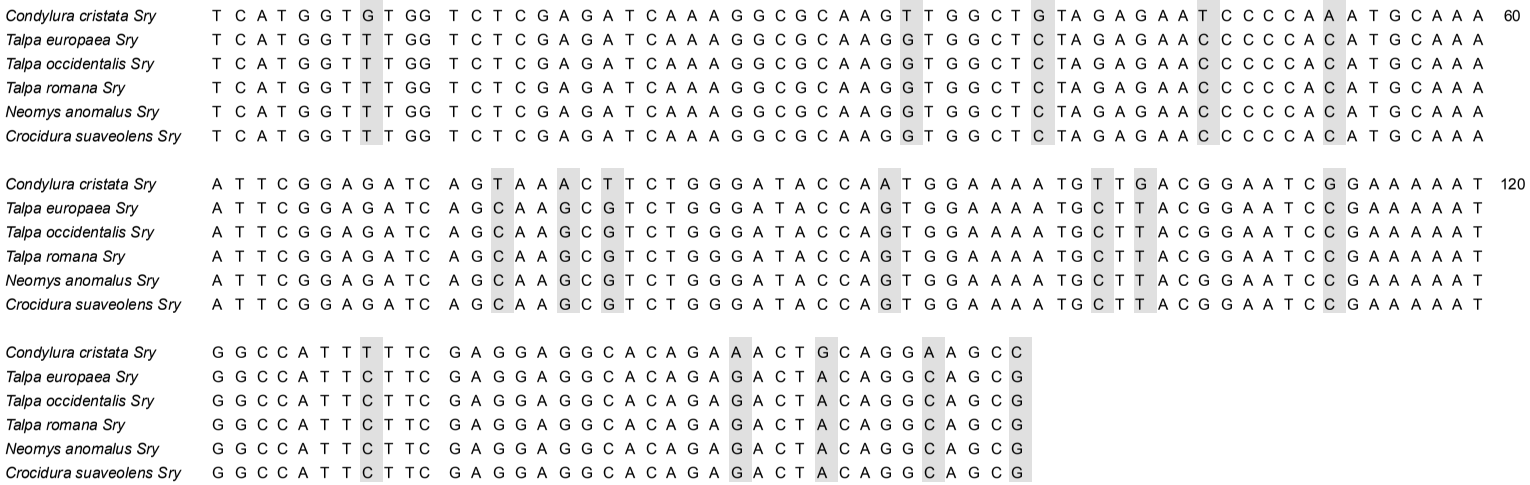
Condylura cristata Sry T C A T G G T G T G G T C T C G A G A T C A A A G G C G C A A G T T G G C T G T A G A G A A T C C C C A A A T G C A A A 60

Cc-SRY-R1 ***A A I G I I G A C G G A A I C G G A A***

Condylura cristata Sry A T T C G G A G A T C A G T A A A C T T C T G G G A T A C C A A T G G A A A A T G T T G A C G G A A T C G G A A A A A T 120

Condylura cristata Sry G G C C A T T T T T C G A G G A G G C A C A G A A A C T G C A G G A A G C C

Figure 5. Multiple sequence alignment of the *Sry* HMG box from *Condylura cristata*, *Talpa europaea*, *T. occidentalis*, *T. romana*, and the Soricidae species *Neomys anomalus* and *Crocidura suaveolens* (Sánchez et al. 1996). Nucleotides highlighted by grey shading show nucleotide differences between *C. cristata* and the other species.


Condylura cristata Sry
Talpa europaea Sry
Talpa occidentalis Sry
Talpa romana Sry
Neomys anomalus Sry
Crocidura suaveolens Sry
60
120

these primers on DNA samples from other talpid species (see methods above), amplification results were negative for all samples except for the positive control DNA sample from a male star-nosed mole (data not shown), suggesting that the primers are star-nosed mole specific.

I obtained an 822 bp sequence fragment of the star-nosed mole *Zfx* locus (Figure 6) that precisely matched the corresponding region in the female star-nosed mole genome. This sequence covered most of the sixth intron and the first 23 bp of the seventh exon of the *Zfx* gene (Appendix 1). I also obtained 830 bp of the star-nosed mole *Zfy* gene sequence (Figure 7). Alignment with the corresponding region of the *Zfx* locus revealed that intron 6 of the *Zfy* locus (746 bp) was shorter than *Zfx* (808 bp) by 62 bp, due to multiple indel sites within the intron for both alleles (Figure 8). I designed the forward *Zfx* primer (Cc-ZFX-F1) at a site where there were four nucleotide polymorphisms and two nucleotide insertions in the aligned *Zfy* sequence (Figure 8). I designed the forward *Zfy* primer (Cc-ZFY-Inter-F1) at a site where there were seven nucleotide deletions and five nucleotide polymorphisms in the aligned *Zfx* sequence (Figure 8). I designed the reverse primer (Cc-ZFX-R1) at a site where the sequence was identical for both *Zfx* and *Zfy* (Figure 8). When I combined these primers in PCRs to amplify star-nosed mole DNA, a 349 bp fragment representing *Zfx* and a 236 bp fragment representing *Zfy* were amplified from male DNA, while only the 349 bp fragment for *Zfx* was amplified from female DNA. As with the *Sry* results (above), all amplification reactions with other talpid mole species were negative (data not shown), suggesting that these primers are also star-nosed mole specific.

After I combined all of the above primers in a multiplex PCR, amplification of all three fragments occurred for DNA samples extracted from liver or kidney samples of known males (*Zfx*, *Zfy* and *Sry*) and amplification of only the *Zfx* fragment occurred for DNA samples extracted from liver samples of known females (Figure 9).

Figure 6. Sequenced fragment (822 bp) of the *Condylura cristata Zfx* gene.

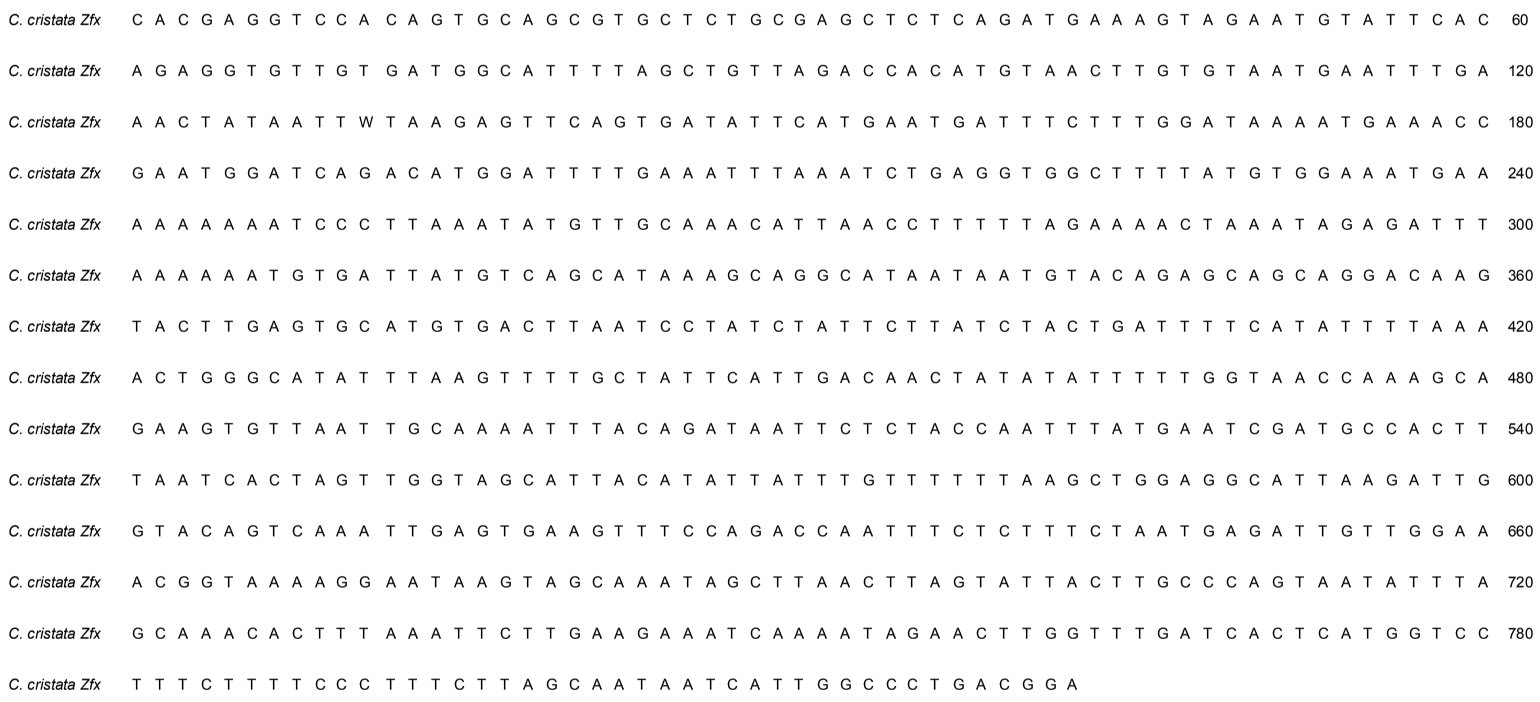

C. cristata Zfx C A C G A G G T C C A C A G T G C A G C G T G C T C T G C G A G C T C T C A G A T G A A A G T A G A A T G T A T T C A C 60
C. cristata Zfx A G A G G T G T T G T G A T G G C A T T T T A G C T G T T A G A C C A C A T G T A A C T T G T G T A A T G A A T T T G A 120
C. cristata Zfx A A C T A T A A T T W T A A G A G T T C A G T G A T A T T C A T G A A T G A T T T C T T T G G A T A A A A T G A A A C C 180
C. cristata Zfx G A A T G G A T C A G A C A T G G A T T T T G A A A T T T A A A T C T G A G G T G G C T T T T A T G T G G A A A T G A A 240
C. cristata Zfx A A A A A A A T C C C T T A A A T A T G T T G C A A A C A T T A A C C T T T T T T A G A A A A C T A A A T A G A G A T T T 300
C. cristata Zfx A A A A A A T G T G A T T A T G T C A G C A T A A A G C A G G C A T A A T A A T G T A C A G A G C A G C A G G A C A A G 360
C. cristata Zfx T A C T T G A G T G C A T G T G A C T T A A T C C T A T C T A T T C T T A T C T A C T G A T T T T C A T A T T T T T A A A 420
C. cristata Zfx A C T G G G C A T A T T T A A G T T T T G C T A T T C A T T G A C A A C T A T A T A T T T T T T G G T A A C C A A A G C A 480
C. cristata Zfx G A A G T G T T A A T T G C A A A A T T T A C A G A T A A T T C T C T A C C A A T T T A T G A A T C G A T G C C A C T T 540
C. cristata Zfx T A A T C A C T A G T T G G T A G C A T T A C A T A T T A T T T G T T T T T T A A G C T G G A G G C A T T A A G A T T G 600
C. cristata Zfx G T A C A G T C A A A T T G A G T G A A G T T T C C A G A C C A A T T T C T C T T T C T A A T G A G A T T G T T G G A A 660
C. cristata Zfx A C G G T A A A A G G A A T A A G T A G C A A A T A G C T T A A C T T A G T A T T A C T T G C C C A G T A A T A T T T A 720
C. cristata Zfx G C A A A C A C T T T A A A T T C T T G A A G A A A T C A A A A T A G A A C T T G G T T T G A T C A C T C A T G G T C C 780
C. cristata Zfx T T T C T T T T C C C T T T C T T A G C A A T A A T C A T T G G C C C T G A C G G A

Figure 7. Sequenced fragment (830 bp) of the *Condylura cristata Zfy* gene.

C. cristata Zfy T A G A C C T G A T T C C A G A C A G T A C C A A A C A G G T G A G G G C A C A T G A G T T C A C A G T G C A G C A T G 60

C. cristata Zfy T T C T G T G A G C T C T C A G A T G A A A G T A A T A T G T A T T C A C A G A G G T G T T A T G A T A G C A T T T T A 120

C. cristata Zfy A T T G C T A G A C C A T A T G T A A C T T G T G T A T T G A A T T T G A A A T T A T A A T T T T A A G A A T G C A G T 180

C. cristata Zfy G A T A T T C A T G A A T G A T T T C C T T C A A T A A A A G A A A A T A A G T T T T G A A T T A A C C A C T T A G G A 240

C. cristata Zfy A A A A T A A T T A A A A T C C A G G T G G T C T T A G T A G A A A T A A A A T C C T A A A A C A A A G T T T A A A T G 300

C. cristata Zfy A A G A A C A A A T T T T T T T T C C C A A C T A A A G C T A G T G T T T A T A A A G C A C A A T A T A A A T A C T T A 360

C. cristata Zfy A A C T G T T C A T T G G A C A T T T T C T T T T T T T T T C C T A G T C T T T T G T A T T C A T T T G C A C A C T A A G 420

C. cristata Zfy C T T C T T A G A T A T A T G A A T G A A C T A T A T G C T T A T T T T G A C T T G G C A T T A C A A A A G C T G C A G 480

C. cristata Zfy T G T T G A A T T T A T C C T T A T G A A T C T G T C C T A T A G T A C C T G A A T G T C T T G G C T T T G A G A T C A 540

C. cristata Zfy C T A T T T T A A A G A A G A G A A T A A G T T T G C T A C A A T T A C T A A C C A T A T T A A A T A A G G T C C T A A 600

C. cristata Zfy G A T A T A A T T T C T T C A C T G T T A G A T T T T T A T G T C T A C T C A A A A A C A G T A A A A G A A G T G C T A 660

C. cristata Zfy A A G A A C C T C A T C T A G G C A T A A A T G T G C A G T A G T A G T C A A A A C A A A C T T T T T A C A T T C C T A 720

C. cristata Zfy A A G G A A C C A A A A T A T T T A G T T T G A T C A C A C A T G C T C C T T T C T T T T C T T T T C T T A G C A A T A 780

C. cristata Zfy A T C A T T G G C C C T G A T G G A C A T C C C T T G A C T G T C T A T C C T T G C A T G A T T T G

Figure 8. Alignment of *Condylura cristata Zfy* and *Zfx* sequence fragments with star-nosed mole specific primers, Cc-ZFX-F1, Cc-ZFY-Inter-F1 and Cc-ZFX-R1 (indicated in bold, italics and underlining). Nucleotides highlighted in grey shading show nucleotide polymorphisms (substitutions) between *Zfy* and *Zfx* while dashes indicate nucleotide indels. Exons six and seven are outlined by boxes.

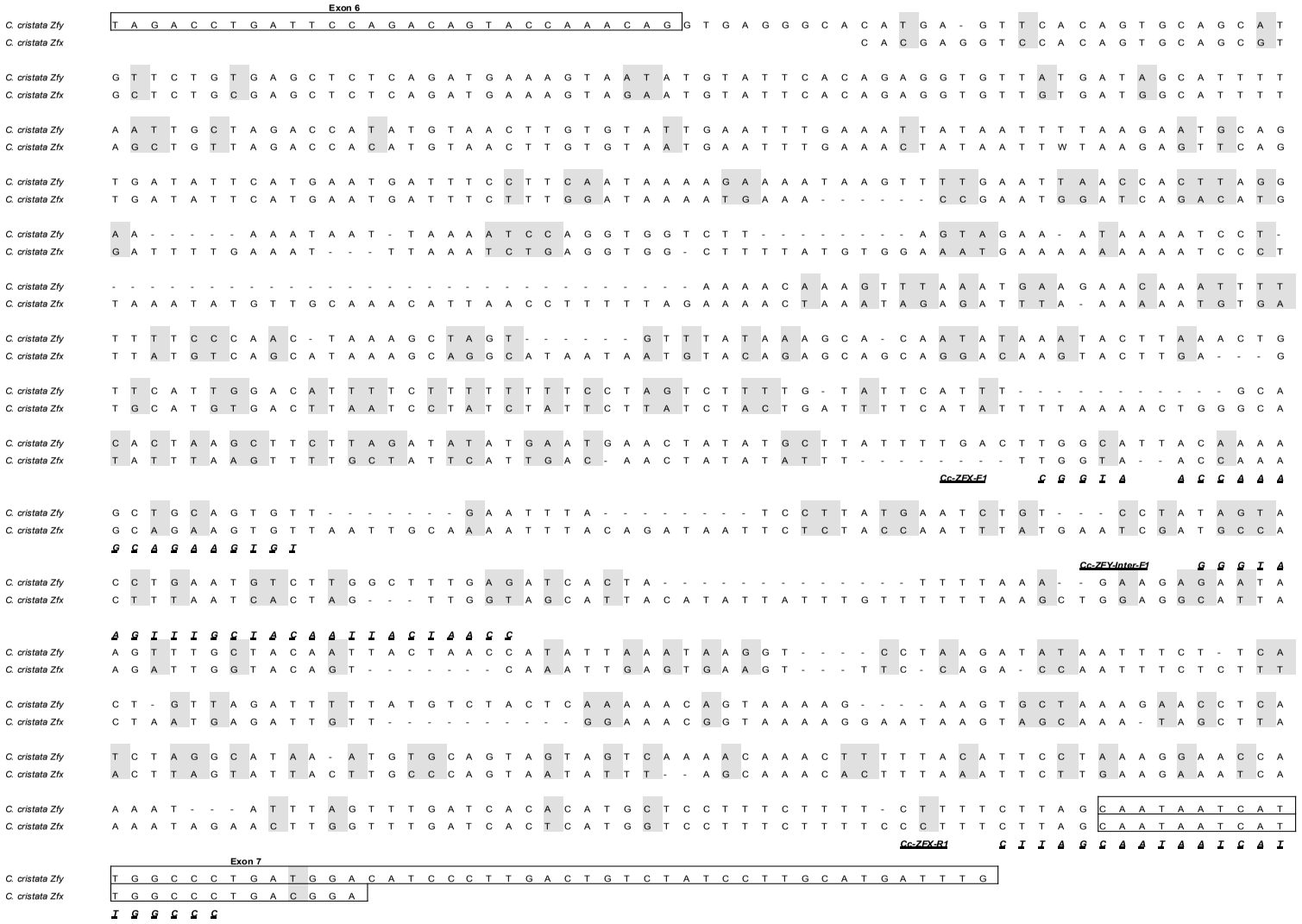
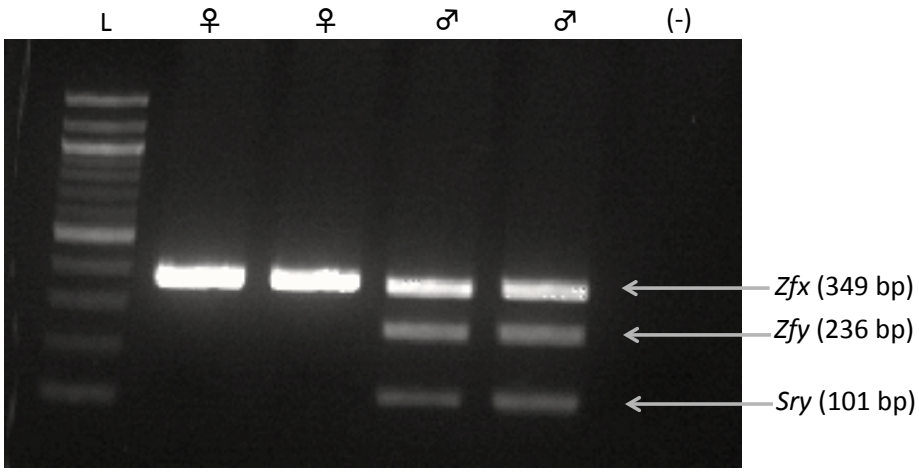

Exon 6
C. cristata Zfy
C. cristata Zfx
Cc-ZFX-F1
Cc-ZFY-Inter-F1
Cc-ZFX-R1
Exon 7

Figure 9. Agarose gel showing amplification of the *Zfx* fragment (349 bp) from two female (♀) star-nosed moles and co-amplification of the *Zfx*, *Zfy* (236 bp) and *Sry* (101 bp) fragments from two male (♂) star-nosed moles. The ladder (L) starts at 100 bp. Negative PCR control (-) also shown.


