RESOLUTION OF PHYLOGENETIC RELATIONSHIPS AND CHARACTERIZATION OF Y-LINKED MICROSATELLITES WITHIN THE BIG CATS, *PANTHERA*

A Thesis

by

BRIAN WILLIAM DAVIS

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 2009

Major Subject: Biomedical Sciences

RESOLUTION OF PHYLOGENETIC RELATIONSHIPS AND CHARACTERIZATION OF Y-LINKED MICROSATELLITES WITHIN THE BIG CATS, *PANTHERA*

A Thesis

by

BRIAN WILLIAM DAVIS

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	William J. Murphy
Committee Members,	Terje Raudsepp
	Tiffani L. Williams
Head of Department,	Evelyn Castiglioni

August 2009

Major Subject: Biomedical Sciences

ABSTRACT

Resolution of Phylogenetic Relationships and Characterization of Y-Linked Microsatellites within the Big Cats, *Panthera*.

(August 2009)

Brian William Davis, B.S.; B.S., Texas A&M University

Chair of Advisory Committee: Dr. William J. Murphy

The pantherine lineage of cats diverged from the remainder of modern Felidae less than 11 million years ago. This clade consists of the five big cats of the genus *Panthera,* the lion, tiger, jaguar, leopard, and snow leopard, as well as the closely related clouded leopard, which diverged from *Panthera* approximately 6 million years ago. A significant problem exists with respect to the precise phylogeny of these highly threatened great cats. Within the past four years, despite multiple publications on the subject, no two studies have reconstructed the phylogeny of *Panthera* with the same topology, showing particular discordance with respect to sister-taxa relationships to the lion and the position of the enigmatic snow leopard. The evolutionary relationship among these cats remains unresolved partially due to their recent and rapid radiation 3-5 million years ago, individual speciation events occurring within less than 1

million years, and probable introgression between lineages following their divergence.

We assembled a 47.6 kb dataset using novel and published DNA sequence data from the autosomes, both sex chromosomes and the mitochondrial genome. This dataset was analyzed both as a supermatrix and with respect to individual partitions using maximum likelihood and Bayesian phylogeny inference. Since discord may exist among gene segments in a multilocus dataset due to their unique evolutionary histories, inference was also performed using Bayesian estimation of species trees (BEST) to form a robust consensus topology. Incongruent topologies for autosomal loci indicated phylogenetic signal conflict within the corresponding segments. We resequenced four mitochondrial and three nuclear gene segments used in recent attempts to reconstruct felid phylogeny. The newly generated data was combined with available GenBank sequence data from all published studies to highlight phylogenetic disparities stemming either from the amplification of a mitochondrial to nuclear translocation event, or errors in species identification. We provide an alternative, highly supported interpretation of the evolutionary history of the pantherine lineage using 39 single-copy regions of the felid Y chromosome and supportive phylogenetic evidence from a revised mitochondrial partition. These efforts result in a highly corroborated set of species relationships that open up new avenues for the study of speciation genomics

and understanding the historical events surrounding the origin of the members of this lineage.

DEDICATION

First, to my wife, my life and very best friend, Crystal Nichole Davis, who supported us financially during my years of full-time study, and who shared equally with me all the emotional and financial burdens involved. Her unwavering encouragement and understanding is unparalleled. Her interest in genetics drew me into the field, which firmly grabbed hold and never let go. For these and innumerable other small things I am eternally thankful.

To my parents William E. and Dava Davis whose encouragement to pursue academic achievement and knowledge from a young age planted the seeds of interest that blossomed with their help as I matured. The concrete work ethic and loyalty of my father and my mother's care and penchant for detail are forever ingrained in my being.

For my son, Christian St. John Davis, who has become a driven and intelligent young man, bent on conquering the world through any means available. I constantly see such immense potential and am anxious to see the impact he will make on this world. I am prouder of him than he could possibly know. The sky is not the limit; it is only the very beginning. Fly even higher.

And finally, to my grandfather, the late James Everett Davis, Jr. He, with my father, worked long hours to build some of the first greenhouses on the Texas A&M campus. His colorful stories as a technician in the peanut breeding programs at A&M and his work as an apiarist initially excited me towards the potential of scientific discovery as a young man.

ACKNOWLEDGEMENTS

I would like to express sincere appreciation to my mentor, William J. Murphy, for tireless ongoing discussion and support throughout the course of this project. I would like to thank the remainder of my committee for their support and fruitful discussion: Tiffany L. Williams and Terje Raudsepp. In addition, I thank Jan Janečka for invaluable work on microsatellite genotyping, Alison Pearks-Wilkerson for the feline cDNA library, and Gang Li for his work on the cauxin gene. For their unselfish donation of time and knowledge that helped me derive conclusions and strategies throughout this project, I am indebted. I would also like to express my gratitude for the funding I have received from Texas A&M, the National Institutes of Health and Morris Animal Foundation.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

LIST OF TABLES

TABLE Page

TABLE Page

CHAPTER I

INTRODUCTION: *PANTHERA* AND THE Y CHROMOSOME

Historical and Modern Status of Genus *Panthera*

The cat family Felidae is reaching a critical point in its history. Almost every one of the 38 species is denoted as endangered or threatened by international endangered species monitoring bodies such as the Convention of International Trade of Endangered Species of Wild Fauna and Flora (CITES), the International Union for Conservation of Nature (IUCN), and the U.S. Endangered Species Act. There is at least one population from every species of the Felidae on Appendix I or II of CITES or on the IUCN Red List of threatened or endangered species. Entire species are on one or both of these lists [1]. This case is especially pronounced for the great pantherine cats, all of which possess protected status. The pantherine lineage consists of the five big cats of the genus *Panthera, P. leo* (lion), *P. tigris* (tiger), *P. onca* (jaguar), *P. pardus* (leopard), and *P. uncia* (snow leopard), as well as the closely related *Neofelis nebulosa* (clouded leopard), which diverged from *Panthera* approximately 6 million years ago [2]. Topping the list of threatened great cats are the tiger, the largest pantherid, and the snow leopard, whose range spans the highest altitudes. In 1998 there was estimated to be between 5,166 – 7,436 tigers (*endangered*) in the wild [3], many more surviving only in captivity, and ever

This thesis follows the style of Genomics.

decreasing wild populations [4]. There are even fewer of the elusive snow leopards (*endangered*), estimated to be between 4,500 and 7,350 in 2003 [5]. Close behind are the 23,000 lions (*vulnerable*), 50,000 jaguars (*near threatened*), and 300,000 of the most prolific pantherine, the leopard (*near threatened*) that are still present in the wild [6]. The clouded leopard (*N. nebulosa*) was recently split into two species based on microsatellite and karyological data, defining a new species (*N. n. diardi*) restricted to the islands of Sumatra and Borneo [7; 8; 9]. The numbers of both species combined are estimated to be less than 10,000 [10], with the mainland clouded leopard (*N. n. nebulosa*) ranging from the Himalayan foothills in Nepal through mainland Southeast Asia into China. The possible extinction of a third Taiwan subspecies (*N. n. brachyurus*) [11] exemplifies that action is required to maintain these charismatic animals.

Being one of the most threatened of all carnivore groups, it is imperative that we understand all we can about these increasingly rare great cats. It is well known that mankind has had a large influence on the dwindling numbers of these wild cats, and conservationists are increasingly utilizing phylogenetic trees to formulate conservation action plans for both land and marine mammals [12; 13]. For instance, extinction risks can be formulated in order to project expected losses to a specific phylogenetic group [14; 15]. These species loss estimates can be combined with other quantified considerations to assist in selecting species populations for conservation that allow a maximal retention of

total phylogenetic variation [16; 17]. In this way, researchers are able to recognize when the loss of certain subsets of species correspond to the loss of a greater proportion of evolutionary history.

Members of family Felidae, particularly those within genus *Panthera* are often the top predator in an ecosystem, existing in comparatively low density to other species, thus requiring a larger territory. In this way, they are under increasingly consistent threat of eradication from their historical range as human expansion constrains their range, forcing their density to increase. This reduces the availability of food and confluent habitat, often resulting in the predation of livestock by these desperate carnivores, and subsequent hunting of the offenders to preserve the livelihoods of humans. Without intervention, this cycle is fated to eradicate the once widespread members of *Panthera*.

The lion (*Panthera leo*) is perhaps the most recognizable and widely known member of the big cats. Being the only social felid, lions live in groups to facilitate breeding success and cooperative stalking of prey [18]. They are currently listed as vulnerable due to an apparent population reduction of at least 30% in the past 20 years [19]. Historically, lions ranged from northern Africa through southwest Asia, west into Europe where it became extinct during Roman rule, and east into India [6; 20]. Lions are extinct in North Africa, with the most recent population disappearing 70 years ago [6]. Genetic analysis definitively divides this group into two subspecies, Asiatic (*P. l. persica*) and

African (*P. l. leo*) [21; 22]. Estimates of current wild African lion population sizes vary widely from 18,000 to 47,000 individuals and are found only in sub-Saharan countries [23; 24]. Their range is estimated at over 4.5 million km², only 22% of their historical range [6; 19]. There is only one isolated extant population of wild Asiatic lion in the Gir Forest National Park and Wildlife Sanctuary in India, with roughly 250-350 individuals [21].

The leopard (*Panthera pardus*) is widely distributed across southern Asia from Turkey, west into the Himalayan foothills, India, China and Russia, south to the islands of Java and Sri Lanka and sub-Saharan Africa with remnant populations in North Africa and the Arabian peninsula [6; 20]. They are the most abundant pantherine with an approximate count of 300,000, concentrated in eastern and southern Africa, but becoming very scarce in other regions [25]. Despite their broad, multicontinental distribution and the most robust tolerance for varying habitat of any Old World felid, they are considered near threatened and declining in numbers in North Africa, the Middle East, and their historically wide Asian ranges [25; 26].

The jaguar (*Panthera onca*) can measure up to 2 meters nose to tail and weigh up to 120 kilograms, making it the largest felid in the Neotropics and the Western Hemisphere, and the only representative of *Panthera* therein [27]. Middle Pleistocene fossils have been found north from Nebraska [28], Washington [29], and West Virginia [30], and south to Bolivia [31]. The jaguars'

current range has been reduced approximately 46% to 8.75 million km^2 in Central and South America, from southern Mexico to the southern Pampas region, east of the Andes, concentrating in the rainforest of the Amazon basin [32; 33]. [33]There is no evidence for discrete subspecies, but extant populations are clustered in phylogeographic groupings fragmented by major geographic barriers, yet not in completely genetically isolated populations [34]. Though estimated at less than 50,000 individuals, the total number of jaguars varies by sampling range and is unknown since current research has only targeted a relatively small area [6; 33]. They are still classified as near threatened due to deforestation, prey loss and population fragmentation [32].

The tiger (*Panthera tigris*) highlights one of the most tenuous situations affecting felid species. Though the lion may be instantly recognizable as the mammalian 'king', the tiger has been crowned as the "world's favorite animal", beating out all human companion animals such as the dog, horse and domestic cat [35]. They are the largest of all extant felids, with males weighing in at up to an impressive 600 kg [36]. Despite its individual power and majesty, it is one of the most endangered animals in the world [37]. Once ranging from Turkey, east to the Pacific coast of Russia, tigers have become extinct in southwest and central Asia, large areas of Southeast and Eastern Asia, and on the islands of Bali and Java within the past century [6]. They are now restricted to an estimated area less than 1.2 million km^2, only 7% of their historical range [4; 38]. Within the past decade, their range has decreased as much as 41%, primarily

due to poaching and habitat loss [4; 37; 39]. Individual population estimates by country shows a global wild population of about 3,400 to 5,140 adults with no subpopulation having an effective population size larger than 250 [37], however estimates show as many as 12,000 captive tigers (4,000 in Texas alone) are present in the United States [40]. The Siberian tiger (*P. t. altaica*) is estimated to have between 400 and 500 wild individuals, the Bengal tiger (*P. t. tigris*) between 1300 and 1500 and the Indochinese Tiger (*P. t. corbetti*) under 1500. The Chinese Tiger (*P. t. amoyensis*) now only exists in captivity and may soon meet the fate of the Javan tiger (*P. t.* sondaica) which became extinct in the 1980's, the Persian tiger (*P. t. virgata*) which disappeared in the 1970s, and the diminutive Balinese Tiger (*P. t. balica*), which was eradicated in 1925.

The range of the endangered snow leopard is restricted to 12 countries across Central Asia, primarily in the Himalayas, the Tibetan plateau, and the mountains surrounding the western border of China into Mongolia [5; 41]. They prefer steep, rocky and rough terrain at altitudes of 3.0 – 4.5 km, but can be found on the flat plateau in the presence of adequate cover. Using an altitude based topographical approach, the estimated potential species range is over 3 million km² with physical evidence of occupation in about 1.8 million km², centered around the mountainous ranges with little evidence in the flatlands of the plateau [5; 41; 42]. Habitat and prey declination as well as poaching and hunting for livestock preservation have led to an apparent decrease in their numbers by at least 20% to an estimated 4,080 to 6,590 individuals in the last 16

years [5; 41; 43]; with much of this decline occurring in the former Soviet Union [44].

The clouded leopard has been called the world's only living saber-toothed cat [45] since it has the longest canine teeth of any living cat proportionate to its size [20]. The smallest of the *Panthera* lineage, it ranges from the Nepalese Himalayan foothills through Southeast Asia into China and Bangladesh [6]. Historically, the clouded leopard ranged across most of southern China; however it has been consistently threatened by rampant poaching and its current population is largely unquantified [20; 46]. Analysis of mitochondrial DNA, microsatellites, chromosomes, and morphological characters, indicates that *N. nebulosa* is currently restricted to mainland Southeast Asia, whereas the recently recognized *N. diardi* lives only on the islands of Sumatra and Borneo [7; 8; 9]. Populations on the island of Taiwan, though present historically, no longer exist [11].

The great cats are both individually powerful and undeniably fragile in the global scheme. Their domestic relatives provide unyielding companionship to humans and have provided us with knowledge of their anatomy, physiology, behavior, pathology, and phylogeny. The survival of each felid species in the wild, including the pantherines is the goal of many international conservation efforts. Clearly, in order to provide for the future preservation of these powerful, unique cat species, immediate action in the form of effective conservation plans

must be built. The foundation for such strategies is rooted in their specific pasts. A primary building block of any conservation technique is an understanding of the species relationships and their evolutionary history [47]. Important information involving the causative events and resultant consequences of their diversification from common ancestry and subsequent speciation allows researchers to best plan for the maintenance of each species individually. A reliable, robust phylogeny serves as an integral framework to address these issues and is requisite not only for the understanding the historical events surrounding the origin of the members of this lineage, but will open up new avenues for the study of pantherine speciation genomics.

Evolutionary History of Family Felidae

The carnivore lineage originated from their closest relatives, the pangolins, approximately 78 million years ago (MYA) [48]. The arrival of the Paleocene 60 MYA saw the first appearance of carnivorous species in the fossil record, including the Nimravidae, or false sabre-tooths, with a later bifurcation of feliform (cat-like) and caniform (dog-like) groups 55 MYA. Within the past 30-40 million years, many iterations of saber-toothed cats came and went in the fossil record since the Oligocene nimravides [29], all of which faced extinction [49]. The nimravids physically resembled the true sabre-toothed cats of genus *Smilodon* through parallel evolution. The cat-like ecomorph enjoyed much

success throughout carnivore history, changing relatively little over the span of evolutionary time. Their success can be seen in the independent evolution of cat-like ecomorphs in at least three caniform and one feliform family [50]. The most distantly known precursor species to living members of Felidae, *Proailurus lemanensis*, existed during the Oligocene roughly 25 MYA, extending into the Early Miocene [50; 51]. Later in this epoch, roughly 11 MYA, a second felid dubbed *Pseudaelurus* appeared with considerably more fossil evidence, primarily isolated lower jaws [51]. It has been thought for more than a century that *Proailurus lemanensis* from the late Oligocene, and *Pseudaelurus* of the Miocene, share a common ancestor with extant felid species [52]. Often argued to be the first true big cat resembling those we know today, *Pseudaelurus* shares much morphological similarity to the modern cheetah [53]. The fossil record indicates that this felid lived in Eurasia during the Miocene and spawned the Asian ancestor of modern cats 10.8 MYA [2].

Rapid speciation events initiated the divergence of extant species of modern felids [54; 55; 56]. The recent and speedy nature of these events, the subsequent parallel evolution of resulting species, underrepresentation in the fossil record, and morphological ambiguities resulting from relative morphological uniformity in skeletal character compared to other carnivores, complicate the refinement of an accurate relationship between extant species [30; 31; 57]. The 38 currently accepted felid species arose from eight major ancestral lineages and have evolved into one of the world's most widespread

carnivore families, occurring on all the continents except Australia and Antarctica [58; 59; 60].

The most recent phylogenetic and paleogeographic hypothesis outlined by Johnson et al. (2006) places felid origins in Asia ~10.8 MYA. The initial feline divergence led to the common ancestor of the five great cats of the genus *Panthera* and the closely related clouded leopard, which separated from *Panthera* about 6 MYA. The tiger, clouded leopard and snow leopard were historically distributed only in Asia, where they remain to this day, albeit over a greatly reduced region. During the Plio-Pleistocene, a pseudo-period roughly 2.5 to 1.5 MYA bridging the Pliocene and Pleistocene, the leopard remained in Asia and migrated into Africa. Uphyrkina et al. (2001) combined phylogeographic and population diversity estimates based on 727 bp of mitochondrial sequence and 25 polymorphic microsatellite loci to investigate leopard origins. Results indicated that modern leopard lineages arose 470,000 to 825,000 years ago in Africa followed by their migration into and across Asia 170,000 to 300,000 years ago [61]. However, recent studies placed leopard divergence from lion and jaguar in Asia 2.87 MYA with subsequent migration into Africa in the Late Pleistocene [2].

The fossil record dates the most ancient felid with resemblance to modern lions to the Late Pliocene in Eastern Africa, roughly 1.8 to 5.0 MYA [29; 62; 63]. 1,000,000 to 800,000 years ago in the Middle Pleistocene, lions colonized the

entire Holarctic ecozone including Eurasia and North America [64; 65]. Fossil records of lions are found not only in Africa (*P. l. shawi*), but also in Europe and Britain by 500,000 MYA (*P. l. fossilis*) and the more derived Cave Lion (*P. l. spelaea*), which appeared about 300,000 years ago [66], becoming the most prevalent large terrestrial mammal during the Late Pleistocene [29; 65; 67]. During the Early Pleistocene, the lion spread across Berengia to North and South America as the American Lion (*P. l. atrox*) [29]. At the end of the last glaciation about 10,000 years ago, a major Pleistocene extinction of mammals eliminated 40 North American species. This affected 75% of the large vertebrates, including American lions, saber-tooth cats, puma and cheetahs that met the fate of many other megafuana such as the mammoth, mastodon, and the short-faced bear [68; 69].

Jaguars also made their way to the Americas during the Plio-Pleistocene. They escaped the major Pleistocene extinction, making their way into South America via the isthmus of Panama. South America had been completely isolated from all major land masses since the separation of the supercontinent Gondwanaland over 100 MYA. The jaguar was able to outcompete many of the endemic carnivorous marsupial species for their habitat and ecological niches, contributing to the success of the species. Using 715 bp of the mtDNA control region and 29 polymorphic microsatellite loci, it was observed that jaguars show evidence of a recent demographic expansion, with modern lineages arising

280,000–510,000 years ago, more recent than the dates inferred from the fossil record [34].

Fossil and molecular data show extant cats have a recent evolution with the oldest fossil record of modern cat species dating to a mere 3–5 million years and many appearing less than 100,000 years ago [30]. The definitive resolution of the interspecific relationships within the five recognized pantherid species has proven to be problematic. Analyses thus far have revealed short internodes that are likely a result of very recent speciation events occurring within 1–2 MYA [2; 67; 70]. Relationships within Felidae have been evaluated with multiple measures of morphologic and molecular evolutionary methods that serve as a framework for tracking gene divergence during brief evolutionary periods. Understanding the evolutionary history of modern felids begins with deciphering the relationships evident between extant species. As with any migration reconstruction, events must be evaluated with respect to the fossil data, but the history of *Panthera* has the potential to be one of the most specific dispersal stories for any group.

Phylogenetic Literature Survey

Before beginning any research to assist the preservation of these charismatic and highly threatened cats based on evolutionary history, such as

phylogeographic, or population genetic studies, it is paramount to fully understand these historical relationships between them. All life on the earth, present and past, originated from a common ancestor. Proceeding along a succession of speciation events, a pattern of evolutionary relationships arises among extant species. Modern analysis seeks to deduce the pattern of relatedness between organisms by using a phylogeny, an idea borne directly from the theory of evolution presented by Charles Darwin in *The Origin of Species* [71]. Interestingly, the only illustration contained in this landmark publication is the first phylogenetic tree.

Linnaeus first conceptualized the genus *Felis* in 1758, ushering in the era of taxonomic classification for this group of mammals [72]. At the inception of felid phylogenetic reconstruction, characters suitable for taxonomic inference were limited strictly to observable morphological characteristics easily compared with members of other clades including skin patterning and pelage coloration [73], the composition of the tongue and lingual structures [74], zygomatic arch morphology [75], and general body structural characters [76]. More recently, morphological studies targeting diversity solely from within the felids have focused on more specific and complex characters such as dentition [77], endocranial casts depicting precise neural structures [78], and comprehensive total skull morphology [79]. Despite the wealth of comparative characters from all these approaches at the disposal of researchers, they have only proven truly useful above the species level in felids, hampered by the fact that felids have

changed relatively little over their evolutionary history [80], resulting in a strong anatomical homogeneity within the family [79; 81]. For instance, the morphological similarities of the hyoid apparatus and the pharynx distinguishes the species of the Pantherinae from the rest of the felids, but does not confer any distinct relationships within the genus [82]. This is interesting in that the only five felid species with a partially ligamentous and incompletely ossified hyoid belong to *Panthera* and are the only cats able to roar but not to purr [83]. This contrasts with a fully ossified hyoid structure in all other felid species, conferring a purring ability without roaring potential [82; 83]. "Roaring" has been defined as a low pitch vocalization with a low fundamental frequency, and lowered formant frequencies [82]. In the case of the tiger, this is an infrasound around 18 hertz, a frequency too low for detection by human hearing [84]. This contrasts to purring, which is the result of very rapid twitching of the vocalis muscle within the vocal folds [85]. The only non-roaring member of *Panthera* is the snow leopard, which lacks a large pad of fibroelastic tissue in the rostral portion of each of the very large undivided vocal folds, present in the other four members [86]. It has been hypothesized that the addition of this structure to the already present incompletely ossified hyoid would allow the snow leopard to roar as the other great cats [82]. Despite the lack of unambiguous morphological characters, many studies have partially resolved the phylogeny of *Panthera* (Figure 1A-1C). All grouped lion, leopard, and jaguar as a monophyletic

trichotomy to the exclusion of the remainder of *Panthera* [79; 87; 88], with one study constructing a sister relationship between lion and leopard [88].

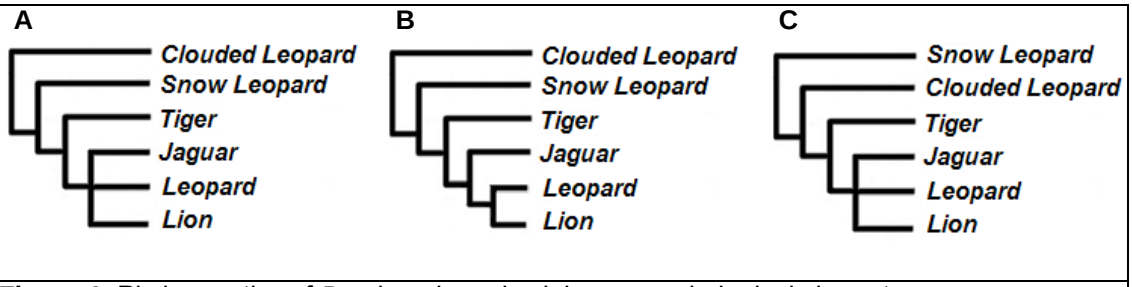


Figure 1. Phylogenetics of *Panthera* based solely on morphological characters. (A) Hemmer (1978), (B) Herrington (1986), (C) Salles (1992).

In contemporary research, some argue for a more restricted use of primarily morphological data [89]. For instance, it has been shown that without molecular support of morphological information as a way to evaluate the accuracy of fossil specimens, as many as 72% of living placental orders move to a different superordinal group, indicating that morphological studies of eutherian interordinal relationships do not sufficiently distinguish between homology and homoplasy [90]. Homoplasy can be described as phylogenetic signal discordant from true evolutionary history. For example, the inference of homology between a bird and bat wing would be homoplastic, as they are analogous structures originating from independent evolutionary processes. Recent results from a study of mustelid species suggest that incongruence between molecular and morphological trees may partly arise from misleading effects of adaptive convergence in morphological characters [91]. In addition to the lack of

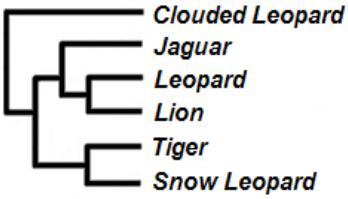
complete fossil data, it was for these reasons that we reconstructed the evolutionary relationships between pantherine species using only molecular data. However, some observed molecular variation can be polymorphic or plesiomorphic, and homoplasies are present throughout the genome. Since convergent evolution and parallelism can exist in any type of phylogenetic data, we recognize the validity of morphological support for molecular phylogenies. Regardless of the procedural methodology underlying evolutionary inference, it is universally accepted that the best phylogenetic hypothesis is the one supported by the greatest amount of independent lines of evidence [92; 93; 94; 95]. It was using this approach that we sought to resolve the phylogenetic relationships of the genus *Panthera* with multiple sources of corroborative evidence.

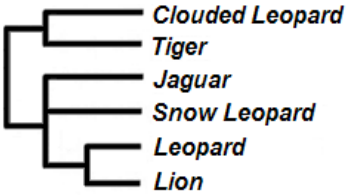
The molecular era ushered in modern approaches to phylogenetic inference, relying on specific cellular characters to deduce relationships rather than a macroscopic, externalized comparison. As technology allowed taxonomic research to adopt a microscopic approach, a completely new realm of data was ripe for exploration, providing much more discernable comparisons with less potential for ambiguity [91; 96]. Chromosomal banding patterns in felids were first revealed by trypsin-Giemsa banding, a method pioneered in these species by Doris Wurster-Hill. This allowed the differentiation and comparison of 30 felid species at the chromosomal level [97; 98; 99; 100]. These efforts found no variation in banding for 13 of the 19 chromosomes across these cat species.

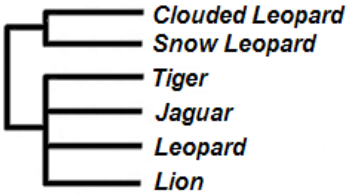
Based on the remaining discernable karyological differences, they found only 16 distinct karyotypes. These broadly divided the cat family into three clades: the ocelot lineage, the domestic cat lineage, and the pantherine lineage.

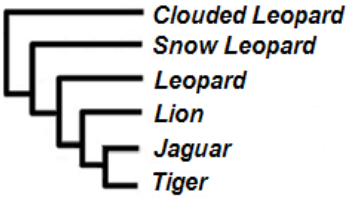
Collier and O'Brien (1985) applied the first biochemical techniques to deduce a molecular phylogeny of the felids using the micro-complement fixation assay to obtain serum albumin immunological distances (AID) for ten representative felid species [101]. This method was used to determine amino acid sequence divergence, as the relative proportionality of this immunological distance and evolutionary time is a component of the foundation of the molecular clock hypothesis [102; 103]. Using AID, they were able to divide felids into three major splits. The oldest divergence occurred roughly 12 MYA and produced the small South American cats of the ocelot lineage. This result was supported by another study utilizing two-dimensional protein electrophoresis on 548 proteins and 40 allozyme loci [104]. The second divergence was that of the domestic cat lineage at 8-10 MYA. The third group was that of the pantherine lineage at 4-6 MYA, which at the time included the big cats of *Panthera*, as well as cheetah, golden cats, serval, and cougar. Within the pantherine lineage, the clouded leopard diverged first, followed by the cougar, the cheetah and then the serval. A final pantherine split at 2 MYA divided lynx species from modern genus *Panthera*. This study used only one *Panthera* member, jaguar, as an outgroup and therefore did not offer any phylogenetic information within the genus. A comprehensive summary of all phylogenetic topologies constructed using

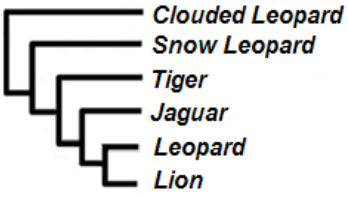
biochemical or molecular techniques with resolution within *Panthera* is shown in Figure 2A-2K.

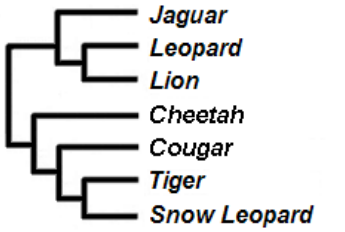
<table>
<tr>
<td>A
 

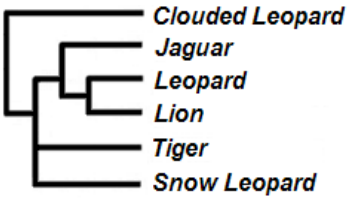
RFLP of complete mitochondrial genomes using 28 restriction endonucleases (Johnson et al. 1996) [105]</td>
<td>B
 

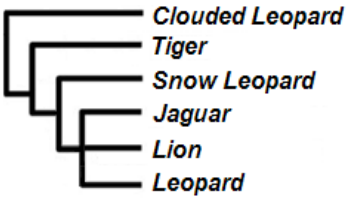
2 Mitochondrial [647 bp] (Janczewski et al. 1995) [59]</td>
<td>C
 

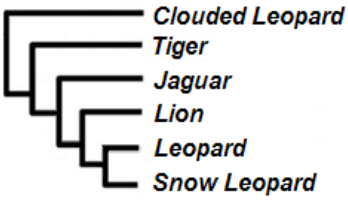
2 Mitochondrial [697 bp] (Johnson & O'Brien 1997) [106]</td>
</tr>
<tr>
<td>D
 

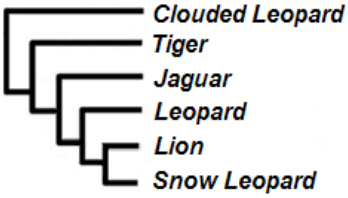
4 Mitochondrial [1,435 bp] (Mattern & McClennan 2000) [60]</td>
<td>E
 

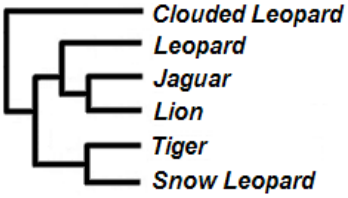
40 source trees, 282 elements (Bininda-Emonds et al. 1999) [56]</td>
<td>F
 

1316 chemical characters (Bininda-Emonds et al. 2001) [107]</td>
</tr>
<tr>
<td>G
 

Variation within the hypervariable mitochondrial CR and RS2 (Kim et al. 2001) [108]</td>
<td>H
 

3 Y-linked [3,604 bp] (Slattery et al. 2004) [109]</td>
<td>I
 

6 Mitochondrial and 3 Autosomal [6,500 bp] (Yu et al. 2005) [110]</td>
</tr>
<tr>
<td>J
 

7 Mitochondrial [3,816 bp] (Wei et al. 2009) [111]</td>
<td>K
 

19 Autosomal, 5 X, 4 Y, 6 Mitochondrial [23,920 bp] (Johnson et al. 2006) [2]</td>
<td></td>
</tr>
<tr>
<td colspan="3">Figure 2. Prior phylogenetic hypotheses of the genus Panthera from biochemical or molecular studies</td>
</tr>
</table>

To further subdivide the three felid lineages defined by AID, a second biochemical approach was taken involving the use of 28 different restriction endonucleases on the mitochondrial genomes of 21 different felid species [105]. In this way, they qualified the restriction fragment length polymorphism (RFLP) between cats by comparing the genomic fragmentation pattern produced by each enzyme. The resultant RFLP patterns were analyzed using minimum evolution in PHYLIP [112] and Dollo parsimony in PAUP [113]. This method was the first to distinctly resolve *Panthera*, finding a sister relationship between tiger and snow leopard, as well as for the monophyly of lion and leopard (Figure 2A). These findings of a close lion and leopard, and tiger-snow leopard relationship were supported by morphological evidence [87; 88]. *Panthera* divergence times of lion and leopard were consistent with the fossil record, indicating a split at around 2 MYA. The 1.5 MYA date inferred via AID was slightly younger than the 1.8 MYA estimate for tiger and snow leopard divergence from the fossil record [105].