L
♀
♀
♂
♂
(-)
Zfx (349 bp)
Zfy (236 bp)
Sry (101 bp)

Nondestructive molecular sex determination

I found no significant difference (Table 2) between amplification success rates for DNA from claw vs. fecal samples (Figure 10). I was not able to statistically compare the amplification success rates for DNA from claw and fecal samples to that from hair samples because the variance for hair was 0 (Figure 10). However all tissues clearly had comparable amplification success rates.

Similarly, I found no significant difference (Table 3) between the proportions of hair vs. claw samples for which sex was assigned correctly (i.e. accuracy rate). I could not statistically compare these proportions to results for fecal samples because there was no variance for fecal samples (Figure 11). Incorrect sex assignments for hair (2/16) and claw (1/12) samples occurred because at least one of the Y fragments amplified in a sample from a known female mole. I did not find these amplification products in negative PCR controls or blank extracts run along with any of these samples.

Of the correctly assigned samples, I found that DNA samples of both claw and feces were significantly more likely to be of "high quality" (i.e., amplified in their first PCR) than DNA samples from hair (Table 4, Figure 12). I found no significant difference (Table 4) between the proportions of "high quality" DNA from claw vs. fecal samples (Figure 12). One fecal sample and one claw sample failed to amplify after repeated PCRs from the first DNA extraction and I did not have enough sample left to repeat extractions.

Of the successfully amplified samples from known males, I found no significant differences (Table 5) between the proportions of fecal, claw, or hair samples resulting in "low quality" DNA as evidenced by allelic dropout of *Zfx* (Figure 13). Although these

Table 2. Statistical comparison between amplification success rates for DNA from claw (n=13) vs. fecal (n=10) samples of star-nosed moles as determined by generalized estimating equations.

Tissue comparison	Log odds ratio estimate	95% Confidence interval		*P*-value
		Lower	Upper	
Claw-Feces	0.2877	-2.6157	3.1911	0.8460

Figure 10. Amplification success rates (+ 1 SD) of DNA samples extracted from fecal, claw and hair tissues of star-nosed moles.

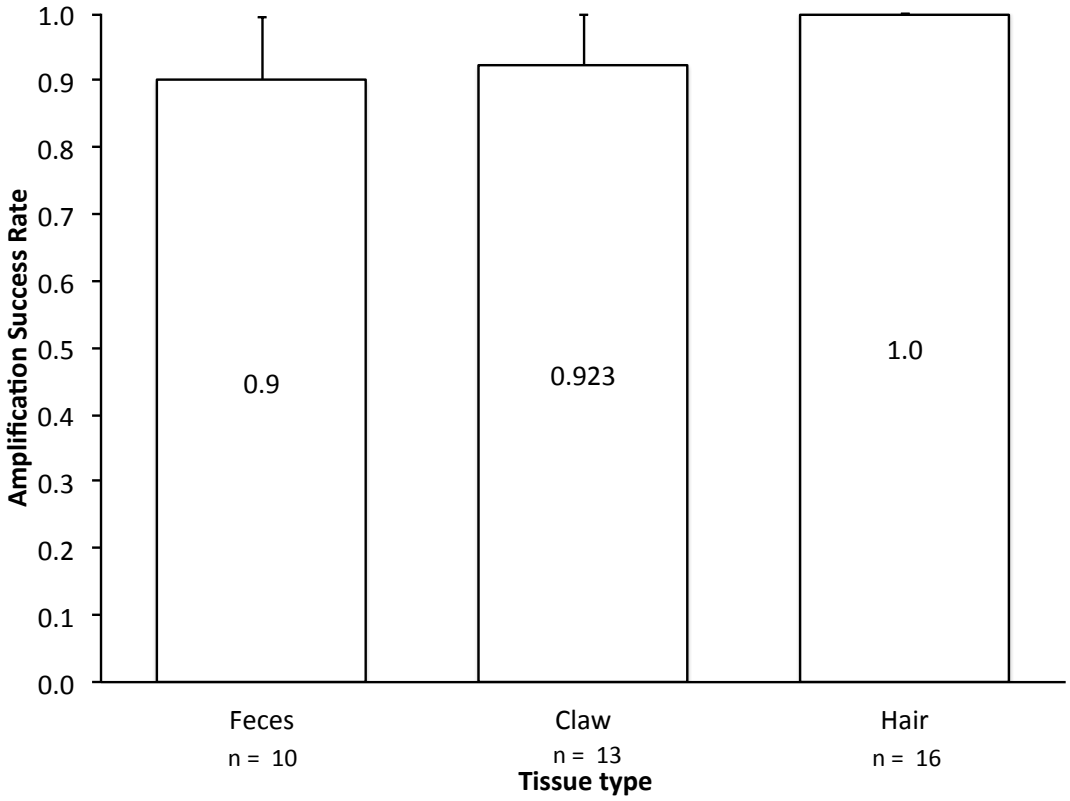

1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Amplification Success Rate
0.9
0.923
1.0
Feces
n = 10
Claw
n = 13
Hair
n = 16
Tissue type

Table 3. Statistical comparison between the proportions of DNA from claw (n=12) vs. hair (n=16) samples of star-nosed moles assigned correct sex as determined by generalized estimating equations.

Tissue comparison	Log odds ratio estimate	95% Confidence interval		*P*-value
		Lower	Upper	
Claw-Hair	0.4913	-2.1958	3.1785	0.7201

Figure 11. Proportion of DNA samples extracted from fecal, claw and hair tissues of star-nosed moles assigned sex correctly (+ 1 SD).

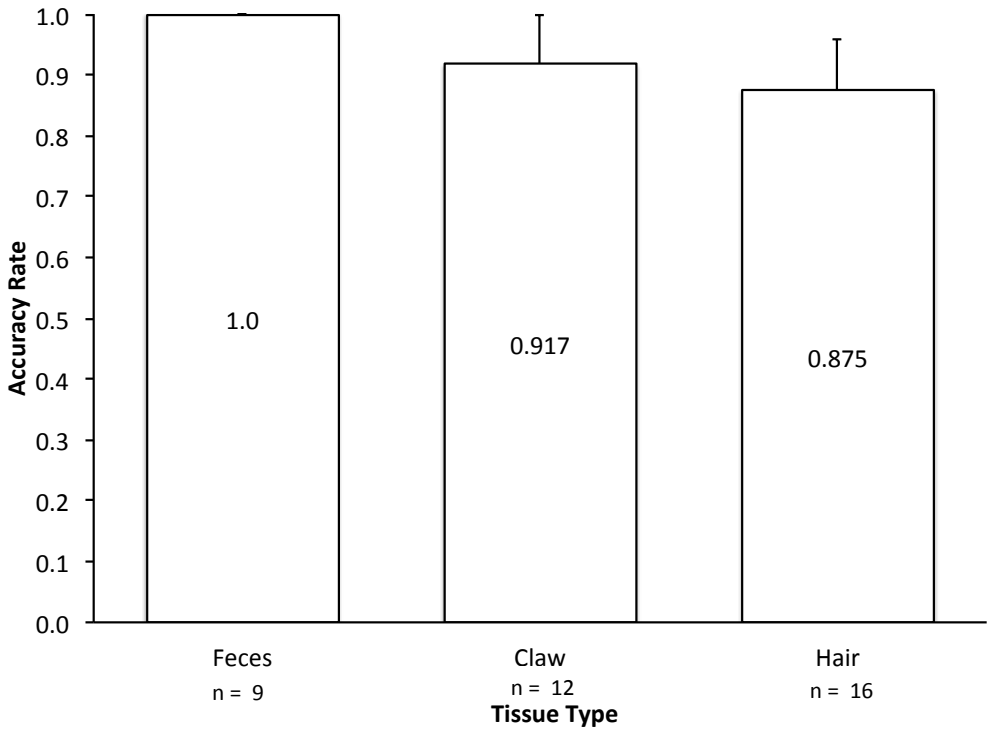

1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Accuracy Rate
1.0
0.917
0.875
Feces
n = 9
Claw
n = 12
Hair
n = 16
Tissue Type

Table 4. Statistical comparisons between first PCR amplification success rates of DNA from claw (n=11), fecal (n=9) and hair (n=14) samples of star-nosed moles as determined by generalized estimating equations.

Tissue comparison	Log odds ratio estimate	95% Confidence interval		*P*-value
		Lower	Upper	
Claw-Feces	-0.3765	-1.7239	0.971	0.5840
Claw-Hair	1.5285	0.2300	2.8270	0.0210
Feces-Hair	1.9050	0.4095	3.4005	0.0125

Figure 12. Amplification success rates (+ 1 SD) of DNA extracted from fecal, claw and hair samples of star-nosed moles in the first PCR trials. Asterisks denote significant differences between tissue types.

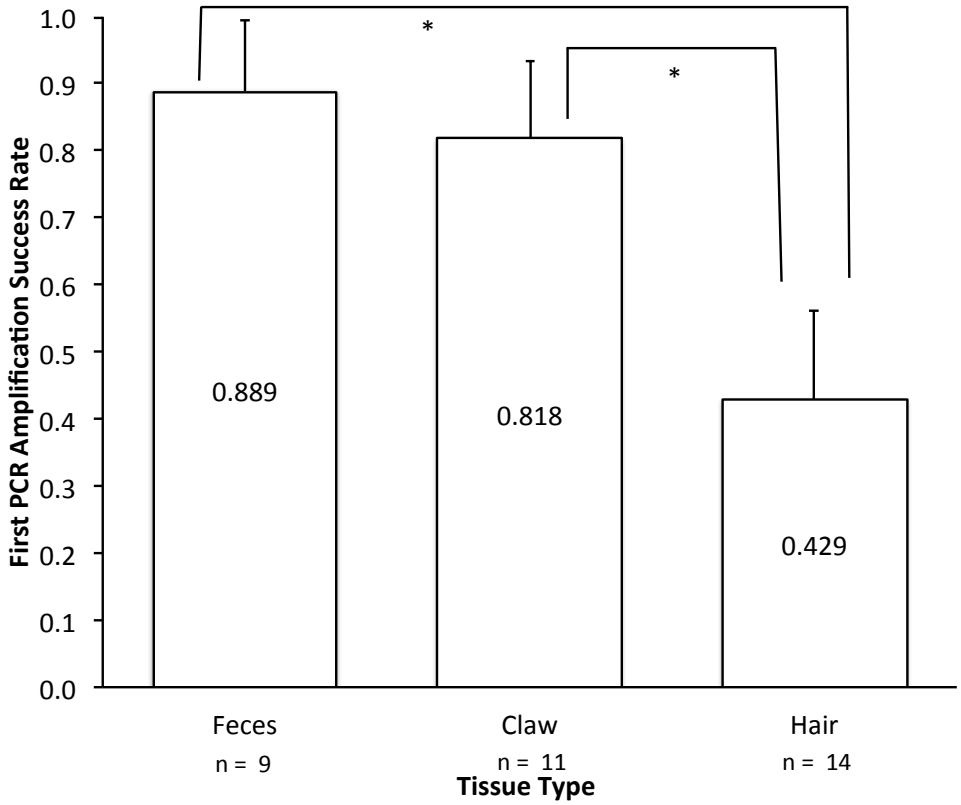
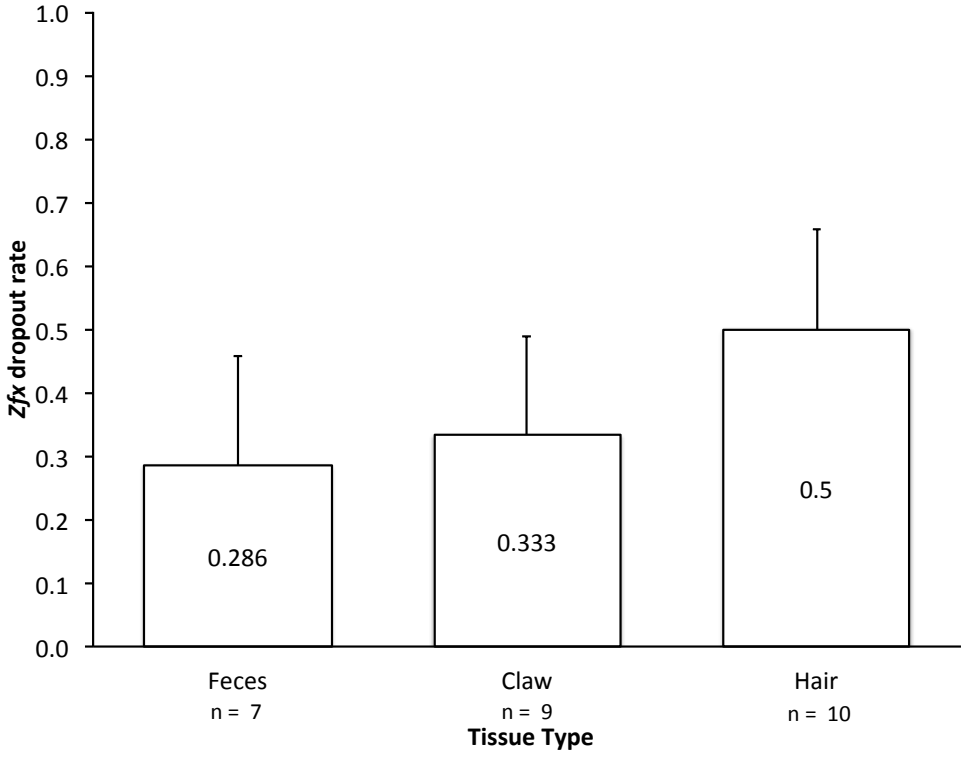

1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
First PCR Amplification Success Rate
*
*
0.889
0.818
0.429
Feces
n = 9
Claw
n = 11
Hair
n = 14
Tissue Type

Table 5. Statistical comparisons between *Zfx* dropout rates of DNA from claw (n=9), fecal (n=7) and hair (n=10) samples of star-nosed moles as determined by generalized estimating equations.

Tissue comparison	Log odds ratio estimate	95% Confidence interval		*P*-value
		Lower	Upper	
Claw-Feces	0.2001	-1.9428	2.3431	0.8548
Claw-Hair	-0.7679	-2.5068	0.9709	0.3867
Feces-Hair	-0.9681	-2.295	0.3588	0.1527

Figure 13. *Zfx* dropout rates (+ 1 SD) of DNA samples extracted from fecal, claw and hair tissues of star-nosed moles.


1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Zfx dropout rate
0.286
0.333
0.5
Feces
n = 7
Claw
n = 9
Hair
n = 10
Tissue Type

samples failed to amplify *Zfx*, they all successfully amplified *Sry* and some amplified both *Sry* and *Zfy*.

Using DNA extracted from all three nondestructive tissues collected from the 11 frozen carcasses, I correctly identified all males (n = 8) in the multiplex PCR. Of the three female samples, I correctly identified one from the PCR results of all three nondestructive tissue types, the second from fecal and claw DNA samples (the hair sample result was contaminated), and the third from only the hair DNA sample (the claw DNA sample was contaminated while the fecal DNA sample failed to amplify).

The DNA I extracted from one of two claw samples successfully amplified allowing me to determine the sex for one of two live moles of unknown sex. The DNA I extracted from four out of five hair samples also successfully amplified and I was able to determine the sex for four of the five moles of unknown sex. The PCR results from the fifth hair sample were inconsistent when run in triplicate. The DNA I extracted from both claw and hair samples of one mole amplified showing the same sex. The DNA I extracted from both claw and hair samples of a second mole resulted in amplification failure from the claw sample but I was able to determine sex from DNA extracted from the hair sample.

Discussion:

My study supports the application of NDS for the molecular sex determination of star-nosed moles. I designed and validated use of a multiplex PCR that co-amplifies *Sry*, *Zfx* and *Zfy* loci and demonstrated that this method works reliably on DNA extracted from NDS of hair, claw and fecal tissue. Nondestructive tissues did not vary in terms of

amplification success, accuracy, or allelic dropout rates for *Zfx*, although DNA extracted from hair required repeated PCRs and DNA extractions more often to achieve successful amplification.

Sequencing and new primers

To my knowledge, this is the first study to obtain a sequence fragment of the *Sry* locus from a New World talpid species. This sequence differed from the known *Talpa Sry* sequence (Sánchez et al. 1996) within the same HMG box region, exhibiting about 89% sequence similarity (Figure 5). This similarity is consistent with what is known about the conserved nature of the *Sry* HMG box across mammalian species (Nishida et al. 2003; Yu et al. 2011). Surprisingly, the published shrew *Sry* sequences precisely matched those of *Talpa* (Sánchez et al. 1996; Figure 5). This suggests that the shrew sequences may have been contaminated with *Talpa* DNA (or vice versa). The Sánchez et al. (1996) study used universal primers that worked for DNA from a variety of species, therefore increasing the risk of contamination between DNA samples from different species. To avoid this problem, I designed new primers from the *Condylura Sry* sequence and checked that they were species-specific.

This study is also the first to sequence a portion of the *Zfy* gene for any insectivore and to determine the exon-intron structure of *Zfx* for talpid moles. After aligning the *Zfx* and *Zfy* sequence fragments, I identified a length polymorphism within intron 6 of these loci, similar to *Zfx* and *Zfy* in other mammals (Shaw et al. 2003, Williams et al. 2004, Han et al. 2010). The original primers I used to obtain these sequences (LGL335 and LGL331) amplified large fragments (892 bp for *Zfx* and 830 bp

for *Zfy*) and worked for a variety of mammalian species (Shaw et al. 2003). I therefore designed new primers from these sequences to amplify much shorter products (349 bp for *Zfx* and 236 bp for *Zfy*) to increase amplification success rates from the lower quality DNA samples typically extracted from nondestructive tissue samples (Durnin et al. 2007; Lindsay and Belant 2008). Additionally, I checked that these primers were star-nosed mole-specific to avoid future contamination problems with DNA from other species.

My results illustrate the benefit of using two genes, rather than just one, for molecular sex determination tests. When the species-specific primers I designed were combined in a multiplex PCR and tested on high quality DNA samples (extracted from liver or kidney tissue) all three expected fragments (*Zfx*, *Zfy* and *Sry*) amplified for males while only one fragment (*Zfx*) amplified for females. A sex determination test relying on *Sry* alone could lead to misidentification of males as females if, for example, a PCR failure occurred (Robertson and Gemmell 2006). The use of two Y-specific markers provides both verification of amplification and confirmation of sex, making this a robust sex determination test (Gokulakrishnan et al. 2012; Bidon et al. 2013), which is an especially important consideration when using low quality DNA samples.

Nondestructive molecular sex determination

Contrary to what I predicted, I found that hair, claw and fecal samples had comparably high amplification success rates (90 – 100%). Indeed, I found that the amplification success rate reported here for fecal DNA (90%; Figure 10) is at the high end of the wide range published in the literature (11%, Vidal et al. 2010 to 100%, Moran et al. 2008). The only other study that measured amplification success from fecal DNA

samples of talpid mole species (Pyrenean desman) achieved a very low amplification success rate (11%; Vidal et al. 2010). Factors such as the DNA extraction method and freshness of the sample when collected can affect amplification success of fecal DNA (Vallet et al. 2008; Reddy et al. 2012; Barbosa et al. 2013). Vidal et al. (2010) suggested that they might need to do more frequent field collections of feces in order to obtain fresher fecal samples that yield higher quality DNA and optimize their protocol to improve the success rate of fecal DNA.

As with the fecal DNA results, the amplification success rate for hair DNA samples in my study (100%; Figure 10) is at the high end of literature values (15 – 100%; Müller et al. 2007; Opel et al. 2008; Amory et al. 2007; Vidal et al. 2010). A comparably high success rate (100%) was also obtained from Pyrenean desman hair samples (Vidal et al. 2010). Similar to fecal and hair DNA samples, the amplification success rate of the claw DNA samples in my study (92%; Figure 10) was at the high end of published values (0 – 98%; Anderson et al. 1999; Hedmark and Ellegren 2005; Casas-Marce et al. 2010; Klassen et al. 2012). The high amplification success rates that I obtained from fecal, hair and claw samples in this study are likely a result of my optimized DNA extraction methods and PCR protocols (Broquet et al. 2007; Müller et al. 2007), which together contributed to the high reliability of all of these tissues for sex determination.

The perfect accuracy rate for fecal DNA (Figure 11) matched sex determination accuracy rates achieved in previous studies on mammals (99 – 100%, Bradley et al. 2001; Lachish et al. 2011; Vanpé et al. 2013). I took careful precautions to avoid contamination as evidenced by the observation that all negative extractions and negative PCR controls were blank. However, samples of this type tend to yield DNA of lower quality and

quantity thus making them more susceptible to contamination (Taberlet et al. 1999). Perhaps unsurprisingly then, I concluded that two hair samples and one claw sample from known (or assumed) female moles, all amplified in triplicate, were contaminated. In each case, one PCR result indicated female while at least one more PCR amplification indicated male. Robertson and Gemmell (2006) suggested verifying the sex of an individual with a second independent sex test. I tested one contaminated hair and one contaminated claw DNA sample in a PCR using only the star-nosed mole-specific *Zfx* and *Zfy* primers. Only the *Zfx* locus amplified for both samples, indicating that they were both female and contamination-free. It is possible that DNA from these hair and claw samples were of lower quality than the fecal DNA but a correct result from the *Zfx* and *Zfy* sex test and the absence of a difference in accuracy rate for these tissue types (Table 3) suggests otherwise. To my knowledge, no studies are available that have measured the accuracy rate of hair shaft or claw clippings for molecular sexing of wild mammals. However, my accuracy rates are comparable (Figure 11) to the high accuracy rates reported from genotyping similar samples in some studies (100%; Nozawa et al. 1999; Hedmark and Ellegren 2005). In future, more replicates and larger sample sizes may help verify the reliability of each tissue type.

I had greater amplification success for the first PCR with fecal and claw samples compared to hair samples (Table 4; Figure 12). These results presumably relate to the presence of PCR inhibitors in hair, which will reduce the quality of the extracted DNA (Eckhart et al. 2000). The fecal DNA samples likely yielded higher quality DNA due to the method in which I collected most of them, i.e. by dissecting them out of the intestinal tracts of deceased moles. This may have resulted in the inadvertent collection of intestinal

cells and therefore higher quality and quantity of DNA in the extract than is normal for an excreted fecal DNA sample. However, the fact I also had amplification success with the one excreted fecal sample on the first PCR suggests that fecal samples have good potential for sex determination based on nondestructive samples. While feces generally contain higher levels of PCR inhibitors than other nondestructive tissues (Oliveira et al. 2010), it is possible that the DNA extraction method I used may have been efficient in removing these inhibitors thereby allowing for a greater amplification success rate during the first PCR (Vallet et al. 2008). Although both claw and hair samples undergo a similar keratinization process, which results in DNA degradation and therefore lower quality and quantity of DNA as the tissues grow outward (Bengtsson et al. 2012), the claw samples likely had greater amplification success after the first PCR as they did not contain melanin - the PCR inhibitor present in the hair samples (Eckhart et al. 2000). It is also possible that some of the claw samples I used for DNA extraction incorporated tissue from the claw base, where a higher quality and quantity of DNA is found (Bengtsson et al. 2012). It is of note that the hair samples did not possess hair roots as I clipped the ends of hairs from the mole's body, which would further explain the reduced amplification success of these samples from the first PCR.

Despite potential drawbacks with working with nondestructive samples, I improved the amplification success for each of the tissue types by using second PCR, replicating PCRs on the same samples and repeating DNA extractions when necessary. Similar methods have been used in previous NDS studies to improve amplification success rates. Although slightly different variations of the second PCR method have been used in previous studies, this method has been shown to improve amplification success

for both fecal DNA (Arandjelovic et al. 2009) and hair shaft DNA (Nozawa et al. 1999). The second PCR likely improves success rates by diluting PCR inhibitors while also increasing the amount of the target DNA template (Nozawa et al. 1999; Arandjelovic et al. 2009). Repeating DNA extractions and performing PCR replicates on the same sample can improve amplification success and the reliability of results for fecal (Vanpé et al. 2013), hair (Amory et al. 2007), nail (Klassen et al. 2012) and claw DNA (Hedmark and Ellegren 2005), as they allow for results from the same individual to be compared (Amory et al. 2007; Vanpé et al. 2013) and may also help to avoid false negatives (Klassen et al. 2012). Likewise, I found repeating DNA extractions for a few hair samples and one claw sample improved overall amplification success and helped avoid false negatives from those samples, while the amplification failure of one fecal sample and one claw sample after repeated PCR attempts may have been avoided if enough sample was left to repeat DNA extractions.

Contrary to the overall high amplification success rates for my nondestructive samples, I observed notable allelic dropout rates of *Zfx* for each tissue type. I predicted that hair samples would exhibit a higher allelic dropout rate than the other tissue types as melanin has been shown to inhibit the amplification of larger gene fragments more than smaller PCR products (Eckhart et al. 2000). However, while hair samples showed the highest allelic dropout rates (Figure 13), this result was not significant. Indeed, each of the three tissue types exhibited relatively high allelic dropout rates (28.6 – 50%) compared to some values found in the literature for feces (0 – 9%, Bradley et al. 2001; Vanpé et al. 2013; Arandjelovic et al. 2009), hair (8 – 10%, Amory et al. 2007; Opel et al. 2008) and claw (<1%, Hedmark and Ellegren 2005). Allelic dropout rates tend to

increase with fragment size, with dropout rates of hair shaft DNA reaching 75% found for fragments greater than 300 bp in length (Amory et al. 2007). This result supports the recommendation of Waits and Paetkau (2005) to choose an amplicon size of no greater than 300 bp for noninvasive samples. Hence, the relatively high allelic dropout rates could in part reflect the fact that the *Zfx* product amplified in this study was 349 bp. However, this larger product could not be avoided: I was unable to find better gene regions for designing workable primers to amplify a smaller *Zfx* fragment.

Despite the relatively high allelic dropout rate of *Zfx*, this did not severely affect the accuracy of the sex determination test, as the two Y fragments I amplified (*Zfy*: 236 bp, *Sry*: 101 bp) in the multiplex PCR were less than 300 bp in length. These smaller fragments presumably increased amplification success for even the lowest quality DNA samples. This allowed for accurate sex determination and avoided the problem of incorrect sex assignment if presence or absence of only one locus or allele was involved in the PCR (Robertson and Gemmell 2006). As Y fragments are only detected in males, however, degraded DNA samples from females may fail to amplify more often due to the larger size of the *Zfx* amplicon I used. Only two female samples in my study failed to amplify and might have been successful if I had collected more tissue to attempt a second DNA extraction. Either way, I found that using multiple loci and alleles avoided the problem of misidentifying sex from a failed amplification, providing support for this multiplex sex determination test.

Conclusions and recommendations

I developed the first nondestructive method for molecular sex determination of star-nosed moles. All three nondestructive tissues I used performed approximately equally (90 – 100% success rate), which suggests that the DNA extraction methods and PCR protocols I developed are well optimized. This work adds to the limited body of literature available on NDS molecular approaches for small mammals and will likely be of value, in particular, for future ecological studies of wild moles.

Although DNA from some of my samples occasionally failed to amplify the target genes or was contaminated, I found that using multiple sources of DNA from one individual was beneficial for confirming sex. In terms of ease of collection and reducing stress on the animals, feces seem to be the best tissue for future studies. Collection of feces from a star-nosed mole involves simply placing the animal in a container and waiting for it to defecate with minimal handling (Campbell KL, pers. comm.), and large samples can be obtained if the animal is provided food (e.g. earthworms). Collection of hair or claw clippings requires physical restraint of the animal or anesthesia. If a second source of DNA is required to supplement the first source, hair would be the next best option, as collection of this tissue will likely result in the least stress for the animal compared to collection of claw clippings. Star-nosed mole claws are small and difficult to manipulate, therefore increasing the chance of clipping too close to the base and inflicting pain. Any source of nondestructive DNA should be collected in excess if possible to allow for multiple DNA extractions from the same individual. This will increase chances of amplification success and accuracy of sex assignment.

There are a number of approaches that could help further verify and improve the success of the sex determination method I designed. Most importantly, fecal samples should be collected using the conventional method (i.e. after defecation instead of during dissection) to validate the success of this tissue as a DNA source. In addition, testing this method on a greater number of individual moles will help confirm the reliability of the multiplex PCR. The number of PCR replicates needed to verify an individual's sex must also be determined. Indeed, the first successful PCR result can often be incorrect as I observed in this study with the few contaminated samples. Using blind control samples, where the sex is unknown before observing the results, to test these methods will help further verify and improve the reliability of this new nondestructive sex determination test for star-nosed moles. In the future, determining sex of star-nosed moles from degraded hair or fecal DNA should be attempted through the use of targeted hybridization capture and next generation sequencing methods instead of traditional PCR. This newer method increases the accuracy and reduces mistakes in sex identification from contamination in samples from females (Skoglund et al. 2013). This method allows for millions of base pairs of sequence to be enriched simultaneously, and would thus facilitate studies on the use of alternate X and Y chromosome genes for sex determination in star-nosed moles and may even facilitate studies on the genetic diversity of this species (Green et al. 2006).

The nondestructive molecular sex determination test I developed for the star-nosed mole could facilitate future research into the ecology of this little-studied species. Similar to other New World talpids, virtually nothing is known about the mating system and social structure of star-nosed moles. Combining radio-telemetry techniques with

NDS methods will enable studies of seasonal patterns and sex-specific differences in social behaviour, activity patterns, home range and space use, male-female interactions, territoriality, and even kinship relationships. Additionally, combining DNA barcoding techniques (Hebert et al. 2003; Hajibabaei et al. 2007) with NDS methods will enable researchers to study the diet of free-ranging star-nosed moles from prey remains in fecal samples alone (e.g. Dodd et al. 2012). Furthermore, the use of hair samples from star-nosed moles and their potential prey items in stable isotope analysis will allow researchers to understand more about the star-nosed mole's seasonal foraging habits (Tieszen et al. 1983; Macko et al. 1999). A greater knowledge of the ecology and resources needed by star-nosed moles may also reveal habitat requirements of this species and therefore provide much needed insight into the population status of star-nosed moles. Understanding sex-specific differences and further developing and using NDS methods for star-nosed moles will improve our understanding of this species in the wild, help with development of management plans incorporating star-nosed moles, and contribute to development of nondestructive methods for studying other poorly understood members of the Talpidae family.

References:

Aitchison, C.W. 1979. Notes on low temperature activity of oligochaetes, gastropods and centipedes in southern Canada. Am. Midl. Nat. **102**: 399-400.

Alasaad, S., Sánchez, A., Marchal, J.A., Píriz, A., Garrido-García, J.A., Carro, F., Romero, I., and Soriguer, R.C. 2011. Efficient identification of *Microtus cabrerae* excrements using noninvasive molecular analysis. Conserv. Genet. Resour. **3**: 127-129.

Amory, S., Keyser, C., Crubézy, E., and Ludes, B. 2007. STR typing of ancient DNA extracted from hair shafts of Siberian mummies. Forensic Sci. Int. **166**: 218-229.

Anderson, T.D., Ross, J.P., Roby, R.K., Lee, D.A., and Holland, M.M. 1999. A validation study for the extraction and analysis of DNA from human nail material and its application to forensic casework. J. Forensic Sci. **44**: 1053-1056.

Arandjelovic, M., Guschanski, K., Schubert, G., Harris, T.R., Thalmann, O., Siedel, H. and Vigilant, L., 2009. Two-step multiplex polymerase chain reaction improves the speed and accuracy of genotyping using DNA from noninvasive and museum samples. Mol. Ecol. Resour. **9**: 28-36.

Barbosa, S., Paupiero, J., Searle, J.B., and Alves, P.C. 2013. Genetic identification of Iberian rodent species using both mitochondrial and nuclear loci: application to noninvasive sampling. Mol. Ecol. Resour. **13**: 43-56.

Belant, J.L., Etter, D.R., Friedrich, P.D., Cosgrove, M.K., Williams, B.W., and Scribner, K.T. 2011. Comparison of techniques for sex determination of American martens. J. Wildl. Manage. **75**: 256-260.

Bellemain, E., Swenson, J.E., Tallmon, D., Brunberg, S., and Taberlet, P. 2005. Estimating population size of elusive animals with DNA from hunter-collected feces: four methods for brown bears. Conserv. Biol. **19**: 150-161.

Bengtsson, C.F., Olsen, M.E., Brandt, L.Ø., Bertelsen, M.F., Willerslev, E., Tobin, D.J., Wilson, A.S., and Gilbert, M.T.P. 2012. DNA from keratinous tissue. Part I: Hair and nail. Ann. Anat. **194**: 17-25.