Janczewski *et al.* (1995) was the first study to rely on nucleotide sequences to resolve *Panthera* phylogeny [59]. They utilized 647 bp of mitochondrial DNA (mtDNA) consisting of 358 bp from *12S* ribosomal rRNA (*12S)* and 289 bp of cytochrome b (*CYTB*) and offered a relatively unresolved phylogeny within *Panthera*, grouping lion as sister to leopard and tiger as sister to clouded leopard (Figure 2B). A similar molecular phylogenetic approach was

taken by Johnson and O'Brien (1997) using partial *16S* rRNA (379 bp) and *ND5* (318 bp) genes [106]. Topologically, there was less resolution with these genes, with snow leopard replacing tiger as sister to clouded leopard, and the remaining taxa collapsing into an unresolved polytomy (Figure 2C).

Mattern and McLennan combined the previously published molecular data in a novel way [60]. They compiled the *16S* rRNA and *ND5* gene segments from Johnson and O'Brien [106] and *CYTB* and *12S* rRNA from Janczewski et al. [59] with 68 karyological [100; 114; 115] and morphological [79; 116] characters resulting in a 1504 character matrix (1438 nucleotides). This offered resolution of the polytomies plaguing the previous studies, and placed tiger as sister to jaguar as the most recent divergence (Figure 2D),

Bininda-Emonds et al. (1999) utilized a newer concept than standard phylogenetic inference called supertree construction. This supertree technique, matrix representation using parsimony analysis (MRP), incorporated data sources spanning 25 years of systematic research relying on source trees and 'elements' [56]. MRP categorizes relationships within a source tree as binary elements describing each node. Descendants of a node are scored as `1', all others as `0', thus amounting to a parsimony majority rule consensus of previous research efforts [117; 118]. In this way, they were able to offer a resolved phylogeny for *Panthera* based on the entirety of previous research, placing lion and leopard as sister taxa and the most recent species divergence (Figure 2E).

In a later study, Bininda-Emonds utilized a non-traditional approach to phylogenetic evaluation involving the characterization of chemical signals from felid anal sac secretions. Such characters involved glycolipids, neutral lipids, phospholipids, and glycophospholipids and have the ability to influence behavioral or ecological traits. Since chemical signaling can form behavioral isolation prezygotic barrier to reproduction, they can therefore be highly informative about the evolutionary relationships between species. This alternative method was useful to resolve incompatibilities when molecular and morphological phylogenies did not corroborate one another [107]. The combination of these characters provided good resolution though included novel anomalies (Figure 2F). The study did not include clouded leopard, but the most parsimonious topology from the each of the four classes of chemicals, both separately and in the 50% majority rule consensus tree, supported both the monophyly of lion and leopard and a monophyletic snow leopard and tiger. However, unlike previous analyses, the consensus and the neutral lipids topologies positioned cougar (*Puma concolor*) and cheetah (*Acinonyx jubatus*) between the two *Panthera* clades, a relationship not supported by any other study.

Kim et al. [108] used mitochondrial enrichment with centrifugation in a sucrose gradient to separate the mitochondrial genome from the nuclear fragment of all five great cats. Subsequently, they sequenced the complete mitochondrial control region (CR), the most rapidly evolving region of the mtDNA

molecule [119]. In addition to classifying the conservation and variability of the CR within *Panthera*, Kim et al. characterized two CR repetitive segments, dubbed RS-2 and RS-3. Phylogenetic analysis encompassing the sequences in the 80-bp repeats in RS-2 and the patterns of variation in the two CR hypervariable segments (HVS-1 and HVS-2) suggested that snow leopard was the closest relative of tiger, and that the lion and leopard were sister taxa, and share a common ancestor to jaguar (Figure 2G). These findings supported the two monophyletic relationships determined by the RFLP analysis of Johnson et al. (1996) and the chemical signal characterization of Bininda Emonds (2001).

Pecon-Slattery et al. utilized 3,604 bp of intronic sequence and SINE insertions / deletions in three genes (*SMCY*, *UBE1Y*, and *ZFY*) on the Y chromosome for phylogenetic reconstruction. They were unable to fully resolve the relationships within the genus using these segments, but produced a familiar trichotomy with respect to the lion-leopard-jaguar clade (Figure 2H) [109]. The study also highlighted the importance of using the Y chromosome for phylogenetic reconstruction in situations where recent, rapid speciation has occurred. They found that the Y-linked introns contained a very high percentage of lineage-specific informative substitutions with a very low incidence of homoplasy.

In the next year, Yu et al. [110] suggested that snow leopard and leopard were sister species based on six mitochondrial genes (*ND2*, *ND4*, *ND5*, *CYTB*,

12S, and *16S*rRNA) and three nuclear segments (β-Fibrinogen gene intron 7, IRBP, and *TTR*), though jaguar was not included in the nuclear partition. Unique to this study, their analyses shows the following relationships: (((((leopard, snow leopard), lion), jaguar), tiger), clouded leopard); (Figure 2I).

The most recent molecular phylogenetic study by Wei et al. was published in 2009 was included with the first full sequencing of the snow leopard mitochondrial genome [111]. The phylogenetic component of the study utilized seven mtDNA genes (*12S* rRNA, *16S* rRNA, *ND2*, *ND4*, *ND5*, *CYTB*, and *ATP8*) totaling 3,816 bp. This study relied on previously published sequences from multiple studies [59; 106; 110; 120] for all taxa other than snow leopard. The resultant inference strongly supported the close affinity between snow leopard and lion (Figure 2J). This was yet another novel sister taxa association not found in any previous molecular study, and further obscured the precise relationships within *Panthera*.

The most comprehensive molecular phylogenetic study encompassing the greatest amount of sequence data over all inheritable regions of the felid genome (autosomes, mitochondria, X and Y chromosome) was completed by Johnson et al. in 2006 [2]. This study assessed a total of 19 independent autosomal, five X-linked, six Y-linked, and nine mitochondrial gene segments. This work also included independent sequencing of portions of the mitochondrial data from the Janczewski et al. 1995 [59], Johnson et al. 1997 [106], Pecon-

Slattery et al. 2004 [109], and the Yu and Zhang 2005 [110] studies, therefore encompassing virtually all previous molecular work. A total evidence approach produced a phylogeny that resolved a monophyletic lion-leopard-jaguar clade, with leopard as the most basal member (Figure 2K). Further, this phylogenetic study reconstructs the bifurcation between the lion clade and the snow leopard-tiger clade, supporting the findings made by previous RFLP analysis [105] and anal sac secretion characterization studies [107].

Taking in to account all previous studies that sought to provide definitive evolutionary resolution for *Panthera*, the basal placement of clouded leopard relative to *Panthera* was the only consist finding across studies [56; 60; 105; 106; 107; 108; 110; 121]. The only discrepant association was that of clouded leopard as a possible sister species to tiger, a hypothesis put forward during early behavioral studies [122]. This was supported by one molecular phylogenetic study [59], however this analysis only involved mitochondrial *12S* and *CYTB* genes, totaling less than 650 bp. The overwhelming evidence for a sister-group relationship of the clouded leopard to *Panthera* supports our use of this species as an outgroup for phylogenetic reconstruction. However, outside of this association, the evolutionary relationships within the genus and the reasons for published discordances remain obscure.

Selection of the Y Chromosome and Mitochondria for Phylogenetic Inference

The Y chromosome is a highly underutilized nuclear molecule with largely untapped phylogenetic potential. When compared to mitochondrial, autosomal, or X-linked markers, polymorphism on the Y-chromosome has been shown to yield more information for the resolution of rapid felid speciation [2; 109]. The male-specific, haploid Y chromosome offers a unique vantage for phylogeny estimation. The pseudoautosomal region, comprising only a small percentage of the total Y chromosome, contains the only sequences subject to meiotic recombination. This facilitates the conservation of paternally inherited allelic states by constraining virtually all sequence changes to mutation events rather than large-scale rearrangement. As a mate-pair, one generation of cats possesses four copies of autosomes, three X chromosomes, but only one Y. Thus as a whole, the effective population size (N_e) of the Y chromosome is less than that of any other nuclear molecule. For heterogametic, diploid mammals, N_e for Y-linked and mitochondrial loci is ¼(N_e) of autosomal loci. This is due to the twofold reduction of N_e related to the haploidy of the molecules and another twofold reduction due to uniparental patrilineal transmission of the Y-chromosome and uniparental matrilineal transmission of the mitochondria. For X-linked loci, N_e is ¾ that of autosomal loci since mammalian males are hemizygous. When mutation mechanisms are applied equally across all chromosomes, the Y chromosome contains lower sequence diversity compared

to the remainder of the nuclear genome [123; 124]. Despite this low number of variable sites, the X-degenerate genes in the male specific region of the Y (MSY) confer a remarkably large amount of phylogenetic signal, both lineage and species specific, within felids [109]. This level of informative signal is significant due to the recent and rapid evolution of the cat family. The Y chromosome is relatively gene poor and has a very low incidence of convergent, parallel, or reversal nucleotide substitutions, the primary sources for molecular homoplasy [109]. This indicates that a significant majority of observed substitutions are phylogenetically informative and not as prone to a low signal-to-noise ratio. These features make the MSY a highly suitable region for phylogenetic inference. Even so, there is currently very little published genomic information on the felid Y chromosome, making it the least explored area of the feline genome. It therefore has the potential to yield a large amount of previously underutilized sequence data.

Within the MSY, introns provide the greatest source of variation while still tracking the genes within each lineage. They can contain many repetitive and highly variable elements and have a higher incidence of genomic change relative to coding sequences, therefore providing an increased level of variation and lower selective constraint than exonic and even intergenic regions [125]. Some introns do contain regulatory elements and all contain post-transcriptional splicing sites, however many comparative genomic efforts suggest that introns evolve largely free from selective constraints [126; 127; 128; 129].

The mitochondrial genome is also a desirable genomic region for phylogenetic inference, partially due to their high copy number per cell, wealth of published primers, and low rates of recombination [130; 131]. Mitochondrial DNA (mtDNA) has an estimated mutation rate that is up to 20 fold higher than nuclear DNA [132]. There may be many causes for this, including the close proximity of mtDNA to reactive oxygen species generated by the respiratory chain, and its lack of protection by histones [133; 134]. It is therefore more susceptible to damage than nuclear DNA. MtDNA has been used in the majority of felid phylogenetic studies, and many other mammalian species [2; 59; 60; 106; 110; 111; 121; 135]. However, the use of this marker is prone to misamplification of heterologous sequences in the form of mitochondrial to nuclear translocation events (numt), resulting in nuclear sequences of mitochondrial origin [136]. A 7.9kb nuclear insertion of a large portion of the mitochondrial genome, tandemly repeated 38–76 times on domestic cat chromosome D2 has been identified by Lopez et al. (2004, 2006) and will be referred to herein as the Lopez-numt [137; 138]. Numt instances have also been reported in *Panthera* based on mtDNA RFLP data [105], and sequence analysis [139]. A study comparing the sequence of cytoplasmic mtDNA with the homologous numt sequence on tiger chromosome F2 that found the two shared about 90% sequence identity [140]. This high sequence identity is adequate to inadvertently amplify numt sequence using primers designed from the mitochondrial sequence. A more recent evaluation of felid numts characterized

these translocation events in the domestic cat genome and constructed a catalog of cat numts. This was done by using MegaBLAST queries against the 1.9x domestic cat whole-genome shotgun sequence to compare both the contigs and unplaced reads in the cat nuclear genome sequences with that of the mtDNA genome [141; 142]. Only about 15% of the 298,320 bp of numts could be mapped onto cat chromosomes due to a lack of flanking sequence availability. They also characterized distinct insertions from the Lopez-numt, showing portions of cytoplasmic mitochondrial (cymt) sequence not present in the Lopez-numt. This recent study clearly outlines the pervasive nature of numt translocations and highlights the care with which phylogenetic inference must be performed to control for such homoplastic interference. Experimentally, amplification of numts can be controlled by performing a mitochondrial enrichment procedure prior to performing PCR [137; 138]. This involves the separation of the nuclear fraction from the mitochondria by using a gradient medium such as cesium chloride or sucrose, and multiple centrifugation steps to pellet out each fraction. Though numts can be phylogenetically informative when properly identified and separated from cymt amplification [143; 144], for the purpose of this study, they were identified and removed *in silico* from phylogenetic consideration.

Microsatellite loci (microsats), also known as simple sequence repeats (SSRs), variable number tandem repeats (VNTRs), or simple tandem repeats (STRs) are 2 to 6-base pair tandem motifs repeated numerous times and

bookended by specific flanking sequences. Multiple studies have demonstrated a very strong sequence conservation of these microsatellite-flanking regions across related species [145; 146]. Microsats are present in occurrences ranging from approximately 100,000 to 200,000 per organism with ubiquitous and virtually stochastic appearances densely distributed within eukaryotic genomes [147]. Microsats confer a high mutability associated with slippage of DNA polymerase over these highly repetitive regions during replication [148]. They are considered selectively neutral DNA markers and are thought to play a role in chromatin organization, gene regulation, recombination, DNA replication, cell cycle, mismatch repair [149; 150]. Their propensity for mutation during replication leads to high variability within populations and the speedy accumulation of new alleles [151]. Thus, they have become invaluable markers for parentage and relatedness determination, gene mapping, genetic individual identification, and genetic diversity monitors. Their genomic abundance, conservation of their distinctive flanking sequence across closely related species, apparent selective neutrality, and high heterozygosity contribute to their utility in detecting historic demographic events in natural populations [152].

There are currently about 600 microsats characterized in the domestic cat alone[153]. The quantification of these variable polymorphic markers in other felid species have been an invaluable tool for resolving modern felid population demographics with respect to bobcats [154], cheetahs [155], and jaguars [156], identifying *Panthera* species in India [157], as well as for discerning historic

population contractions in wild cats [158] and the identification of individual snow leopards [159].

CHAPTER II

MOLECULAR PHYLOGENY OF GENUS *PANTHERA:* MULTILOCUS SUPERMATRIX AND GENE TREE METHODS

Introduction

A significant problem exists with respect to the phylogeny of the highly threatened great cats, the felids within the pantherine lineage. This clade consists of the five big cats of the genus *Panthera,* the lion, tiger, jaguar, leopard and snow leopard, as well as the closely related clouded leopard, which diverged from *Panthera* approximately 6 million years ago. Multiple groups have attempted to resolve this problem with varied and discordant results. The evolutionary relationship among these cats is difficult to resolve, in part due to their recent and rapid radiation 3-5 million years ago (MYA), individual speciation events occurring within less than 1 million years, and probable introgression between lineages following their divergence. Disparity between every previously published phylogenetic result for genus *Panthera*, stems from this fact and both the amplification of mitochondrial to nuclear translocation events (numt), and errors in species identification. This is evidenced by multiple, conflicting hypotheses and the lack of any corroborative conclusions between previous phylogenetic studies. The resolution of this ambiguity is of primary importance for these highly charismatic cats. Without a defined phylogeny, definitive

conclusions cannot be made concerning speciation, introgression, and migration events as well as deductions about speciation genetics.

Family Felidae has been described with multiple measures of morphologic and molecular evolutionary methods that serve as a framework for tracking character evolution during brief evolutionary periods. Understanding the evolutionary history of all felids begins with deciphering the relationships evident between extant species. A better understanding of the phylogenetic relationships within *Panthera* not only enhances our knowledge of their evolutionary history, it allows us to better plan for their future through conservation efforts, a paramount goal as many of these species are nearing the brink of extinction. A complete phylogenetic understanding contributes a strong historical foundation for future population genetic and phylogeographic studies and opens up new avenues for the study of speciation genomics and understanding the historical events surrounding the origin of the members of this lineage. From this point, we are able to track the evolution of clade and species-specific traits that contribute to the success of these graceful, yet powerful apex predators.

The constitutively haploid Y chromosome has a uniparental, male-specific inheritance, passing only from father to son. It is almost totally unaffected by meiotic recombination events experienced by all other chromosomes [160]. The exception is the pseudoautosomal (PAR) region, a segment on eutherian Y

chromosomes which synapses with a homologous region on the X to facilitate meiotic crossover and to ensure accurate chromosomal segregation in males. However, all genes to be screened for the purpose of this study are located outside the PAR [161] and within the boundaries of male-specific Y (MSY) region. The escape of these genes from recombination is of primary importance for phylogeny in that Y-specific haplotypes will typically pass intact through generations, changing only by mutation, therefore preserving a simpler record of their evolutionary history [109]. The combination of this property in addition to the avoidance of potential numt amplification associated with phylogenies based on mitochondrial sequences makes this an effective region for phylogenetic reconstruction.

Here, we provide an alternative evaluation of the evolutionary history of the pantherine lineage using intronic sequences contained within single-copy genes on the felid Y chromosome. This information was combined with previously published data, and newly generated sequence for four mitochondrial and three nuclear genes, highlighting any incongruence. *In silico* evaluation and identification of putative numt sequences, together with a thorough phylogenetic exploration of the complete dataset provided a highly supported topology, consistent with previous studies not involving nucleotide sequence phylogeny [105; 107]. The results of these comprehensive methods are summarized and compared to outline the complex history of *Panthera*.

Methods

Sequence Generation

The complete expressed sequence for each Y-linked, X-degenerate, putative single-copy gene was obtained from published *Felis catus* cDNA sequences. These genes of interest included *SMCY* (EU879977), *EIF1AY* (EU879973), *EIF2S3Y* (EU879975), *DDX3Y* (EU879971), *USP9Y* (EU879980), *UBE1Y* (DQ329521), *UTY* (EU879982), and *ZFY* (EU879984) [162; 163]. Since these are post-transcriptionally modified sequences, the exon-intron boundary was defined by using the BLAST [164] tool to query the published cDNA sequences to the *Homo sapiens* Y chromosome sequence contained in the genome build 36.3 [165]. In the cases of *EIF2S3Y* and *UBE1Y*, comparative Y data was unavailable for human. The cDNA sequences for these genes were aligned with the published data on the *Mus musculus* Y chromosome assembly in build 37.1 of the mouse genome [166]. Gaps in the feline cDNA sequence alignment relative to the human or mouse genomic sequence were used to define approximate locations of intron-exon boundaries. Sequences between each gap in the cDNA/genomic sequence alignment were identified as exons, and were used to design primers that amplify each intron. The observed intron length in the human or mouse genome was used to approximate the estimated size for the associated felid intron, as intron length tends to covary with genome size across species. According to this estimate, intron amplicon size classes

were defined as 250bp-1kb, 1kb-3kb, 3kb-5kb, 6kb-8kb and 8-11kb to allow efficient utilization of thermal cycler programming extension times and enhance the specificity of PCR reactions. Each intron size class contained multiple exonic targets from multiple genes.

EPIC (exon-primed, intron crossing) primers were designed using Primer3 software to target the BLAST-defined exonic flanks and extend into the intronic region [167]. Optimal conditions adopted for each PCR reaction were a 95°C initial hot start for 60 seconds, with a total of 35 cycles of 94°C denaturation for 15 seconds, a 58°C annealing for 30 seconds, a 72°C extension of 60 seconds per 1kb corresponding to the largest estimated amplicon size in each category, and a final extension of 300 seconds. PCR reactions were performed on an Applied Biosystems GeneAmp 9700 thermal cycler. Each reaction utilized all Invitrogen reagents and contained 2.5µL 10X PCR buffer, 0.75µL 50mM MgCl2, 2ul 10mM dNTPs, and 2µL for each of the 5µM forward and reverse primer. Amplicon classes below 6kb were amplified using 0.1µl Platinum Taq DNA polymerase and classes from 6kb to 11kb with AccuPrime high fidelity DNA polymerase. The remainder of each 25µl reaction contained 1µL - 3µL of template DNA, depending on the spectrophotometric concentration, with the balance adjusted with molecular biology grade water. Y chromosome BAC clone DNA from the RPCI-86 10X male domestic cat was used to ensure amplification of actual Y chromosomal sequence, rather than a highly similar X-linked or autosomal paralog [168]. Figure 3 depicts a tentative map of the

single-copy, Y-linked genes of interest and the corresponding BAC clone DNAs lines used for PCR amplification. Those primers most successful in amplifying intronic regions from the BAC clones (single banded, high amplification) were sequenced on an Applied Biosystems 3730 DNA Analyzer to provide targets for subsequent intronic primer design. These primers were initially tested in male and feline genomic DNAs from both jaguar and clouded leopard to assess their cross-species, Y-specificity. Successful amplicons were sequenced. Their chromatograms were trimmed, parsed for quality, and assembled into contigs using Sequencher 4.7 [169]. The primers with high quality sequence reads in both species were selected for amplification and sequencing in lion, leopard, tiger, and snow leopard, resulting in a total of 39 successfully sequenced amplicons in all six species totaling 15,392 bp.

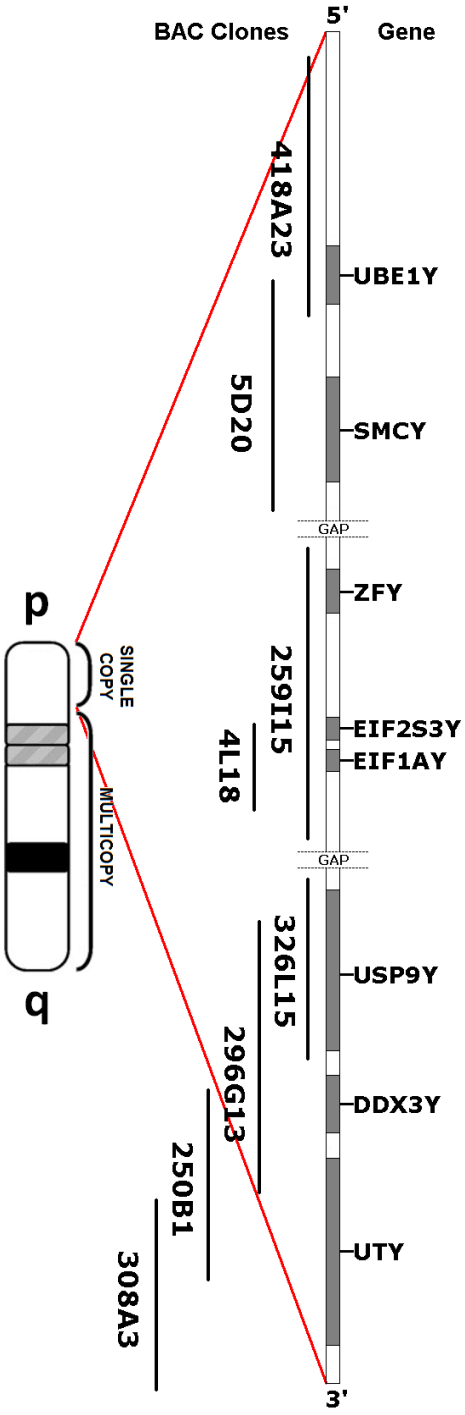


Figure 3. Tentative map of a portion of the single-copy X-degenerate region of the domestic cat Y chromosome and selected RPCI-86 10X BAC library clones used in this study.

Microsatellite Characterization

In addition to assembling molecular data for phylogenetic analysis, sequences showing di-, tri-, or tetra-nucleotide repeat character indicative of microsatellite were notated. The number of repeats for these putative microsatellites varied from 8 to 16. Figure 4 shows six of the total 24 sequences determined to have the potential to be useful in further population genetic or phylogeographic studies involving the Y chromosome within the cat family. Primers were designed to amplify these putative microsatellites using Primer3 with target amplicon sizes under 350bp and are listed in Table 1.

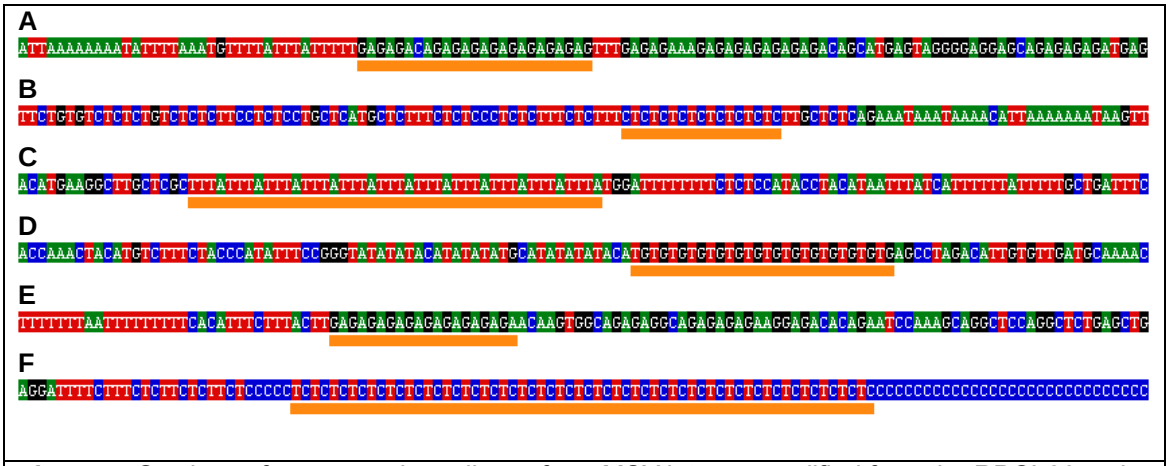


Figure 4. Sections of sequenced amplicons from MSY introns amplified from the RPCI-86 male domestic cat library with potentially variable microsatellites underlined in orange. (A) UTY intron 12, (B) SMCY intron 7, (C) SMCY intron 2, (D) UTY intron 24, (E) EIF2S3Y intron 6, (F) USPY intron 10.

These 24 sets of primers were amplified in 11 domestic cat and 75 ocelot individuals (Janečka, unpublished). Amplification of these segments

incorporated one of four fluorescent M13 labels onto the forward primer (PET, FAM, VIC, NED) and utilized the same reagents and thermocycler protocol as those submitted for sequencing. Amplicons were checked for successful amplification on a 1.5% agarose gel stained with ethidium bromide. Genotyping analysis was performed on an Applied Biosystems 3730 DNA Analyzer and compared to a LIZ-500 size standard. Results of this reaction were analyzed using GeneMapper [170] to determine the polymorphic character of each of these microsatellites in the species sampled.

Table 1
Forward and reverse primers used for genotyping microsatellite markers in 11 domestic cats and 75 ocelots. Microsatellite naming involves the gene name and intron number.

Microsatellite	Forward	Reverse
DDX3Y-02	TGGCACAGTCAGTAGGATGG	TGCAAAACATTTCAGACAAAGC
DDX3Y-04	TTCCTCTTCTCTGCCCTCCT	TCCCCCAGTATCTCTCTCTCTC
DDX3Y-05	CGACTGTGCTCTCTCTCTCTCA	GATTATCTCCCTTTACCTCTCTCC
DDX3Y-06	TGTGCCAATAAAACTGTGTGTG	TGTGTTTCCCTCTCTTTTTGC
DDX3Y-11	AAGGGTGCTGATTCTCTGGA	CTTTTTCCTGAGCGGAACTG
EIF2S3Y-04	GAGGATATTGCCTCCACCAA	CTAATGCACAGGGGCAGAAT
EIF2S3Y-06	TTCCAGAATTGGAACTGATGG	AGAGCCTGGAGCCTGCTT
EIF2S3Y-07	GACAAATATGTCGGTGGCTAA	GAGTCTGGAGGCTGTTTCC
EIF2S3Y-09a	TCTCCATCTGACCCTCTGCT	CCACAACCACCACCATGTAT
EIF2S3Y-09b	GGGTAAGCTAAATACTAGCAAAGTTTA	GCCTGGAACCTGTTTCTGAT
SMCY-2	GAGGACATGAAGGCTTGCTC	TGCTCTGAACTTTGCTTCTATGA
SMCY-7	CTGCTCATGCTCTTTCTCTCC	TGTCCAGCCACTGTTGCTTA
UBE1Y-18	GTCACCAGGTGGGGACATAC	ACCTGGTATGGGGCAGTTTT
USP9Y-10f	TCAAGTAATACAAATTCACATCACAA	GAATAGAAGATGGGGAGAGAGAG
USP9Y-10r	TCTGAGGTTATGGAGAATTCTTTT	GAATCTGCTTCAGGTTCTGTG
USP9Y-12	TTTTGGAAAGTATGTAATTGTGAAGG	CATCATTTGTTTAACCCAGGTG
USP9Y-16	TTTTGGAAAGTATGTAATTGTGAAGG	CATCATTTGTTTAACCCAGGTG
USP9Y-17	GCTTTCTTTTGTGTATTATTAGTGAG	ACGATCAAAGCAAGACTGGA
USP9Y-23	CCTGAATTCACTATCTTTCTTTTCTTG	GGCTCTGTGCTGAACATGG
USP9Y-41	TCTGCATGACTGCTTCACTTG	ATCTCACTCTGCTCCCCTCA
UTY-21	GACAGAGTGTGATGCTAAATTTCTAA	CCTGGAGCCTGTTTCTGAT
UTY-24	CCAGATGGCACCAAACTACA	TGTTTTGCATCAACACAATGTC
ZFY-1	GCTAGACAAGGATTGGAGCA	GCTCCAGAGTAGCTTCGGTTT
ZFY-6	AGGAGAAATTAAAATAAATCCCTAAAA	GGAGCCTGGAGCCTGTTT

To build upon previous published work and increase the amount of data available for supermatrix analysis, previously published sequences available across all six species were obtained from GenBank and aligned manually in WordPad++, a Windows text editor. The Johnson et al. (2006) dataset utilized previously published Y chromosome sequences [109; 120] and generated new sequences for the mitochondrial and autosomal partitions [59; 106], resulting in 19 autosomal (11,360 bp), 5 X-linked (3,267 bp), 4 Y-linked (5,357 bp) and six mitochondrial (3,936 bp) gene segments. The Yu and Zhang (2005) dataset also utilized previously published datasets [59; 106; 121; 171] and contributed six mitochondrial (3,472 bp) and three nuclear (2,767 bp) gene segments, however they did not generate nuclear gene sequences for jaguar. In addition, the complete mitochondrial genome sequences are available for leopard, tiger, snow leopard and clouded leopard from Wei et al. (2009). Many gene sequences were available from published phylogenetic, population genetic and phylogeographic studies. See Appendix Table 1 for a list of accession numbers for sequences used in the described analyses. Accession numbers not listed in GenBank are referenced by primary author and publication year. Some sequences were manually entered from within the body text of the Janczewski (1995) study and are denoted ‘Jancz95’.

Phylogenetic Analyses

All computation-intensive analyses were performed within the Windows XP 64-bit environment on an Intel platform with a quad-core QX9775 processor overclocked to 3.4 Ghz, 12MB of L2 cache, a 1333 Mhz front-side bus, and 8GB of RAM. Sequence alignments were performed using ClustalX 2.0.3 [172] (gap opening penalty of 10, gap extension penalty of 0.2) with subsequent by-eye verification and manual-editing done with BioEdit 7.0.9.0 [173]. Data was partitioned based on the mode of inheritance: (Y chromosome, autosomes, X chromosome, and mitochondria. Combinations of partitions were also defined as nuclear (Y chromosome, X chromosome, and autosomes) and uniparental (Y chromosome and mitochondria). Henceforth, all six of these data subsets will be referred to as partitions. There are a total of $(2n - 5)!!$ unrooted topologies for any phylogenetic tree with n taxa [174]. Since this dataset contains only six taxa, exhaustive maximum likelihood (ML) tree searches were performed using PAUP 4.0b10 [175]. Bayesian inference was implemented using MrBayes 3.0.4 [176].

Maximum likelihood exhaustive searches with TBR branch-swapping were performed based on the parameter vales obtained using ModelTest [177; 178] for each partition as well as each individual gene segment utilizing the Akaike Information Criteria (AIC) hierarchical test statistic (Appendix Table 2 for the details of specific parameter values). MrModeltest v2, also implementing

AIC [179], was used to determine the optimal substitution model for each locus and partition for use in Bayesian phylogenetic inference, shown in Appendix Table 3. Gene trees were estimated independently for each locus by exhaustive ML searches in PAUP* using the models in Appendix Table 2, as was done for the combined partitioned and the supermatrix trees. For bootstrapping, 1,000 iterations were implemented using TBR branch-swapping. Bayesian inference was also utilized on the same gene segments and partitions with MrBayes, implementing the models in Appendix Table 3. For the individual genes, MrBayes ran for 1,500,000 generations, saving every 100th tree, and discarding the first 250,000 as burn-in. For the partitions, the MCMC algorithm ran for 3,000,000 generations, with every 100th tree saved and the first 750,000 generations were discarded as burn-in. For both BI analyses, one cold and seven hot chains explored treespace to converge on the best phylogenetic tree and parameters given the data.

Homoplasy can be described as phylogenetic signal discordant from true evolutionary history. This can be caused by nucleotide reversal, parallelism, or convergence. A single homoplastic event contains at least 2 independent origins of a character state or the combination of at least 1 origin and 1 reversal. Four indices were applied to the data to quantify the quality of the phylogenetic signal. The first is the consistency index (CI), which is a measure of homoplasy for a given tree. It is calculated as the number of character state changes on the tree divided by the smallest possible number of steps. It is therefore a ratio of the

number of characters in the data set to the length of the tree and ranges values from 0 to 1, with 1 denoting no homoplasy and 0 as completely homoplastic. The complement of this value is the homoplasy index (HI) and is interchangeable with CI. Another metric implemented was the retention index (RI). This measures the proportion of synapomorphy expected from a data set that is retained as synapomorphy on a given tree, or the proportion of similarities on a tree and RI results follow the same range and interpretation as CI. The product of the CI ad RI indices produces the rescaled consistency index (RCI). All indices were calculated using PAUP* for each partition.

A partition homogeneity test was performed both within and between the partitions to test the combinability of the data. Also called the incongruence-length difference test (ILD) [180], this statistic was implemented in PAUP to determine if different partitions of the data, either loci within a partition, or partitions within the supermatrix, have significantly different signals. The number of ILD replicates was set at 1000, and the number of random taxon addition replicates at 10 per ILD replicate. If the test is non-significant, the partitions cannot be combined. We followed the precedent set by Sullivan (1996)[181] and Cunningham (1997)[182] in implementing a significance threshold of $P<0.05$ as too conservative for ILD and set the threshold of combinability for the data to $P<0.01$.

To compare the topologies generated by the ML analyses on each partition and quantify their similarity, a Shimodaira-Hasegawa (SH) test was performed. The SH test uses a statistic that is the likelihood score difference between the specified ML tree and every other tree compared. The null hypothesis (H_0) is that all trees are equally good explanations of the data. The alternate hypothesis (H_1) is that some or all trees are not equally good explanations of the data. These tests were conducted by RELL (resampling estimated log-likelihood) method and 10,000 bootstrap replicates by comparing the scores of each tree generated from each partition with one another [183].

A gene jacknifing approach was applied to the supermatrix to investigate if the removal of a specific gene segment from the concatenated alignment influenced a specific topology. Using PAUP* to perform a ML exhaustive search with 1,000 bootstrap replicates and TBR branch swapping, the ModelTest parameters were again implemented. Individual genes were removed from the dataset and the ML exhaustive search was performed both on the supermatrix as well as the specific partition where the segment resided. Only the Y-chromosome, X-chromosome, autosomal, and mitochondrial partitions were examined in this fashion.

In addition to the traditional supermatrix-based phylogenetic methods, the BEST method was used to construct a species tree from the individual gene trees generated using Bayes. In this way we were able to estimate the posterior distribution of species trees using the multilocus dataset. To do this, the dataset

was then analyzed in BEST, a modified MrBayes package, using the locus-specific models detailed in Appendix Table 3. This program implements the Bayesian Estimation of Species Trees (BEST) approach, a Bayesian method that facilitates the estimation of species trees, allowing for heterogeneous gene trees among loci under the multi-species coalescent model [184; 185]. This method has been shown to consistently estimate species trees, even when the species tree is in the "anomaly zone", a class of species trees whose most common gene tree is topologically different due to very short branches in the species tree as measured in coalescent units [186]. Species tree approaches are advantageous over strict supermatrix approaches which assume homogeneous tree topologies across loci especially within or near the anomaly zone [185], a region that does not possess such homogeneity [187]. This model-based method utilizes the complete information content of the data to estimate the species tree and is very computationally intensive. Thus, this method may not be applicable to large numbers of taxa or partitioned loci, but is useful for small datasets such as the one herein, with six taxa and less than 50 partitions. It has not been determined whether increased sampling within each species leads to a higher level of confidence in the species tree [188]. The data was partitioned into 29 "genes" with the mitochondria and the Y chromosome genes combined into a single partition, respectively, since they do not undergo recombination and are inherited as a complete unit. The X chromosome and autosomal genes were separated into individual loci. Parameters specifically

implemented in the BEST analysis involved specifying the haploid nature of the mitochondrial and MSY segments and setting the prior for the mutation rate across loci to a relative variance between 0.2 and 2, with the average as 1 as suggested by the documentation. The joint posterior distribution of the gene trees was first estimated from the DNA sequences with an approximate joint prior for all 29 gene segments. This prior indicates the joint probability of gene trees and coalescent times across loci. The probability distribution of gene trees is constructed using coalescent theory, based on a Baysian approximation of the species tree with the caveat that all species pairs diverged after their respective gene pairs. This results in an approximate joint posterior distribution of gene trees. Next, BEST utilizes this distribution to estimate the posterior distribution of the species tree, again under coalescent theory. This Markov Chain Monte Carlo algorithm ran for 10 million generations, saving every 1000^{th} gene tree, and discarding the first 1,000,000 as burn-in. For each of the 10 million cycles, we sampled 1000 species trees. The posterior distribution of the species tree topology, including branch lengths, is weighted through importance sampling as it is simultaneously compared to the desired posterior distribution under the true prior of the gene trees. The result was a final species tree topology with support values in terms of the posterior probabilities.

Testing for signals of positive selection in the mitochondrial protein coding genes *ND1*, *ND2*, *ND4*, *ND5*, and *CYTB* was done using the codeml component of PAML [189]. The ratio of nonsynonymous nucleotide substitutions per

nonsynonymous site (dN) to synonymous substitutions per synonymous site (dS), called the dN/dS ratio or ω was estimated as a metric for Darwinian selective pressure. Using the method of Goldman and Yang (1994), values greater than one are indicative of positive selection and less than 0 of negative purifying selection [190]. This statistic was calculated for each gene segment and each taxa, but showed values between 0.0 and 0.2, indicating no positive selection in any of the protein coding genes within any of the taxa.

Since the ω method does take into account that adaptation may come in the form of very few amino acid changes, an alternative method for detecting selection was implemented using TreeSAAP 3.2 [191]. The program identifies regions with significant physicochemical changes by optimizing the in-frame nucleotide sequence, codon by codon, onto the specified supermatrix phylogenetic tree. It then compares the observed, empirical distribution of physicochemical changes for each amino acid with the null hypothesis of an expected random distribution based on the assumption of selective neutrality. The inferred amino acid replacement pattern is then analyzed using the percent difference between expected and observed mean changes [192] and goodness-of-fit of observed to expected applied as a $\chi 2$-distribution [193]. Each amino acid site was classified as positive and negatively destabilizing using 31 physiochemical properties [191]. The magnitude of the physiochemical change was binned into eight categories (1 – 8), with one being the most conservative change and eight the most radical. We looked at amino acid differences

correlated to radical physicochemical changes, indicating positive-destabilizing selection and a corresponding significant change in function. Thus, to detect evidence of strong directional selection, we followed the precedent from da Fonseca (2008) [194] and only changes binned into categories 7 and 8 at the $P \leq 0.001$ (z-score > 3.09) level were investigated. Those genes with scores indicating positive-destabilizing selection across taxa were analyzed using a sliding window approach in TreeSAAP in order to visualize the regions affected by radical amino acid substitutions.

Molecular Dating

A molecular clock test (MCT) was performed in PAUP* to classify the nucleotide substitution behavior of the data as clocklike or non clock-like. Using the appropriate likelihood models selected by ModelTest, the likelihood value is calculated for each gene segment as well as each partition under both the null and alternative hypotheses. Under the null hypothesis of clocklike behavior, the branch lengths on the rooted phylogeny were constrained to fit an ultrametric tree. Under the alternative hypothesis of non-clocklike behavior, each branch corresponded to the actual branch lengths of the tree generated from the data. The difference between these two likelihood scores was doubled and applied to a χ^2 test with degrees of freedom = 4 (ie. the number of taxa minus 2). Though the critical value used to reject the molecular clock hypothesis varies from

P<0.05 to P<0.01 [195; 196], a conservative p-value of P<0.05 was used to for this study.

To infer the divergence times for each species within *Panthera*, we began by deducing ML estimates of sequence divergence parameters generated by the ESTBRANCHES component of PAML 3.15 [189] followed by Bayesian relaxed clock dating in the MULTIDIVTIME program, part of the MULTIDISTRIBUTE package [197]. As with the BEST analysis, the data was segmented into 29 "genes". For nucleotide evolution model parameter estimation, the BASEML component of PAML was used for each section of the data. We estimated branch lengths under the six-species tree obtained herein by all phylogenetic methods using the ESTBRANCHES algorithm, using the specific parameters generated by the BASEML algorithm. The tree was calibrated using three published fossils: 1) a minimum of 1.6 MYA [59; 198] for the base of the lion-leopard-jaguar clade; 2) a minimum of 1.8 MYA [59; 199] for the base of the tiger-snow leopard clade; and 3) the minimum for earliest *Panthera* species based on leopard fossils from African Villafranchian deposits was 3.8 MYA [2; 200]. With all of these parameters defined for each gene, MULTIDIVTIME was used to perform a Bayesian MCMC approximation of the posterior distributions of substitution rates and divergence times. To evaluate the sensitivity of the derived divergence dates to the removal of constraints, the MULTIDIVTIME MCMC analysis was performed in four replicates, each time removing one constraint.

Results

Matrix Analysis

The fully assembled pantherine dataset includes DNA sequences sampled from 43 loci for the five species of *Panthera*: lion, leopard, jaguar, tiger, snow leopard, and a clouded leopard (*Neofelis nebulosa*) as the outgroup taxa. To compile the complete dataset for phylogenetic inference, the validity of both the new and previously published data must be scrutinized. Since the majority of all previous studies relied at least partially on mitochondrial data, the published existence of *Panthera* numts introduced the possibility of numt amplification for these studies. Two of the most recent publications, Yu and Zhang (2005)[110], and Johnson *et al*. (2006)[2] sample virtually all mitochondrial gene segments used in previous molecular phylogenies for *Panthera*. Therefore these sequences were extensively scrutinized prior to assembly into our final supermatrix.

Levels of molecular homoplasy (nucleotide reversal, parallelism, or convergence) were determined for each partition using four metrics: consistency index (CI), homoplasy index (HI), retention index (RI), and rescaled consistency index (RCI) (Table 2). The high CI, RI, and RCI vales along with low HI confirmed the results of Pecon-Slattery et al. (2004) that the Y chromosome had significantly less homoplasy than all other regions of the genome [109].

Table 2. Consistency Index (CI) Retention Index (RI), Homoplasy Index (HI) and Rescaled Consistency Index (RCI) values for the autosomal, Y chromosome, and mitochondrial partitions, as well as the complete matrix.

	CI	RI	HI	RCI
Autosomes	0.6287	0.5118	0.3713	0.4465
Y Chromosome	0.9838	0.8889	0.1364	0.8745
Mitochondria	0.5565	0.3238	0.4435	0.2478
Total	0.5661	0.3624	0.4339	0.2930

Table 3. Accession numbers for mitochondrial gene segments used in prior phylogenetic analysis for *Panthera*. Those segments without accession numbers are referenced by their publication date and primary author.

		CYTB	ND1	ND2	ND4	ND5	12S	16S
Lion	Janczewski et al. [illegible]	Jancz.1995					Jancz.1995	
	Johnson et al. 1997					AF006458		AF006457
	Mattern et al. 2000	Jancz.1995				AF006458	Jancz.1995	AF006457
	Yu & Zhang 2005	Jancz.1995		AY170043	AY634398	AF006458	Jancz.1995	AF006457
	Johnson et al. 2006	John.2006	John.2006	John.2006	John.2006	John.2006	John.2006	John.2006
	Wei et al. 2009	S79302		AY170043	AY634398	AF006458	S79300	AF006457
Leopard	Janczewski et al. [illegible]	Jancz.1995					Jancz.1995	
	Johnson et al. 1997					AF006444		AF006443
	Mattern et al. 2000	Jancz.1995				AF006444	Jancz.1995	AF006443
	Yu & Zhang 2005	Jancz.1995		AY634383	AY634395	AF006444	Jancz.1995	AF006443
	Johnson et al. 2006	John.2006	John.2006	John.2006	John.2006	John.2006	John.2006	John.2006
	Wei et al. 2009	EF437590		AY634383	AY634395	AF006444	AM779888	AF006443
Jaguar	Janczewski et al. [illegible]	Jancz.1995					Jancz.1995	
	Johnson et al. 1997					AF006442		AF006441
	Mattern et al. 2000	Jancz.1995				AF006442	Jancz.1995	AF006441
	Yu & Zhang 2005	Jancz.1995		AY634391	AY634403	AF006442	Jancz.1995	AF006441
	Johnson et al. 2006	John.2006	John.2006	John.2006	John.2006	John.2006	John.2006	John.2006
	Wei et al. 2009	EF437582		AY634391	AY634403	AF006442	AY012151	AF006441
Tiger	Janczewski et al. [illegible]	Jancz.1995					Jancz.1995	
	Johnson et al. 1997					AF006460		AF006459
	Mattern et al. 2000	Jancz.1995				AF006460	Jancz.1995	AF006459
	Yu & Zhang 2005	Jancz.1995		AY634384	AY634396	AF006460	Jancz.1995	AF006459
	Johnson et al. 2006	John.2006	John.2006	John.2006	John.2006	John.2006	John.2006	John.2006
	Wei et al. 2009	EF437588		DQ151550	AY634396	AF006460	DQ151550	DQ151550
Snow Leopard	Janczewski et al. [illegible]	Jancz.1995					Jancz.1995	
	Johnson et al. 1997					AF006450		AF006449
	Mattern et al. 2000	Jancz.1995				AF006450	Jancz.1995	AF006449
	Yu & Zhang 2005	Jancz.1995		AY634382	AY634394	AF006450	Jancz.1995	AF006449
	Johnson et al. 2006	John.2006	John.2006	John.2006	John.2006	John.2006	John.2006	John.2006
	Wei et al. 2009	EF551004		EF551004	EF551004	EF551004	EF551004	EF551004
Clouded Leopard	Janczewski et al. [illegible]	Jancz.1995					Jancz.1995	
	Johnson et al. 1997					AF006426		AF006425
	Mattern et al. 2000	Jancz.1995				AF006426	Jancz.1995	AF006425
	Yu & Zhang 2005	Jancz.1995		AY634385	AY634397	AF006426	Jancz.1995	AF006425
	Johnson et al. 2006	John.2006	John.2006	John.2006	John.2006	John.2006	John.2006	John.2006
	Wei et al. 2009	DQ257669		DQ257669	DQ257669	DQ257669	DQ257669	DQ257669

The sequence data common to both the Yu and Zhang (2005) and the Johnson et al. (2006) datasets included a segment from the mitochondrial cytochrome b (*CYTB*) , NADH dehydrogenase subunit 5 (*ND5*) and *16S* rDNA genes, where Yu and Zhang utilized the previously published sequences available in GenBank, including those from Johnson (1997) [106] and Janczewski (1995) [59]; whereas Johnson (2006) generated new sequences for the same segments. Also common to both were the mitochondrial NADH dehydrogenase subunit 2 (*ND2*) and the nuclear transthyretin (*TTR*) gene, generated independently in each study. Accession numbers for mitochondrial segments and publication where they are utilized are listed in Table 3.

In order to compare the similarities between the homologous sequences generated independently by both publications, the DNAdist component of PHYLIP 3.67 [201] was used to compute a LogDet distance matrix between each species for *ND2* (Table 4 and Figure 5). Evidence for a discrepant phylogenetic relationship involving this gene segment was seen in the high level of intraspecies dissimilarity within the lion and snow leopard between these two publications. Conversely, the interspecies difference between the lion in the Yu and Zhang publication and the snow leopard in the Johnson et al. study shows a 0.01 pairwise difference, much lower than all other interspecies comparisons, and at a level consistent with intraspecies distance calculations.

Table 4. LogDet pairwise distances between published Johnson et al. (2006) and Yu and Zhang (2005) *ND2* gene sequences.

	Johnson Leopard										
Yu Leopard	0.007	Yu Leopard									
Johnson Lion	0.089	0.087	Johnson Lion								
Yu Lion	0.104	0.096	0.091	Yu Lion							
Johnson Jaguar	0.099	0.097	0.089	0.105	Johnson Jaguar						
Yu Jaguar	0.101	0.094	0.092	0.102	0.003	Yu Jaguar					
Johnson Tiger	0.129	0.127	0.124	0.126	0.117	0.12	Johnson Tiger				
Yu Tiger	0.126	0.117	0.125	0.116	0.116	0.113	0.009	Yu Tiger			
Johnson Snow L.	0.093	0.091	0.088	0.01	0.097	0.1	0.118	0.115	Johnson Snow L.		
Yu Snow L.	0.106	0.098	0.093	0.116	0.084	0.081	0.126	0.116	0.118	Yu Snow L.	
Johnson Clouded	0.159	0.157	0.155	0.169	0.177	0.181	0.184	0.196	0.159	0.168	Johnson Clouded
Yu Clouded	0.16	0.155	0.156	0.168	0.179	0.179	0.185	0.194	0.16	0.166	0.005

Similar Intraspecies Comparison (< 1%)

Dissimilar Intrapecies Comparison (> 1%)

Low Interspecies Pairwise Difference

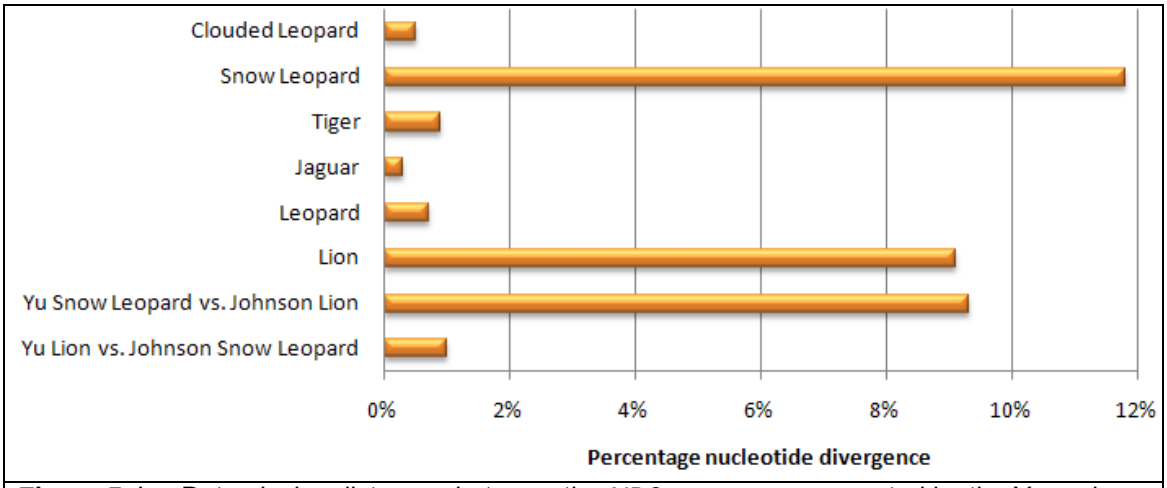


Figure 5. LogDet pairwise distances between the *ND2* sequences generated by the Yu and Johnson publications show a very high disparity within snow leopard and lion sequences. There is more interspecies similarity between Yu's lion sequence and Johnson's snow leopard sequence than within each species.

Transthyretin

There was no jaguar sequence for *TTR* in the Yu and Zhang publication, so an alternate approach was taken to investigate phylogenetic dissimilarity between studies. Intron 1, the region amplified and sequenced in both studies, was sequenced in our laboratory for all six taxa using primers designed from the consensus sequence from all taxa in both studies. These new *TTR* intron 1 sequences were added to the Johnson et al and Yu and Zhang sequences as well as one tiger sequence from a third phylogenetic study [135]. The lion *TTR* intron 1 sequence used in this latter study was also used by Yu and Zhang in their analysis. A ML tree was constructed using RaxML VI HPC [202], a general time reversible (GTR + Γ) model, and TBR branch swapping. The resultant tree not only indicates the same problem with the species identification of the Yu

snow leopard sequence, as observed with the *ND2* LogDet distance calculation, but shows a very unique topology. Figure 6 shows the topology with all available sequences combined and Figure 7 that of each *TTR* sequencing effort separately with associated bootstrap support values. The tiger and snow leopard sequences show 100% identity with one another, as do the lion, leopard, and jaguar sequences. Within the 779 bp sequenced from *TTR* intron 1, there are a total of 10 informative sites, 5 supporting each of the two clades.

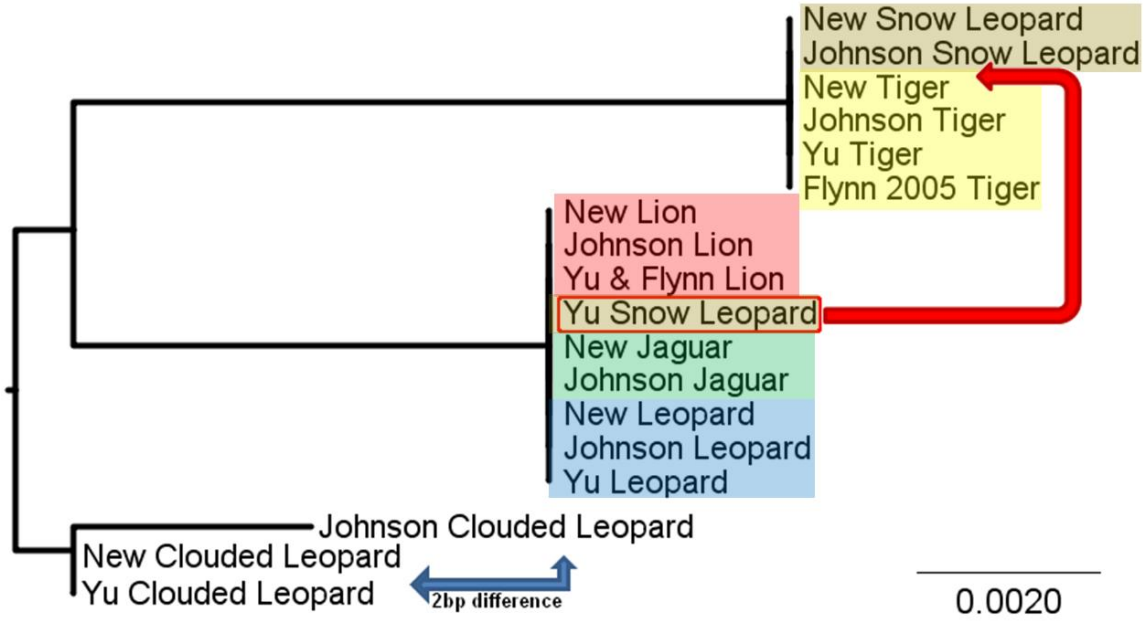


Figure 6. Maximum likelihood clustering analysis for *TTR* shows putative misidentification of the snow leopard sequence in the Yu publication (boxed in red). Evident is a unique phylogenetic topology tracking the divergence of the tiger-snow leopard and lion-leopard-jaguar clade with a lack of any subsequent changes since that event. Each clade has a total of 10 diagnostic sites differentiating each branch.

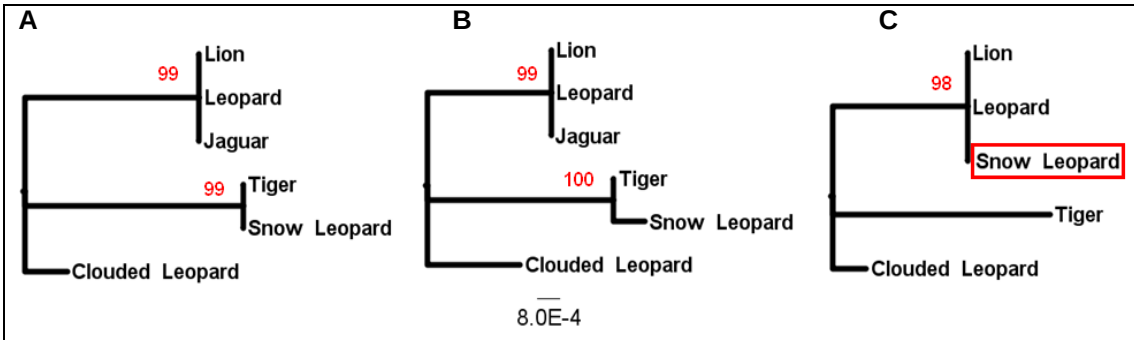


Figure 7. Topology of the transthyretin gene (*TTR*), intron 1 as determined by three separate publications with nonparametric bootstrap support (1000 replicates) indicated in red. (A) Newly generated sequences. (B) Johnson et al. (2006). (C) Yu and Zhang (2005).

Mitochondrial Vetting

To further determine whether the phylogenetic discrepancies observed in the *ND2* gene between the two independent publications could be attributed to species misidentification, or the amplification of a mitochondrial pseudogene present in the nuclear genome (numt), subsequent sequencing of *12S*, *CYTB*, *ND2*, and *ND4* was performed on in-house DNAs (provided by Dr. Oliver Ryder and Leona Chemnick at the Beckman Center for Conservation Research and Center for the Reproduction of Endangered Species, San Diego Zoo) using the previously described reagent and thermal cycler protocols. Direct sequencing of the lion *ND2* PCR amplicon produced unexplained heterozygosity in a portion of the sequence traces. Subsequent cloning and sequencing of this PCR product showed one clone to have a complete sequence when compared to other pantherines, however one possessed a 4 bp insertion at base 48 and a 70 bp

deletion at base 141, confirming a numt co-amplification. When newly sequenced and previously published *Panthera ND2* sequences (Appendix Table 1) were aligned, all sequences from the Johnson et al. (2006) dataset showed a 2bp deletion at base 369 for all six taxa. When unaligned with the remainder of the data, the sequence after the deletion was out of frame and created stop codons. This 2bp region was excluded in all analyses.

A complete collection of published mitochondrial sequences for each of the six pantherine taxa was obtained from GenBank and assembled into aligned matrices. Sequences that did not span the same regions used in the Johnson et al. 2006 and Yu and Zhang (2005) phylogenetic studies were excluded. Many sequences were from population genetic, phylogeographic, and genomic studies, including those used to characterize a mitochondrial to nuclear translocation in the tiger [140]. Five sequences [DQ151550, DQ257669, EF551002, EF551003, EF551004] were obtained from PCR products amplified from an enriched mitochondrial fraction, drastically reducing the likelihood of amplifying a numt [111; 140; 184]. RaxML was used to generate a ML tree for all individuals for each gene segment, and separated them into clades to identify numt outliers. By including all published and newly discovered numts, as well as purified mitochondrial DNA-based sequences, the numts could be readily identified (See Figure 8A for the *ND2* segment). This method has recently been used to remove numt sequences of the ATP8 gene segment from a phylogeographic study of both extant and extinct lions [203]. All mitochondrial

segments utilized in this study were vetted in this way and the resultant trees are depicted in Appendix Figure 1. Following removal of putative numt segments, the matrices were reanalyzed using RaxML. The resulting topologies clearly highlight species misidentifications (Figure 8B) and allow the selection of putatively cymt sequences. Specifically, the sequences from the Wei et al. (2009) publication [EF551004] were used to represent snow leopard in the final mitochondrial matrix, as well as unpublished GenBank sequences from the same [EF551002, EF551003, DQ257669] for leopard, tiger, and clouded leopard respectively, as their 2009 methods include enrichment for the mitochondrial DNA fraction, greatly reducing the possibility of numt amplification.

As criteria for inclusion of jaguar and lion sequences in the final matrix, the sequence could not be part of a putative numt clade, and it must be corroborated as an accurate species identification by at least one other sequence. Once these criteria were met, the sequences with the greatest length were selected first, followed by those with the most recent publication date. For both lion and jaguar, newly generated sequences were selected for *12S* and *CYTB*, Johnson et al. (2006) sequences were chosen for *ND1* and *ND5,* and the Johnson et al. (1997) sequences were selected for the *16S* segments of the final matrix. Newly generated lion sequences were included for *ND2* and *ND4*, and the sequences published by Yu and Zhang (2005) were selected for jaguar. Prior to performing phylogenetic analyses, all sequences were translated to verify that they did not contain stop codons. The anti-guanine bias for each

segment averaged 13.75% (Table 5) versus 32.5% for adenine, 27.9% for cytosine, and 25.85% for thymine, a further indication of true mitochondrial sequence. The only exception to this strong bias was that of the 16S sequence, which still retained an anti-G bias (20.66%), though not as pronounced.

Table 5. Base frequencies for each segment of the final mitochondrial matrix show a distinct anti-G bias that is characteristic of true mitochondrial DNA sequences.

	A	C	G	T
12S	0.3575	0.2360	0.1815	0.2250
16S	0.3280	0.2240	0.2066	0.2414
CYTB	0.2878	0.2970	0.1400	0.2752
ND1	0.3165	0.3457	0.0878	0.2500
ND2	0.3602	0.2938	0.0981	0.2479
ND4	0.3140	0.2894	0.1245	0.2722
ND5	0.3119	0.2669	0.1233	0.2980
Average	0.3251	0.2790	0.1374	0.2585

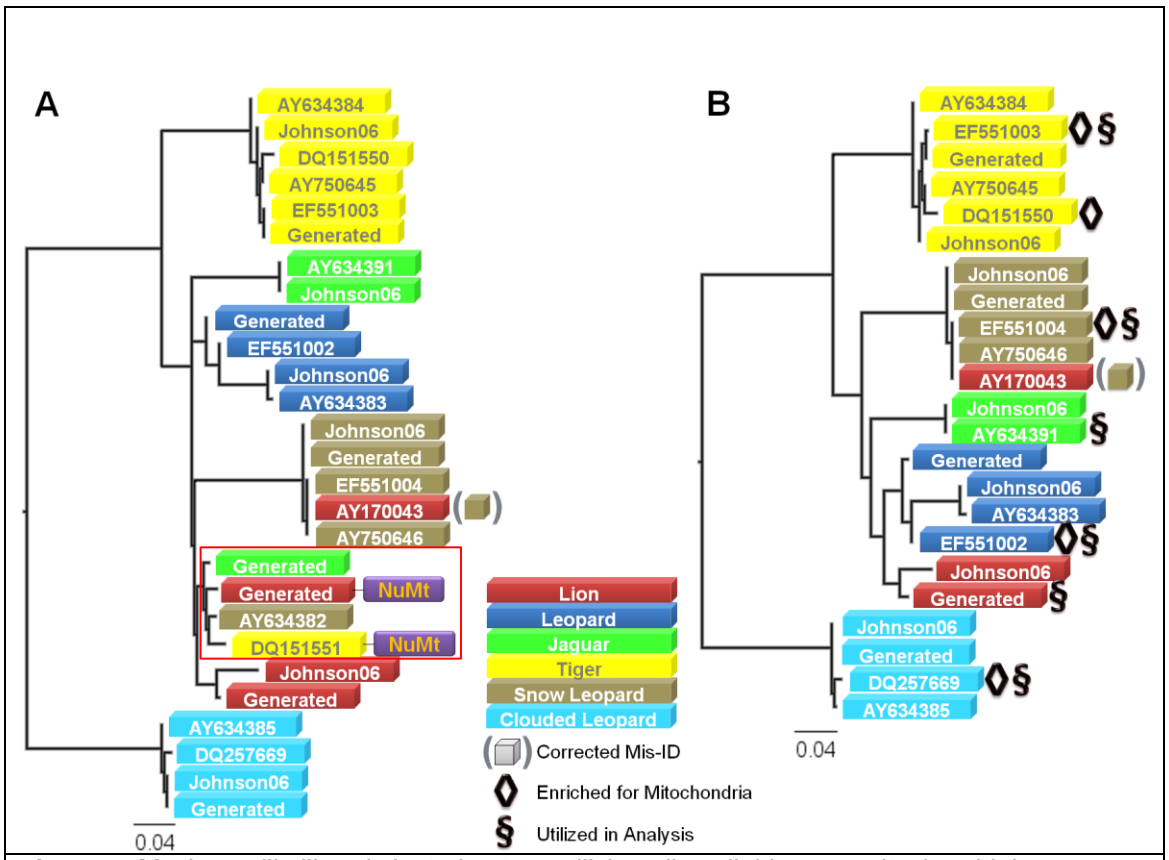


Figure 8. Maximum likelihood clustering tree utilizing all available *ND2* mitochondrial gene segments. (A) Clear evidence of numt amplification within the red box. (B) Clustering tree with numts removed. Evidence for species misidentification is shown in (B) for the lion sequence within the snow leopard clade. Corrected species identification is indicated in parentheses.