Beolchini, F., and Loy, A. 2004. Diet of syntopic moles *Talpa romana* and *Talpa europaea* in central Italy. Mamm. Biol. **69**: 140-144.

Bidon, T., Frosch, C., Eiken, H.G., Kutschera, V.E., Hagen, S.B., Aarnes, S.G., Fain, S.R., Janke, A., and Hailer, F. 2013. A sensitive and specific multiplex PCR approach for sex identification of ursine and tremarctine bears suitable for non-invasive samples. Mol. Ecol. Resour. **13**: 362-368.

Bodmer, R.E. 1991. Influence of digestive morphology on resource partitioning in Amazonian ungulates. Oecologia. **85**: 361-365.

Boston, E.S.M., Puechmaille, S.J., Scott, D.D., Buckley, D.J., Lundy, M.G., Montgomery, I.W., Prodöhl, P.A., and Teeling, E.C. 2012. Empirical assessment of non-invasive population genetics in bats: comparison of DNA quality from faecal and tissue samples. Acta Chiropt. **14**: 45-52.

Bradley, B.J., Chambers, K.E., and Vigilant, L. 2001. Accurate DNA-based sex identification of apes using non-invasive samples. Conserv. Genet. **2**: 179-181.

Bradley, B.J., Doran-Sheehy, D.M., and Vigilant, L. 2008. Genetic identification of elusive animals: re-evaluating tracking and nesting data for wild western gorillas. J. Zool. **275**: 333-340.

Broquet, T., Ménard, N., and Petit, E. 2007. Noninvasive population genetics: a review of sample source, diet, fragment length and microsatellite motif effects on amplification success and genotyping error rates. Conserv. Genet. **8**: 249-260.

Bryja, J., and Konečný, A. 2003. Fast sex identification in wild mammals using PCR amplification of the *Sry* gene. Folia Zool. **52**: 269-274.

Bullejos, M., Sánchez, A., Burgos, M., Jiménez, R., and Díaz de la Guardia, R. 2000. The SRY gene HMG-box in micro- and megabats. Cytogenet. Cell Genet. **88**: 30-34.

Campbell, K.L., McIntyre, I.W., and MacArthur, R.A. 1999. Fasting metabolism and thermoregulatory competence of the star-nosed mole, *Condylura cristata* (Talpidae: Condylurinae). Comp. Biochem. Physiol. **123A**: 293-298.

Campbell, G.M., Pauli, J.N., Thomas, J.G., and McClean, T. 2010. Accuracy in molecular sexing of martens (*Martes americana* and *Martes caurina*) varies among sample types. Mol. Ecol. Resour. **10**: 1019-1022.

Carraway, L.N. 2009. Determining sex of *Sorex* shrews (Soricomorpha: Soricidae). Am. Midl. Nat. **162**: 87-97.

Casas-Marce, M., Revilla, E., and Godoy, J.A. 2010. Searching for DNA in museum specimens: a comparison of sources in a mammal species. Mol. Ecol. Resour. **10**: 502-507.

Catania, K.C. 1999. A nose that looks like a hand and acts like an eye: the unusual mechanosensory system of the star-nosed mole. J. Comp. Physiol. A, **185**: 367-372.

Catania, K.C., and Remple, F.E. 2005. Asymptotic prey profitability drives star-nosed moles to the foraging speed limit. Nature. **433**: 519-522.

Catania, K.C. 2006. Underwater 'sniffing' by semi-aquatic mammals. Nature. **444**: 1024-1025.

Catania, K.C. 2012. Evolution of brains and behavior for optimal foraging: a tale of two predators. Proc. Natl. Acad. Sci. USA. **109**: 10701-10708.

Cathey, J.C., Bickham, J.W., and Patton, J.C. 1998. Introgressive hybridization and nonconcordant evolutionary history of maternal and paternal lineages in North American deer. Evolution **52**: 1224-1229.

Choate, D.M., Wolfe, M.L., and Stoner, D.C. 2006. Evaluation of cougar population estimators in Utah. Wildl. Soc. Bull. **34**: 782-799.

Crawford, J.C., Liu, Z., Nelson, T.A., Nielsen, C.K., and Bloomquist, C.K. 2008. A comparison of field and molecular techniques for sexing beavers. J. Wildl. Manage. **72**: 1805-1807.

Crawford, K., McDonald, R.A., and Bearhop, S. 2008. Applications of stable isotope techniques to the ecology of mammals. Mammal Rev. **38**: 87-107.

Croin Michielsen, N. 1966. Intraspecific and interspecific competition in the shrews *Sorex araneus* L. and *S. minutus* L. Arch. Néerland. Zool. **17**: 73-174.

Curtis, C., Stewart, B.S., and Karl, S.A. 2007. Sexing pinnipeds with *ZFX* and *ZFY* loci. J. Hered. **98**: 280-285.

Dahl, T.E. 1990. Wetlands losses in the United States 1780's to 1980's. U.S. Department of the Interior, Fish and Wildlife Service, Washington, D.C. 13 pp.

Danks, H.V. 2007. How aquatic insects live in cold climates. Can. Entomol. **139**: 443-471.

DeYoung, R.W., and Honeycutt, R.L. 2005. The molecular toolbox: genetic techniques in wildlife ecology and management. J. Wildl. Manage. **69**: 1362-1384.

Dodd, L.E., Chapman, E.G., Harwood, J.D., Lacki, M.J., and Rieske, L.K. 2012. Identification of prey of *Myotis septentrionalis* using DNA-based techniques. J. Mammal. **93**: 1119-1128.

Ducks Unlimited Canada. 2010. Southern Ontario wetland conversion analysis. 23 pp.

Durnin, M.E., Palsbøll, P.J., Ryder, O.A., and McCullough, D.R. 2007. A reliable genetic technique for sex determination of giant panda (*Ailuropoda melanoleuca*) from non-invasively collected hair samples. Conserv. Genet. **8**: 715-720.

Eadie, W.R., and Hamilton, W.J., Jr. 1956. Notes on reproduction in the star-nosed mole. J. Mammal. **37**: 223-231.

Eckhart, L., Bach, J., Ban, J., and Tschachler, E. 2000. Melanin binds reversibly to thermostable DNA polymerase and inhibits its activity. Biochem. Biophys. Res. Commun. **271**: 726-730.

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**:1792-1797.

Ferguson-Smith, M.A. and Rens, W. 2010. The unique sex chromosome system in platypus and echidna. Russ. J. Genet. **46**: 1160-1164.

Fontanesi, L., Tazzoli, M., Pecchioli, E., Hauffe, H.C., Robinson, T.J., and Russo, V. 2008. Sexing European rabbits (*Oryctolagus cuniculus*), European brown hares (*Lepus europaeus*) and mountain hares (*Lepus timidus*) with *ZFX* and *ZFY* loci. Mol. Ecol. Resour. **8**: 1294-1296.

Frank, L.G. 1986. Social organization of the spotted hyaena *Crocuta crocuta*. II. Dominance and reproduction. Anim. Behav. **34**: 1510-1527.

García-López de Hierro, L., Moleón, M., Lupiáñez, D.G., Virgós, E., and Jiménez, R. 2013. Positive and negative unintended human-induced effects on Iberian mole abundance at the edge of its distribution area. Mamm. Biol. **78**: 276-282.

García-Moreno, J., and Mindell, D.P. 2000. Rooting a phylogeny with homologous genes on opposite sex chromosomes (gametologs): a case study using avian CHD. Mol. Biol. Evol. **17**: 1826-1832.

Gokulakrishnan, P., Kumar, R.R., Sharma, B.D., Mendiratta, S.K., and Sharma, D. 2012. A duplex PCR assay for sex determination of cattle meat by simultaneous amplification of SRY, AMELX and AMELY genes. Food Biotechnol. **26**: 75-84.

Goldberg, C.S., Woodruff, K., Toldness, R., and Waits, L.P. 2011. Robust molecular sex identification of beaver (*Castor canadensis*) from non-destructive samples. Conserv. Genet. Resour. **3**: 729-731.

González-Esteban, J., Villate, I., and Castién, E. 2003a. A comparison of methodologies used in the detection of the Pyrenean desman *Galemys pyrenaicus* (E. Geoffroy, 1811). Mamm. Biol. **68**: 387-390.

González-Esteban, J., Villate, I., and Castién, E. 2003b. Sexual identification of *Galemys pyrenaicus*. Acta Theriol. **48**: 571-573.

Gorman, M.L., and Stone, R.D. 1990. The natural history of moles. Cornell University Press, Ithaca, NY.

Green, R.E., Krause, J., Ptak, S.E., Briggs, A.W., Ronan, M.T., Simons, J.F., Du, L., Egholm, M., Rothberg, J.M., Paunovic, M., and Pääbo, S. 2006. Analysis of one million base pairs of Neanderthal DNA. Nature. **444**: 330-336.

Grigione, M.M., Burman, P., Bleich, V.C., and Pierce, B.M. 1999. Identifying individual mountain lions *Felis concolor* by their tracks: refinement of an innovative technique. Biol. Conserv. **88**: 25-32.

Guerier, A.S., Bishop, J.M., Crawford, S.J., Schmidt-Küntzel, A., and Stratford, K.J. 2012. Parentage analysis in a managed free ranging population of southern white rhinoceros: genetic diversity, pedigrees and management. Conserv. Genet. **13**: 811-822.

Hajibabaei, M., Singer, G.A.C., Hebert, P.D.N., and Hickey, D.A. 2007. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. Trends Genet. **23**: 167-172.

Hamilton, W.J., Jr. 1931. Habits of the star-nosed mole, *Condylura cristata*. J. Mammal. **12**: 345-355.

Han, S.-H., Yang, B.-C., Ko, M.-S., Oh, H.-S., and Lee, S.-S. 2010. Length difference between equine *ZFX* and *ZFY* genes and its application for molecular sex determination. J. Assist. Reprod. Genet. **27**: 725-728.

Hänni, C., Brousseau, T., Laudet, V., and Stehelin, D. 1995. Isopropanol precipitation removes PCR inhibitors from ancient bone extracts. Nucleic Acids Res. **23**: 881-882.

Hartman, G.D., Whitaker, J.O., Jr., and Munsee, J.R. 2000. Diet of the mole *Scalopus aquaticus* from the Coastal Plain region of South Carolina. Am. Midl. Nat. **144**: 342-351.

Hebert, P.D.N., Cywinska, A., Ball, S.L., and deWaard, J.R. 2003. Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B **270**: 313-321.

Hedmark, E., and Ellegren, H. 2005. Microsatellite genotyping of DNA isolated from claws left on tanned carnivore hides. Int. J. Legal Med. **119**: 370-373.

Henry, P., and Russello, M.A. 2011. Obtaining high-quality DNA from elusive small mammals using low-tech hair snares. Eur. J. Wildl. Res. **57**: 429-435.

Hibert, F., Sabatier, D., Andrivot, J., Scotti-Saintagne, C., Gonzalez, S., Prévost, M.-F., Grenand, P., Chave, J., Caron, H., and Richard-Hansen, C. 2011. Botany, genetics and ethnobotany: a crossed investigation on the elusive tapir's diet in French Guiana. PLoS One. **6**: e25850.

Hoyle, S.D., Horsup, A.B., Johnson, C.N., Crossman, D.G., and McCallum, H. 1995. Live-trapping of the northern hairy-nosed wombat (*Lasiorhinus krefftii*): population-size estimates and effects on individuals. Wildl. Res. **22**: 741-755.

Igea, J., Aymerich, P., Fernández-González, A., González-Esteban, J., Gómez, A., Alonso, R., Gosálbez, J., and Castresana, J. 2013. Phylogeography and postglacial expansion of the endangered semi-aquatic mammal *Galemys pyrenaicus*. BMC Evol. Biol. **13**: 115-133.

IUCN 2013. IUCN Red List of Threatened Species. Version 2013.1. <www.iucnredlist.org>. Downloaded on 10 September 2013.

Janečka, J.E., Blankenship T.L., Hirth, D.H., Tewes, M.E., Kilpatrick, C.W., and Grassman Jr, L.I. 2006. Kinship and social structure of bobcats (*Lynx rufus*) inferred from microsatellite and radio-telemetry data. J. Zool. **269**: 494-501.

Just, W., Rau, W., Vogel, W., Akhverdian, M., Fredga, K., Graves, J.A., and Lyapunova, E. 1995. Absence of *Sry* in species of the vole *Ellobius*. Nat. Genet. **11**: 117-118.

Kashimura, A., Moteki, K., Shinohara, A., Tsuchiya, K., Takahashi, T., and Morita, T. 2010. Seasonal changes in the diet and diet selectivity of large Japanese moles, *Mogera wogura* (Temminck, 1842). Jpn. J. Environ. Entomol. Zool. **21**: 147-153.

Kelly, J.F. 2000. Stable isotopes of carbon and nitrogen in the study of avian and mammalian trophic ecology. Can. J. Zool. **78**: 1-27.

Kishore, R., Hardy, W.R., Anderson, V.J., Sanchez, N.A., and Buoncristiani, M.R. 2006. Optimization of DNA extraction from low-yield and degraded samples using the BioRobot® EZ1 and BioRobot® M48. J Forensic Sci. **51**: 1055-1061.

Klassen, T.L., von Rüden, E.-L., Drabek, J., Noebels, J.L., and Goldman, A.M. 2012. Comparative analytical utility of DNA derived from alternative human specimens for molecular autopsy and diagnostics. J. Mol. Diagn. **14**: 451-457.

Kohn, M.H., and Wayne, R.K. 1997. Facts from feces revisited. Trends Ecol. Evol. **12**: 223-227.

Koper, N., and Manseau, M. 2012. A guide to developing resource selection functions from telemetry data using generalized estimating equations and generalized linear mixed models. Rangifer. **20**: 195-203.

Kühn, R., Schwab, G., Schröder, W., and Rottman, O. 2002. Molecular sex diagnosis in Castoridae. Zoo Biol. **21**: 305-308.

Lachish, S., Passmore, A., and Jones, M. 2011. A new PCR assay for reliable molecular sexing of endangered Tasmanian devils (*Sarcophilus harrisii*) from non-invasive genetic samples. Conserv. Genet. Resour. **3**: 279-281.

Lanyon, J.M., Sneath, H.L., Ovenden, J.R., Broderick, D., and Bonde, R.K. 2009. Sexing Sirenians: Validation of visual and molecular sex determination in both wild dugongs (*Dugong dugon*) and Florida manatees (*Trichechus manatus latirostris*). Aquat. Mamm. **35**: 187-192.

Lindsay, A.R., and Belant, J.L. 2008. A simple and improved PCR-based technique for white-tailed deer (*Odocoileus virginianus*) sex identification. Conserv. Genet. **9**: 443-447.

Lynch, Á.B., and Brown, M.J.F. 2006. Molecular sexing of pine marten (*Martes martes*): how many replicates? Mol. Ecol. Notes. **6**: 631-633.

Macdonald, D.W., Atkinson, R.P.D., and Blanchard, G. 1997. Spatial and temporal patterns in the activity of European moles. Oecologia. **109**: 88-97.

Macko, S.A., Engel, M.H., Andrusevich, V., Lubec, G., O'Connell, T.C., and Hedges, R.E.M. 1999. Documenting the diet in ancient human populations through stable isotope analysis of hair. Phil. Trans. R. Soc. Lond. B **354**: 65-76.

Matsubara, K., Ishibashi, Y., Ohdachi, S., and Matsuda, Y. 2001. A new primer set for sex identification in the genus *Sorex* (Soricidae, Insectivora). Mol. Ecol. Notes. **1**: 241-242.

McIntyre, I.W., Campbell, K.L., and MacArthur, R.A. 2002. Body oxygen stores, aerobic dive limits and diving behavior of the star-nosed mole (*Condylura cristata*) and comparisons with non-aquatic talpids. J. Exp. Biol. **205**: 45-54.

Melero, Y., Aymerich, P., Luque-Larena, J.J., and Gosàlbez, J. 2012. New insights into social and space use behaviour of the endangered Pyrenean desman (*Galemys pyrenaicus*). Eur. J. Wildl. Res. **58**: 185-193.

Merriam, C.H. 1884. The star-nosed mole amphibious. Science. **4**: 429.

Meyer, J., Kohnen, A., Harf, R., Froeschke, G., and Brandl, R. 2006. Molecular markers for some small mammals of southern Africa. Folia Zool. **55**: 444-447.

Moran, S., Turner, P.D., and O'Reilly, C. 2008. Non-invasive genetic identification of small mammal species using real-time polymerase chain reaction. Mol. Ecol. Resour. **8**: 1267-1269.

Müller, K., Klein, R., Miltner, E., and Wiegand, P. 2007. Improved STR typing of telogen hair root and hair shaft DNA. Electrophoresis. **28**: 2835-2842.

Nishida, S., Pastene, L.A., Goto, M., and Koike, H. 2003. *SRY* gene structure and phylogeny in the cetacean species. Mamm. Study. **28**: 57-66.

Nozawa, H., Yamamoto, T., Uchihi, R., Yoshimoto, T., Tamaki, K., Hayashi, S., Ozawa, T., and Katsumata, Y. 1999. Purification of nuclear DNA from single hair shafts for DNA analysis in forensic sciences. Legal Med. **1**: 61-67.

Ojeda, G.N., Amavet, P.S., Rueda, E.C., and Siroski, P.A. 2012. DNA extraction from skins of wild (*Hydrochoerus hydrochaeris* and *Pecari tajacu*) and domestic (*Sus scrofa domestica*) species using a novel protocol. Genet. Mol. Res. **11**: 672-678.