Phylogenetic Reconstruction

The final set of cytoplasmic mitochondrial sequences was assembled into the final supermatrix along with the newly generated *IRBP*, β fibrinogen (*FGB*) and Y chromosome gene segments, as well as the autosomal, Y chromosome and X chromosome loci from Johnson et al. 2006. An unpublished set of *CES7* gene sequences (~6161 bp) in all pantherines was also included in the final phylogeny (Li et al. unpublished). One small segment of *SMCY* overlapped between the newly generated and Johnson et al. (2006) dataset and the newly generated segment was used. The complete supermatrix consisted of 47,628 nucleotides (974 sites excluded as ambiguous) partitioned as follows: Y chromosome (19,140 bp), autosomes (19,124 bp), X chromosome (3,223 bp) and mitochondria (6,141 bp). Additional partitions used in analyses included: uniparental: (Y chromosome and mitochondria); and nuclear (X chromosome, Y chromosome, and autosomes).

To test for combinability between and within these partitions, the incongruence length difference test was used. The null hypothesis (H_0) is that the partitions are congruent enough to be combined in a concatenated matrix. The alternate hypothesis (H_1) is that the partitions contain statistically significant incongruent phylogenetic signal and should not be concatenated in order to avoid signal conflict. The suggested conservatism of the P-value to determine the threshold of congruence varies with the publication [204]. For the purpose of

this study, a value less than 0.01 was considered a rejection of combinability [181; 182]. The results of this test are listed in Table 6. The test indicated sufficient congruence for the combinability of individual gene segments within each partition, with the exception of the autosomal partition (ILD = 0.002). This partition was incongruent with the Y chromosome (ILD = 0.001), and was very close to incongruence with the X chromosome (ILD 0.016). Within the nuclear partition (autosomes, X chromosome, Y chromosome) there was also significant incongruence (ILD = 0.008). When the mitochondrial partition was added to any partition combination, the congruence increased above statistical significance due to the heterogeneity of the phylogenetic signal within this partition. The complete supermatrix passed ILD when partitioning each gene segment separately (43 partitions) and with the mitochondria and Y chromosome as individual partitions respectively (29 partitions). However, when partitioned only into the four inheritable regions (autosomes, mitochondria, X chromosome and Y chromosome, it was statistically incongruent and was therefore analyzed with multiple phylogenetic methods and varying partitioning schemes to ensure an accurate species tree topology.

Table 6. Partition homogeneity results (α=0.01). The supermatrix is internally incompatible when partitioned into the four genomic regions. Incongruence is observed for the autosomal, autosomal + Y and nuclear partitions. The autosomal + X partition was very close to statistical incongruence.

Partition	Sequence Length	# Partitions	P-value
Supermatrix	47,628 bp	4	P = 0.001*
Supermatrix	47,628 bp	29	P = 0.346
Supermatrix	47,628 bp	43	P = 0.981
Autosomes	19,124 bp	22	P = 0.002*
Y Chromosome	19,140 bp	9	P = 1.000
X Chromosome	3,223 bp	5	P = 1.000
Mitochondria	6,141 bp	7	P = 0.893
Uniparental	25,281 bp	16	P = 0.936
Nuclear	41,487 bp	36	P = 0.008*
Mitochondria + X	9,364 bp	12	P = 0.999
Mitochondria + Autosomes	25,265 bp	29	P = 0.955
Autosomes + Y	38,264 bp	31	P = 0.001*
Autosomes + X	22,347 bp	27	P = 0.016
Chromosomes X + Y	22,363 bp	14	P = 0.780
Mitochondria + X + Y	28,504 bp	21	P = 0.998
Mitochondria + Autosomes + Y	44,405 bp	28	P = 0.907
Mitochondria + Autosomes + X	28,488 bp	34	P = 0.998

Maximum likelihood and Bayesian phylogenetic inference produced identical rooted topologies for the complete supermatrix, as well as the Y chromosome and uniparental partitions (Figure 9). This topology places lion and leopard as sister taxa with jaguar as the basal member of this monophyletic clade. Tiger was placed as sister to snow leopard in a separate monophyletic group, with clouded leopard as the outgroup to *Panthera*. These associations were also constructed by ML for the autosomal and nuclear partitions, however BI produced lion and jaguar as sister taxa rather than lion and leopard (Figure 10). As Figure 9 and Table 7 shows, ML bootstrap values and Bayesian posterior probabilities (BPP) for lion-leopard monophyly were high for the rooted

uniparental (ML: 100, BPP: 1.0), and mitochondrial (ML: 99.8, BPP: 1.0) partitions; and moderate for the Y chromosome partition (ML: 70.0, BPP: 0.87). Support increased for the unrooted Y chromosome (ML 82.2, BPP: 0.99), remained static for the uniparental (ML: 100, BPP: 1.0), and virtually static for the mitochondrial partition (ML: 99.3, BPP: 1.0). Individual maximum likelihood topologies with branch lengths and clade support values are included as appendices for the partitions in Appendix Figure 2.

There was lower bootstrap support for lion-leopard monophyly in the autosomal (rooted: 66.6, unrooted 54.8) and nuclear (rooted 62.5, unrooted 55.5) partitions. Bayesian results recreated lion-jaguar monophyly for nuclear (BPP rooted: 0.97, BPP unrooted 0.99) and autosomal (BPP rooted: 0.59, BPP unrooted: 0.79) partitions, indicating varying histories of their individual loci. The Bayesian posteriors for each node showed significantly greater support than their likelihood counterparts, consistent with the liberal nature of the Bayesian method [205]. Despite this, it was clear that the lion-leopard monophyly was not well supported using only these partitions. The X chromosome partition recapitulated the alternative topology constructed by the autosomal and nuclear BI methods, placing lion and jaguar as a monophyletic clade to the basal exclusion of leopard. The support for this topology was not high for maximum likelihood (rooted: 61.1, unrooted: 85.4) but was high for Bayesian inference (rooted: 0.97, unrooted: 1.0).

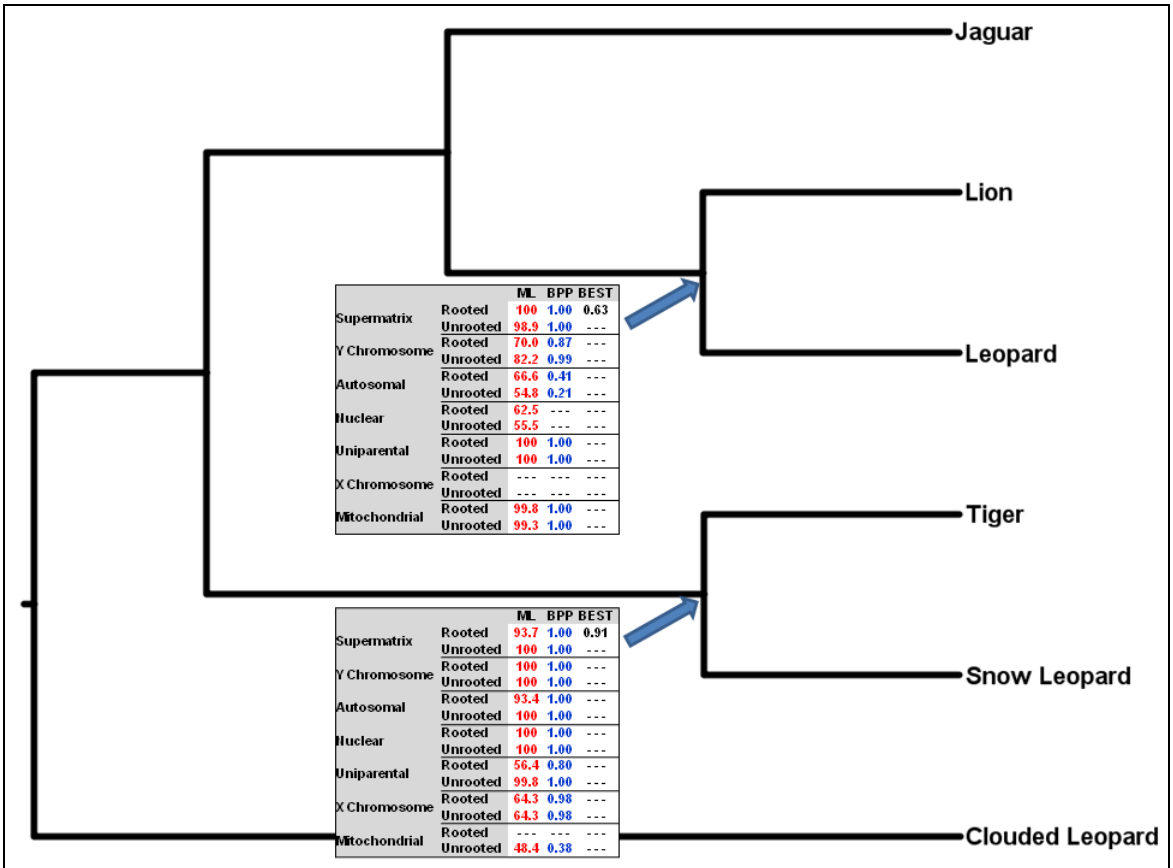
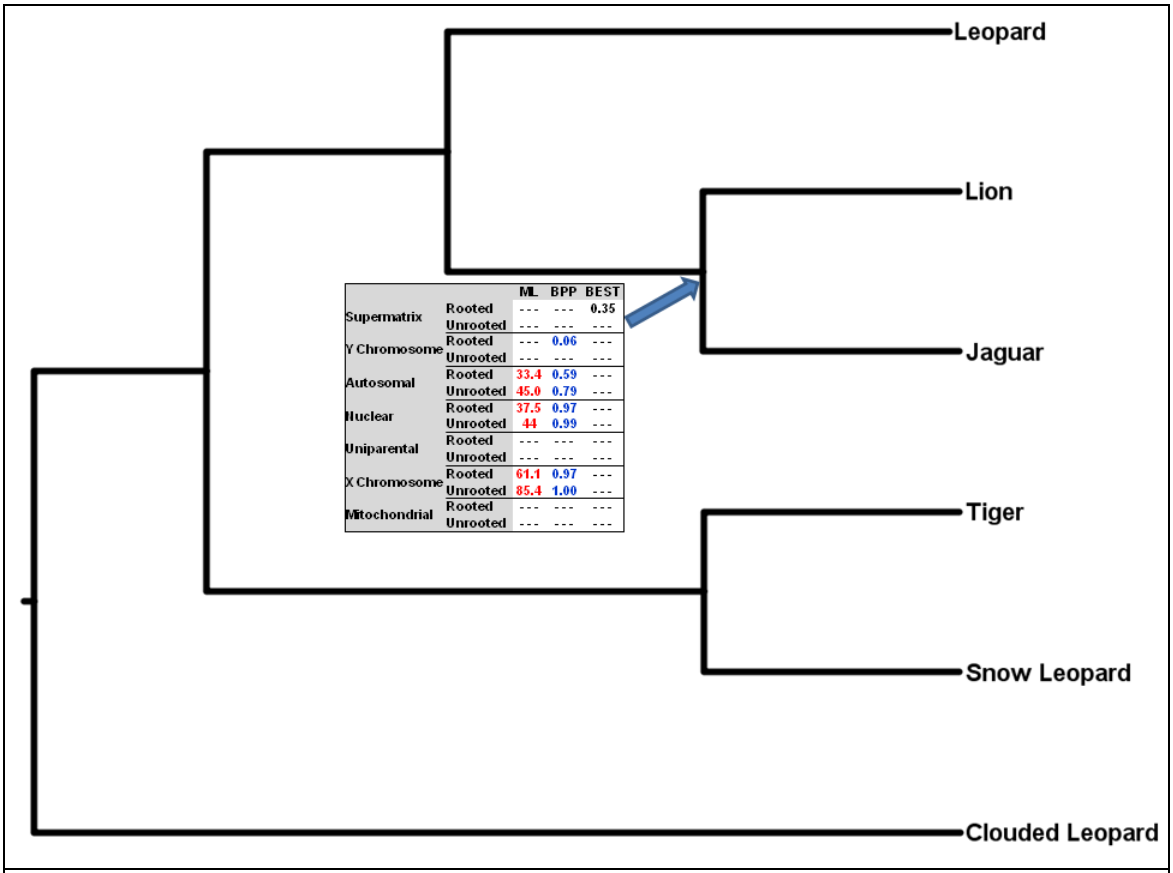


Figure 9. Cladogram depicting the maximum likelihood (ML) phylogenetic topology for *Panthera* generated by the supermatrix, Y chromosome, and autosomal partitions in this study. ML bootstrap values (1000 replicates) shown in red. Bayesian posterior probabilities (BPP) are in blue and BEST clade support values in black for the two primary incongruent nodes from prior studies.



		ML	BPP	BEST
Supermatrix	Rooted	---	---	0.35
	Unrooted	---	---	---
Y Chromosome	Rooted	---	0.06	---
	Unrooted	---	---	---
Autosomal	Rooted	33.4	0.59	---
	Unrooted	45.0	0.79	---
Nuclear	Rooted	37.5	0.97	---
	Unrooted	44	0.99	---
Uniparental	Rooted	---	---	---
	Unrooted	---	---	---
X Chromosome	Rooted	61.1	0.97	---
	Unrooted	85.4	1.00	---
Mitochondrial	Rooted	---	---	---
	Unrooted	---	---	---

Figure 10. Cladogram depicting the phylogenetic topology for *Panthera* generated by the maximum likelihood (ML) X chromosome, Bayesian autosomal / nuclear partitions from this study and Johnson et al. (2006). ML bootstrap values (1000 replicates) shown in red. Bayesian posterior probabilities (BPP) are in blue and BEST clade support values.

Table 7. Support for species relationships within genus Panthera for supermatrix and partitioned analyses. Nonparametric bootstrap values in red. Bayesian posterior probabilities in blue. Rooted analyses included all six taxa. Unrooted analyses performed without clouded leopard.

	Tiger Snow	Lion Leopard	Lion Jaguar	Lion Leopard Jaguar	Tiger Jaguar	Jaguar Snow	Tiger Clouded	Lion Jaguar Clouded
Rooted	93.7	100	- - -	100	- - -	- - -	6.0	- - -
Supermatrix 43 Paritions	1.00	1.00	- - -	1.00	- - -	- - -	- - -	- - -
29 Paritions	1.00	1.00	- - -	1.00	- - -	- - -	- - -	- - -
4 Paritions	0.70	1.00	- - -	1.00	- - -	- - -	- - -	- - -
Unrooted	100	98.9	- - -	100	- - -	- - -	N / A	N / A
Supermatrix	1.00	1.00	- - -	1.00			N / A	N / A
Rooted	100	62.5	37.5	100	- - -	- - -	- - -	- - -
Nuclear	1.00	- - -	0.97	1.00	- - -	- - -	- - -	- - -
Unrooted	100	55.5	44.4	100	- - -	- - -	N / A	N / A
Nuclear	1.00	- - -	0.99	1.00			N / A	N / A
Rooted	93.4	66.6	33.4	99.8	- - -	- - -	- - -	- - -
Autosomal	1.00	0.41	0.59	1.00	- - -	- - -	- - -	- - -
Unrooted	100	54.8	45.0	100	- - -	- - -	N / A	N / A
Autosomal	1.00	0.21	0.79	1.00	- - -	- - -	N / A	N / A
Rooted	56.4	100	- - -	100	- - -	- - -	34.0	- - -
Uniparental	0.80	1.00	- - -	1.00	- - -	- - -	- - -	- - -
Unrooted	99.8	100	- - -	100	- - -	- - -	N / A	N / A
Uniparental	1.00	1.00	- - -	1.00	- - -	- - -	N / A	N / A
Rooted Y	100	70	- - -	100	- - -	- - -	- - -	- - -
Chromosome	1.00	0.87	- - -	1.00	- - -	- - -	- - -	- - -
Unrooted Y	100	82.2	- - -	- - -	- - -	- - -	N / A	N / A
Chromosome	1.00	0.99	- - -	1.00	- - -	- - -	N / A	N / A
Rooted X	64.3	- - -	61.1	- - -	- - -	- - -	- - -	66.0
Chromosome	0.98	- - -	0.97	- - -	- - -	- - -	- - -	0.98
Unrooted X	64.3	- - -	85.4	- - -	- - -	- - -	N / A	N / A
Chromosome	0.98	- - -	1.00	- - -	- - -	- - -	N / A	N / A
Rooted	- - -	99.8	- - -	71.2	- - -	- - -	94.3	- - -
Mitochondrial	- - -	1.00	- - -	0.86	- - -	0.07	1.00	- - -
Unrooted	48.4	99.3	- - -	- - -	36.9	12.6	N / A	N / A
Mitochondrial	0.38	1.00	- - -	0.39	- - -	0.23	N / A	N / A

With respect to tiger-snow leopard monophyly, there was complete support (ML: 100, BPP: 1.0) from the Y chromosome, and nuclear partitions for both rooted and unrooted topologies. The autosomal partition showed identical support values with the exception of a slight drop in support in the rooted tree (ML: 93.4). The uniparental partition shared this high level of support in the unrooted conformation (ML: 99.8, BPP: 1.0), but support dropped (ML: 56.4, BPP: 0.8) when the outgroup was added. Support for tiger-snow leopard monophyly in the X chromosome partition was moderate for both rooted and unrooted topologies (ML: 64.3, BI: 0.98). The rooted X partition placed leopard as basal to the tiger-snow leopard clade, as opposed to the lion-leopard-jaguar clade as was supported by all other partitions, albeit with low support (ML: 0.66).

The rooted and unrooted topologies constructed by the mitochondrial partition are depicted in Figure 11. The topology of the rooted mitochondrial partition did not support tiger-snow leopard monophyly, but did support lion-leopard monophyly (ML: 99.8, BPP: 1.0). When the outgroup was removed for the unrooted conformation, the split between the two monophyletic clades was supported, but at a low level (ML: 48.4, BPP: 0.38). When removing the mitochondrial segments not resequenced by this study (*ND1*, *ND5* and *16S*), unrooted bootstrap support for the relationship between tiger and snow leopard increases to 63.2 (data not shown).

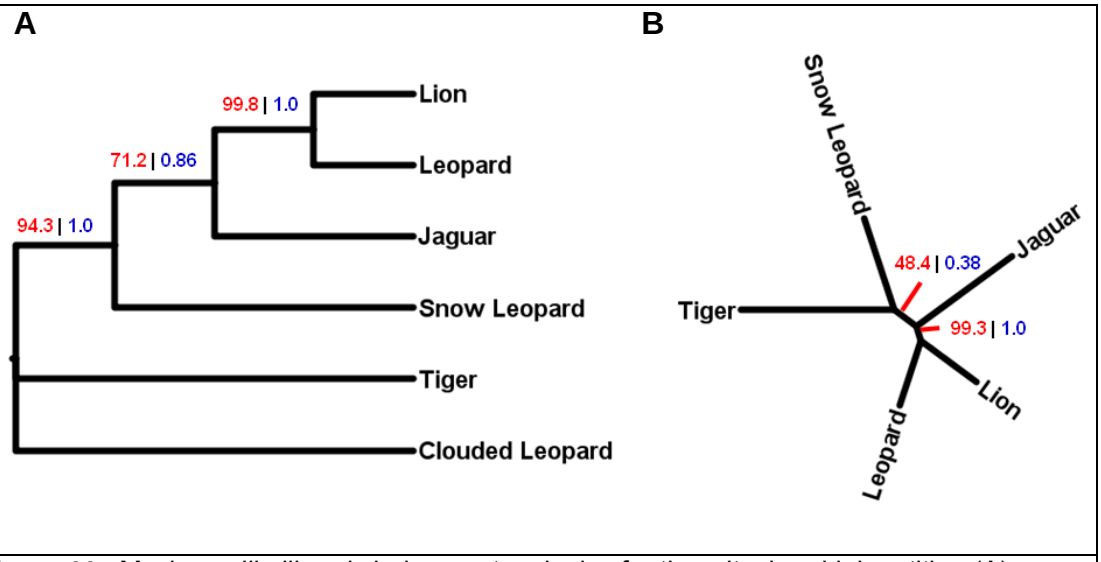


Figure 11. Maximum likelihood cladogram topologies for the mitochondrial partition (A) rooted and (B) unrooted. Maximum likelihood bootstrap values (1000 replicates) are shown in red, Bayesian posterior probabilities in blue.

The BEST method was implemented on the total matrix and also was able to reconstruct the same topology as the ML and BI supermatrix approaches, however with lower support for the lion-leopard monophyly (ML: 100, BPP: 1.0, BEST: 0.63) and tiger-snow leopard monophyly (ML: 100, BPP: 94, BEST: 91). The maximum likelihood topology with support values for all three analyses is shown in Figure 12.

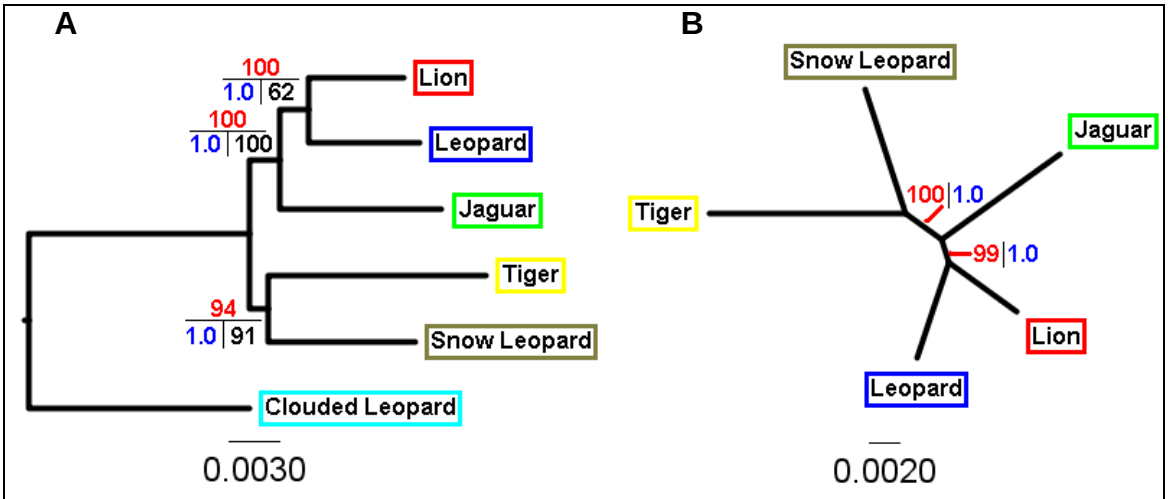


Figure 12. Maximum likelihood total evidence tree for the complete supermatrix. 1000 bootstrap replicate percentages depicted in red, Bayesian posterior probabilities in blue, and BEST posterior probabilities in black. (A) Rooted with clouded leopard as outgroup. (B) Unrooted topology.

In order to compare the topologies generated by these different approaches, a Shimodaira-Hasegawa (SH) test was performed. The SH test uses a statistic that is the likelihood score difference between the specified ML tree and every other tree compared. The null hypothesis (H_0) is that all trees are equally good explanations of the data. The alternate hypothesis (H_1) is that some or all trees are not equally good explanations of the data [183]. These comparisons were made between the topology supported by the supermatrix and each tree generated from each partition with a threshold for statistical significance of P>0.05.. The results of the SH test are shown in Table 8. The only trees with topologies discordant from that generated from the supermatrix were that of the mitochondrial and the X chromosome partition. Neither the

mitochondrial topology, nor the other partition topologies were significantly different than the supermatrix tree. However the X chromosome partition was significantly different than every other partition with the exception of the autosomal topology, though at P=0.574, the topological difference was virtually incongruent.

Table 8. Results of the Shimodaira-Hasegawa test. Significant topological differences (α=0.05) between partitions indicated in yellow. The likelihood data used is listed in the vertical column with the compared maximum likelihood topology in the uppermost horizontal row.

	Supermatrix	Y Chrom.	Autosomal	Mitochondrial	X Chrom.	Uniparental	Nuclear
Supermatrix		1.0000	0.5411	0.2948	0.1210	1.0000	1.0000
Y Chromosome	1.0000		0.5411	0.2948	0.1210	1.0000	1.0000
Autosomal	1.0000	1.0000		0.2948	0.1210	1.0000	1.0000
Mitochondrial	0.3169	0.1067	0.2485		0.0880	0.6023	0.0776
X Chromosome	0.0001	0.0132	0.0574	0.0090		0.0010	0.0073
Uniparental	1.0000	1.0000	0.5411	0.2948	0.1210		1.0000
Nuclear	1.0000	1.0000	0.5411	0.2948	0.1210	1.0000	

Tiger Lineage Acceleration

When examining the topologies in each gene segment within the mitochondrial partition, it was seen that the tiger lineage is accelerated with respect to every other species in many of the segments (Figure 13). To mathematically examine the possible acceleration of the tiger lineage, multiple tests were used to determine the presence of outliers. Statistical theory indicates that skewness and the Grubb's test are among the most powerful that can be used on such a small dataset [206] (Table 9). Skewness is the degree to

which a distribution departs from symmetry about its mean. Symmetric data should have a skewness near zero with significant positive skewness critical value ≈ 1.73 (n=5, α=0.05). The Grubb's test [207; 208] is one of the most popular tests for outliers with a critical value for statistical significance = 1.71 (n=5, α=0.05). Both tests resulted in positive outlier values at the border of statistical significance, with most of the signal originating from the NADH dehydrogenase subunit genes. The interpretation of these values should be performed with caution and are not concrete. The statistical power of this sample distribution is low, since the sample size is 5, the lowest end of applicability for both of these tests [206]. However, as outlier detection is ultimately subjective, a graphical representation of the acceleration and deceleration for each gene segment in each taxa based on the variance from the mean branch length measured from the terminal branch to the outgroup (Table 10) was constructed to visually identify potential lineage acceleration (Figure 14).

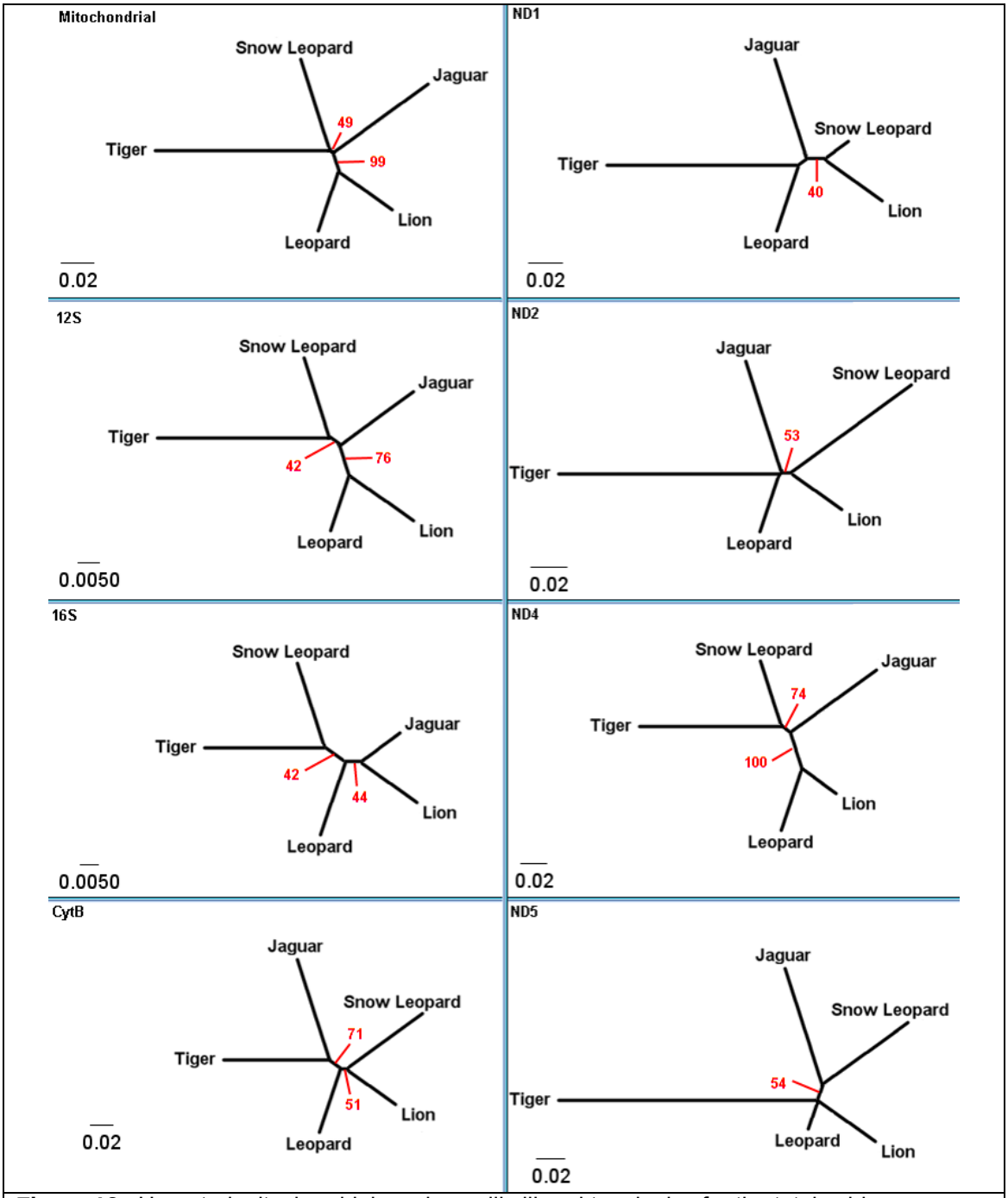


Figure 13. Unrooted mitochondrial maximum likelihood topologies for the total evidence mitochondrial data set and each component gene segment. 1,000 bootstrap replicate percentages shown in red.

Table 9. Branch lengths measured from each taxon to the outgroup in each gene segment with skewness and Grubb's tests for outliers results. Positive skewness critical value ≈ 1.73 (n=5, α=0.05) and Grubb's critical value for statistical significance = 1.71 (n=5, α=0.05). Values with the largest departure from normal distribution are in orange. Values at or approaching significance are in yellow.

	12S	16S	CytB	ND1	ND2	ND4	ND5	ALL
Lion	0.0661	0.2776	0.2147	0.2158	0.3349	0.2406	0.2646	1.6144
Leopard	0.0599	0.2804	0.2204	0.2240	0.3281	0.2565	0.2384	1.6078
Jaguar	0.0637	0.2726	0.2571	0.2393	0.3569	0.2567	0.2950	1.7412
Tiger	0.0763	0.2982	0.2424	0.3030	0.4077	0.2796	0.3984	2.0056
Snow Leopard	0.0541	0.2905	0.2323	0.1824	0.3752	0.2208	0.2851	1.6404
Skewness	0.6208	0.5788	0.4502	1.0031	0.7059	-0.1634	1.5334	1.7393
Grubb's P-value	1.4930	1.3868	1.3913	1.5797	1.4614	1.3781	1.6724	1.6950

Table 10. Variance in branch lengths from the mean, measured from each taxon to the outgroup in each gene segment. Nominally positive values (accelerated substitution) are in orange. Values with the highest values (highest rate of substitution) are in yellow. A graphical representation for the data is depicted in Figure 14.