Oliveira, R., Castro, D., Godinho, R., Luikart, G., and Alves, P.C. 2010. Species identification using a small nuclear gene fragment: application to sympatric wild carnivores from South-western Europe. Conserv. Genet. **11**: 1023-1032.

Opel, K.L., Fleishaker, E.L., Nicklas, J.A., Buel, E., and McCord, B.R. 2008. Evaluation and quantification of nuclear DNA from human telogen hairs. J. Forensic Sci. **53**: 853-857.

Petersen, K.E., and Yates, T.L. 1980. *Condylura cristata*. Mamm. Species. **129**: 1-4.

Piggott, M.P., and Taylor, A.C. 2003. Remote collection of animal DNA and its applications in conservation management and understanding the population biology of rare and cryptic species. Wildl. Res. **30**: 1-13.

Pilgrim, K.L., McKelvey, K.S., Riddle, A.E., and Schwartz, M.K. 2005. Felid sex identification based on noninvasive genetic samples. Mol. Ecol. Notes. **5**: 60-61.

Pomp, D., Good, B.A., Geisert, R.D., Corbin, C.J., and Conley, A.J. 1995. Sex identification in mammals with polymerase chain reaction and its use to examine sex effects on diameter of day-10 or -11 pig embryos. J. Anim. Sci. **73**: 1408-1415.

Proctor, M., McLellan, B., Boulanger, J., Apps, C., Stenhouse, G., Paetkau, D., and Mowat, G. 2010. Ecological investigations of grizzly bears in Canada using DNA from hair, 1995-2005: a review of methods and progress. Ursus. **21**: 169-188.

Puechmaille, S.J., and Petit, E.J. 2007. Empirical evaluation of non-invasive capture-mark-recapture estimation of population size based on a single sampling session. J. Appl. Ecol. **44**: 843-852.

Ray, D.K., Freshwater, D.W., and Webster, W.M.D. 2002. Multiplex PCR identification of eastern and Allegheny woodrats in zones of contact. J. Wildl. Manage. **66**: 86-90.

Reddy, P.A., Bhavanishankar, M., Bhagavatula, J., Harika, K., Mahla, R.S., and Shivaji, S. 2012. Improved methods of carnivore faecal sample preservation, DNA extraction and quantification for accurate genotyping of wild tigers. PLoS ONE. **7**: e46732.

Reed, D.H., Nicholas, A.C., and Stratton, G.E. 2007. Genetic quality of individuals impacts population dynamics. Anim. Conserv. **10**: 275-283.

Reiners, T.E., Encarnação, J.A., and Wolters, V. 2011. An optimized hair trap for non-invasive genetic studies of small cryptic mammals. Eur. J. Wildl. Res. **57**: 991-995.

Robertson, B.C., and Gemmell, N.J. 2006. PCR-based sexing in conservation biology: wrong answers from an accurate methodology? Conserv. Genet. **7**: 267-271.

Rodgers, T.W., and Janečka, J.E. 2013. Applications and techniques for non-invasive faecal genetics research in felid conservation. Eur. J. Wildl. Res. **59**: 1-16.

Rowe, L.E., and Dawson, S.M. 2009. Determining the sex of bottlenose dolphins from Doubtful Sound using dorsal fin photographs. Mar. Mamm. Sci. **25**: 19-34.

Rubenstein, N.M., Cunha, G.R., Wang, Y.Z., Campbell, K.L., Conley, A.J., Catania, K.C., Glickman, S.E., and Place, N.J. 2003. Variation in ovarian morphology in four species of New World moles with a peniform clitoris. Reproduction. **126**: 713-719.

Rust, C.C. 1966. Notes on the star-nosed mole (*Condylura cristata*). J. Mammal. **47**: 538.

Sánchez, A., Bullejos, M., Burgos, M., Hera, C., Jiménez, R., and Díaz de la Guardia, R. 1996. High sequence identity between the SRY HMG box from humans and insectivores. Mamm. Genome. **7**: 536-538.

Sawaya, M.A., Ruth, T.K., Creel, S., Rotella, J.J., Stetz, J.B., Quigley, H.B., and Kalinowski, S.T. 2011. Evaluation of noninvasive genetic sampling methods for cougars in Yellowstone National Park. J. Wildl. Manage. **75**: 612-622.

Scheppers, T.L.J., Frantz, A.C., Schaul, M., Engel, E., Breyne, P., Schley, L., and Roper, T.J. 2007. Estimating social group size of Eurasian badgers *Meles meles* by genotyping remotely plucked single hairs. Wildl. Biol. **13**: 195-207.

Schmidt, F.J.W. 1931. Mammals of Western Clark County, Wisconsin. J. Mammal. **12**: 99-117.

Seddon, J.M. 2005. Canid-specific primers for molecular sexing using tissue or non-invasive samples. Conserv. Genet. **6**: 147-149.

Shaw, C.N., Wilson, P.J., and White, B.N. 2003. A reliable molecular method of gender determination for mammals. J. Mamm. **84**: 123-128.

Shinohara, A., Campbell, K.L., and Suzuki, H. 2003. Molecular phylogenetic relationships of moles, shrew moles, and desmans from the new and old worlds. Mol. Phylogenet. Evol. **27**: 247-258.

Skoglund, P., Storå, J., Götherström, A., and Jakobsson, M. 2013. Accurate sex identification of ancient human remains using DNA shotgun sequencing. J. Archaeol. Sci. **40**: 4477-4482.

Stewart, R.L., Jr., Kindlin, C.M., and Delis, P.R. 2008. An inventory of the small mammals of Letterkenny Army Depot, Franklin County, Pennsylvania. J. PA Acad. Sci. **81**: 59-65.

Stone, R.D., and Gorman, M.L. 1985. Social organization of the European mole (*Talpa europaea*) and the Pyrenean desman (*Galemys pyrenaicus*). Mammal Rev. **15**: 35-42.

Streatfeild, C.A., Mabry, K.E., Keane, B., Crist, T.O., and Solomon, N.G. 2011. Intraspecific variability in the social and genetic mating systems of prairie voles, *Microtus ochrogaster*. Anim. Behav. **82**: 1387-1398.

Sutou, S., Mitsui, Y., and Tsuchiya, K. 2001. Sex determination without the Y chromosome in two Japanese rodents *Tokudaia osimensis osimensis* and *Tokudaia osimensis* spp. Mamm. Genome. **12**: 17-21.

Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., Waits, L.P., and Bouvet, J. 1996. Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic Acids Res. **24**: 3189-3194.

Taberlet, P., Waits, L.P., and Luikart, G. 1999. Noninvasive genetic sampling: look before you leap. Trends Ecol. Evol. **14**: 323-327.

Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res. **17**: 6463-6471.

Tieszen, L.L., Boutton, T.W., Tesdahl, K.G., and Slade, N.A. 1983. Fractionation and turnover of stable carbon isotopes in animal tissues: implications for $\delta^{13}C$ analysis of diet. Oecologia. **57**: 32-37.

Vallet, D., Petit, E.J., Gatti, S., Levréro, F., and Ménard, N. 2008. A new 2CTAB/PCI method improves DNA amplification success from faeces of Mediterranean (Barbary macaques) and tropical (lowland gorillas) primates. Conserv. Genet. **9**: 677-680.

Vanpé, C., Salmona, J., Pais, I., Kun-Rodrigues, C., Pichon, C., Meyler, S.V., Rabarivola, C., Lewis, R.J., Ibouri, M.T., and Chikhi, L. 2013. Noninvasive molecular sexing: an evaluation and validation of the SRY- and Amelogenin-based method in three new lemur species. Am. J. Phys. Anthropol. **150**: 492-503.

Vidal, O., Pérez-Serra, A., and Pla, C. 2010. A sex determination protocol for the Iberian desman (*Galemys pyrenaicus*) based on a three primer amplification of *DBX* and *DBY* fragments with non-invasive samples. Conserv. Genet. **11**: 1185-1187.

Villesen, P., and Fredsted, T. 2006. Fast and non-invasive PCR sexing of primates: apes, Old World monkeys, New World monkeys and Strepsirrhines. BMC Ecol. **6**: 8.

Waits, L.P., and Paetkau, D. 2005. Noninvasive genetic sampling tools for wildlife biologists: a review of applications and recommendations for accurate data collection. J. Wildl. Manage. **69**: 1419-1433.

Waldron, G., Rodger, L., Mouland, G., and Lebedyk, D. 2000. Range, habitat, and population size of the Eastern Mole, *Scalopus aquaticus*, in Canada. Can. Field-Nat. **114**: 351-358.

Walker, F.M., Sunnucks, P., and Taylor, A.C. 2006. Genotyping of "captured" hairs reveals burrow-use and ranging behavior of southern hairy-nosed wombats. J. Mammal. **87**: 690-699.

Wallis, M.C., Waters, P.D., and Graves, J.A.M. 2008. Sex determination in mammals - before and after the evolution of *SRY*. Cell. Mol. Life Sci. **65**: 3182-3195.

Williams, C.L., Breck, S.W., and Baker, B.W. 2004. Genetic methods improve accuracy of gender determination in beavers. J. Mammal. **85**: 1145-1148.

Williams, B.W., Etter, D.R., DeWitt, P.D., Scribner, K.T., and Friedrich, P.D. 2011. Uncertainty in determination of sex from harvested bobcats. J. Wildl. Manage. **75**: 1508-1512.

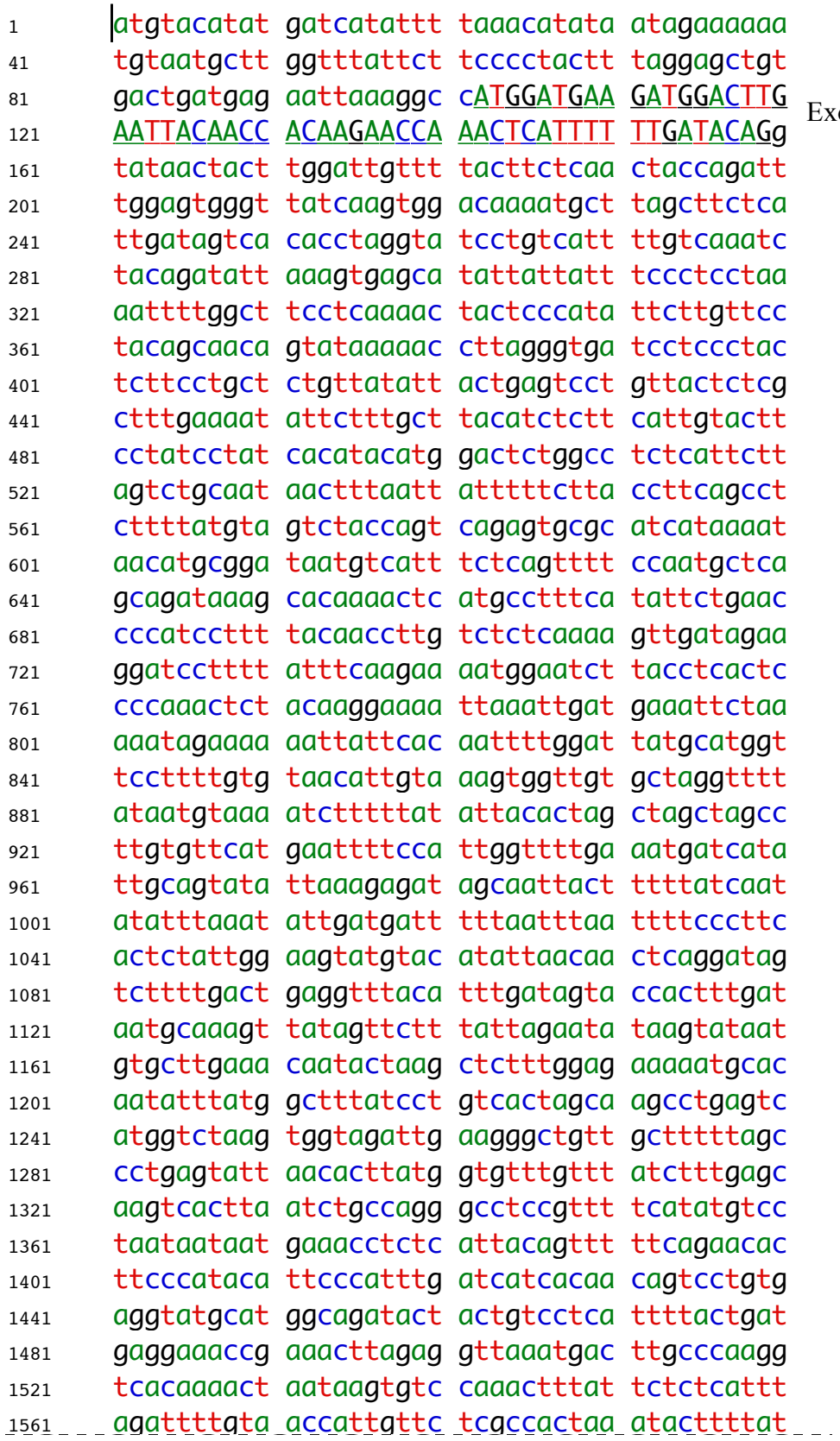
Yamamoto, K., Tsubota, T., Komatsu, T., Katayama, A., Murase, T., Kita, I., and Kudo, T. 2002. Sex identification of Japanese black bear, *Ursus thibetanus japonicus*, by PCR based on amelogenin gene. J. Vet. Med. Sci. **64**: 505-508.

Yu, H.-T., Ma, G.-C., Lee, D.-J., Chin, S.-C., Tsao, H.-S., Wu, S.-H., Shih, S.-Y., and Chen, M. 2011. Molecular delineation of the Y-borne *Sry* gene in the Formosan pangolin (*Manis pentadactyla pentadactyla*) and its phylogenetic implications for Pholidota in extant mammals. Theriogenology. **75**: 55-64.

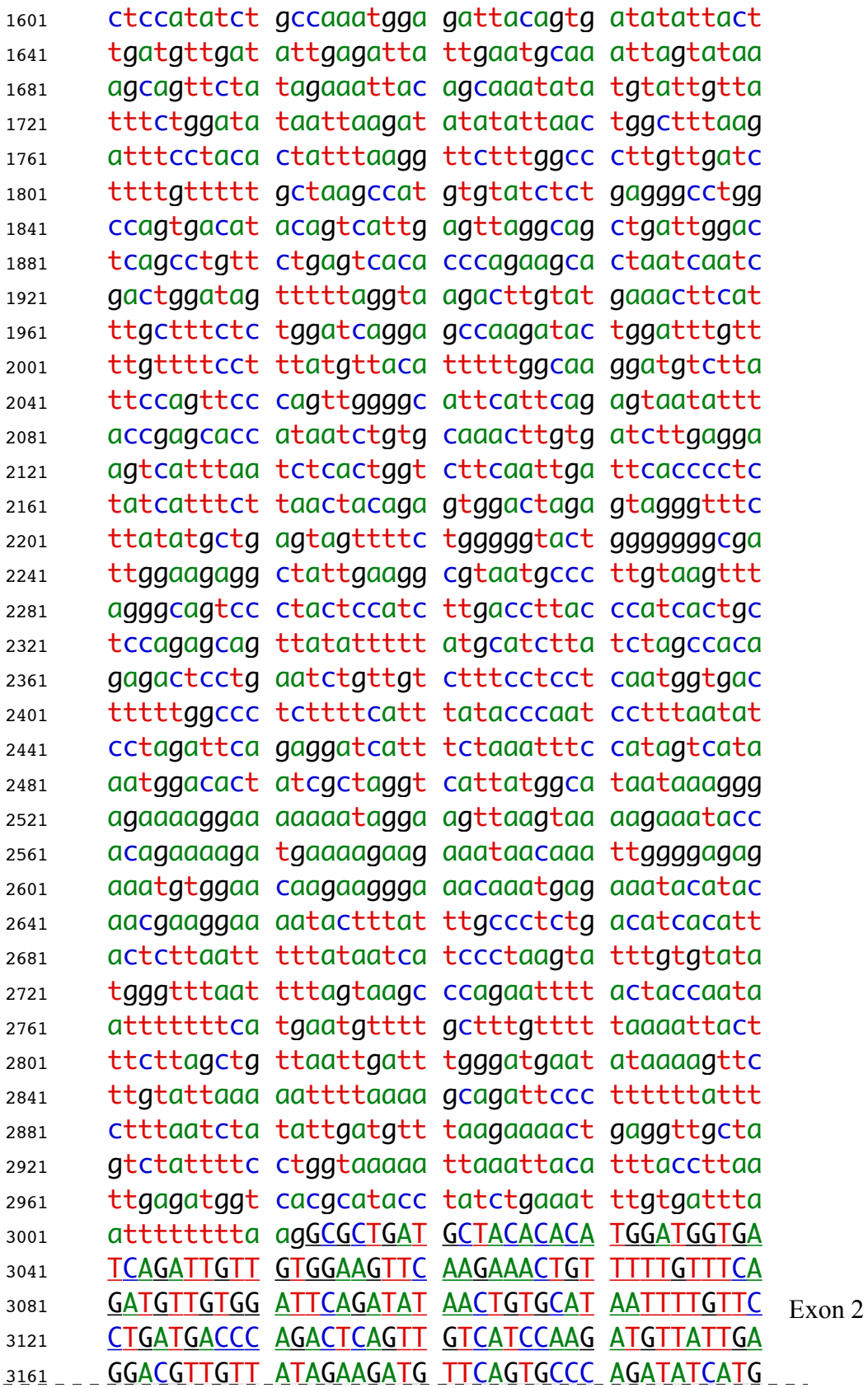
Zurawksa-Seta, E., and Barczak, T. 2012. The influence of field margins on the presence and spatial distribution of the European mole *Talpa europaea* L. within the agricultural landscape of northern Poland. Arch. Biol. Sci. **64**: 971-980.