	12S	16S	CytB	ND1	ND2	ND4	ND5	ALL
Lion	0.0021	-0.0062	-0.0187	-0.0171	-0.0256	-0.0103	-0.0317	-0.1075
Leopard	-0.0041	-0.0035	-0.0130	-0.0089	-0.0324	0.0057	-0.0579	-0.1141
Jaguar	-0.0004	-0.0113	0.0237	0.0064	-0.0037	0.0059	-0.0013	0.0193
Tiger	0.0123	0.0144	0.0090	0.0701	0.0471	0.0287	0.1021	0.2838
Snow Leopard	-0.0099	0.0066	-0.0011	-0.0505	0.0146	-0.0300	-0.0112	-0.0815

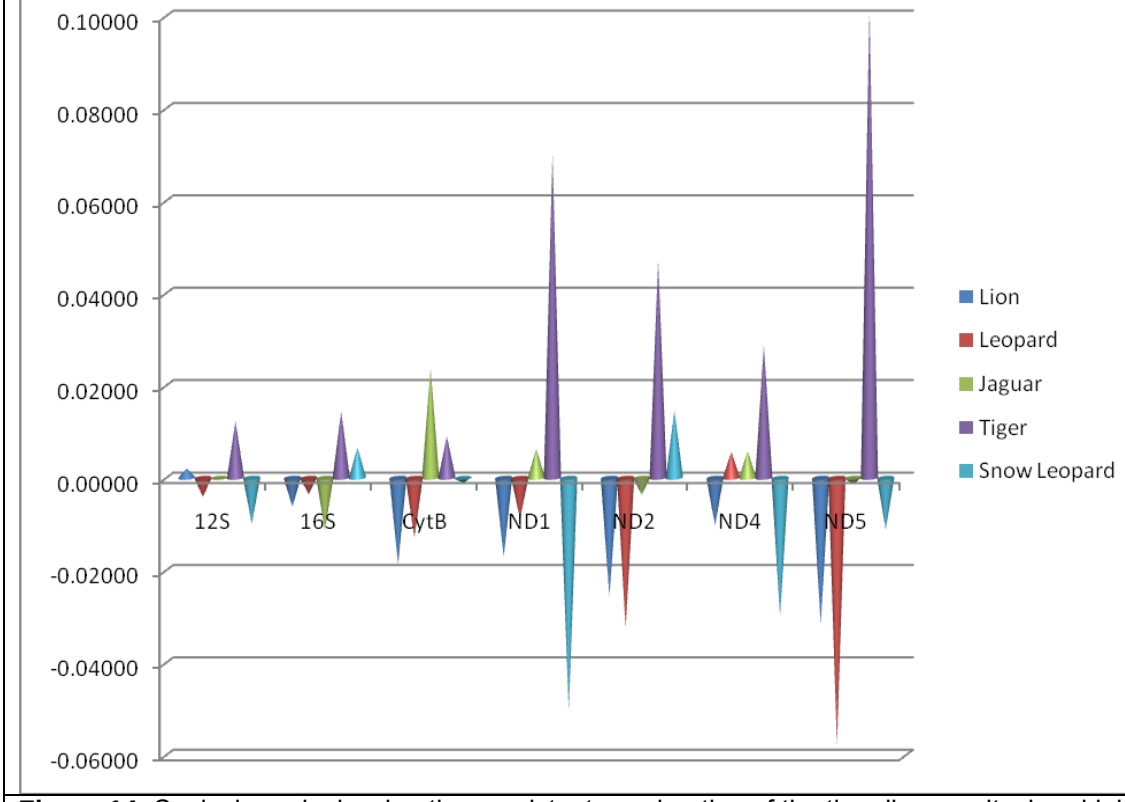


Figure 14. Conical graph showing the consistent acceleration of the tiger linage mitochondrial segments. The x-axis depicts the percentage divergence from the mean outgroup-to-tip branch length. Each gene segment and each taxa are shown. The higher the positive percentage divergence from the mean, the more accelerated the lineage's mutation rate for that gene.

Phylogenetic Signal

To quantify the signal contribution of each gene segment relative to the total support of the matrix, pairwise distance calculations were performed between each pair of all six taxa for each gene segment. Figure 15 graphically represents the average genetic distance between all pairwise comparisons.

This offered a basic metric to quantify the sequence variation, and therefore the amount of general phylogenetic signal per gene segment. There was significantly more sequence variation in the mitochondrial segments than there was in any other region by between 300% and 700%, with *ND2* as the most variable and *12S* as the least. A more detailed listing of the total number of identities and differences and the average percent identity between all pairs of taxa for each gene segment is shown in Appendix Table 5, with a graphic representation of the variation for each pairwise distance calculation within *Panthera* in Appendix Figure 3. This metric was able to indicate the general phylogenetic signal for each gene segment as any variable site between any two species. However, this method did not reflect other sources of phylogenetic signal such as a single nucleotide polymorphism present in only one species.

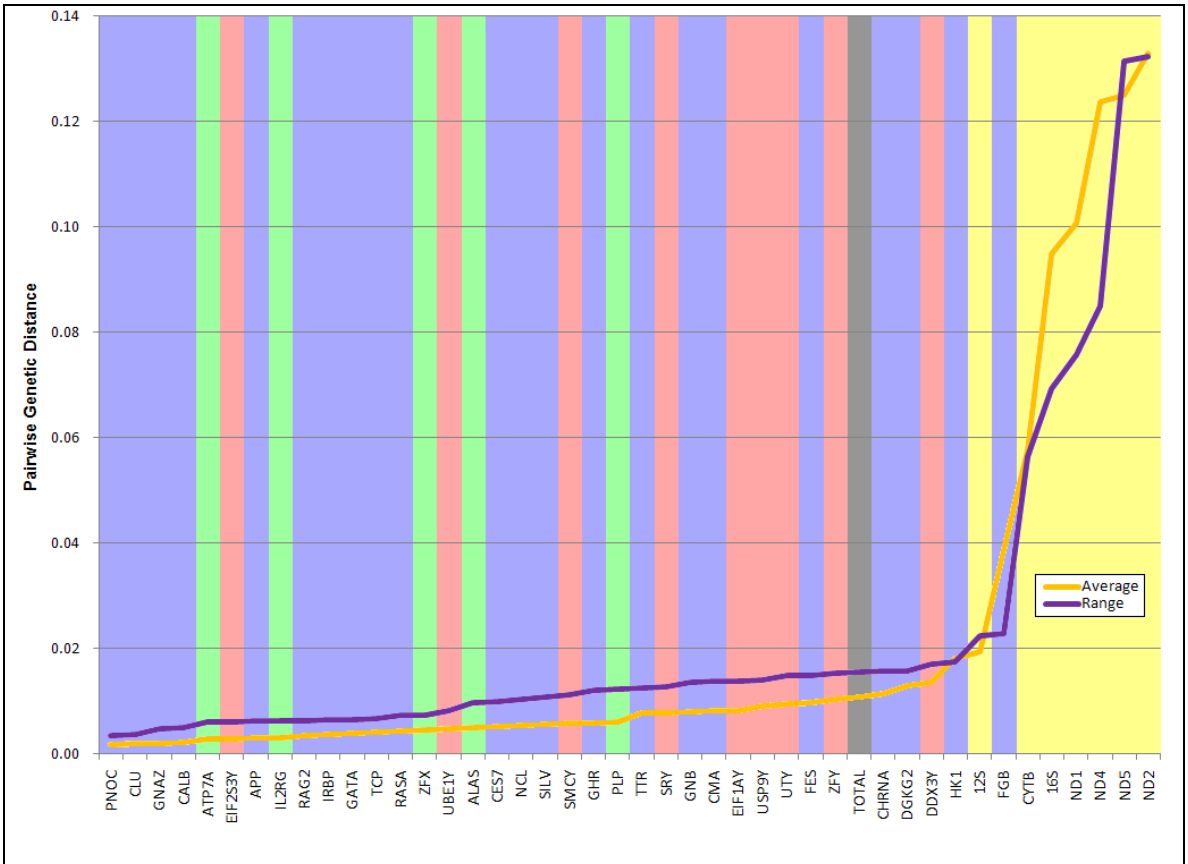


Figure 15. Total genetic variation per gene segment. Pairwise genetic distance between all six taxa shown on x-axis. Orange line denotes average pairwise distance and purple line denotes the range between all taxa. Regions in red denote Y-chromosome, blue: autosomes, green: X-chromosome, yellow: mitochondria.

Figure 16 shows two more refined statistics with the supporting calculations in Appendix Table 6. The ratio of the percentage of parsimony informative (PI) sites to the percentage of the total length of the supermatrix is represented by the blue line. For example, *ZFX* was 823 bp long. This was 1.72% of the supermatrix length and 25.54% of the X chromosome partition length. It contained 0.31% of the supermatrix PI sites and 66.7% of the partition's PI sites. This corresponded to a ratio of 0.18 for the supermatrix

(0.31% / 1.73%) and 2.61 for the ratio of the percentage of PI sites to the percentage of total sequence length in the partition attributed to *ZFX* (66.7% / 25.54%). These values were plotted on the x-axis to visualize the contribution of each gene segment to the total signal within the partition as well as the signal contribution within the supermatrix, normalized to the length of the gene segment. This method removes signal estimation bias towards longer sequence length. The ratio of 0.18 for 823 bp *ZFX* when compared to the 0.17 ratio in the long gene segment *USP9Y* (6,132 bp) demonstrated this bias. They carried roughly the same amount of phylogenetic weight; however site-for-site, *ZFX* had much higher signal content than *USP9Y*, even though it had only two PI sites. Three of the five genes in the X chromosome partition contributed no parsimony informative phylogenetic signal for the resolution of *Panthera* (*ALAS*, *ATP7A*, *IL2RG*). For the two genes that did contribute, *ZFX* drove two thirds of the signal for the entire partition (2 PI sites) and *PLP* the other third (1 PI site). This was useful to determine where the majority of the phylogenetic signal originated, gave approximate signal densities, and related it directly to the total signal contained in the supermatrix as well as the specific partition. However, this statistic still offered no specifics as to which taxonomic associations were being driven by the signal.

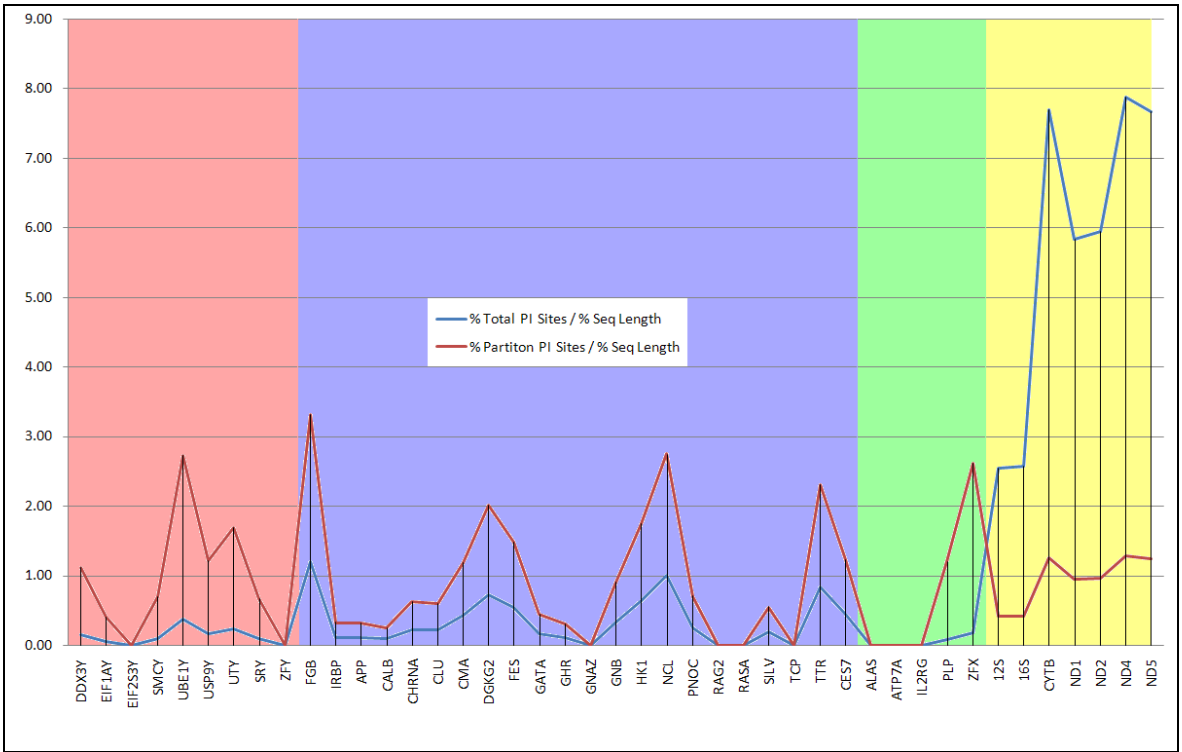


Figure 16. Phylogenetic signal from each gene segment contributing to the total supermatrix signal (blue line) and partition signal (red line) normalized to the gene sequence length to remove segment length bias. Each gene segment can be examined as the total parsimony information content with respect to its partition as well as to the total supermatrix length represented by the gene segment sequence. Regions in red denote Y-chromosome, blue: autosomes, green: X-chromosome, yellow: mitochondria.

The signal contribution of each segment specific to each taxonomic association was investigated in two ways. First, we removed each gene segment partition by jacknifing and recorded the resulting topological rearrangements with changing bipartition support (gene jacknifing, or GJ). The GJ bipartition support data for the supermatrix can be seen in Appendix Table 7 and Figure 17. As is seen in the figure, *w*hen the *ND4* segment was removed, the support for a monophyletic lion and jaguar increased sharply from zero to

37%. Though this was still a very low support value, such a sharp increase indicated strong phylogenetic signal in this segment. The autosomal partition shows a very pronounced spike in the support for the lion-jaguar monophyly and a drop for lion-leopard monophyly when *TTR* is jacknifed out (Figure 18). This was more evident without the unpublished CES7 sequence (Figure 19), since the bias associated with such a large autosomal gene segment is eliminated. Y chromosome partition GJ results decrease in support for lion-leopard monophyly when SMCY is removed (Figure 20). The moderate support for lion-jaguar monophyly drops to zero when PLP is removed, as did support for tiger-snow leopard monophyly and the spurious relationship of leopard-tiger-snow leopard when ZFX was removed from the X chromosome partition (Figure 21). The mitochondrial GJ plot (Figure 22) also shows significant topological rearrangement when *ND4* was removed, lowering support for monophyletic lion-leopard-jaguar as well as for the sister relationship of lion-leopard. Jacknifing *ND4* also increased support for monophyletic lion-leopard-snow leopard, lion-snow leopard, and jaguar-snow leopard; associations not supported by any other partition. This result further demonstrated the signal heterogeneity within the mitochondrial genome. The remainder of the jackknifed bootstrap support values are listed in Appendix Tables 8-12.

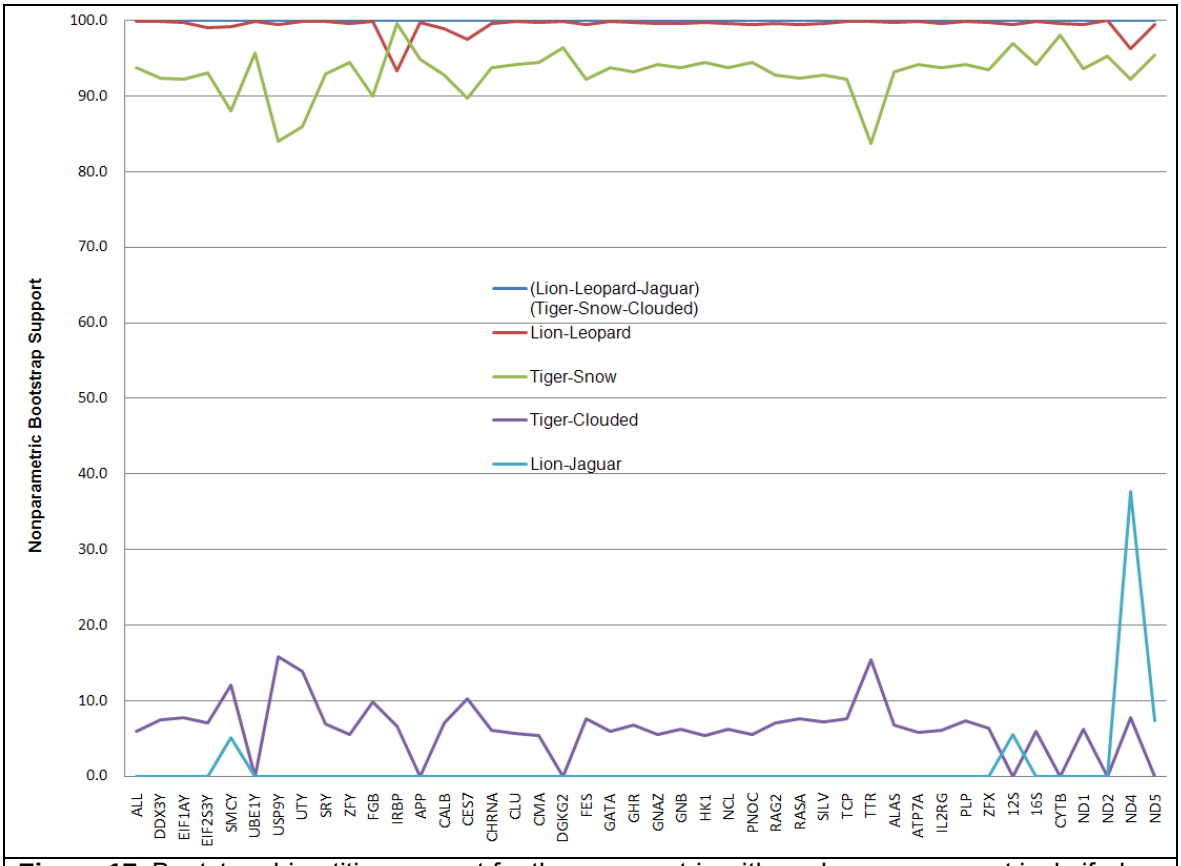


Figure 17. Bootstrap bipartition support for the supermatrix with each gene segment jacknifed out. The Y-axis shows bootstrap percentages based on 1000 replicates.

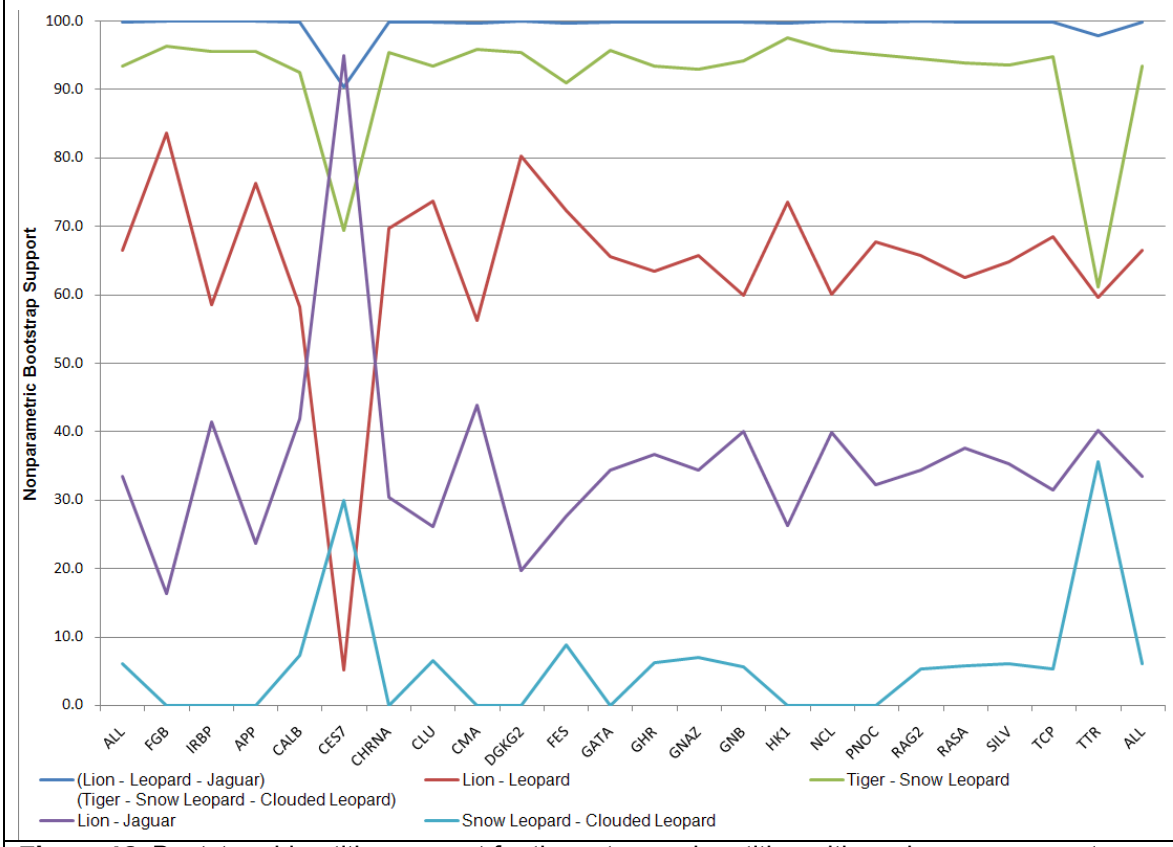


Figure 18. Bootstrap bipartition support for the autosomal partition with each gene segment jacknifed out. The Y-axis shows bootstrap percentages based on 1000 replicates. Note the topological change when both *TTR*, and *CES7* gene partitions are removed from the dataset.

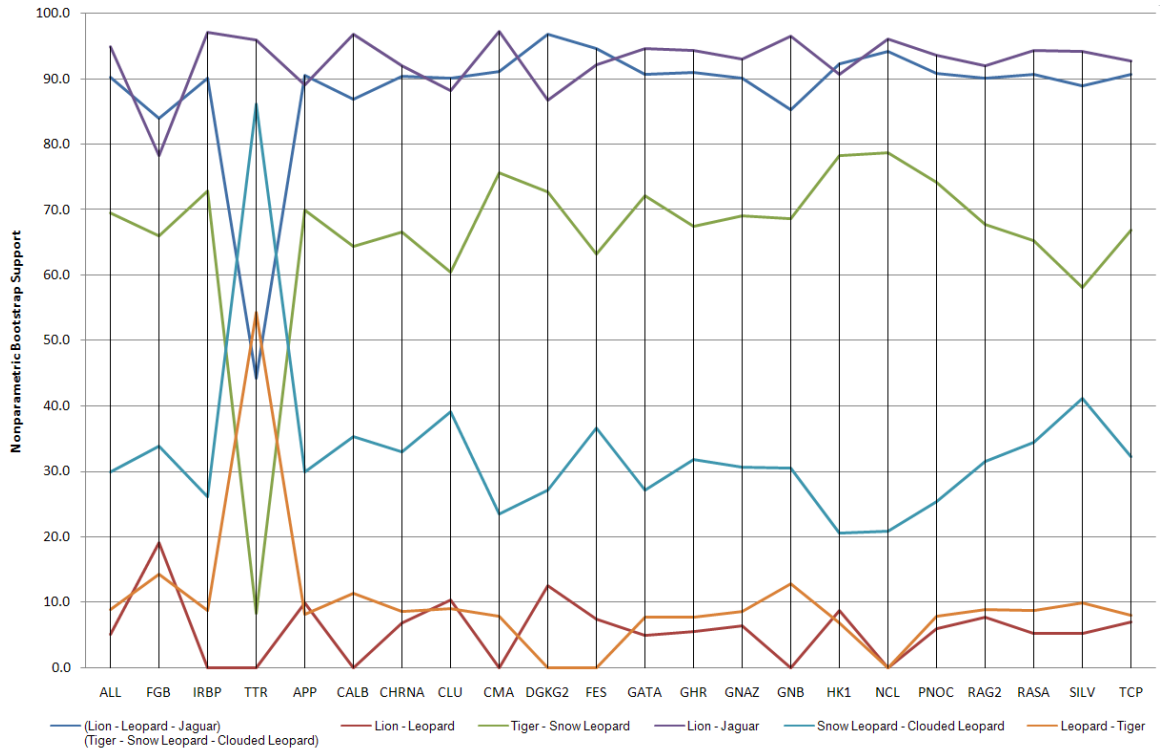


Figure 19. Bootstrap bipartition support for the autosomal partition (no *CES7*) with each gene segment jacknifed out. The Y-axis shows bootstrap percentages based on 1000 replicates. Note the topological change present when *TTR* is removed from the dataset.

Figure 20. Bootstrap bipartition support for the Y chromosome partition with each gene segment jacknifed out. The Y-axis shows bootstrap percentages based on 1000 replicates. Note the bootstrap support change present when *SMCY* is removed from the dataset.

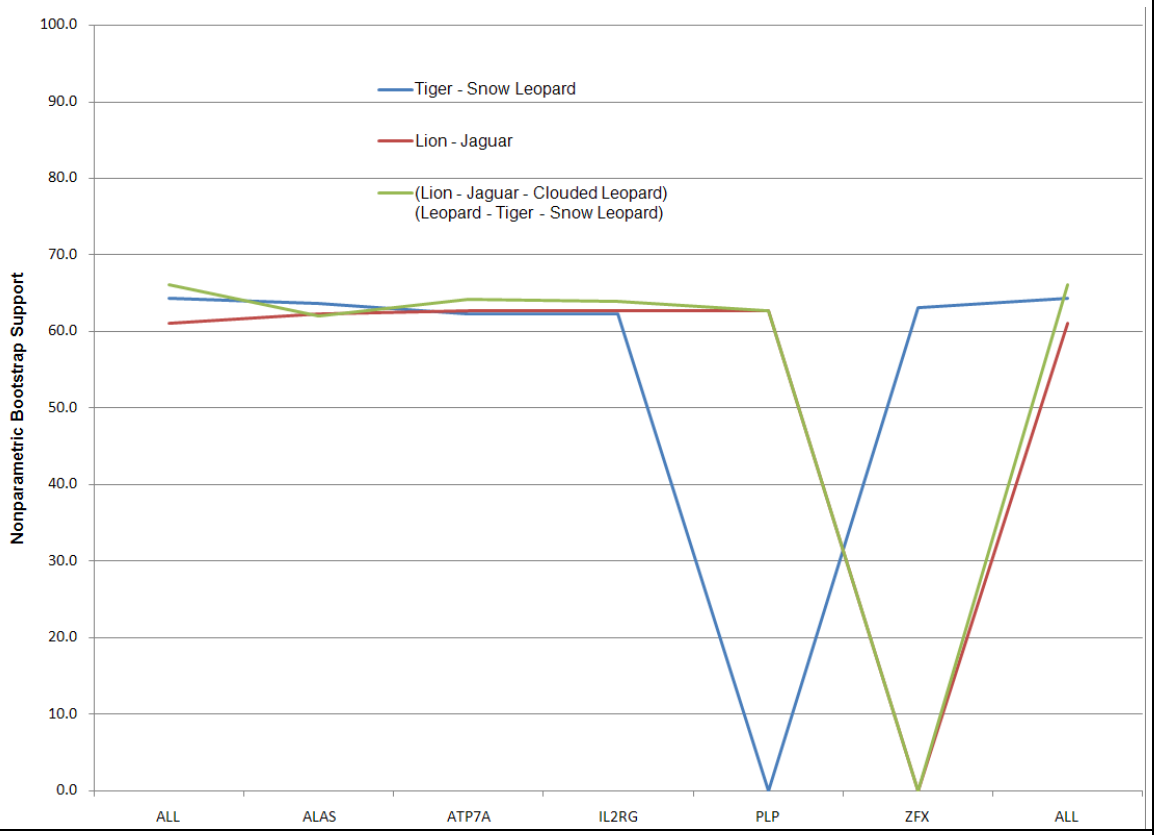


Figure 21. Bootstrap bipartition support for the X chromosome partition with each gene segment jacknifed out. The Y-axis shows bootstrap percentages based on 1000 replicates. Note the topological change in bootstrap support when both *ZFX and PLP are* removed from the dataset.

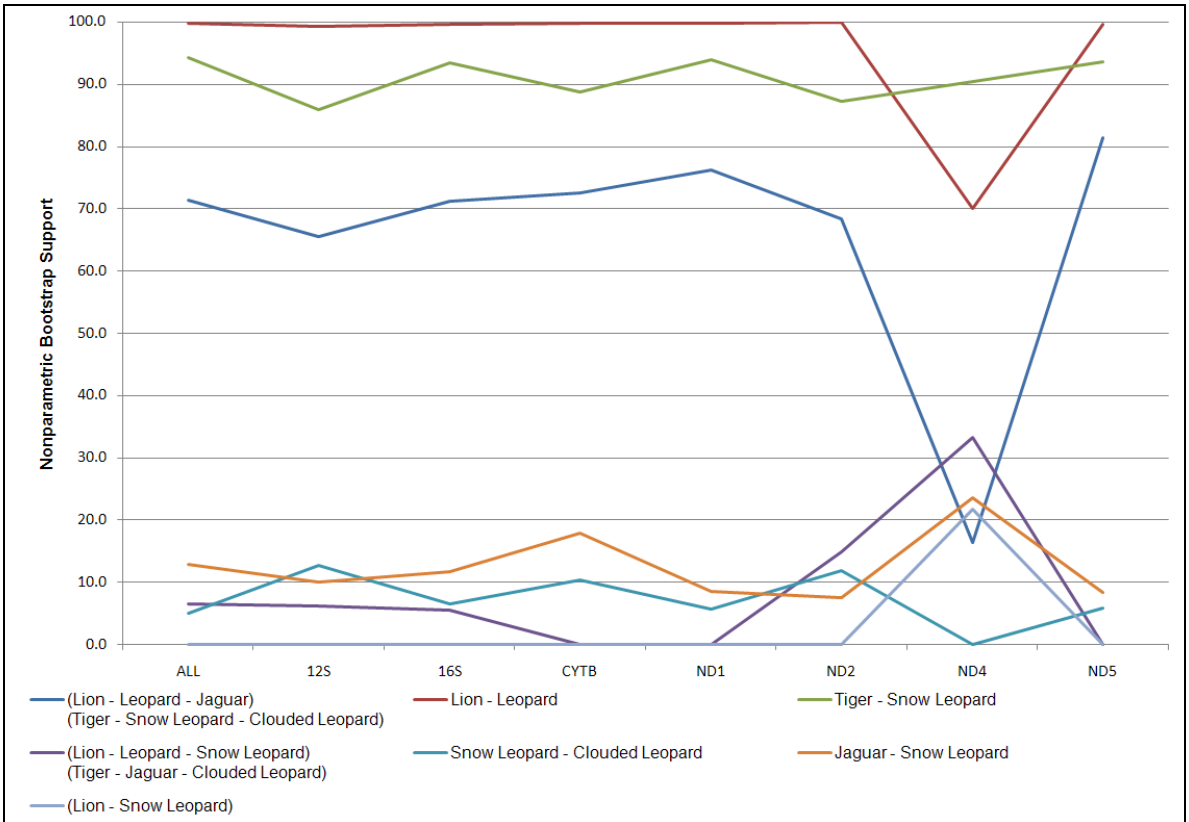


Figure 22. Bootstrap bipartition support for the mitochondrial partition with each gene segment jacknifed out. The Y-axis shows bootstrap percentages based on 1000 replicates. Note the topological change present when *ND4* is removed from the dataset.

To further explore the phylogenetic signal supporting each taxonomic relationship, the total PI sites for each specific bipartition was tablulated for each gene segment across the supermatrix with the results in Appendix Tables 13-15. A more summarized partitioned result is shown in Table 11. Both tables show large signal differences between partitions, and even within the autosomes, X chromosome and mitochondrial partitions. The mitochondrial partition alone contained almost 80% of the total PI sites in the supermatrix. The high percentage of phylogenetic signal originating from the mitochondrial partition highlights the care with which these sequences should be prepared and vetted to deter anomalous species associations. Aberrant signal can be attributed to many natural causes including homoplasy and a low signal-to-noise ratio for these segments; as well as procedural errors such as erroneous amplification of numts.

Table 11. Percentage of PI sites within each partition supporting each species relationship. Monophyly supported by the supermatrix topology in this study is highlighted in green. The topology from Johnson et al. 2006 is in orange.