Appendix 1:

Annotated *Zfx* gene sequence from the publically available star-nosed mole genome (GenBank accession #: AJFV00000000.1). Capitalized and underlined nucleotides represent the exon regions while lower case nucleotides represent the intron (and flanking) regions. The poly-A tail, found near the end of the sequence, is represented by the underlined lower case nucleotides "aataaa". The gap of unknown length, found in the intron region between exons two and three, is represented by a series of underlined lower case n's. Finally, the sequenced region of *Zfx*, found between exons six and seven, is highlighted in grey.


1 atgtacatat gatcatattt taaacatata atagaaaaaa
41 tgtaatgctt ggtttattct tcccctactt taggagctgt
81 gactgatgag aattaaaggc cATGGATGAA GATGGACTTG
121 AATTACAACC ACAAGAACCA AACTCATTTT TTGATACAGg
161 tataactact tggattgttt tacttctcaa ctaccagatt
201 tggagtgggt tatcaagtgg acaaaatgct tagcttctca
241 ttgatagtca cacctaggta tcctgtcatt ttgtcaaatc
281 tacagatatt aaagtgagca tattattatt tccctcctaa
321 aattttggct tcctcaaaac tactcccata ttcttgttcc
361 tacagcaaca gtataaaaac cttagggtga tcctccctac
401 tcttcctgct ctgttatatt actgagtcct gttactctcg
441 ctttgaaaat attctttgct tacatctctt cattgtactt
481 cctatcctat cacatacatg gactctggcc tctcattctt
521 agtctgcaat aactttaatt atttttctta ccttcagcct
561 cttttatgta gtctaccagt cagagtgcgc atcataaaat
601 aacatgcgga taatgtcatt tctcagtttt ccaatgctca
641 gcagataaag cacaaaactc atgcctttca tattctgaac
681 cccatccttt tacaaccttg tctctcaaaa gttgatagaa
721 ggatcctttt atttcaagaa aatggaatct tacctcactc
761 cccaaactct acaaggaaaa ttaaattgat gaaattctaa
801 aaatagaaaa aattattcac aattttggat tatgcatggt
841 tccttttgtg taacattgta aagtggttgt gctaggtttt
881 ataatgtaaa atctttttat attacactag ctagctagcc
921 ttgtgttcat gaattttcca ttggttttga aatgatcata
961 ttgcagtata ttaaagagat agcaattact ttttatcaat
1001 atatttaaat attgatgatt tttaatttaa ttttcccttc
1041 actctattgg aagtatgtac atattaacaa ctcaggatag
1081 tcttttgact gaggtttaca tttgatagta ccactttgat
1121 aatgcaaagt tatagttctt tattagaata taagtataat
1161 gtgcttgaaa caatactaag ctctttggag aaaaatgcac
1201 aatatttatg gctttatcct gtcactagca agcctgagtc
1241 atggtctaag tggtagattg aagggctgtt gctttttagc
1281 cctgagtatt aacacttatg gtgtttgttt atctttgagc
1321 aagtcactta atctgccagg gcctccgttt tcatatgtcc
1361 taataataat gaaacctctc attacagttt ttcagaacac
1401 ttcccataca ttcccatttg atcatcacaa cagtcctgtg
1441 aggtatgcat ggcagatact actgtcctca ttttactgat
1481 gaggaaaccg aaacttagag gttaaatgac ttgcccaagg
1521 tcacaaaact aataagtgtc caaactttat tctctcattt
1561 agattttgta accattgttc tcgccactaa atacttttat

Exon 1


1601 ctccatatct gccaaatgga gattacagtg atatattact
1641 tgatgttgat attgagatta ttgaatgcaa attagtataa
1681 agcagttcta tagaaattac agcaaatata tgtattgtta
1721 tttctggata taattaagat atatattaac tggctttaag
1761 atttcctaca ctatttaagg ttctttggcc cttgttgatc
1801 ttttgttttt gctaagccat gtgtatctct gagggcctgg
1841 ccagtgacat acagtcattg agttaggcag ctgattggac
1881 tcagcctgtt ctgagtcaca cccagaagca ctaatcaatc
1921 gactggatag tttttaggta agacttgtat gaaacttcat
1961 ttgctttctc tggatcagga gccaagatac tggatttgtt
2001 ttgttttcct ttatgttaca ttttttggcaa ggatgtctta
2041 ttccagttcc cagttggggc attcattcag agtaatattt
2081 accgagcacc ataatctgtg caaacttgtg atcttgagga
2121 agtcatttaa tctcactggt cttcaattga ttcacccctc
2161 tatcatttct taactacaga gtggactaga gtagggtttc
2201 ttatatgctg agtagtttttc tgggggtact gggggggcga
2241 ttggaagagg ctattgaagg cgtaatgccc ttgtaagttt
2281 agggcagtcc ctactccatc ttgaccttac ccatcactgc
2321 tccagagcag ttatattttt atgcatctta tctagccaca
2361 gagactcctg aatctgttgt ctttcctcct caatggtgac
2401 tttttggccc tcttttcatt tatacccaat cctttaatat
2441 cctagattca gaggatcatt tctaaatttc catagtcata
2481 aatggacact atcgctaggt cattatggca taataaaggg
2521 agaaaaggaa aaaaatagga agttaagtaa aagaaatacc
2561 acagaaaaga tgaaaagaag aaataacaaa ttggggagag
2601 aaatgtggaa caagaaggga aacaaatgag aaatacatac
2641 aacgaaggaa aatactttat ttgccctctg acatcacatt
2681 actcttaatt tttataatca tccctaagta tttgtgtata
2721 tgggtttaat tttagtaagc ccagaatttt actaccaata
2761 atttttttca tgaatgtttt gctttgtttt taaaattact
2801 ttcttagctg ttaattgatt tgggatgaat ataaaagttc
2841 ttgtattaaa aattttaaaa gcagattccc ttttttattt
2881 ctttaatcta tattgatgtt taagaaaact gaggttgcta
2921 gtctattttc ctggtaaaaa ttaaattaca tttaccttaa
2961 ttgagatggt cacgcatacc tatctgaaat ttgtgattta
3001 atttttttta agGCGCTGAT GCTACACACA TGGATGGTGA
3041 TCAGATTGTT GTGGAAGTTC AAGAAACTGT TTTTGTTTCA
3081 GATGTTGTGG ATTCAGATAT AACTGTGCAT AATTTTGTTC
3121 CTGATGACCC AGACTCAGTT GTCATCCAAG ATGTTATTGA
3161 GGACGTTGTT ATAGAAGATG TTCAGTGCCC AGATATCATG

Exon 2

3201 GAAGAAGCAG ATGTGTCTGA AACAGTCATC ATTCCTGAGC
3241 AAGTGCTAGA TTCAGATGTA ACCGAAGAAG TTTCTTTAGC
3281 ACATTGCACG GTTCCAGATG ATGTCTTAGC TTCCGACATT
3321 ACTTCAGCCT CAATATCTAT GCCAGAACAC GTCTTGACAA
3361 GTGAATCTAT ACATGTGTCT GATGTTGGAC ATGTTGAACA
3401 TGTGGTTCAT GATAGTGTAG TAGAAGCAGA AATTGTCACT
3441 GATCCTCTAA CAACCGACGT AGTTTCCGAA GAGGTACTAG
3481 TAGCAGATTG TGCCTCTGAA GCAGTCATAG ATGCCAATGG
3521 GATCCCTGTG GACCAGCAAG ATGACGACAA AAGCAACTGT
3561 GAGGACTACC TTATGATTTC CTgtaagtct tggggttata
3601 gtgcttgtca gaggttattt ctgaagtctg tctttccaac
3641 cgtacacctg ggtgaaatgt tcctgacttg tgtaaaatta
3681 aatttcctag acctgcctgc attttagctt cagtttagag
3721 catgtaagat ggctcagaat taagaaaagt gaagtggtac
3761 agtagagaat aaaaacctgt ctctgtgtaa agatgcagta
3801 ggaagatgag agtttttcag tctagaaaga caaagataga
3841 aaggaataga actgaatgtc atgaaatcat taaggattag
3881 gttgctacca gccttcatta gaattaggaa ccctcttaaa
3921 acttaaaagg agtttgactt ggataaaaat atatgacaat
3961 actacttcta ttcttgtata tggagtacaa gaaggtatga
4001 tactgcttat atgatattat gtaatatgtc tcatagataa
4041 tagatgatag atagatagat agatagatag atagatagat
4081 agatagatag atagatagat agatagataa gataagatag
4121 atagatgcta tttatacaac acataccttg gaaaaagggc
4161 acaggcaata ctgttaaaac ataaatgatt ccaaaacctg
4201 atttcaggta tgatggactc attttaggtt attgaaaaaa
4241 agtaagatgt tttggaatac cctcttaatg ttcaacttaa
4281 catctggagg gataactctg tcttcttaca acataaaatc
4321 ctcatttgta tagaccacag gtttgaccta gtgggaaatt
4361 cttacacttt taggagcttt acttcaagca tatacagaat
4401 cacagactat acagggattt aattagcata tattggagcc
4441 taccagcatt ttaagaaata ccttttaaaa gttactgaga
4481 ttttgagagt tggaaaattc tttctgtaaa gagtaagaaa
4521 acaccttcat ttaacacatg ctcattgggc ccctactgca
4561 aatcagacac tgttgtagat agtgcaggca cagctatttg
4601 ccaaattaaa aaaaaagaaa ccattgtcct tggagagaaa
4641 ggtaggttgt cctgtgctgc agaattgtga cagaattgct
4681 gtggagacaa gctggagtag tagctgacct ttgaactact
4721 caaagataca aggaaaaaga tgtggatggg agggagacat
4761 tcctgacaga tagatagata tgtaaaaaag gtcctagtct

4801 tcgctcaggt tatgaatttt aggaaatact tgtatatctg
4841 tatattcttt aacttttag aatacattta tagaaataga
4881 attgctggat caaagggcat acacatttgc gagtctggtt
4921 atgttgccat ggtattttgc ctaatataac tccccatcaa
4961 gtttatgaga aggtacattt tcctgtactc tgatcagtac
5001 agggtatttt cttcgttttt agttgttttc actgtgacag
5041 tgaaaaatat gataccctgt tcaacataga gtttatttta
5081 ggaagataat cctaaaatct ggaataagga tagggggaagg
5121 aaggattact ggcagggagt ggttgaagga ttcatgtatg
5161 agataaagca gttcctaagt accttcagta gtaatgggga
5201 cagaaaggag aagtaatggg gacacaggag tggtcagaag
5241 taaaatgtat gaggaagatt ggatgactga tcatgccatt
5281 gcctgagagg aggaatagat ttggtggaaa gaatgtgatt
5321 taggacctat taaaagtttg agttgtttgt gctgggtgta
5361 ttctggtggg tgtttttcta ggcagctgaa gataagaatt
5401 tggtgctttt gagaatgttt gcagcttaaa aaagaagaaa
5441 tttaggacag aatatatata gactgtactc tatatagaaa
5481 tggatggtgg tccccagaca atgcttgtag aataagaatt
5521 tgaagtcgga ggcctctctg gtggcacagg agttaagtca
5561 cagcactatc aagctatctc tagctactat atcatgagtt
5601 tgatccccaa cctgggagct gacccacatg gcactacaga
5641 cctctcctgt ctcgcctact gctcccctta gccagtgtat
5681 tttagtacat ctgcctgctc tgttcctcac tctgtctagt
5721 gaatcctttt acccctacaa ttctggcctg taccccagtt
5761 cttgtatgca taaataaatc ttaccccaaa atacttttaa
5801 atatatttt aaagattttt gtttatttat gcatatagaa
5841 atggggtaca ggctagaggt gcaggggcgg caagaggact
5881 caccagaaac aaagtgggga gcagaacagg cagactgaca
5921 aagtacactg ctgagggaag cagcggacaa gacagaaaag
5961 gtttattatg tcatatggat cacccagggt ggggatcaaa
6001 cttggactga agtgaccatg gcttggcgtg gcagtgcagt
6041 gctctaacct gctgtgccac cagggaaggc tccccaaaaa
6081 aaactttatt aaaaaaattt ttttactatt gactttggaa
6121 tcaaaccaat tgaatgactt ggaattagca tttctcatat
6161 ttctgcacac cagaaaagaa ctcatctact acattttaac
6201 ttaaaatctg tattgatttc agaaatttac attaacaaga
6241 attttgtggt aagtgtatag aggaactatg gccggacttc
6281 atctttgata gagaatgctg agatgccatt attaacagtc
6321 agtttcaatt ctttatacac aatatgagga gagatggtac
6361 tatctctgaa ttcctgtgat actcattcat cacttcatag

6401 gtcaaagaag aacacgaatg aatgtctaaa tacttcttta
6441 aaatgtcagt aaaccattgg acctatgaac ctaatcaact
6481 tttataaagt ctttagtgtc tctctggagt tcaatttaaa
6521 atgacgaagc agcaatcttc attctttata gtgtcatggt
6561 agatctctct ctctggttct ctctctagct gcctttgaaa
6601 tcaggatgat ctctttagtc ttagcagtct gtatgttatg
6641 actgtcaagt tgattgtcta tttcttgaag acagaaattt
6681 agtcatttat gtatcttctc catcaaacca tatatatatt
6721 tggtgtgaaa tatttgttga atatataaat gctatagctg
6761 aaagtctggt cattatgctt tatccatgct attaatacta
6801 aaattcttcc ctccaagtag ttatttattg actctcagct
6841 tggttctggg cagactagat aatgggggta ctactttatg
6881 gtacttaagc ttttttttt ttnnnnnnnn nnggctctta
6921 attcaaagga ctttttttt tctttgacta aggaacggtt
6961 ttgctcagga attgaactca tgcggaagcg gctttagatt
7001 atcagtgctg tgctttaacc cgctgcgcca ccagggaggg
7041 cctcaagttg tttcttactt ctcctttaat ctgagacaac
7081 tgccagacac ccccagtttc tttgcagttg tcttcactct
7121 cctgaagttt cacttggaag ctggatggac agtcaagatt
7161 ctatttgttt aaaggtcagt ttatacaata gggaaacttg
7201 gaagagattt ggtgatgatg catcagcctt acaacagctg
7241 caggaagtgc ctctgataca aacattatca gcaagatgtc
7281 actttagatc atactgacct ccgcatatct tataattaga
7321 acatttagga tagaagatgg gctggagtta ggggttcttt
7361 atattcttag attcaacttt ttaagtgttg aagtaatttt
7401 ttttgtaat ggctttttaa tcataacata caattcagtg
7441 gttttttttt ttttttttat ttaaagattt atttatttgt
7481 acatacagaa ctggggtaca ggccagagac tcaggggcag
7521 tgagagactt caccagagac agagtgggga gcagaacagg
7561 cagatcaacg aagtacactg ctgaggggag cagtgggtgg
7601 gtcaggagag gtctgccacg ccatgtcgtc tcctctgggg
7641 caggcgggga tcgaactcat gcggaagtgg ctttagcttc
7681 tcagtgcagt gctttaaccc gctgcgccac cagggagggc
7721 cacaattcag tggttcttaa taaattcatt aatttgtgca
7761 actgtcatca cattcaattt tggaaaattt tcatcgcccc
7801 aaggagaaac cctgtgtcct ttagctgtca tgcccatcaa
7841 ccttctccat cccttccaac tctaagcaac ctctaatcta
7881 ctttctatct ttatggattg gcttttctgg acattttgtg
7921 cgaatggaga cctataatat gtggtctttt atgactgact
7961 tcttttatat ttacaaggtt catccatatt gcacatgtat

8001 cagtacttca ttccttttaa ttactgaatt aatatttagt
8041 caatatacat ctatacattt tagttcatgc atcaatggat
8081 gggcatttgg gttatttcta ctttttgact gttaagacta
8121 taatactaat atgggcggag agggggggtgt tttagttaat
8161 tcattatagc aaaaactacc cagacagctt ccaaacatta
8201 ctgcattatt tttctaccta aaagaatgtt aattttcaaa
8241 tttatgttta catgagaaaa gaagccatat ataagctata
8281 atttttactt tttgcctatg tttacttttc ctgtgtttcc
8321 tttattcagc cattattaag ccagcgatga gctaggaacc
8361 ttgctagcag ttaaacagaa aaaaatgctt gtctttttag
8401 agtttatagt catagagggga atgcaaacac aaataaacaa
8441 ttacagctga gtgtgaaaag atgcaacttg atatgcagtg
8481 tatttatttt cgttctattt tttttaaaaa aatttattta
8521 tacatataag aactgggcca aaatcataaa agcccctgga
8561 tccactaaga agcatgcttt actgtttaaa cgtgggctaa
8601 tgataaaggg tttgagcaga ttaataataa tttagattta
8641 tagtttaaag agagaatctt catgtctgtg ttgaaagcaa
8681 agttaagagt aaaatcagaa acagaaacta cagtaaggag
8721 gcttagtcca ggtgagaaac agattttgat catcaacagt
8761 gttggcagaa actgttttca gttacatgtt agcaagtaga
8801 ctttggatta aatttagaat tgaggaagag gaaggaccca
8841 aagttccctc catggtttcc aaaaccattc ccagagcagt
8881 acacgctgac ttgctgacga tctgctctag ggttgttgca
8921 tacacctata ccctggacac tgtagtatgg atggcacccc
8961 tgcagatgtc agggcttgag cctactagtt tcagtctgtt
9001 ggctcccagt taagcacaat tttgttcata gtagacattc
9041 agtaagtaga ttagttcatg gaatggatga gtaaagtcct
9081 gtgtttaatg cataaacttа gaggatcagt gcaattgagt
9121 gatttattat ttatttattc catggatgtg gcaaaatatt
9161 tatttgtccc tttgttctat gagctggtgg ttagatgtgt
9201 atgtatctta ccattatttt tatatataca tcctatcact
9241 tttgtgtttt aaaccatttg ttggtaagcc attgaataaa
9281 attgtggaat tttaaaatta cacacccaca taggagcatt
9321 tctgaaatgt gagaaaagtc ttatttactg aactaaggat
9361 gttttagaga atatgtattg gacaggcttc aaggttattt
9401 ctgagaaaat tcaaaaaggt ttctgttcat acgtctaaat
9441 gtattatctt cattgtaggt actctgaagt ttatttagaa
9481 atcaacacta aagttaatga gtttagagtt ttcagtatac
9521 tgttcaatct gagcattaga tttacaccag gaataaagta
9561 ctcaagtaag gcaaagatgg tatagtcttc ccctctgcac