	Y Chrom.	Autosomes	X Chrom.	Mitochondria	Supermatrix	Total #
(Lion-Leopard) (Jaguar-Tiger-Snow)	2.99%	14.94%	0.00%	11.16%	11.13%	118
(Tiger-Snow) (Lion-Leopard-Jaguar)	88.06%	36.36%	33.33%	10.68%	19.43%	206
(Lion-Jaguar) (Leopard-Tiger-Snow)	1.49%	11.69%	66.67%	5.64%	6.60%	70
(Leopard-Tiger) (Lion-Jaguar-Snow)	0.00%	9.09%	0.00%	8.52%	8.02%	85
(Leopard-Jaguar) (Lion-Tiger-Snow)	0.00%	5.84%	0.00%	6.48%	5.94%	63
(Jaguar-Snow) (Lion-Leopard-Tiger)	0.00%	5.19%	0.00%	10.32%	8.87%	94
(Lion-Tiger) (Leopard-Jaguar-Snow)	0.00%	2.60%	0.00%	5.28%	4.53%	48
(Leopard-Snow) (Lion-Tiger-Jaguar)	0.00%	2.60%	0.00%	6.60%	5.57%	59
(Tiger-Jaguar) (Lion-Leopard-Snow)	0.00%	0.00%	0.00%	8.88%	0.00%	106
(Lion-Snow) (Leopard-Tiger-Jaguar)	0.00%	0.00%	0.00%	2.52%	0.00%	42

Microsatellite Characterization

Within the introns utilized for phylogenetic inference, 24 segments showed repetitive sequence character indicative of potential microsatellites. These characters included di-, tri-, or tetra-nucleotide blocks repeated from 8 to 16 times in one contiguous sequence. These markers were evaluated in 11 domestic cat and 75 ocelot individuals to quantify their polymorphic character in these species. Of the 24 markers examined, only two showed polymorphic character: SMCY-7, $(ATTT_{11})$ a tetra-nucleotide microsatellite repeated 11 times in the domestic cat BAC clone sequenced and SMCY-2, (TC_9) a di-nucleotide

microsatellite repeated 9 times in domestic cat. In ocelot, both of these markers demonstrated polymorphic character (Janečka, unpublished). SMCY-2 had three alleles, with 6 individuals possessed a TC_8 allele and one a TC_7 allele. SMCY-7 also had three alleles in ocelot with four individuals carrying the $ATTT_{12}$ allele and three the $ATTT_{10}$ allele. Within the 11 domestic cats sampled, only SMCY-7 existed in multiple alleles, with two individuals possessing the $ATTT_{10}$ allele.

Molecular Dating

Multiple approaches can be taken to date the divergences of species within *Panthera*. To investigate if this large dataset can be dated using a molecular clock approach, a relative rate test was performed in PAUP for each partition within the concatenated dataset (supermatrix, nuclear, X chromosome, Y chromosome, autosomes, mitochondria, and uniparental) with degrees of freedom = 4. The results shown in Table 12 show that at $\alpha=0.05$, no partitions behave in a clocklike manner.

Table 12. Molecular clock test results. L_0 is the unconstrained likelihood value and L_1 is the likelihood value when the topology is constrained using the molecular clock hypothesis.

		-ln L	L_0-L_1	2(L_0-L_1)	P value (df=4)
Supermatrix	L_0	79776.83	344.004	688.01	0.00000
	L_1	80120.84			
Nuclear	L_0	61892.37	219.756	439.51	0.00000
	L_1	62112.13			
Uniparental	L_0	45148.36	299.84	599.68	0.00000
	L_1	45448.20			
Y	L_0	27591.79	198.425	396.85	0.00000
	L_1	27790.22			
X	L_0	4775.93	8.79662	17.593	0.00148
	L_1	4784.73			
Mitochondria	L_0	15818.45	139.333	278.67	0.00000
	L_1	15957.78			
Autosomal	L_0	29010.00	60.332	120.66	0.00000
	L_1	29070.33			

In lieu of using a strict molecular clock to date the divergences within *Panthera*, a bayesian relaxed clock approach was implemented with MULTIDIVTIME using a probabilistic model to quantify the change in evolutionary rate over time. As in other Bayesian algorithms, MULTIDIVTIME uses a Markov chain Monte Carlo algorithm to derive the posterior distribution of mutation rates and divergence times. The results of the dating analysis are shown with 3 fossil calibrations and one molecular-based maximum in Table 13. The 95% Bayesian credibility interval for the basal divergence time of genus *Panthera* was 3.80 – 4.31 MYA, corresponding to the initial bifurcation of the tiger-snow leopard clade and the lion-leopard-jaguar clade. The jaguar split from the lion-leopard lineage was between 2.56 and 3.66 MYA, while the lion and leopard diverged 1.95 – 3.10 MYA. The snow leopard and tiger diverged

from one another roughly 2.70 – 3.70 MYA. As shown in Table 13, the removal of each individual internal calibration point did not significantly affect the divergence time or the 95% credibility interval, with the exception of the minimum for the base of *Panthera*. Removal of this minimum reduced the divergence times at each node by roughly 40-50%.

Table 13. Divergence time estimates for *Panthera* calculated by PAML and the MULTIDISTRIBUTE software packages. Effects of removing each fossil calibration individually and combined are shown along with the standard error and Bayesian 95% highest posterior densities.

	Date (MYA)	Std. Error	95% CI	Constraint
Leopard–Lion	2.51791	0.29980	1.94543 – 3.09747	None
(Leopard-Lion)–Jaguar	3.11937	0.28788	2.55964 – 3.65988	Min = 1.6
Tiger – Snow	3.19020	0.25875	2.69694 – 3.70713	Min = 1.8
(Tiger-Snow)–(Leopard-Lion-Jaguar)	3.93854	0.14132	3.80334 – 4.31548	Min = 3.8
Leopard–Lion	1.42900	0.12029	1.21738 – 1.68751	None
(Leopard-Lion)–Jaguar	1.77558	0.10976	1.61313 – 2.02800	Min = 1.6
Tiger–Snow	1.87336	0.07335	1.80196 – 2.06906	Min = 1.8
(Tiger-Snow)–(Leopard-Lion-Jaguar)	2.08457	0.13506	1.87767 – 2.39847	No Min
Leopard–Lion	2.52895	0.29429	1.95579 – 3.09698	None
(Leopard-Lion)–Jaguar	3.12846	0.27900	2.56703 – 3.65990	Min = 1.6
Tiger–Snow	3.19831	0.25813	2.70868 – 3.71997	No Min
(Tiger-Snow)–(Leopard-Lion-Jaguar)	3.93983	0.14295	3.80351 – 4.32539	Min = 3.8
Leopard–Lion	2.52864	0.29452	1.96272 – 3.10415	None
(Leopard-Lion)–Jaguar	3.12868	0.28137	2.57948 – 3.67079	No Min
Tiger–Snow	3.19702	0.25972	2.70633 – 3.72084	Min = 1.8
(Tiger-Snow)–(Leopard-Lion-Jaguar)	3.94055	0.14360	3.80344 – 4.32649	Min = 3.8
Leopard–Lion	1.33691	0.08812	1.17724 – 1.52900	None
(Leopard-Lion)–Jaguar	1.66255	0.06424	1.60147 – 1.83505	Min = 1.6
Tiger–Snow	1.61874	0.11499	1.40940 – 1.86766	No Min
(Tiger-Snow)–(Leopard-Lion-Jaguar)	1.88204	0.11608	1.70390 – 2.15993	No Min
Leopard–Lion	1.40496	0.13563	1.14014 – 1.68243	None
(Leopard-Lion)–Jaguar	1.74801	0.13199	1.49718 – 2.01808	No Min
Tiger–Snow	1.86972	0.07289	1.80168 – 2.06325	Min = 1.8
(Tiger-Snow)–(Leopard-Lion-Jaguar)	2.07479	0.13531	1.86172 – 2.38737	No Min

Discussion

In this study, we provide an independent assessment of the pantherine phylogeny by supplementing previously published datasets with newly generated sequences from 39 Y-linked segments, three autosomal genes, and four mitochondrial genes. Phylogenetic inference provided by maximum likelihood, Bayesian phylogenetic inference, and Bayesian Estimation of Species Trees all recreate the same topology. This phylogenetic tree, depicted in Figure 12 represents the most comprehensive dataset for *Panthera*, spanning the largest number of genomic regions, including the most sequence data, and the strictest vetting process for determining bonafide mitochondrial genes as opposed to numt heterologs.

The Y chromosome had the most consistent signal across all gene segments, indicating its usefulness as a reliable marker with little heterogeneity within its single-copy loci (Y Chromosome HI: 0.1364, RCI: 0.8745; Autosomes HI: 0.3713, RCI: 0.4465; Mitochondria HI: 0.4435, RCI: 0.2478). This is consistent with the conclusions of Pecon-Slattery et al. (2004), which indicated the high quality of phylogenetic signal present in the felid MSY and very low incidence of convergent, parallel, or reversal substitutions [109]. It escapes recombination as does the mitochondria, but is not subject to the high mutation rate associated with the cytoplasmic organelle. This makes it a prime candidate for future phylogenetic study. Though its evolutionary history may differ from the

autosomes or the mitochondria, comparing the sequence in this region with data from other inheritance regions as was done in this study is the most comprehensive way to address rapid, recent speciations.

The lack of complete congruency of the autosomal partition from the ILD test (Table 6) and the presence of highly varied support for multiple topologies in the signal quantification for each autosomal gene segment (Appendix Table 14) demonstrate that the autosomal partition is subject to a large amount of signal heterogeneity. As such, gene trees in this partition and within the reminder of the supermatrix exhibited heterogeneity as to the internal phylogenetic relationships between these six species. In such instances, concatenation in a Bayesian phylogenetic inference framework as is implemented by MrBayes may overestimate the posterior probability [188]. The BEST analysis was performed to estimate the final species tree from individual gene trees, avoiding the Bayesian posterior overestimation and allowing for heterogeneity among loci. The results of the BEST analysis are lower in confidence than that of ML or BI. This may be the result of after-speciation gene flow (hybridization), which the author cautions may result in decreased confidence levels in the species tree estimation [209]. Though the support levels are lower, the overall topology has been reconstructed with moderate confidence. Despite the extreme heterogeneity of the dataset, the reconstruction of the final topology depicted in Figure 12 shows a highly corroborated phylogeny utilizing data from every inheritable portion of the genome.

Phylogenetic Reconstruction

One aspect of the topology generated in this study, the true sister taxon of the lion, has been a topic of debate in previous studies. This study places lion as sister to leopard, with jaguar as the immediately basal species to this clade. This aspect of the topology is supported by multiple other studies, including a recent analysis of 45 osteological, and 13 soft tissue, and behavioral characters [210], which places leopard as a closer relative to lion than both the extinct American lion (*P. l. atrox*) and cave lion *P. l. spelaea*). [188]. Other recent research by Barnett et al. (2009) utilized a median-joining network analysis of the mitochondrial hypervariable region 1 and ATP synthase F0 subunit 8 (*ATP8*) which placed the two extinct lions closer to the extant lion, but maintained the monophyly of lion and leopard [203]. This sister relationship is also solidified by other studies utilizing mitochondrial information such as RFLP analysis of the complete *Panthera* mitochondrial genomes [105], and characterization of the variability of the mitochondrial control region. These two publications have been the only phylogenetic studies for *Panthera* thus far to control for the amplification of numts and therefore are more reliable than those that did not. In addition, the lion-leopard sister relationship is supported by non-molecular studies such as the characterization of the chemical components of anal sac secretions [107], as well as older morphological studies [88]. As is shown in Figure 23, the signal attributed to this sister relationship was present within each partition, with the bulk originating in the mitochondrial partition. The Johnson et al. (2006) study

recreated jaguar as sister to lion, the only study to support this relationship. When the partitioned PI signal supporting each specific topology was compared to the total PI signal supporting either topology, the total amount of PI signal supporting lion-leopard monophyly in the mitochondrial partition (49%) was greater than the total PI signal supporting the lion-jaguar monophyly (37%). In addition, lion-leopard monophyly was supported by a higher percentage of the total PI sites in every partition with the exception of the X chromosome as seen in Table 11. In the Y chromosome (lion-leopard: 3%, lion-jaguar 1.5%) and mitochondrial partitions (lion-leopard: 11.16%, lion-jaguar: 5.64%), as well as in the supermatrix (lion-leopard: 11.13%, lion-jaguar: 6.6%), there was roughly twice the support for lion-leopard than there was for lion-jaguar. Within the autosomes, support favored lion-leopard at a lower margin (lion-leopard: 14.94%, lion-jaguar: 11.69%), with the level of support for lion-jaguar intermediate between lion-leopard and the monophyly of leopard and tiger (9.09%). This comparison and the remainder of the statistics in Table 11 highlighted the varied phylogenetic signal and partially explained the ILD test approaching statistical incongruence within this partition (Table 6). Within the 3,223 bp of the X chromosome partition, there were only three parsimony informative sites. Only one of these sites supported lion-jaguar, with the same site supporting leopard-tiger-snow leopard, with the other two supporting tiger-snow leopard monophyly (Appendix Table 15).

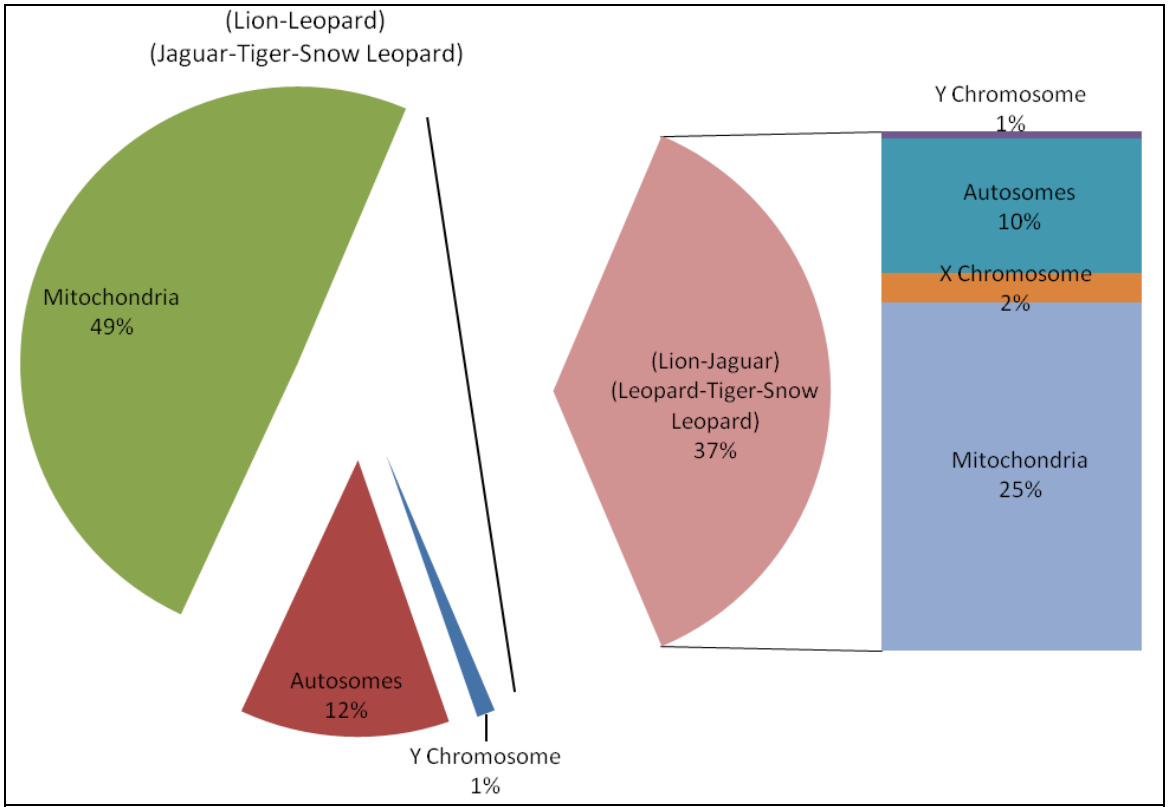


Figure 23. Percentage of parsimony informative sites by partition supporting the bipartition (lion-leopard)-(jaguar-tiger-snow leopard) and the bipartition (lion-jaguar)-(leopard-tiger-snow leopard). Percentages based on the 188 total informative site-relationships supporting either partition.

A second aspect of the topology generated in this study that has been contested in past studies is the monophyly of tiger and snow leopard. Maximum likelihood support for this was high within the supermatrix (rooted: 93.7, unrooted: 100). The BI clade support values calculated for each partition were high for the rooted supermatrix (BPP=1.0) both partitioning every gene segment separately, as well as combining the Y chromosome and mitochondrial segments into a single partition, respectively. This support decreases (BPP=0.70) when the data is divided into four partitions (Y chromosome, X

chromosome, autosomes and mitochondria). This reflects underpartitioning of the matrix, which can cause the selection of specific model parameters to be inapplicable across the entire partition. In two separate studies, when looking at morphological, ethological, and physiological features, Hemmer indicated that *Panthera* appeared to divide into two distinct clades [211; 212]. According to his studies, lions, leopards and jaguars share a specific set of common characters and can be separated from the second large cat clade containing the tiger, supporting the pantherine bifurcation reconstructed in this study. In addition this is supported by the Johnson et al. (1996) mitochondrial RFLP analysis [105], as well as the Bininda-Emonds et al. (2001) characterization of *Panthera* excretory chemical signals [107]. The comprehensive Johnson et al. (2006) fully supports this association with high support values, however all other published molecular phylogenetic studies either relied heavily on mitochondrial sequences that have not been vetted as true cymt amplifications, or failed to fully resolve the pantherine phylogeny [59; 60; 106; 109; 110; 111].

As seen in Figure 9 and Table 7, unlike lion-leopard monophyly, the sister relationship of tiger and snow leopard is supported throughout most partitions with high levels of clade support, with the exception of the mitochondrial partition. This is not surprising given the amount of homoplasy seen in the mitochondrial partition (Table 2) parsimony informative sites supporting virtually every interspecies relationship (Appendix Table 15) and differing phylogenetic reconstructions depending on the gene segment (Figure 13). This could likely

be compounded by long branch attraction (LBA) between an accelerated tiger lineage and the divergent clouded leopard [213] or heterotachy between the tiger lineage and the remainder of the *Panthera* clade. Andersson and Swofford define LBA as “any situation in which similarity due to convergent or parallel changes produces an artifactual phylogenetic grouping of taxa due to an inherent bias in the estimation procedure” [213]. This could also be an aberrant association brought about by heterotachy: the variance through time of the evolutionary rate at a given position [214; 215]. This is a significant violation of existing maximum likelihood models and decreases the accuracy of ML estimation. The prognostic step to correct this anomaly was to remove the clouded leopard from the analysis. When doing so, the unrooted supermatrix shows total support for the pantherine bifurcation of the lion-leopard-jaguar and the tiger-snow leopard clade. In all other rooted analyses, this relationship is supported in every partition with the exception of the X chromosome. However, in all unrooted analyses this relationship is supported.

The strength of the molecular phylogenetic signal driving the lion-leopard and tiger-snow leopard monophylies in this matrix was evident in all analytical methods. The variation in signal between partitions for these two relationships as was shown by the ILD test, and the variation in clade support by partition, strongly indicates a species history that differs from the history of each inheritable portion of the genome. The same was true within the autosomal and nuclear partition. Each gene segment within these partitions had a unique

history that may or may not track speciation events. This could be due to incomplete lineage sorting, varying evolutionary rates among different loci, the introgression of populations following speciation, or any combination of the three. The most pronounced example of this was the recreation of the lion-jaguar monophyly in the X chromosome and with BI for the autosomal and nuclear partitions. Since lion, leopard, and jaguar underwent recent and rapid speciation events, it is highly likely that incomplete lineage sorting is a significant confounding factor. Another very probable cause for incongruent gene and species histories is the likelihood of introgression by hybridization, a highly plausible scenario for sympatric species due to recent speciation events [2] and high colinearity of felid chromosomes [216].

Recent theory suggested by Per Christiansen (2008) says that the extinct American lion (*P. l. atrox*), was a species that evolved in the late Pleistocene from the 'paleo-jaguar' (*P. gombaszoegensis*) lineage that entered the Americas in the early-mid Pleistocene [210]. This is earlier than the late Pleistocene expansion of the lion lineage from Eurasia, across Berengia into North America [65; 211]. In constrast, a recent molecular study by Barnett et al. (2009) suggests P. l. *atrox* originated from a Beringian population that migrated across Berengia into North America and was subsequently isolated around 337 thousand years ago (KYA) with the most recent common ancestor dating *to* 200 KYA [203]. This isolation mechanism is undetermined, since no evidence for a barrier to gene flow exists between Beringia and North America for other

terrestrial species such as the bison and horse. The American lion persisted in North America until the end of the Pleistocene [198] with the most recent fossil found in North America 11.35 KYA.

In either interpretation of the evolutionary history of the American lion, populations would have been sympatric with the migration route of the jaguar in North America. As the jaguar migrated south into South America across the Isthmus of Panama, ample opportunity for introgression between these two species existed. This would have allowed portions of the lion genome to introduce novel alleles into that of the jaguar. Since the American lion had been living in North America for generations, accumulating positively selected alleles, some would likely have had a selective advantage to the newly introduced jaguar and would subsequently have been retained in some populations during the two species' sympatry. Evidence for this process has recently been seen in the introduction of the domestic dog *Melanocortin 1 receptor* ligand, the melanistic K locus, into North American wolf populations [217]. After its introduction, this allele rapidly increased in frequency for forest-dwelling populations as a positively selection adaptive trait.

These unique gene histories, whether caused by incomplete lineage sorting or introgression, can be partially compensated by using the BEST method to recreate the species tree. The BEST implemented on the total matrix reconstructed the ML and BI supermatrix topology, with lower support for the

lion-leopard monophyly (ML: 100, BPP: 1.0, BEST: 0.63) and tiger-snow leopard monophyly (ML: 100, BPP: 94, BEST: 91). This is likely due to the considerably more liberal nature of the Bayesian gene tree reconstruction approach. The results of the complete phylogenetic efforts showed that there can be a definitive monophyly established containing tiger and snow leopard. The monophyly of lion-leopard is supported with very high support, though the BEST analysis did not place a high level of confidence in this association due to the large amount of conflicting gene histories between the lion, leopard and jaguar.

Transthyretin Anomaly

As seen in Figure 6, the *TTR* intron 1 produced a unique topology for the relationships within the pantherine lineage. Lion, leopard, and jaguar were monophyletic and shared 100% identity. The same was true for tiger and snow leopard. Within the 779 bp region sequenced of *TTR* intron 1, there were a total of 10 informative sites supporting each of the two clades, a greater proportion than any other gene segment other than those in the mitochondrial partition. Within the big cats, this gene segment appears to precisely follow the divergence of these two clades based on the results of the other two data partitions, with no interspecies divergence within each clade.

Jacknifing analysis showed that when the *TTR* segment was not present, the bipartition support in the autosomal partition changed significantly, seen in Figure 19. The support for the pantherine bifurcation splitting the tiger and snow leopard clade from the lion, leopard and jaguar plummeted: bootstrap values for the monophyly of lion, leopard, and jaguar fall from around 90% to below 50%, and from 70% to below 10% for tiger and snow leopard monophyly. In addition, the support for a sister relationship for leopard and tiger increased from around 10% to above 50%, and the support for snow leopard and clouded leopard monophyly increased from approximately 30% to close to 90%. *TTR* was the only gene segment that had such a profound effect on topological rearrangement when removed from the autosomal partition without the unpublished *CES7* sequence. Resequencing of the *TTR* intron sequences verified that the initial published sequences were accurate [2].

There are several explanations for this result. When more than one lineage of an ancestral population contributes to the new species in a series of speciation events, it is possible that a later extinction pattern will contribute to a tree topology that does not correspond to the general population splitting trend. Perhaps there is a much different level of selection acting upon this single gene. This can be an indication of an ancestral state for this gene that has underwent lineage sorting into the lion-leopard-jaguar and tiger-snow leopard clades prior to speciation [218]. Subsequent purifying selection within each of the two clades would maintain the assorted allele and therefore skew the effect this gene has

on resolving the associated phylogeny. The protein produced by this gene has been identified as a major urinary protein (MUP) in the urine fraction of the marking fluid of the male Bengal tiger[219]. It has been characterized as a carrier protein for many different molecules, including retinol [220]. Multiple studies have suggested that urinary proteins highly similar to *TTR* are involved in chemical communication in mammals. Excreted proteins in the same urinary fraction of the domestic cat have been shown to perform an enzymatic role in the synthesis of putative pheromone precursor proteins [221]. It therefore may be involved in the territorial marking behavior of the great cats. This may explain its unique phylogenetic signal in tracking the monophyly of lion-leopard-jaguar and tiger-snow leopard as was supported by the supermatrix phylogeny as well as the Y chromosome, X chromosome, autosomal, uniparental, and nuclear partitioned topologies.

Tiger Lineage Acceleration

Mitochondrial DNA divergence is estimated to accrue at approximately 2% sequence divergence per million years between pairs of lineages separated for less than 10 million years in animals, corresponding to $20x10^{-9}$ substitutions per site per year [222]. Deviation from this general rule is evidence for acceleration or deceleration in a particular lineage. Within this study, the tiger mitochondrial DNA sequences often appeared to have an increased rate of

nucleotide substitution relative to the other pantherids, the hallmark of lineage acceleration. In order to identify regions within the mitochondrial partition subject to significant acceleration, outliers in branch lengths within this partition were sought by quantifying variation in branch length from a normal distribution based on mean and standard deviation. As was seen in the Grubb's test for outliers and the skewness test, the tiger lineage is at the threshold for statistically significant acceleration.

When examining the branch lengths measures from the outgroup to each terminal species, there were some slower lineages such as lion and leopard, the consistent positive outlier is tiger. It was the only species that showed acceleration across the entire mitochondrial partition. Figure 13 shows the topology for each individual gene segment in the mitochondrial partition with their associated bootstrap values. It is clear to see both in Figure 14 and the topologies in Figure 13 that there is a noticeable acceleration with respect to tiger in the NADH dehydrogenase genes. The complex assembled from the products of these genes is the first and largest enzyme in the respiratory electron transport chain. Though the exact structure and mechanisms for eukaryotic NADH dehydrogenase are not well understood, it is known that it translocates 4 protons from oxidised NADH across the inner mitochondrial membrane and provides electrons for reduction of quinone to quinol. This helps build a transmembrane electrochemical potential used to produce ATP [223]. As the tiger is the largest extant felid, acceleration in this lineage for this

complicated enzyme may be due to differing oxidative requirements owing to metabolic differences.

The program TreeSAAP was used to determine regions of positive-destabilizing selection within the set of reference mitochondrial protein coding genes that would correspond to significant changes in amino acid function. Of the five genes in the dataset, only *ND1* showed significant radical amino acid variation in categories 7 and 8 at $\alpha = 0.001$ (z-score > 3.09) across the genus. Each amino acid residue (henceforth referred to as site) for these two genes was examined in a sliding window analysis to indicate regions of the polypeptide under adaptive evolution. *ND2* (window size 20) had three well defined regions indicating positive-destabilizing selection (Figure 24). One region stretched from site windows 91-111 to 99-119 and corresponded only to solvent access ratio. Another region correlated to three physiochemical changes from site window 249-269 to 268-288 (isoelectric point, α-helical tendencies, and compressibility) and another from windows 299-319 to 318-388 showed significant changes in another three properties (buriedness, isoelectric point, and composition. Thus it appears this region is under significant positive destabilizing selection for these properties.

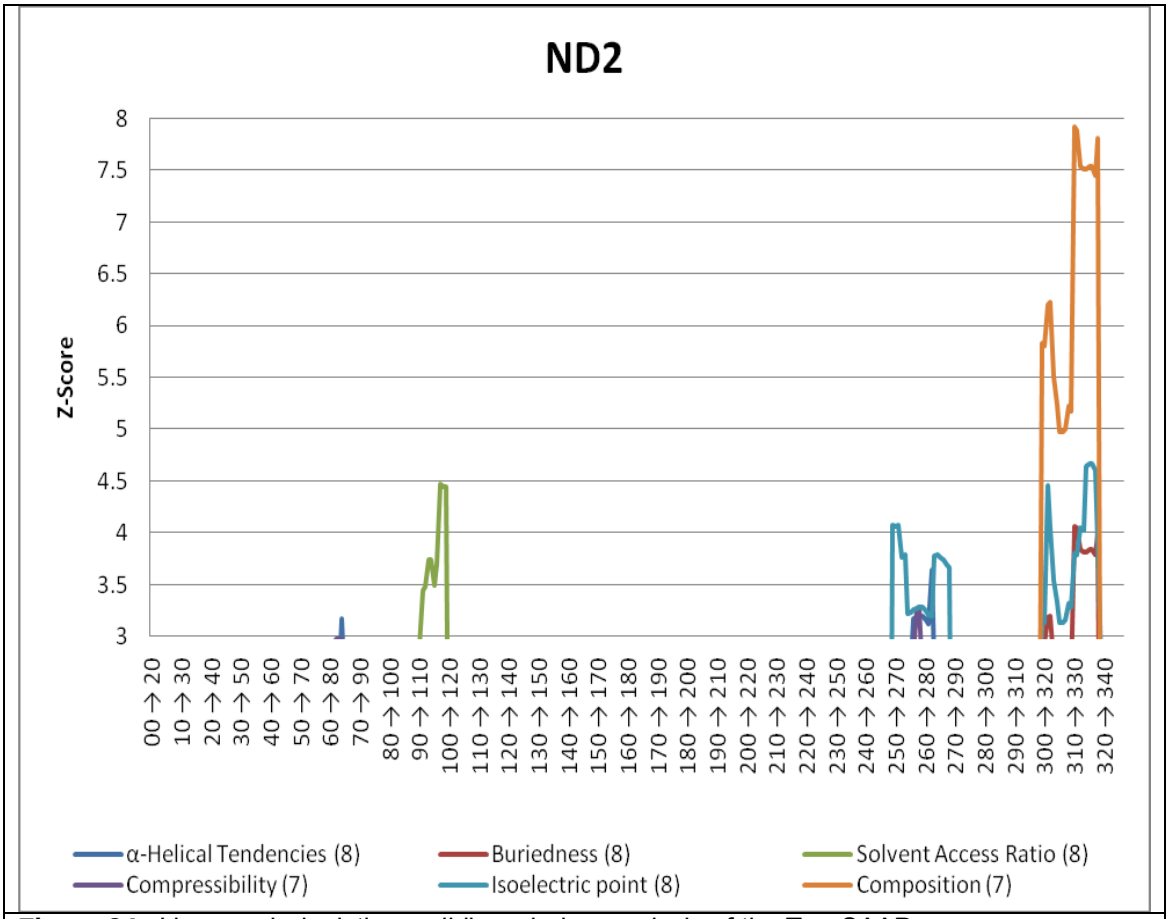


Figure 24. Line graph depicting a sliding window analysis of the TreeSAAP z-scores corresponding to physiochemical amino acid variation in *ND2*. Z-scores based on comparison of empirical values at each amino acid site compared to those expected under neutral evolution. Sliding window size set to 20 and shown on the x-axis. Magnitude category in parenthesis adjacent to amino acid physiochemical change.

All radical nonsynonymous amino acid changes were quantified by branch to determine which species may be under selection for a particular mitochondrial gene segment (Figure 25). Tiger did not show significant acceleration with respect to radical amino acid substitutions, with roughly the same number as its sister taxon, snow leopard. The lack of a signature for selection in both the calculations of ω and TreeSAAP analyses indicate that the

acceleration of the tiger lineage is not due to an excess of nonsynonymous amino acid substitutions. However, since these are only partial mitochondrial sequences vetted *in silico* and do not encompass all NADH dehydrogenase subunit genes, a complete sequencing of these segments in *Panthera* with methods to control for numt amplification should be performed. Doing so would allow a full characterization of the specific mitochondrial variations within the lineage and indicate potential adaptations associated with each species.

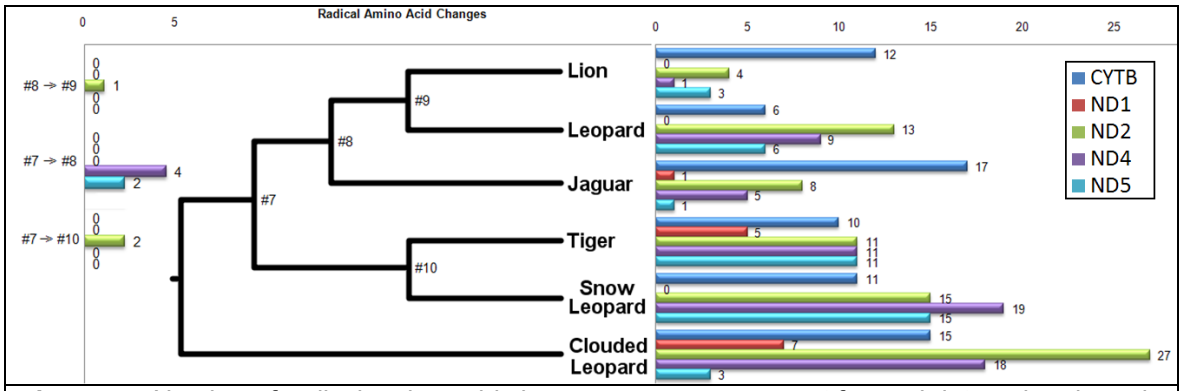


Figure 25. Number of radical amino acid changes per gene segment for each internal node and terminal species in *Panthera*.

Though the reliability of phylogenetic inference relies on model selection and methods used, the reference data set is of paramount importance. Judging by the discrepancies outlined in the trees created using new and previously published mitochondrial data; it appeared that the multiple groups mistakenly amplified a numt instead of the targeted mitochondrial genes in their own reference data. RFLP analysis performed by Johnson et al. (1996) identified at least twenty variants of numt sequences in the pantherine lineage and provided

the first methodology to avoid relying on signal originating from numts. Yet the only publications that provide a methodology to purify the mitochondrial fraction are the initial study characterizing the numt present on the F2 chromosome of tiger and its identity to the cymt sequence [140], and research sequencing the entire snow leopard, tiger, leopard and clouded leopard mitochondrial genomes [111; 184]. No other studies provide any information regarding a method utilized or suggested to alleviate such a discrepancy.

Comparison of the pairwise LogDet distances between the Yu and Johnson datasets shows a solid interspecies difference of about 10-12% and an intraspecies distance of less than 1%. This holds true for all taxa except lion and snow leopard which have an intraspecies distance of 9.1% and 11.8% respectively. This is comparable to the interspecies difference seen in other taxa and much greater than the interspecies comparison between the two of 1%. The clustering analysis of each mitochondrial segment indicated multiple incidences of possible species misidentification in many different publications. Since the ML clustering only allowed us to exclude and identify misidentification by consensus, this is not a certainty. The combination of multiple possible numt translocation events increased the chance that some or all of these sequences were misamplifications rather than misidentifications. Research has shown when looking at species barcoding based on mitochondrial sequence, the removal of numts using careful examination of indels, in-frame stop codons, and nucleotide composition does not eliminate their presence, increasing species

number estimation [224]. The only proven way to ensure mitochondrial amplification is to enrich for mitochondrial DNA by ultracentrifucation in a gradiented medium. Until such methods are performed on all of *Panthera*, published mitochondrial segments should be viewed as putative.

Upon closer examination of the sequences used in the tiger mitochondrial partition, there appears to be a discrepancy in the two publications with verifiable mitochondrial tiger sequences. Mitochondrial segments for Kim's numt paper [DQ151550] and Wu's unpublished mitochondrial genome GenBank sequence for tiger [EF551003] are identical for the *12S* segment and the *16S* region differs at only 1 base of the total 376. For the *ND2* gene, the newly generated sequence is identical to Wu's but the Kim sequence differs by 147 of the total 1,038 bp, amounting to 14%. This calls into question one of the two sequences. According to the clustering diagram in Figure 8, they both are within the putative cymt tiger clade. This does not alleviate the possibility of amplifying a more recent numt translocation event that has not had adequate time to diverge according to the nuclear mutation rate. The sequence chosen for inference was that of [EF551003]. Though there was a complete mitochondrial genome published for this sequence, there are no methods associated with it since it has no linked publication (at the date of this publication: May, 2009). The methods used by the same investigators to sequence the snow leopard genome indicate that the mitochondrial fraction was purified. This information was extrapolated by this study to include the other three sequences from the same group, perhaps

erroneously. The only sequences that are certain to have been enriched for mitochondria are associated with [DQ151550], and unfortunately only cover the genes *12S*, *16S*, and *ND2*. Despite these uncertainties, much has been done to alleviate aberrant signal associated with numt sequences including careful assessment of in-frame stop codons and nucleotide composition, producing the most comprehensive and carefully vetted mitochondrial data set for *Panthera* to date. With the potential for the complete sequencing of purified mitochondrial genomes on the horizon, this is one short step away from a comprehensive mitochondrial phylogeny for these cats.

Molecular Dating

The dates inferred by this study were similar with those from Johnson et al (2006); however, the topological difference between the two studies, that of the monophyly of lion and leopard proposed by this research versus the previous monophyletic association of lion and jaguar, altered the divergence times within this clade. The basal divergence time of *Panthera* was 3.80 – 4.31 MYA calculated in this study is slightly older than the 3.72 MYA published in Johnson et al. (2006) [2]. The divergence time of snow leopard and tiger herein was 2.70 – 3.70 MYA, consistent with the 2.88 MYA divergence deduced by

Johnson et al. (2006). We determined the jaguar split from the lion-leopard lineage to be between 2.56 and 3.66 MYA, concordant with the leopard divergence in Johnson et al. (2006) of 2.87 MYA, though obviously associated with a different species. The lion and leopard diverged 1.95 – 3.10 MYA according to our study, also congruent with the 2.06 MYA divergence date placed on the lion-jaguar divergence in Johnson et al (2006). Very rarely is speciation an immediate process. Probable introgression by hybridization of sympatric species during and following speciation events contribute to the prolonged process. There has been no evidence uncovered for sudden allopatry in *Panthera.* Even today, the colinearity of felid chromosomes and low sequence divergence contribute to the hybrid compatibility of the great cats, albeit subject to Haldane's rule: the preferential sterility or inviability of hybrids of the heterogametic sex [225].

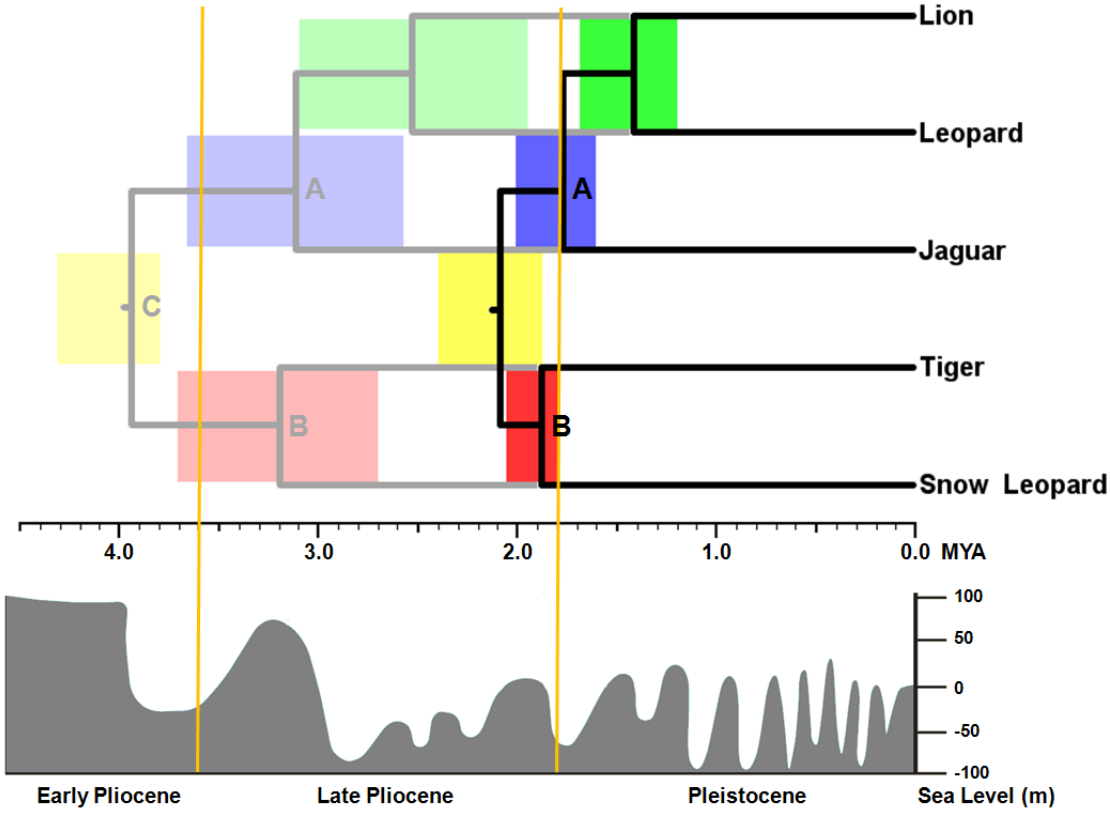


Figure 26. Divergence times estimated using MULTIDIVTIME with fossil calibrations: **A**) 1.6 MYA minimum **B**) 1.8 MYA minimum **C**) 3.8 MYA minimum. Colors indicate node 95% confidence intervals. Black tree / darker colors represent times estimated with fossil calibration (**C**) removed and grey tree / lighter colors estimated with fossil calibration (**C**) included.

As seen in Table 13 and Figure 26, the removal of each individual internal calibration point did not significantly affect the divergence time or the 95% credibility interval. However, when the minimum for the base of *Panthera* is removed, the divergence times at each node were decreased by roughly 40-50%, indicative of the likely unreliability of this fossil calibration. Without this constraint, the lion-leopard divergence dates to 1.22 - 1.69 MYA and the basal divergence of jaguar from this monophyly at 1.61 - 2.03 MYA. The split of tiger

and snow leopard dates to 1.80 – 2.07 MYA with the base of *Panthera* at only 1.88 – 2.40 MYA. This places the divergence of lion and leopard into the Pleistocene and the split of jaguar from this clade as well as the divergence of snow leopard and tiger at the border of the Pliocene and Pleistocene. Though the precise migration story remains occluded in the absence of adequate population sampling and comprehensive marker coverage within *Panthera*, these more recent dates further emphasize the possibility of introgression between lion and jaguar during their sympatry in North America. Each of these divergences still coincides with decreases in sea levels [226] that would have facilitated the migration of lions and jaguars from Asia, over Berengia into North America and jaguars south across the isthmus of Panama into South America.

CHAPTER III

CONCLUSIONS

The genus *Panthera* has undergone decades of intense research with respect to biogeography, habitat, prey behaviors, and other ecological factors in an effort to provide knowledge and assistance to those who wish to help these noble animals stem the pressure placed upon them by human expansion. Many studies have involved the description of the phylogenetic lines drawn between the great cats, but none have offered corroborative evidence to another study. Therefore the precise nature of the relationship between these cats has been shrouded in conflict. Within this study, we built upon all previous phylogenetic work and generated our own novel sequence data. We verified sequence from the underutilized Y chromosome and resequenced sections of three nuclear and four mitochondrial genes that were previously published in order to verify taxon labeling and to help control for the amplification of numts. By using maximum likelihood to build trees of all available mitochondrial sequence data, we were able to determine which sequences appeared to be part of a numt and not to be used in phylogenetic inference. This also allowed the identification of species misidentifications that previously contributed to confounding species relationships.

When constructing the supermatrix phylogeny containing the vetted mitochondrial sequence, we generated a highly corroborated topology with high

clade support using ML and BI methods. Due to the conflicting signal from each partition and even within the autosomal and mitochondrial partitions as shown by the ILD test, and represented by the various signal quantification calculations performed, we sought to construct the species tree using a newly conceptualized BEST method, tailored to handle datasets with incongruent signal. This method also recreated the topology generated by multiple phylogenetic methods. As expected, the more conservative nature of the BEST method deduced lower clade support values for the lion-leopard sister relationship due to gene trees, particularly in the autosomes, offering signal for an alternative topology placing jaguar as sister to lion. However, we can see that the bulk of the data strongly supports the monophyly of lion and leopard, a unique molecular phylogenetic relationship uncovered by this study, supported by multiple other studies not directly utilizing nucleotide sequence as the basis for phylogenetic inference. As genomes increase and more sequence data becomes readily available, we will be able to more fully uncover the genomic regions contributing to specific topologies. This study advocates the use of the Y chromosome and the mitochondria as the key sources of sequence data for molecular based studies reconstructing future phylogenies with recent and rapid speciation events. Data should also be garnered from the autosomal and X chromosome, though recombination and heterogeneous gene histories complicate species inference. These molecular resources should be the primary source for the inference of evolutionary history with supportive evidence from

morphological, ecological and biochemical studies. Though use of mitochondrial sequence data for phylogenetics is plagued with heterologous amplification of numts, laboratory protocols should evolve to reflect this knowledge if the goal is an accurate representation of evolutionary history. By utilizing larger sequence datasets, using the robust phylogenetic signal in Y chromosomes for sequencing, controlling numt amplification, and utilizing methodologies that can control for heterogeneous gene tree histories, scientific discovery will be able to resolve troublesome species at the tips of the intricate mammalian evolutionary Tree of Life.

Now that the precise relationships within *Panthera* have been determined, research can concentrate on the specifics of the introgression events occurring during the migrations of extant species. Further genomic analyses of which regions were involved in such events, and which regions were unaffected would allow a much more detailed story of the evolutionary history of *Panthera*. In addition, this allows researchers a window into the study of speciation genetics. By tracking which genes are segregated into each lineage during speciation and are subsequently maintained, the genetic mechanisms behind speciation can begin to be unraveled.

REFERENCES

[1] J. Baillie, and B. Groombridge, IUCN Red List of Threatened Animals, International Union for Conservation of Nature and Natural Resources, Gland, Switzerland, 1996.

[2] W.E. Johnson, E. Eizirik, J. Pecon-Slattery, W.J. Murphy, A. Antunes, E. Teeling, and S.J. O'Brien, The late Miocene radiation of modern Felidae: a genetic assessment. Science 311 (2006) 73-77.

[3] P. Jackson, Current status of the tiger. Cat News 28 (1998) 11.

[4] E. Dinerstein, C. Loucks, E. Wikramanayake, J. Ginsberg, E. Sanderson, J. Seidensticker, J. Forrest, G. Bryja, A. Heydlauff, S. Klenzendorf, P. Leimgruber, J. Mills, T.G. O'Brien, M. Shrestha, R. Simons, and M. Songer, The fate of wild tigers. BioScience 57 (2007) 508–514.

[5] T.M. McCarthy, and G. Chapron, Snow leopard survival strategy, Seattle, WA, 2003.

[6] K. Nowell, and P. Jackson, Status Survey and Conservation Action Plan, Wild Cats, International Union for the Conservation of Nature and Natural Resources (IUCN), Gland, Switzerland, 1996.

[7] V.A. Buckley-Beason, W.E. Johnson, W.G. Nash, R. Stanyon, J.C. Menninger, C.A. Driscoll, J. Howard, M. Bush, J.E. Page, M.E. Roelke, G. Stone, P.P. Martelli, C. Wen, L. Ling, R.K. Duraisingam, P.V. Lam, and S.J. O'Brien, Molecular evidence for species-level distinctions in clouded leopards. Curr. Biol. 16 (2006) 2371-2376.

[8] A.C. Kitchener, M.A. Beaumont, and D. Richardson, Geographical variation in the clouded leopard, *Neofelis nebulosa*, reveals two species. Curr. Biol. 16 (2006) 2377-2383.

[9] A. Wilting, V.A. Buckley-Beason, H. Feldhaar, J. Gadau, S.J. O'Brien, and K.E. Linsenmair, Clouded leopard phylogeny revisited: support for species recognition and population division between Borneo and Sumatra. Front. Zool. 4 (2007) 15.

[10] J. Sanderson, J. Khan, L. Grassman, and D. Mallon, *Neofelis nebulosa*, IUCN Red List of Threatened Species, International Union for Conservation of Nature and Natural Resources, Gland, Switzerland, 2008.

[11] K.B. Pei, and P.J. Chiang, Present status and conservation of Formosan clouded leopard and other medium-to-large mammals at Tawu Nature Reserve and vicinities, Report Conservation Research Series No. 92-02, Council of Agriculture, Taiwan Forestry Bureau, 2004.

[12] D.R. Brooks, R.M. Mayden, and D.A. McLennan, Phylogeny and biodiversity: conserving our evolutionary legacy. Trends Ecol. Evolut. 7 (1992) 55-59.

[13] J. Schipper, J.S. Chanson, F. Chiozza, N.A. Cox, M. Hoffmann, V. Katariya, J. Lamoreux, A.S.L. Rodrigues, S.N. Stuart, H.J. Temple, J. Baillie, L. Boitani, T.E. Lacher, Jr., R.A. Mittermeier, A.T. Smith, D. Absolon, J.M. Aguiar, G. Amori, N. Bakkour, R. Baldi, R.J. Berridge, J. Bielby, P.A. Black, J.J. Blanc, T.M. Brooks, J.A. Burton, T.M. Butynski, G. Catullo, R. Chapman, Z. Cokeliss, B. Collen, J. Conroy, J.G. Cooke, G.A.B. da Fonseca, A.E. Derocher, H.T. Dublin, J.W. Duckworth, L. Emmons, R.H. Emslie, M. Festa-Bianchet, M. Foster, S. Foster, D.L. Garshelis, C. Gates, M. Gimenez-Dixon, S. Gonzalez, J.F. Gonzalez-Maya, T.C. Good, G. Hammerson, P.S. Hammond, D. Happold, M. Happold, J. Hare, R.B. Harris, C.E. Hawkins, M. Haywood, L.R. Heaney, S. Hedges, K.M. Helgen, C. Hilton-Taylor, S.A. Hussain, N. Ishii, T.A. Jefferson, R.K.B. Jenkins, C.H. Johnston, M. Keith, J. Kingdon, D.H. Knox, K.M. Kovacs, P. Langhammer, K. Leus, R. Lewison, G. Lichtenstein,

L.F. Lowry, Z. Macavoy, G.M. Mace, D.P. Mallon, M. Masi, M.W. McKnight, R.A. Medellin, P. Medici, G. Mills, P.D. Moehlman, S. Molur, A. Mora, K. Nowell, J.F. Oates, W. Olech, W.R.L. Oliver, M. Oprea, B.D. Patterson, W.F. Perrin, B.A. Polidoro, C. Pollock, A. Powel, Y. Protas, P. Racey, J. Ragle, P. Ramani, G. Rathbun, et al., The status of the world's land and marine mammals: diversity, threat, and knowledge. Science 322 (2008) 225-230.

[14] D.P. Faith, Phylogenetic diversity and conservation. in: S. Carroll, and C. Fox, (Eds.), Conservation Biology: Evolution in Action, Oxford University Press, New York, 2008, pp. 99-115.

[15] L. Witting, and V. Loeschcke, The optimization of biodiversity conservation. Biol. Cons. 71 (1995) 205-207.

[16] M.L. Weitzman, The Noah's Ark problem. Econometrica 66 (1998) 1279-1298.

[17] K. Hartmann, and M. Steel, Maximizing phylogenetic diversity in biodiversity conservation: greedy solutions to the Noah's Ark problem. Syst. Biol. 55 (2006) 644-651.

[18] P.E. Stander, Cooperative hunting in lions: the role of the individual. Behav. Ecol. Sociobio. 29 (1992) 445-454.

[19] H. Bauer, K. Nowell, and C. Packer, *Panthera leo*, IUCN Red List of Threatened Species, International Union for Conservation of Nature and Natural Resources, Gland, Switzerland, 2008.

[20] M.E. Sunquist, and F. Sunquist, Wild Cats of the World, University of Chicago Press, Chicago, IL, 2002.

[21] Asiatic lion - history, Asiatic Lion Information Centre, Wildlife Conservation Trust of India, 2006.

[22] S.J. O'Brien, J.S. Martenson, C. Packer, L. Herbst, V.D. Vos, P. Joslin, J. Ott-Joslin, D.E. Wildt, and M. Bush, Biochemical genetic variation in geographic isolates of African and Asiatic lions. Natl Geogr. Res. 3 (1987) 114-124.

[23] P. Chardonnet, Conservation of African lion: contribution to a status survey, International Foundation for the Conservation of Wildlife, Paris, France, 2002.

[24] H. Bauer, and S.V.D. Merwe, The African lion database. Cat News 36 (2002) 41-53.

[25] S.J. O'Brien, and W.E. Johnson, Big cat genomics. Ann. Rev. Genom. Hum. Genet. 6 (2005) 407-429.

[26] U. Breitenmoser, C. Breitenmoser-Wursten, P. Henschel, and L. Hunter, *Panthera pardus*, IUCN Red List of Threatened Species, International Union for Conservation of Nature and Natural Resources, Gland, Switzerland, 2008.

[27] A. Kitchener, The Natural History of the Wild Cats, Cornell University Press, Ithaca, NY, 1998.

[28] K.L. Seymour, *Panthera onca*. Mamm. Spec. 340 (1989) 1-9.

[29] A. Turner, and M. Antón, The Big Cats and Their Fossil Relatives, Columbia University Press, New York, 1997.

[30] B. VanValkenburgh, F. Grady, and B. Kurtén, The Plio-Pleistocene cheetah-like cat *Miracinonyx inexpectatus* of North America. J. Vert. Paleont. 10 (1990) 434-454

[31] H. Hemmer, Fossil history of the living Felidae. in: R.L. Eaton, (Ed.), The World's Cat, Carnivore Research Institute, Burke Museum, Seattle, WA, 1976, pp. 1-14.

[32] E.W. Sanderson, K.H. Redford, C.-L.B. Chetkiewicz, R.A. Medellin, A.R. Rabinowitz, J.G. Robinson, and A.B. Taber, Planning to save a species: the jaguar as a model. Conserv. Biol. 16 (2002) 58-72.
[33] A. Caso, C. Lopez-Gonzalez, E. Payan, E.E.T.d. Oliveira, R. Leite-Pitman, M. Kelly, and C. Valderrama, *Panthera onca*, IUCN Red List of Threatened Species, International Union for Conservation of Nature and Natural Resources, Gland, Switzerland, 2008.
[34] E. Eizirik, J.-H. Kim, M. Menotti-Raymond, P.G.C. Jr., S.J. O'Brien, and W.E. Johnson, Phylogeography, population history and conservation genetics of jaguars (*Panthera onca*, Mammalia, Felidae). Mol. Ecol. 10 (2001) 65-79.
[35] IndependentOnline, "Tiger tops dog as world's favourite animal", http://www.int.iol.co.za/index.php?newslett=1&em=28164a99a20041206ah&click_id=29&art_id=qw1102325040750B216&set_id=1, April 22, 2009.
[36] J. Seidensticker, On the ecological separation between tigers and leopards. Biotropica 8 (1976) 225-234.
[37] R.S. Chundawat, B. Habib, U. Karanth, K. Kawanishi, J.A. Khan, T. Lynam, D. Miquelle, P. Nyhus, Sunarto, R. Tilson, and S. Wang, *Panthera tigris*, IUCN Red List of Threatened Species, International Union for Conservation of Nature and Natural Resources, Gland, Switzerland, 2008.
[38] E. Dinerstein, E. Wikramanayake, J. J. Robinson, U. U. Karanth, A. Rabinowitz, D. Olson, T. Mathew, P. Hedao, M. Connor, G. Hemley, and D. Bolze, A framework for identifying high priority areas and actions for the conservation of tigers in the wild, World Wildlife Fund, Wildlife Conservation Society, Washington DC, New York, 1997.
[39] E. Sanderson, J. Forrest, C. Loucks, J. Ginsberg, E. Dinerstein, J. Seidensticker, P. Leimgruber, M. Songer, A. Heydlauff, T. O'Brien, G. Bryja, E. Wikramanayake, and S. Klenzendorf, Setting priorities for the conservation and recovery of wild tigers: 2005-2015, the technical assessment, Wildlife Conservation Society, World Wildlife Fund, Smithsonian, National Fish and Wildlife Foundation - Save The Tiger Fund, Washington, DC, 2006.
[40] SaveChinasTigers.com, "Tiger culture | save china's tigers", http://english.savechinastigers.org, April 22, 2009.
[41] R. Jackson, D. Mallon, T. McCarthy, R.A. Chundaway, and B. Habib, Panthera uncia, IUCN 2008. 2008 IUCN Red List of Threatened Species, 2008.
[42] R.M. Jackson, and D.O. Hunter, Snow leopard survey and conservation handbook, International Snow Leopard Trust and U.S. Geological Survey, Biological Resources Division, Seattle, WA, 1996.
[43] S. Theile, Fading footsteps: the killing and trade of snow leopards, TRAFFIC International, Cambridge, United Kingdom, 2003.
[44] E.P. Koshkarev, and V. Vyrypaev, The snow leopard after the break-up of the Soviet Union. Cat News 32 (2000) 9-11.
[45] Humphrey, and Bain, Endangered Animals of Thailand, Sandhill Crane Press, Gainesville, FL, 1990.
[46] S.C. Austin, and M.E. Tewes, Ecology of the clouded leopard in Khao Yai National Park, Thailand. Cat News 31 (1999) 17-18.
[47] S.J. O'Brien, Conservation genetics of the Felidae. in: J.C. Avise, and J.L. Hamrick, (Eds.), Conservation Genetics, Chapman & Hill, New York, 1996, pp. 50-74.

[48] S.J. O'Brien, and W.E. Johnson, Big Cat Genomics. Annual Review of Genomics and Human Genetics 6 (2005) 407-429.
[49] L.D. Martin, Fossil history of the terrestrial Carnivora. in: J.L. Gittleman, (Ed.), Carnivore Behavior, Ecology, and Evolution, Cornell University Press, Cornell, NY, 1989, pp. 536-568.
[50] B. VanValkenburgh, Déjà vu: the evolution of feeding morphologies in the Carnivora. Integr. Compar. Biol. 47 (2007) 147-163.
[51] T. Rothwell, Phylogenetic systematics of North American *Pseudaelurus* (Carnivora Felidae). Am. Mus. Novit. 3403 (2003) 1-63.
[52] G.I. Adams, The extinct Felidae of North America. Am. J. Sci. 4 (1897) 145-149.
[53] A.P. Allred, Cats' Most Wanted, Brassey's, London, 2005.
[54] R.M. Nowak, Walker's Mammals of the World, 6th edition, Johns Hopkins University Press, Baltimore, MD, 1999.
[55] M.H. Hunt, Biogeography of the Order Carnivora. in: J.L. Gittleman, (Ed.), Carnivore Behavior, Ecology, and Evolution, Cornell University Press, Ithaca, NY, 1996, pp. 485-541.
[56] O.R.P. Bininda-Emonds, J.L. Gittleman, and A. Purvis, Building large trees by combining phylogenetic information: A complete phylogeny of the extant carnivora (mammalia). Biol. Rev. 74 (1999) 143-175.
[57] L. Werdelin, Morphological patterns in the skulls of cats. Biol. J. Linn. Soc. 19 (1983) 375-391.
[58] J. Pecon-Slattery, and S.J. O'Brien, Patterns of Y and X chromosome DNA sequence divergence during the Felidae radiation. Genetics 148 (1998) 1245-1255.
[59] D.N. Janczewski, W.S. Modi, J.C. Stephens, and S.J. O'Brien, Molecular evolution of mitochondrial 12S RNA and cytochrome b sequences in the pantherine lineage of Felidae. Mol. Biol. Evol. 12 (1995) 690-707.
[60] M.Y. Mattern, and D.A. McLennan, Phylogeny and speciation of felids. Cladistics 16 (2000) 232-253.
[61] O. Uphyrkina, W.E. Johnson, H. Quigley, D. Miquelle, L. Marker, M. Bush, and S.J. O'Brien, Phylogenetics, genome diversity and origin of modern leopard, Panthera pardus. Mol. Ecol. 10 (2001) 2617–2633.
[62] R.B.N. Yamaguchi, I. Barnes, and A. Cooper, The origin, current diversity and future conservation of the modern lion (*Panthera leo*). P. Roy. Soc. Lond. B. Bio. 273 (2006) 2119–2125.
[63] G. Petter, Carnivores Pleistocenes du Ravin d'Olduvai (Tanzanie). in: L.S.B. Leakey, R.J.G. Savage, and S.C. Coryndon, (Eds.), Fossil Vertebrates of Africa, Academic Press, London, 1973, pp. 43-100.
[64] C. Stringer, Modern human origins: progress and prospects. Philos. T. Roy. Soc. B. 357 (2002) 563-579.
[65] N. Yamaguchi, A. Cooper, L. Werdelin, and D.W. Macdonald, Evolution of the mane and group-living in the lion (*Panthera leo*): a review. J. Zool. 263 (2004).
[66] J. Burger, W. Rosendahl, O. Loreille, H. Hemmer, T. Eriksson, A. Gotherstrom, J. Hiller, M.J. Collins, T. Wess, and K.W. Alt, Molecular phylogeny of the extinct cave lion *Panthera leo* spelaea. Mol. Phylogenet. Evol. 30 (2004) 841-849.

[67] B. Kurtén, and E. Anderson, Pleistocene Mammals of North America, Columbia University Press, New York, NY, 1980.
[68] M. Hofreiter, Pleistocene extinctions: haunting the survivors. Curr. Biol. 17 (2007) R609–R611.
[69] A.D. Barnosky, P.L. Koch, R.S. Feranec, S.L. Wing, and A.B. Shabel, Assessing the causes of late pleistocene extinctions on the continents. Science 306 (2004) 70-75.
[70] R.K. Wayne, R.E. Benveniste, D.N. Janczewski, and S.J. O'Brien, Molecular and biochemical evolution of the carnivora. in: J.L. Gittleman, (Ed.), Carnivore Behavior, Ecology, and Evolution, Cornell Univ. Press, Ithaca, NY, 1989, pp. 465-495.
[71] C. Darwin, On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life, Murray, London, 1859.
[72] C. Linnaeus, Systema Naturae, per Regna Tria Naturae, Secundum Classes, Ordines, Genera, Species, cum Characteribus, Differentiis, Synonymis, Locis, 10th ed., Laurentii Salvil, Stockholm, Sweden, 1758.
[73] R.L. Pocock, The classification of existing Felidae. Ann. Mag. Nat. Hist. 9 (1917) 375-384.
[74] C. Sonntag, The comparative anatomy of the tongues of Mammalia VIII. Carnivora. P. Zool. Soc. Lond. 9 (1923) 129-153.
[75] E. Lonnberg, Notes on some cats from eastern Asia with description of a new subgenus. Arkiv for Zoologi 18 (1926) 1-23.
[76] R.F. Ewer, The Carnivores, Cornell University Press, Ithaca, NY, 1973.
[77] G.E. Glass, and L.D. Martin, A multivariate comparison of some extant and fossil Felidae. Carnivore 1 (1978) 80-87.
[78] L. Radinsky, Evolution of the felid brain. Brain Behav. Evolut. 11 (1975) 214-254.
[79] L.O. Salles, Felid phylogenetics: extant taxa and skull morphology (Felidae, Aeluroidea). Am. Mus. Novit. 3047 (1992) 1-67.
[80] L.B. Radinsky, Evolution of skull shape in carnivores. 3. The origin and early radiation of the modern carnivore families. Paleobiology 8 (1982) 177-195.
[81] L. Werdelin, Carnivoran ecomorphology: a phylogenetic perspective. in: J.L. Gittleman, (Ed.), Carnivore Behavior, Ecology, and Evolution, Cornell University Press, Ithaca, NY, 1996, pp. 582-624.
[82] G.E. Weissengruber, G. Forstenpointner, G. Peters, A. Kübber-Heiss, and W.T. Fitch, Hyoid apparatus and pharynx in the lion (*Panthera leo*), jaguar (*Panthera onca*), tiger (*Panthera tigris*), cheetah (*Acinonyx jubatus*) and domestic cat (*Felis silvestris f. catus*). J. Anat. 201 (2002) 195-209.
[83] G. Peters, and M.H. Hast, Hyoid structure, laryngeal anatomy, and vocalization in felids (Mammalia: Carnivora: Felidae). Zeitschrift für Säugetierkunde 59 (1994) 87-104.
[84] P. Jackson, The secret of a tiger's roar. Cat News 34 (2001) 12.
[85] D.E. Frazer-Sissom, D.A. Rice, and G. Peters, How cats purr. J. Zool. 223 (1991) 67-78.
[86] M.H. Hast, The larynx of roaring and non-roaring cats. J. Anat. 164 (1989) 117-121.
[87] H. Hemmer, The evolutionary systematics of living Felidae: present status and current problems. Carnivore 1 (1978) 71-79.
[88] S.J. Herrington, Phylogenetic relationships of the wild cats of the world. PhD dissertation, University of Kansas, Lawrence, 1986.
[89] R.W. Scotland, R.G. Olmstead, and J.R. Bennett, Phylogeny reconstruction: the role of morphology. Syst. Biol. 52 (2003) 539-548.