9601 tacacaagaa agctgggctt aagaagtcta catcttaaag
9641 gtcttcacaa gcaattggag ctttagcttg ggcttgtgac
9681 tcttacttgc tgttaatctg cttaagacta agacctgctc
9721 ataggacagc atacctttga acagatctgt gattactttc
9761 tgagacagag gggcttggga agggaagaga gggttggggg
9801 tgcgagacca tgcctttcct ttcagggtct tgtagaaaca
9841 atgtggtgca tggagacagt gcctacactt agtcagctag
9881 ccttccgtct ggatgccagc cctgcttctc actagctgta
9921 tcagtttgca caaggtactg gtttgtctgg ccttcatttc
9961 cttatttgtg agatggcagt aataccacct gacttcaggg
10001 tggttgtgag agttagtgat caggagtcca gcaggtatcc
10041 agtacagagc ctgcctttca ggaactgcta gctgttctta
10081 gcattgtggt ctttatgtcg gtggcattgg gctagcaaga
10121 ataattctgg atggaggagc ttcactggag accttcttgt
10161 tccatcagga ttcttgggta agggcaggca ttggttagat
10201 tataatagtc tttccttta ccccttgaatt tgcagtgtta
10241 aaaagtcagt tcttggttaa ctatgtttgg tttgctaggt
10281 ttttagacta gtacaatttg gcaggctcct ttcattcatt
10321 atatctctga ttgacactat gctagatatc actgacttt
10361 aggacaggca cttgaatgac tgcaatcctt taattctttg
10401 tctgcagaat ggcactaatt cattaaacct ttgcattttt
10441 ccagggcagt ttcatagatt aaagaagata agatagtaac
10481 tgagttttgg gtatctttta tagactatag aataaaaata
10521 aggtgaaatt ttaaacccaa aaagagggtt tctgaggagg
10561 aagatataag agagggaagt ctaaagaaga attttaagtg
10601 tgatttatct tctaagaaca tttgcctcaa aggcagaatg
10641 ctctctgcaa gccacttagt tctttaggtg cttctatttc
10681 acatctgtaa agtggagata aacatctaga tcatggggtt
10721 gttgtcaccc atttcatggt agaatatagg aatcccttac
10761 cgcagtgcct gagacatagc cgtcgctcat aaactcaggt
10801 gtccttaata gtgtctgtcc atagtgtctg caccttgca
10841 tagcagcctg cccacatgga aacccgtaga tgtcaaataa
10881 tagtttctct tctagttttt cagttgacag ggtttagggc
10921 ttggagtatt aaaagaaagc atgagatgat tgggtagtta
10961 gactcttgca ctcagagtaa gttaggatgg ctcttaattc
11001 aaaggacttt ttttttctt tgtcagcaat actgagagta
11041 caaaaaccag tttcccttca catcggggaa tcttgttaaa
11081 tattcatgta ttcctcttga agtgacttac ttaggatgta
11121 aaaatcttct ctgtttttat ttctaattag ccagtttcaa
11161 gatcttgaca gataacttct agttctagaa tactagaagt

11201 tagtgaccga actcctgtcc actttgtggg tactgatcat
11241 agtctcatag atgtcaaacc taatcagact ggaaagacta
11281 tgtgatccca tttatgtgac gtgttcccaa tagcagttgc
11321 ctgggacttg ggggagagag aagttgggcc tcagtgctgg
11361 ttgttacagg actttctttt ggttagggaa taaaatagtc
11401 gtgatggtat agtacacatg gtgactgtat actaaaacca
11441 tttaactgta acttggaaat ggggaatttt aagttatgtg
11481 atttatatct ccagaatgac agtagtaaaa acgatgtgga
11521 ccaagcagta ataccttaga gccctcagtg ttaacgtttc
11561 tgtttaaaac tgaagtgcct aatgtagttt tacctttaa
11601 gcagttattt actgactgca tgaaacgtcc cccagtatca
11641 tgggtgattt tttttttaa agatttattt atttattcgt
11681 acagaactgg ggtacatgcc agagatgcag agggtgagag
11721 gattcaccag agatagagtg gggagcagaa caggcagatc
11761 gatgaagtac actgctgagg ggagcagcgg gcaggtcaga
11801 agaggtctgt gggcggatca gaagaggttt gcctcaccat
11841 gtggatctcc tccagggcag gcggggatcg aactcatgca
11881 gaagtggctt tagcttctta gtgctgcgct ttaacctgct
11921 gcgccaccag ggagggcccc tcatgggtga tgttttaagg
11961 agaggtgatg gttgcggaat tatgccaaac gctttatttt
12001 tcttaccatt tttacatggt ggtgaaatac aaatcaccaa
12041 ttttaagtcc atagctcagt ggcattgttg cattcacact
12081 gttgtgcaac cattaccaac atcaccctcc acaactcttc
12121 atcttgcaca atcaaagttg tacctgtaac aacttagcct
12161 ccccctccac ccatcaccat tttattttct gtctatagct
12201 gccactactc gaaatgcctt acattggtgg aatcacacaa
12241 tagttatctt tctgtgactg aattatttca cttagcatca
12281 tatcttcagg acttgcccat ggtgtagcct gtatcaaaaa
12321 cttcattcct ttttaagctg agtaatattc cattatgtgg
12361 atctaccaca tactgctcct ccagagtttt attgtaaagc
12401 agtaatcctt tgccttacct actccttgct tcccttccct
12441 ctctccccat ataactattt ctctttagtg gcacttatag
12481 tattttaaat aatggacttt cttaatttaa aaaaattttt
12521 gttttatata ttcattcttg actttctact gtgatagatg
12561 tcgatttttа ttctagcctc ctttccccac aactgcttag
12601 gatagttatg atacattttt ttattttact tttcaaaaat
12641 tgagtcatat gttaatatct ttatttcctg ttccatcata
12681 tggacagtgt tccttaactc ctgcttttta agataaggat
12721 attaatattc ttatcctcct tcctttaagt tgttctacct
12761 gccttccaac ctgtttgtgt tttgttgttg aagctacatt

12801 acattctgta accattagtc ttccagtcaa gttgaaactc
12841 agtatgcagt gcttacatta ttttgattga gaatctattg
12881 tttgcagcat aattattgat cactttgatt atatttatgt
12921 ataaagcttt tttccctcct agaaatgtaa ttgccttttg
12961 ctctatttct aagaaagatg cccctctaaa aatttttttt
13001 ttctaaagat ttctttatgc atacagaatt ggggtacagg
13041 ccagaggtgc aagggcagtg agaggattca ccagagatag
13081 tagggagcag aatagacaga tcaacgaagt acactgctga
13121 ggggagcagt gagcgagaca ggagaggtct gctgcgccat
13161 gtaggtcacc tccaaggtgg ggattgaact ctcactgaag
13201 cagctgcagt tagcttttag tgctgtgtcc taacccattg
13241 caccacccaa gcttctttct ttatgctccc tcagaagttg
13281 ctttctaggc ttattgtaat agacttttt cctgaaattt
13321 tccttttctt acttcctgag atttgtcatt ttctggatac
13361 ctgtgccacc tttttttctt gaattacccc cttattttac
13401 tagaccacat aataaaatga cttcctaagg aatataaatt
13441 tgtttctaag aatatttta ttctaccta aactcagttg
13481 acagtttggt ggggcacaga attctaaatt gaaattaatt
13521 ttcccttaga atttgtgtta acatgtttgt cttcctgcat
13561 tcagtggaag acaatttagg agcctgttgc cattttgatt
13601 cacattcctc cgtaggtact gcttttattt ggaaaagctt
13641 taaagctctt ctttttcag atcctgctga tctgcatctt
13681 atgttagtgc ctaattaggt gtcttttttc atacattgtg
13721 ttcttgtgca cggtaaattg tttgaggatg catgttcttt
13761 agctttctga aagaaaaaca gcaacagtca atcgcagcag
13801 ttttctaagc ctgtgcattc cagtatggtg gttactaagc
13841 gacatgtggt aactagattt aagtttcaca aaaatgaaat
13881 gatttaaaaa aaagatacaa aagttattat ggataacata
13921 gatacaggag acaggcaaag gacatacaaa tacggaattc
13961 cagaagaaga gaataaccaa tgaattagaa gtatatttgt
14001 aataaagaaa tttcctgaaa tacagactcc ttcttattcc
14041 agtgtcttaa atgaatgatg aaagaaccaa gtcataagca
14081 tggtgatgtt gagtgaactt tagggataaa gtattgtagg
14121 ggcatccaag caaacataca ctgggctgaa acagtagaac
14161 tatgtctgca aagttctgag aaactataac ctaagaactt
14201 tatatccagc cattttgttc ctcaagttaa agccatcaga
14241 ctaacatttt cagatattca ggaactatag tatctattaa
14281 ccattcttta cagttctatt tttaatgaat ctagcagacc
14321 aaaggatcaa tttataaaac tctagaatgg ggaaacagta
14361 ataaaaagat cctcagagta agtattgaat taatttaata

14401 taaaactata aaatatttgg caagggaagg tatgtaaact
14441 ctgttatctt tccacatctt tcatagcagg gtgacaatag
14481 atagcagctt tatttatcta tttttattat ggcaacatag
14521 gtttacaatg ctgtatatta aatttgaggt gtaggaatgg
14561 agcttccatg tgacccagca atcccactcc taggtatcta
14601 cccaaaagac acaaagacac aaaaacactg attcaaaaat
14641 ataccttcac ccccatgttt agtgctgctt tatttacaat
14681 tggaaagata tggaaacaac ccaaatgccc aaaagtggat
14721 gactggatta aaaagatgtg gtacatacac acagtggaat
14761 attactcaga aaccaaaaaa aaaaaaagaa gaaaaacatg
14801 cttttgcag caatatggat ggagctggga aagatcatgc
14841 ttagtgaaat aagtcagaaa gaaaaggaaa aataccagat
14881 ggtctcactc ataacaggag tataacagac caagcgagga
14921 aaaataaata aataaaaaaa ataaacagct atataccaga
14961 gacattttat gggggtggag taggtaatgg gatcagaaaa
15001 agtgggggga gggaaatgca ctggactatg gtgatggcac
15041 acaggtgatt tagtgatggg aagtccatgg ggattacaac
15081 tgagatggaa aaatgcacaa aaggatgtaa aaataaactt
15121 acattaattg ccctccctgg tggcgcaatg ggttaaagca
15161 cagtactgag aagctaaagc cacttccgca tgagttcgat
15201 ccccgcctgc cccggaggtg accacatggc atggcagacc
15241 tctcctgacc cgcccgctgc tcccctcagc agtgtacttc
15281 gtcgatctgc ctgttctgct ccccactctg tctctggtga
15321 agcctcttac cgcccctgag cctctggcct gtaccccagt
15361 tctgtatgca caaataaata aatctttttt aaaaaaaataa
15401 aataaacttа cattattaaa aatatttttt gaggtgtagg
15441 ttcacacttc ttcagttatt tgttaccaca ccagacctac
15481 caccacggca accataaccc tcccataatc cattggaccc
15521 cctttccttc tacctgtggt aacttcagtt ctacagtcaa
15561 gttttctagt gtttgttttg gtttggtttg gtttctctgt
15601 accacataga atgaaaccat ccagtatatg tcctgctgac
15641 ttattttgct taacataaat ttctgttgtc aaatggcaac
15681 atccatcctt tttaacagct gtgtagtgtt ctggggtgtg
15721 tttgtgtgtg tgtgtatgcg cgtgcgcacg cgtgcctgtg
15761 tacatgcacg tacgcacatg tcacacttat ttatccattc
15801 atctgttatt ggtcacttag gttttctcca taaccaggct
15841 attgtaaata ggactgtttt gtatatagtg tacttgtaga
15881 tgtacacata caaacatccc ttcaagttag ttttttgtat
15921 tctttaaatt tatagtctag aaaagtagaa ttgctggatc
15961 tttttttaa taataqttta ttttacatc attttqtccc

16001 ttttttcccag ttgtaatccc cgtggacttc ccatcaccaa
16041 atcacccgtg tgccatcacc atagtccagt gcacttccct
16081 ccccccactg tttctgatcc cattacccac tccaccccca
16121 taaaatgtct ctggtatata gttgattatt attattatta
16161 ttattattat tattatctcg cctggtctgt tatactcctg
16201 ctatgagtga aactatccgg tatttttcct tttctttctg
16241 gcttatttca ttgagcatga tcttttccag ctctatccat
16281 gttgctgtaa aaggcctgtg tttatttttt ttggttgctg
16321 agtaatattc cattgtgtat atgtactaca tctttttaac
16361 tcagtcatcc acttttgggc atttgggttg tttccatgtt
16401 tttccaattg taaatacgga ttttttttt ttaaagatct
16441 ctatttatgc atacaagaac tgtgtgcatg ccagaggtgt
16481 gggggagggga gatggggcat gagaagattc accggagaca
16521 gagtcaggag cagaacagac agatctactg aaatacactg
16561 ctgaggcagc tggagagata ggagaggtct gctgtgccct
16601 gtgagtcaaa cttattctgc tgtggctggt gatagcttga
16641 taatgttgag actttaaccc ctgcacagcc agggaggcct
16681 gaattgctgg atctcatagt agatctactt ttaatatttt
16721 gagaatggat attgttttta atatggattc atgtattata
16761 aacaaaatat gacttaacaa attcataata gttcttgtct
16801 ttagaggttt aagaactaat aaatatttac ttcaagatca
16841 gcattcatat atttaacaaa caaatgtatg cttttttttt
16881 tttaagatta tttatgtgta cagacacgga tacaggtcag
16921 agttgtgggg aggggtgaaa gggttcacca gagacagaat
16961 ggggagcata acaggtggat tgatgaagta tactgctgag
17001 gggagcaaca ggtgagacag gagaggtctg ctgcaccatg
17041 tgggtcagct ccctggtgga agattgaact tactctgaag
17081 tgactgctca gtggctagct ttgcagtgct gtgttttaac
17121 ccagtgcgcc accagggagg gccctatatg cttattctgt
17161 taaataatgt ttgcacttgg ggatacatca gtgaacaaag
17201 ctgatgaaag tgcctttccc ttgcggatat attagaagta
17241 tgcaggcata aattagtgat acagtgcaag tgggaacaga
17281 gacttaattt ggagatgtct aatagatatc aaacacatgt
17321 gaatctgaac ttggtagatt gggcttgagt agagatgtca
17361 aggtgggagt tggcatatag atgagatttt agctgtgaaa
17401 caaatgagat tcaccagagt gattatcgat atagtgtttt
17441 caacggtcag agacagaagg aagctcttgc aggactgagg
17481 agttgggcac ccagtgagag aggaggagac aaggagactg
17521 ggtccttgag caagtgagga aagtgaatct cagtggaagg
17561 tgcagccaga agtagtcatc agaagagctca gcactgagaa

17601 ctatagaaat ggaggtggtg gagaaatgga atcaagagtt
17641 tttgtttgtt ttaaagtggg tgatagagta gctggtttgt
17681 gtgctgatga aacggattca gtcaagggaa gtttggtggg
17721 aagaagaaga gtaagagggt ctggagcaat aaaaagtagg
17761 cagacacagc aagaacacag atggtttgtc tggcctggtt
17801 agtgactggt agatagctaa gtgtttcttt cagttctcag
17841 tgggtaagaa gtagtaacat catcaactga gaatgaggac
17881 aggaaagtga ctttagaggt ctgaggagaa agtgaaaaat
17921 ggtggaagta gcaggtaagg tccacttctg gttgagtcat
17961 gaatttaaat tgaaactgtg tgtgtccaaa cagatttagc
18001 tgcgtagagg aaaattgcat ttagccaagt ttgtggttct
18041 gccaaagcag tgtgatgaaa tgaaaagagg gggtagggca
18081 aggaactgga gagcagtggg atggaattga ccataaaatt
18121 aaagttgatt tgttgatcag ttcaaataaa ttctaataga
18161 ctaatttttg tctttaaaag gcatgttatc atgaaaaaaa
18201 aaaaagctca ttctgtcctt ttatttataa ataggataga
18241 ggtatatctg aataatgact tataaattct attgatagtt
18281 tcaggttgtg acttcttttt atacagataa aattattttt
18321 tcctagtaat aagccatgta acatgtatat gggggggtg
18361 tgtgtgtgta tgtgtgttta agcaaagttg aaggagcata
18401 aaggagtgga aataaacaca catgtttaag ttcagggcac
18441 atgggctagt ctggggctgc cagggaagag agactggagg
18481 cagaaggcag tttgtctggt caattgctgt tctgtgtgca
18521 gtgcatgttt tacttggtag aacgatgggt cagtgttctg
18561 tgctggcgta aggcaagtga gtgttttagg aaagatggct
18601 aaaggcagtg tctaggggtg tttggtggta acctggagac
18641 agatcaaata ttaaggatgt ctcttttta gatgttaatt
18681 ttgacaaaat gcaagtcatt acaaatacca tgtagcttga
18721 cagaattata ataagagtta tctagaagtt tagctaggtg
18761 agaatagcaa aacaatataa catcacctat tttatgagcc
18801 tgtcatgagt agtaaatgat aatttttta aagactagtc
18841 ctcctgcatg gtgtgtagtc agtaaatgtt cttataaatc
18881 ttatcatgaa gaccatcact ccttattacg cagatttttc
18921 atgaaatgtt cttcgctcag ttttgctgat ttttaattag
18961 atcacattcc ttaaacagtt atttgacacg gtatacattc
19001 ccaccccaat atttcttttc cagaaactac tgagtctact
19041 gaaatattaa tgagctctat ttttagacac catgacgtct
19081 attagcattg caatttttag ccttcaggga attttactga
19121 aggtgtctta aagcagtgta atgtatacct atatgatact
19161 ataqttttqc tttactaatq aqattqtaaq tcaaattttt

19201 attttacttt tgaaatttaa agctggatga gttataaaga
19241 atcttatgta aagcttggaa atggatgcca atttatttta
19281 aaagagacag taaaacatta tttactcata atgtggaatt
19321 tctggttgtt ggggttattg gttttttggg tttttttaaa
19361 ccaaattttg tatagtggat atatatttaa tttaaaaacc
19401 aaaagataca aataagggtc tagattgtaa agacaaacac
19441 atcaccttct ctcccacccc tcattttact agttcactga
19481 ctttgggcaa gtcacatacc atctcagggg tctttatttt
19521 attatctgca aaaggtgaga atcaaatgaa ataaaagtac
19561 atagaaatgc taaggaaatt gcaaaatgcc atgaaaatgg
19601 cagttattat tagaagtagt cttaatttcc agaaattatg
19641 ctcttctatt ttggagagac ttattctaaa gctctagttc
19681 ttcattacta tattagaaga aggctcttat gcaaaaatca
19721 tacttttgat gactcagttt tatcaaaacc tttgtctgtt
19761 tttagaaatt tatttttaat ggtagcttta acatcacagt
19801 aaaattttaa gtacttaaca gtatcatcca agttttattt
19841 taaaacaaat aatttctgca atatttcata tattcacaat
19881 aatttctgag gttgagtcag tggggcacat tttcaaggtg
19921 aaaatttgaa ggtgaattct actttagcct aatacagaat
19961 agtttcattg ggttcattta agcctgtaga tttgaaatgt
20001 tttttttca cacatgtatg aaaacagaga gttcttttat
20041 ttcaagttac agttaaatac ctgacatttt gcaagctcca
20081 aaatgattag attttccagc atcagaggtt ttatcttgtg
20121 agtaccaacc aaccaacgta ggcgaaaaca gctagactga
20161 ccttgttctt tgtaaacaca aggacatttg attttccagt
20201 atctgttaga atcagtttta acaaaccttg ccaaaagagg
20241 ccaagcgtca tcaacaatat caaactgttc agaaatgaat
20281 tctctgaagt tgtgacagac ttattcagga ctctcctgct
20321 aattttgaga gcttcatcct tctcatgcta cttcaatttg
20361 aaattccaaa gctatcttta aggaaataaa attggctcga
20401 tcagcttgat ttgtgataaa aatatttaac tttgaaaaaa
20441 gtttaagtac acatcaggct agaataagac tataatgcca
20481 gttcttcccc ctacaaacaa tttttataaa ccacagttaa
20521 ggtacttgaa tttaggtagg tatattcaga gttcagagac
20561 ctctatttgt ttttttaaac cacaaaaaat ttaaccagct
20601 gttcctttgc atcttcatgt cgttagcttg tcctggatca
20641 gctcattttc tatagcttct ctagctctta accttgcctg
20681 ccattctctc actagtaacc acagcggtca ccccttgttt
20721 catacctatg gttgaagcca ccagtcatca agtcttaaat
20761 ttctgtattg caaccttctc cccttcccta taaataatat

20801 aactgtattc tcatccaggt ttacctactc aaatcttgga
20841 aggatagcaa gtttctcctc ctgtttaaca aatatatctt
20881 catctttgtg cctagtttca tcttcctcct tttcctctga
20921 gatcttgacc caggaatgtc accatgccct ctggcttcag
20961 cctcttattt tctactatat atttttctc atagattctg
21001 tatacttgag tctttttaa ctagaaaagt ctgttctctg
21041 cagcccattg tagtcactgc tctcttgttt tccatttta
21081 aattcaattt cttgaagaaa ttcatacttt tctgttgaat
21121 cctaaacccca ctttggtctg acccatacag cttcaccata
21161 ccacttttca tcaggaacat gagcgacttc tggatacaac
21201 agtcagtgta gacatttcct cctgctgcat tattgccata
21241 gccttgttag tgctattaga acttttaacg ccctaacttc
21281 cctttctac catcctctcc tttctcactt tcctttgtga
21321 actcttgtca ctcactcaca ctctgacatt cccaaggatt
21361 ccctcagcct actttcttct cacactgcag attctcaccg
21401 tgtggtctca aggctcagtt aacacagggt tgatagtgcc
21441 caaatctgtg tcagccttca attttagggc ttgaatattg
21481 ggcagctgaa ttcacatctc agtcattcat tgagtcattg
21521 cttcattcgt attctgagca tatactctca ttggactctg
21561 tcttggaaga ctgaatgatg aaaatacagg ttccactttg
21601 tcaatgtgaa agattgtggg gggacagagg agtacccaaa
21641 tagtgacaat acagtgaact gtatactctg agaagaacca
21681 acagtgggtg tcctgagcac aatgatagga acagttgctc
21721 agcctggggg cttttagggt tgttagcttc tcagactgtc
21761 tattaattta ctaaacttat agaagaaatc cttttcttca
21801 ttctctttta aacttttctt ctccagtaat tcaatcacct
21841 aaatcagaaa ccacatagtt accttggctc catggtcaca
21881 ttcagttttt ttaagtttta tttatttatg cttacagaac
21921 tgaggtacaa gccagaggtc caggggcagt gagaagtttc
21961 accagagaca gagtggggag cagaacaagc agattaatga
22001 agtgcattgc tgaggggagt agcaggcgag acaggagagg
22041 tgtgccaaca ccatgtgggt caccttcggt gccagggatc
22081 gaactcacac tgaagtggct gtggctcgct tgcagggctg
22121 cgtcctaacc cattgcgcca ccagggaggg cccccacatt
22161 cagtttttaa gcagacacca agtcctgctc atttcgacct
22201 caaggtagaa ctcagccctt tggcctgtct gtgtcatctc
22241 tgtcctgatg tgatgccttc tcatttgttg cctgcattat
22281 tgcagtagac tcacaagtgg gatctctgct ttatgtcttg
22321 atcactttgg atttactttt ctcaaaggca gaaggatatt
22361 gtagcaaaaa tgacaaagat ctttgtcaag cactgcttga

```
22401  aagtttctga ttacctacag cagaggttgt tctcaaacct
22441  ctgatcagaa tccctgggga caggagtatc cccatctagc
22481  tgtatatata ttaggctccg ctcctggaga ttctgattct
22521  gtaaaatgga tgtgctgggg cgtgtgcttt tgagttccta
22561  agagattttt aacacctgtc cagatttaga aactactgac
22601  ttacttggat caagtctcag tccctctact tggtgtacca
22641  agaccctctg tcacttgact taagcttagt agtcagccac
22681  actctaccat agtcttcgtg tctacttgtg taagtaaaat
22721  cattttatgg aaatcaagtt agaggtcata ctagtaaaaa
22761  aaacaaaagc tttctgcagt caataacttt tacttaagta
22801  tgtactataa ttattagtaa tagggttctt gccttttag
22841  TGGATGATGC TGGGAAAATC GAGCATGATG GTTCCTCTGG
22881  AATGACCATG GACACAGAGT CAGAAATCGA TCCTTGTAAA
22921  GTGGATGGCA CTTGCCCTGA AGTCATCAAG GTGTACATTT
22961  TTAAAGCTGA CCCTGGAGAA GATGACTTAG gtaagaggaa
23001  accttcagtg tattgcacat cctgatgtaa ttgctttaat
23041  ctgattgctt tcaggtggct tagattgaga acatatgaaa
23081  tcttttctga agataatttt taagcaggga ttccaatgta
23121  gcagtggaaa agtatgtaaa gtgtattaag tgaaacagat
23161  agaaatgaag ttttcacata aattcttgga gctatcgttt
23201  tccctgctag GCGGCACAGT AGACATTGTG GAGAGTGAGC
23241  CTGAGAATGA CCACGGAGTT GAGTTACTTG ACCAGAACAG
23281  CAGTATTCGT GTGCCAAGGG AAAAGATGGT TTATATGACT
23321  GTCAACGACT CTCAACAAGA AGATGAAGAT TTAAgtaagt
23361  aggtggcctt tttgtgggag aaaattttat gttcctgtag
23401  ggtttttttt ttaaagattt attttatttg tacatacaga
23441  actggggtac aggccagagg ctcagagggt gagaggcttc
23481  accagagaca gagtggggag cagaacaggc agatcgacaa
23521  agtacactgc tgaggggagc agcgggcagg tcatgtggtc
23561  acctccgggg cgggcgggga tcgaactcat atggaagtgg
23601  ctttagcttc ttagtactgt gctttaaccc gctgcgccac
23641  cagggagggc cctaatatag gcagtattgg aaagaaacaa
23681  gcattattac agattattat gcagttatag tccatgggaa
23721  acattagtga ctctttatgc ttagaatgtt gctttgaaaa
23761  cagatagtga tgtcctcttt ggtggcgcaa gggttaggat
23801  gcagtactca gtactgcaag ctggtcacaa ctgcttcagt
23841  acaagttcag tccccgaccc cggaggtgaa cacatggtgc
23881  ggcagacttc tcctgaccca cccgctgctc ccctcagctg
23921  tgtactttct cagtctgcct gttctgctcc ccactctctc
23961  tagtaaagcc tgtcactcac tacccctaca cctctagctt
```

Exon 3

Exon 4

```
24001  gtaccccagt tatgtttaca taaataaata aatcttttc
24041  tttttaaaaa aaacagatca tgaaagccaa gaaatgaatg
24081  atttgttcaa caagtattta ttgaatgctt actgtaaaaa
24121  agctatgttt tagatctata acactgaact tgctttctat
24161  atatctgtaa accgagtata aaagtaaaat atattctcgt
24201  gatctctgta aacccagtca ccaacgttta tgtaaactct
24241  ttctttttg attgctataa ctgcatgata acatcataga
24281  agttttaggt tggtatcatt ggaaagcagc agtgtgaaaa
24321  ttatattaag ttaaccaccg gaataatttt gtgcttagAT
24361  GTTGCCGAAA TTGCCGATGA AGTTTATATG GAAGTGATTG
24401  TAGGGGAGGA GGACGCAGCA GCGGCAGCGG CTGCCGCTGC
24441  TGCACATGAA CAGCAGATGG ATGACAGTGA GATCAAGACC
24481  TTCATGCCGA TAGCGTGGGC GGCCGCTTAC Ggtaggtcac
24521  acagcagccc ccaggactga catggctctt ggatagtgtg
24561  tgaattcgtg attgaaaaat ggtttctgta ggtatcagaa
24601  atatgaacta ctagtgtaat gtgaagatta tttgatttga
24641  tcgagagtag aggaaggtta taaagtattt tttgtcttta
24681  ttttcaacag gaaattcctt catgtgtata ttatgtagaa
24721  ggaagtactg caagaagtat acaggctttc aagctgaaac
24761  tgcatcttga ttatatgtgg cccatgactt tataactttg
24801  ggagataaca agggacttca tggggcaaaa tgctgacaag
24841  tgttttttag atttgaagtt gggcatgtcc tgtgattgca
24881  ctataataaa tgtgctcata aaacccattg cctagaatgg
24921  ggacttcttg tcattcatga gtatcatggc ttacttttg
24961  ttgttgtagt taataaagaa tacctaattc tttataattt
25001  ataatactgt ataattttgt tttttaatgc acactgttag
25041  GTAATAATTC TGATGGAATT GAAAACCGGA ATGGCACTGC
25081  AAGTGCCCTC TTGCACATAG ATGAGTCTGC TGGGCTCGGC
25121  AGACTGGCTA AACAAAAACC AAAGAAAAGG AGAAGACCTG
25161  ATTCCAGGCA GTACCAAACA Ggtgagggca cacgaggtcc
25201  acagtgcagc gtgctctgcg agctctcaga tgaaagtaga
25241  atgtattcac agaggtgttg tgatggcatt ttagctgtta
25281  gaccacatgt aacttgtgta atgaatttga aactataatt
25321  ttaagagttc agtgatattc atgaatgatt tctttggata
25361  aaatgaaacc gaatggatca gacatggatt ttgaaattta
25401  aatctgaggt ggcttttatg tggaaatgaa aaaaaaatcc
25441  cttaaatatg ttgcaaacat taacctttt agaaaactaa
25481  atagagattt aaaaaatgtg attatgtcag cataaagcag
25521  gcataataat gtacagagca gcaggacaag tacttgagtg
25561  catgtgactt aatcctatct attcttatct actgattttc
```

Exon 5

Exon 6

```
25601  atattttaaa actgggcata tttaagtttt gctattcatt
25641  gacaactata tatttttggt aaccaaagca gaagtgttaa
25681  ttgcaaaatt tacagataat tctctaccaa tttatgaatc
25721  gatgccactt taatcactag ttggtagcat tacatattat
25761  ttgtttttta agctggaggc attaagattg gtacagtcaa
25801  attgagtgaa gtttccagac caatttctct ttctaatgag
25841  attgttggaa atggtaaaag gaataagtag caaatagctt
25881  aacttagtat tacttgccca gtaatattta gcaaacactt
25921  taaattcttg aagaaatcaa aatagaactt ggtttgatca
25961  ctcatggtcc tttcttttcc ctttcttagC AATAATCATT
26001  GGCCCTGACG GAACATCCCT TGACTGTCTA TCCATGCATG
26041  ATTTGTGGAA AAAAGTTTAA GTCAAGAGGT TTTTTGAAAA
26081  GGCACATGAA AAACCATCCT GAACACCTTA CCAAGAAGAA
26121  GTACCGCTGT ACTGACTGTG ATTACACTAC CAATAAGAAG
26161  ATAAGTTTAC ACAACCACTT GGAGAGCCAC AAGCTAACCA
26201  GCAAGACAGA GAAGGCCATC GAATGTGATG AGTGTGGGAA
26241  GCATTTCTCT CATGCTGGGG CTTTGTTTAC TCACAAAATG
26281  GTGCATAAGG AAAAAGGAGC CAACAAAATG CACAAGTGTA
26321  AATTCTGTGA ATATGAGACA GCTGAACAAG GCTTATTGAA
26361  TCGCCACCTG TTGGCAGTCC ACAGCAAAAA CTTTCCTCAT
26401  ATTTGCGTGG AGTGCGGGAA AGGTTTTCGT CACCCCTCGG
26441  AGCTCAAAAA ACACATGCGA ATCCATACTG GGGAGAAACC
26481  GTACCAGTGC CAGTACTGCG AGTATAGGTC TGCAGACTCT
26521  TCTAACTTGA AAACTCATGT AAAAACTAAG CATAGTAAAG
26561  AGATGCCGTT CAAGTGTGAC ATTTGTCTTC TGACTTTCTC
26601  AGATACCAAA GAGGTGCAGC AACATGCTCT TATCCACCAA
26641  GAAAGCAAAA CACACCAGTG TTTGCACTGC GACCACAAGA
26681  GTTCGAACTC GAGTGATCTG AAACGACACA TAATTTCAGT
26721  TCACACGAAG GACTACCCCC ATAAGTGTGA CATGTGTGAT
26761  AAAGGTTTTC ACAGGCCTTC AGAACTCAAG AAACACGTGG
26801  CTGCCCACAA GGGTAAAAAA ATGCACCAGT GTAGACATTG
26841  TGACTTTAAA ATTGCAGATC CGTTTGTTCT AAGTCGCCAT
26881  ATTCTCTCAG TTCACACAAA GGATCTGCCA TTTAGGTGTA
26921  AGAGGTGTAG AAAGGGATTT AGGCAACAGA ATGAGCTTAA
26961  AAAGCATATG AAGACTCACA GTGGCAGGAA AGTGTATCAG
27001  TGTGAGTACT GTGAGTACAG CACTACAGAC GCCTCAGGCT
27041  TTAAACGACA CGTTATTTCC ATCCATACAA AAGACTATCC
27081  ACACCGATGT GAGTACTGCA AGAAAGGGTT CCGGAGACCT
27121  TCGGAAAAGA ACCAGCACAT AATGCGACAT CATAAAGAAG
27161  TTGGCCTGCC CTAAcaagac tactgcagac atttgtagag
```

Exon 7

27201 atattggcct tgaagcagaa aattatttga aagccaattg
27241 gtctcgttca catacaatac tgtatattga tttatgctgt
27281 gtacaaatag agttactgct tctagttgac ttttttttta
27321 cattttgttc agtagtgtgt tctgaattct attcagtttg
27361 tttaataaat ggggaaaagc agcaacaagc aggttgcttt
27401 taaagataat ccctgattct atactgaatt tttctatctt
27441 agaagtttta tatttattta attatttacc ttgcttacc