[90] M.S. Springer, A. Burk-Herrick, R. Meredith, E. Eizirik, E. Teeling, S.J. O'Brien, and W.J. Murphy, The adequacy of morphology for reconstructing the early history of placental mammals. Syst. Biol. 56 (2007) 673-684.
[91] K.-P. Koepfli, M.E. Gompper, E. Eizirik, C.-C. Ho, L. Linden, J.E. Maldonado, and R.K. Wayne, Phylogeny of the procyonidae (Mammalia: Carnivora): Molecules, morphology and the Great American Interchange. Mol. Phylogenet. Evol. 43 (2007) 1076-1095.
[92] A.G. Kluge, A concern for evidence and a phylogenetic hypothesis of relationships among Epicrates (Boidae Serpentes). Syst. Zool. 38 (1989) 7-25.
[93] M.J. Novacek, Mammalian phylogeny: shaking the tree. Nature 356 (1992) 121-125.
[94] W.W. deJong, Molecules remodel the mammalian tree. Trends Ecol. Evolut. 13 (1998) 270-275.
[95] O.R.P. Bininda-Emonds, Factors influencing phylogenetic inference: a case study using the mammalian carnivores. Mol. Phylogenet. Evol. 16 (2000) 113-126.
[96] S.J. O'Brien, G.E. Collier, R.E. Benveniste, W.G. Nash, A.K. Newman, J.M. Simonson, M.A. Eichelberger, U.S. Seal, D. Janssen, M. Bush, and D.E. Wildt, Setting the molecular clock in Felidae: the great cats, Panthera. in: R.L. Tilson, and U.S. Seal, (Eds.), Tigers of the World, William Andrew Inc., Noyes, NJ, 1987, pp. 10-27.
[97] D.H. Wurster-Hill, and C.W. Gray, Giesma banding patterns in the chromosomes of twelve species of cats (Felidae). Cytogenet. Cell Genet. 12 (1973) 377-397.
[98] D.H. Wurster-Hill, The G-banded chromosomes of the marbled cat *Felis marmorata*. Mamm. Chrom. News. 15 (1974) 14.
[99] D.H. Wurster-Hill, and C.W. Gray, The interrelationships of chromosome banding patterns in procyonids, viverrids and felids. Cytogenet. Cell Genet. 15 (1975) 306-331.
[100] D.H. Wurster-Hill, and W.R. Centerwall, The interrelationships of chromosome banding patterns in canids, mustelids, hyena, and felids. Cytogenet. Cell Genet. 34 (1982) 178-92.
[101] G.E. Collier, and S.J. O'Brien, A molecular phylogeny of the Felidae: immunological distance. Evolution 39 (1985) 437-487.
[102] J.P. Thorpe, The molecular clock hypothesis: biochemical evolution, genetic differentiation and systematics. Annu. Rev. Ecol. Syst. 13 (1982) 139-168.
[103] A.C. Wilson, S.S. Carlson, and T.J. White, Biochemical evolution. Annu. Rev. Biochem. 46 (1977) 573-639.
[104] J. Pecon-Slattery, W.E. Johnson, D. Goldman, and S.J. O'Brien, Phylogenetic reconstruction of South American felids defined by protein electrophoresis. J. Molec. Evol. 39 (1994) 296-305.
[105] W.E. Johnson, P.A. Dratch, J.S. Martenson, and S.J. O'Brien, Resolution of recent radiations within three evolutionary lineages of Felidae using mitochondrial restriction fragment length polymorphism variation. J. Mammal. Evol. (1996) 97-120.
[106] W.E. Johnson, and S.J. O'Brien, Phylogenetic reconstruction of the Felidae using 16S rRNA and NADH-5 mitochondrial genes. J. Molec. Evol. 44 Suppl 1 (1997) S98-116.
[107] O.R.P. Bininda-Emonds, D.M. Decker-Flum, and J.L. Gittleman, The utility of chemical signals as phylogenetic characters: an example from the Felidae. Biol. J. Linn. Soc. 72 (2001) 1-15.

[108] K. Jae-Heup, E. Eizirik, S.J. O'Brien, and W.E. Johnson, Structure and patterns of sequence variation in the mitochondrial DNA control region of the great cats. Mitochondrion 1 (2001) 279-292.
[109] J. Pecon-Slattery, A.J. Pearks-Wilkerson, W.J. Murphy, and S.J. O'Brien, Phylogenetic assessment of introns and SINEs within the Y chromosome using the cat family Felidae as a species tree. Mol. Biol. Evol. 21 (2004) 2299-2309.
[110] L. Yu, and Y.P. Zhang, Phylogenetic studies of pantherine cats (Felidae) based on multiple genes, with novel application of nuclear beta-fibrinogen intron 7 to carnivores. Mol. Phylogenet. Evol. 35 (2005) 483-95.
[111] L. Wei, X. Wu, and Z. Jiang, The complete mitochondrial genome structure of snow leopard *Panthera uncia*. Mol. Biol. Rprts. 36 (2009) 871-878.
[112] J. Felsenstein, PHYLIP, Phylogeny Inference Package (Version 3.5), University of Washington, Seattle, 1993.
[113] D.L. Swofford, PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1.1, Smithsonian Institute, Washington, DC, 1993.
[114] D.H. Wurster-Hill, T. Doi, M. Izawa, and Y. Ono, Banded chromosome study of the Iriomote cat. J. Hered. 78 (1987) 105-107.
[115] B. Dutrillaux, and J. Couturier, Carnivora: comparison with that of platyrrhine monkeys. Cytogenet. Cell Genet. 35 (1983) 200-208.
[116] G. Peters, and M.H. Hast, Hyoid structure, laryngeal anatomy, and vocalization in felids (Mammalia: Carnivora; Felidae). Zeitschrift für Säugetierkunde 59 (1994) 87-104.
[117] B.R. Baum, Combining trees as a way of combining data sets for phylogenetic inference, and the desirability of combining gene trees. Taxon 41 (1992) 3-10.
[118] M.A. Ragan, Phylogenetic inference based on matrix representation of trees. Mol. Phylogenet. Evol. 1 (1992) 53-58.
[119] J.V. Lopez, M. Culver, J.C. Stephens, W.E. Johnson, and S.J. O'Brien, Rates of nuclear and cytoplasmic mitochondrial DNA sequence divergence in mammals. Mol. Biol. Evol. 14 (1997) 277-286.
[120] V. King, P.N. Goodfellow, A.J.P. Wilkerson, W.E. Johnson, S.J. O'Brien, and J. Pecon-Slattery, Evolution of the male-determining gene SRY Within the cat family Felidae. Genetics 175 (2007) 1855-1867.
[121] L. Yu, Q.W. Li, O.A. Ryder, and Y.P. Zhang, Phylogenetic relationships within mammalian order Carnivora indicated by sequences of two nuclear DNA genes. Mol. Phylogenet. Evol. 33 (2004) 694-705.
[122] P. Leyhausen, Cat Behavior: The Predatory and Social Behavior of Domestic and Wild Cats (Garland series in ethology), Garland Publishing Inc., New York, NY, 1979.
[123] M.F. Hammer, A recent common ancestry for human Y chromosomes. Nature 378 (1995) 376-378.
[124] R. Thomson, J.K. Pritchard, P. Shen, P.J. Oefner, and M.W. Feldman, Recent common ancestry of human Y chromosomes: evidence from DNA sequence data. PNAS 97 (2000) 7360-7365.
[125] D.J. Gaffney, and P.D. Keightley, Genomic selective constraints in murid noncoding DNA. PLoS Genet. 2 (2006) 1912-1923.
[126] W.-H. Li, and D. Graur, Fundamentals of Molecular Evolution, Sinauer, Sunderland, MA, 1991.

[127] W.-H. Li, Molecular Evolution, Sinauer, Sunderland, MA, 1997.
[128] N. Jareborg, E. Birney, and R. Durbin, Comparative analysis of noncoding regions of 77 orthologous mouse and human gene pairs. Genome Res. 9 (1999) 815-824.
[129] S.A. Shabalina, A.Y. Ogurtsov, V.A. Kondrashov, and A.S. Kondrashov, Selective constraint in intergenic regions of human and mouse genomes. Trends Genet. 17 (2001) 373-376.
[130] J.C. Avise, Molecular Markers, Natural History, and Evolution, 2nd edition, Sinauer Associates, Sunderland, MA, 2004.
[131] E. Randi, Mitochondrial DNA. in: A.J. Baker, (Ed.), Molecular Methods in Ecology, Blackwell Science, London, 2000, pp. 136-167.
[132] T. Parsons, and M.M. Holland, Response to: Mitochondrial mutation rate revisited: hot spots and polymorphism. Nat. Genet. 18 (1998) 110.
[133] P.A. Mason, and R.N. Lightowlers, Why do mammalian mitochondria possess a mismatch repair activity. FEBS Lett. 554 (2003) 6-9.
[134] P.A. Mason, E.C. Matheson, A.G. Hall, and R.N. Lightowlers, Mismatch repair activity in mammalian mitochondria. Nucleic Acids Res. 31 (2003) 1052-1058.
[135] J.J. Flynn, J.A. Finarelli, S. Zehr, J. Hsu, and M. Nedbal, Molecular phylogeny of the carnivora (mammalia): Assessing the impact of increased sampling on resolving enigmatic relationships. Syst. Biol. 54 (2005) 317-337.
[136] E. Richly, and D. Leister, NUMTs in sequenced eukaryotic genomes. Mol. Biol. Evol. 21 (2004) 1081-1084.
[137] J.V. Lopez, N. Yuhki, R. Masuda, W. Modi, and S.J. O'Brien, Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. J. Mol. Evol. 39 (1994) 174-90.
[138] J.V. Lopez, S. Cevario, and S.J. O'Brien, Complete nucleotide sequences of the domestic cat (Felis catus) mitochondrial genome and a transposed mtDNA tandem repeat (Numt) in the nuclear genome. Genomics 33 (1996) 229-46.
[139] J. Cracraft, J. Feinstein, J. Vaughn, and K. Helm-Bychowski, Sorting out tigers (Panthera tigris): mitochondrial sequences, nuclear inserts, systematics, and conservation genetics. Anim. Conserv. 1 (1998) 139-150.
[140] J.-H. Kim, A. Antunes, S.-J. Luo, J.M.W.G. Nash, S.J. O'Brien, and W.E. Johnson, Evolutionary analysis of a large mtDNA translocation (numt) into the nuclear genome of the Panthera genus species. Gene 366 (2006) 292-302.
[141] Z. Zhang, S. Schwartz, L. Wagner, and W. Miller, A greedy algorithm for aligning DNA sequences. J. Comput. Biol. 7 (2000) 203-214.
[142] J.U. Pontius, J.C. Mullikin, D.R. Smith, K. Lindblad-Toh, S. Gnerre, M. Clamp, J. Chang, R. Stephens, B. Neelam, N. Volfovsky, A.A. Schäffer, R. Agarwala, K. Narfström, W.J. Murphy, U. Giger, A.L. Roca, A. Antunes, M. Menotti-Raymond, N. Yuhki, J. Pecon-Slattery, W.E. Johnson, G. Bourque, G. Tesler, and S.J. O'Brien, Initial sequence and comparative analysis of the cat genome. Genome Res. 17 (2007) 1675-89.
[143] D. Bensasson, D. Zhang, D.L. Hartl, and G.M. Hewitt, Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends Ecol. Evolut. 16 (2001) 314-321.
[144] D. Mishmar, E. Ruiz-Pesini, M. Brandon, and D.C. Wallace, Mitochondrial DNA-like sequences in the nucleus (NUMTs): insights into our African origins and the mechanism of foreign DNA Integration. Hum. Mut. 23 (2004) 125-133.

[145] L. Pépin, Y. Amigues, A. Lépingle, J.-L. Berthier, A. Bensaid, and D. Vaiman, Sequence conservation of microsatellites between *Bos taurus* (cattle), *Capra hircus* (goat) and related species. Examples of use in parentage testing and phylogeny analysis. Heredity 74 (1995) 53-61.
[146] G.A. Wilson, C. Strobeck, L. Wu, and J.W. Coffin, Characterization of microsatellite loci in caribou *Rangifer tarandus*, and their use in other artiodactyls. Mol. Ecol. 6 (1997) 697-699.
[147] K.A. Selkoe, and R.J. Toonen, Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecol. Lett. 9 (2006) 615-629.
[148] C. Schlötterer, and D. Tautz, Slippage synthesis of simple sequence DNA. Nucleic Acids Res. 20 (1992) 211-215.
[149] Y.-C. Li, A.B. Korol, T. Fahima, A. Beiles, and E. Nevo, Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol. Ecol. 11 (2002) 2453-65.
[150] Y.-C. Li, A.B. Korol, T. Fahima, and E. Nevo, Microsatellites within genes: structure, function, and evolution. Mol. Biol. Evol. 21 (2004) 991-1007.
[151] D.B. Goldstein, and D.D. Pollock, Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. J. Hered. 88 (1997) 335-342.
[152] D.B. Goldstein, and C. Schlotterer, Microsatellites: Evolution and Applications, Oxford University Press, Oxford, UK, 1999.
[153] M. Menotti-Raymond, V.A. David, R. Agarwala, S. A.A., R. Stephens, S.J. O'Brien, and W.J. Murphy, Radiation hybrid mapping of 304 novel microsatellites in the domestic cat genome. Cytogenet. Genome Res. 102 (2003) 272-6.
[154] J.E. Janečka, T.L. Blankenship, D.H. Hirth, M.E. Tewes, C.W. Kilpatrick, and L.I.G. Jr, Kinship and social structure of bobcats (*Lynx rufus*) inferred from microsatellite and radio-telemetry data. J. Zool. 269 (2006) 494-501.
[155] L.L. Marker, A.J.P. Wilkerson, R.J. Sarno, J. Martenson, C. Breitenmoser-Würsten, S.J. O'Brien, and W.E. Johnson, Molecular genetic insights on cheetah (*Acinonyx jubatus*) ecology and conservation in Namibia. J. Hered. 99 (2008) 2-13.
[156] M. Ruiz-Garcia, E. Payán, A. Murillo, and D. Alvarez, DNA Microsatellite characterization of the jaguar (*Panthera onca*) in Colombia. Genes Genet. Syst. 81 (2006) 115-127.
[157] A. Singh, A. Gaur, K. Shailaja, B. Satyare Bala, and L. Singh, A novel microsatellite (STR) marker for forensic identification of big cats in India. Forensic Sci. Int. 141 (2004) 143-7.
[158] C.A. Driscoll, M. Menotti-Raymond, G. Nelson, D. Goldstein, and S.J. O'Brien, Genomic microsatellites as evolutionary chronometers: a test in wild cats. Genome Res. 12 (2002) 414-413.
[159] L.P. Waits, V.A. Buckley-Beason, W.E. Johnson, D.E. Ontorato, and T. McCarthy, A select panel of polymorphic microsatellite loci for individual identification of snow leopards (Panthera uncia). Mol. Ecol. Not. 7 (2007) 311-314.
[160] M.A. Jobling, and C. Tyler-Smith, The human Y chromosome: an evolutionary marker comes of age. Nat. Rev. Genet. 4 (2003) 598-612.
[161] W.J. Murphy, S. Shan, Z.-Q. Chen, J. Pecon-Slattery, and S.J. O'Brien, Extensive conservation of sex chromosome organization between cat and human revealed by parallel RH mapping. Genome Res (1999) 1223-1230.

[162] W.J. Murphy, A.J. Pearks-Wilkerson, T. Raudsepp, R. Agarwala, A.A. Schäffer, R. Stanyon, and B.P. Chowdhary, Novel gene acquisition on carnivore Y chromosomes. PLoS Genet. 2 (2006) e43.
[163] A.J. Pearks-Wilkerson, T. Raudsepp, T. Graves, D.A.W. Warren, B.P. Chowdhary, L.C. Skow, and W.J. Murphy, Gene discovery and comparative analysis of X-degenerate genes from the domestic cat Y chromosome. Genomics 92 (2008) 329-338.
[164] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, and D.J. Lipman, Basic local alignment search tool. J. Molec. Biol. 215 (1990) 403-410.
[165] NCBI, *Homo sapiens* (human) genome build 36.3, http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?chr=hum_chr.inf&query=, March 2008.
[166] NCBI, *Mus musculus* (laboratory mouse) genome build 37.1, http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?taxid=10090, March 2008.
[167] S. Rozen, and H.J. Skaletsky, Primer3 on the WWW for general users and for biologist programmers. in: S. Krawetz, and S. Misener, (Eds.), Bioinformatics Methods and Protocols: Methods in Molecular Biology, Humana Press, Totowa, NJ, 2000, pp. 365-386.
[168] RPCI - 86 male feline BAC library, http://bacpac.chori.org/mfeline86.htm, 2008.
[169] Sequencher©, Gene Codes Corporation, Ann Arbor, MI, 2008.
[170] GeneMapper©, Applied Biosystems, Foster City, CA, 2003.
[171] R. Masuda, J.V. Lopez, J.P. Slattery, N. Yuhki, and S.J. O'Brien, Molecular phylogeny of mitochondrial cytochrome b and 12S rRNA sequences in the Felidae: ocelot and domestic cat lineages. Mol. Phylogenet. Evol. 6 (1996) 351-65.
[172] M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, and D.G. Higgins, Clustal W and Clustal X version 2.0. Bioinformatics 23 (2007) 2947–2948.
[173] T.A. Hall, BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41 (1999) 95-98.
[174] F. Yue, and J. Tang, A divide-and-conquer implementation of three sequence alignment and ancestor inference. Proc. 1st IEEE Int. Conf. BIBM. (2007) 143 - 150.
[175] D.L. Swofford, PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods) Version 4, Sinauer Associates, Sunderland, MA, 2002.
[176] F. Ronquist, and J. Huelsenbeck, MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19 (2003) 1572-1574.
[177] D. Posada, and K.A. Crandall, Modeltest: testing the model of DNA substitution. Bioinformatics 14 (1998) 817-818.
[178] D. Posada, and T.R. Buckley, Model selection and model averaging in phylogenetics: advantages of the AIC and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53 (2004) 793-808.
[179] J.A.A. Nylander, MrModeltest v2, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden, 2004.
[180] J.S. Farris, M. Kallersjo, A.G. Kluge, and C. Bult, Testing significance of incongruence. Cladistics 10 (1995) 315-319.
[181] J. Sullivan, Combining data with different distributions of among-site variations. Syst. Biol. 45 (1996) 375-380.

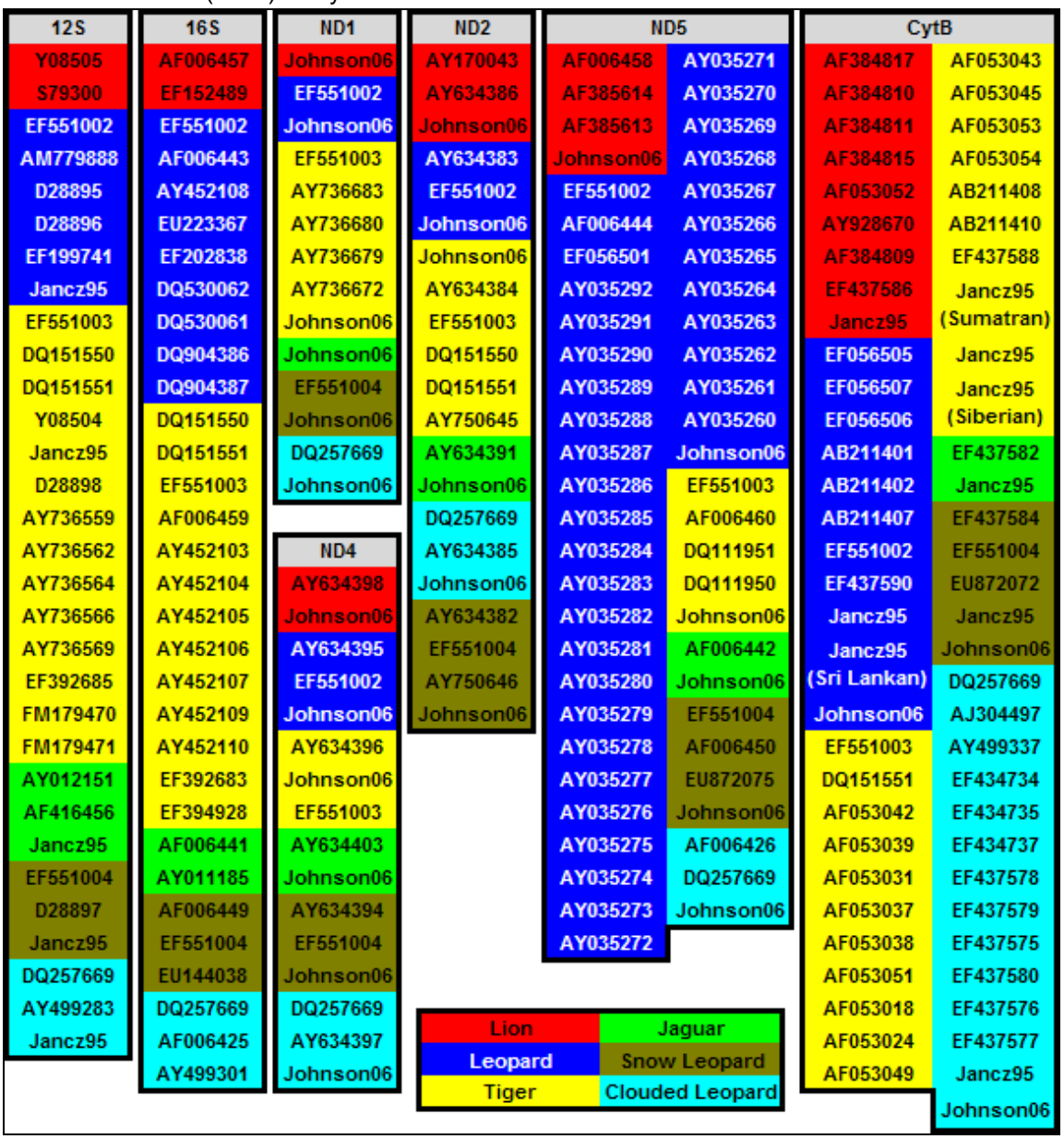
[182] C.W. Cunningham, Can tree incongruence tests predict when data should be combined? Mol. Biol. Evol. 14 (1997) 733-740.
[183] H. Shimodaira, and M. Hasegawa, Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16 (1999) 1114-1116.
[184] X. Wu, T. Zheng, Z. Jiang, and L. Wei, The mitochondrial genome structure of the clouded leopard (*Neofelis nebulosa*) Genome 50 (2007) 252-257.
[185] S.V. Edwards, L. Liu, and D.K. Pearl, High-resolution species trees without concatenation. PNAS 104 (2007) 5936-5941.
[186] J.H. Degnan, M. DeGiorgio, D. Bryant, and N.A. Rosenberg, Properties of consensus methods for inferring species trees from gene trees. Syst. Biol. 58 (2009) 35-54.
[187] L.S. Kubatko, and J.H. Degnan, Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst. Biol. 56 (2007) 17-24.
[188] L. Liu, D.K. Pearl, R.T. Brumfield, and S.V. Edwards, Estimating species trees using multiple-allele DNA sequence data. Evolution 62 (2008) 2080-2091.
[189] Z. Yang, PAML 4: Phylogenetic Analysis by Maximum Likelihood Mol. Biol. Evol. 24 (2007) 1586-1591.
[190] N. Goldman, and Z. Yang, A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol. Biol. Evol. 11 (1994) 725-736.
[191] S. Woolley, J. Johnson, M.J. Smith, K.A. Crandall, and D.A. McClellan, TreeSAAP: selection on amino acid properties using phylogenetic trees. Bioinformatics 19 (2003) 671-672.
[192] X. Xia, and W.H. Li, What amino acid properties affect protein evolution? J. Mol. Evol. 47 (1998) 557-564.
[193] D.A. McClellan, and K.G. McCracken, Estimating the influence of selection on the variable amino acid sites of the cytochrome b protein functional domains. Mol. Biol. Evol. 18 (2001) 917-925.
[194] R. da Fonseca, W. Johnson, S. O'Brien, M. Ramos, and A. Antunes, The adaptive evolution of the mammalian mitochondrial genome. BMC Genom. 9 (2008) 119.
[195] M. Hasegawa, H. Kishino, and T. Yano, Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22 (1985) 160-174.
[196] F. Tajima, Simple methods for testing the molecular evolutionary clock hypothesis. Genetics 135 (1993) 599-607.
[197] J.L. Thorne, and H. Kishino, Divergence time and evolutionary rate estimation with multilocus data. Syst. Biol. 51 (2002) 689-702.
[198] B. Kurten, and E. Anderson, Pleistocene Mammals of North America, Columbia University Press, New York, 1980.
[199] N.A. Neff, The Big Cats: Paintings by Guy Coheleach, Abrams, New York, 1982.
[200] L. Werdelin, and M.E. Lewis, Plio-Pleistocene Carnivora of eastern Africa: species richness and turnover patterns. Zool. J. Linn. Soc. 144 (2005) 121-144.
[201] J. Felsenstein, PHYLIP (Phylogeny Inference Package) version 3.6, Department of Genome Sciences, University of Washington, Seattle, WA, 2004.
[202] A. Stamatakis, RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22 (2006) 2688-2690.
[203] R. Barnett, B. Shapiro, I. Barnes, S.Y.W. Ho, J. Burger, N. Yamaguchi, T.F.G. Higham, H.T. Wheeler, W. Rosendahl, A.V. Sher, M. Sotnikova, T. Kuznetsova, G.F. Baryshnikov, L.D. Martin, C.R. Harington, J.A. Burns, and A. Cooper, Phylogeography of lions (*Panthera*

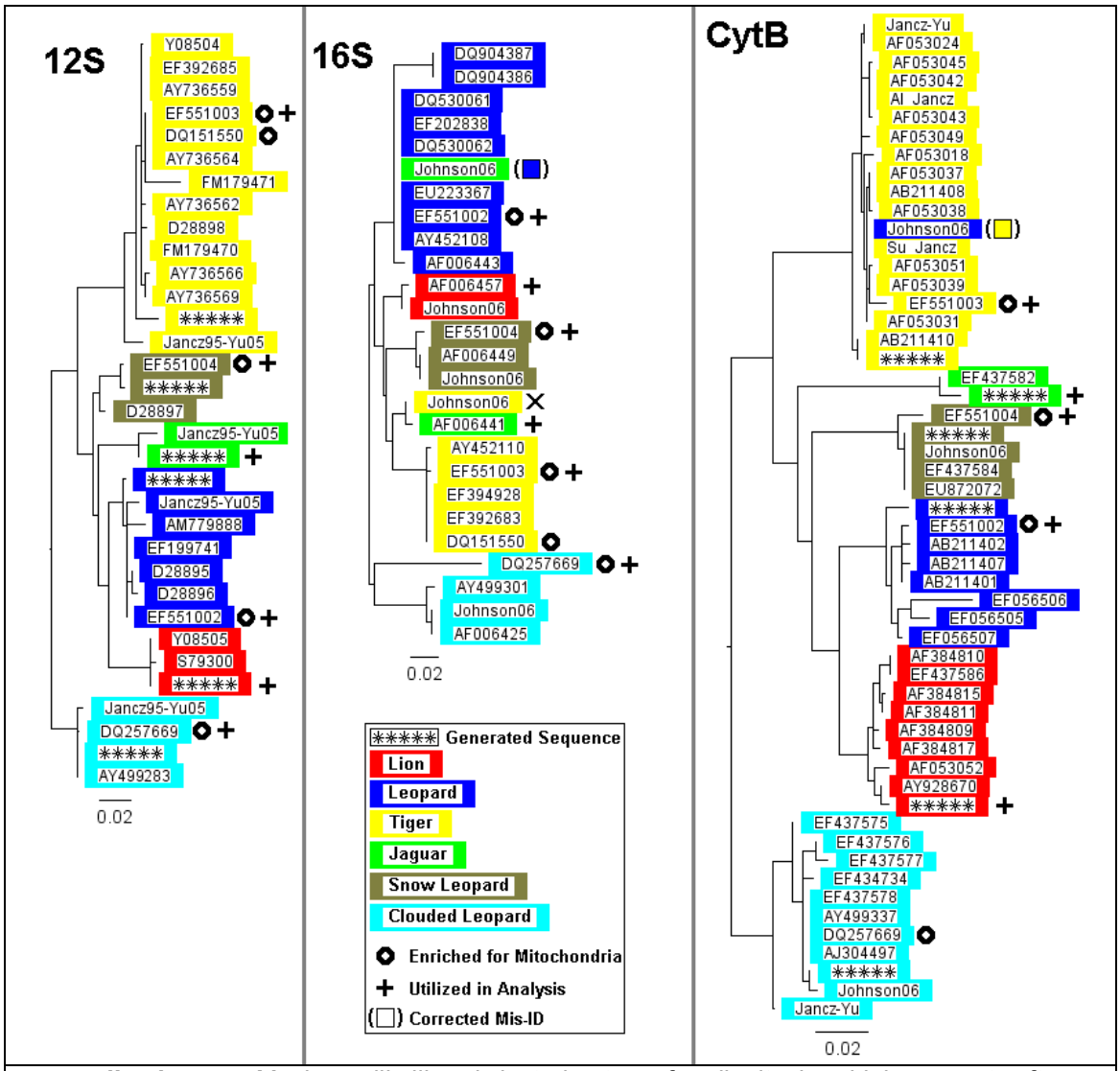
leo ssp.) reveals three distinct taxa and a late Pleistocene reduction in genetic diversity. Mol. Ecol. 18 (2009) 1668–1677.
[204] P.J. Planet, and I.N. Sarka, mILD: a tool for constructing and analyzing matrices of pairwise phylogenetic character incongruence tests. Bioinformatics 21 (2005) 4423-4424.
[205] Y. Suzuki, G.V. Glazko, and M. Nei, Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. PNAS 99 (2002) 16138-16143.
[206] J.H. Livesey, Kurtosis provides a good omnibus test for outliers in small samples Clin. Biochem. 40 (2007) 1032-1036.
[207] V. Barnett, and T. Lewis, Outliers in Statistical Data, 3rd edition, Wiley, West Sussex, England, 1994.
[208] F. Grubbs, Procedures for detecting outlying observations in samples. Technometrics 11 (1969) 1-21.
[209] L. Liu, and D.K. Pearl, Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. Syst. Biol. 56 (2007) 504-514.
[210] P. Christiansen, Phylogeny of the great cats (Felidae: Pantherinae), and the influence of fossil taxa and missing characters. Cladistics 24 (2008) 977–992.
[211] H. Hemmer, Untersuchungen zur stammesgeschichte der Pantherkatzen (Pantherinae). Teil III. Zur Artgeschichte des Löwen, *Panthera leo* (Linnaeus 1758). Veröffentlichungen der Zoologischen Staatssammlung 17 (1974) 167-280.
[212] H. Hemmer, Die evolution der Pantherkatzen: Modell zur überprüfung der brauchbarkeit der hennigschen prinzipien der phylogenetischen systematik für wirbeltierpaläontologische studien. Paläontologische Zeitschrift (1981) 109-116.
[213] F.E. Andersson, and D.L. Swofford, Should we be worried about long-branch attraction in real data sets? Investigations using metazoan 18S rDNA. Mol. Phylogenet. Evol. 33 (2004) 440-451.
[214] W.M. Fitch, and E. Markowitz, An improved method for determining codon variability in a gene and its application to the rate of fixation of mutations in evolution. Biochem. Genet. 4 (1970) 579-593.
[215] H. Philippe, and P. Lopez, On the conservation of protein sequences in evolution. Trends Biochem. Sci. 26 (2001) 414-416.
[216] B.W. Davis, T. Raudsepp, A.J.P. Wilkerson, R. Agarwala, A.A. Schäffer, M. Houck, O.A. Ryder, B.P. Chowdhdary, and W.J. Murphy, A high-resolution cat radiation hybrid and integrated FISH mapping resource for 2 phylogenomic studies across Felidae. Genomics (2009) 299-304.
[217] T.M. Anderson, B.M. vonHoldt, S.I. Candille, M. Musiani, C. Greco, D.R. Stahler, D.W. Smith, B. Padhukasahasram, E. Randi, J.A. Leonard, C.D. Bustamante, E.A. Ostrander, H. Tang, R.K. Wayne, and G.S. Barsh, Molecular and evolutionary history of melanism in North American gray wolves. Science 323 (2009) 1339-1343.
[218] R.D.M. Page, and M.A. Charleston, From gene to organismal phylogeny: reconciled trees and the gene tree/species tree problem. Mol. Phylogenet. Evol. 7 (1997) 231-240.
[219] B.V. Burger, M.Z. Viviers, J.P.I.B. &, M.l. Roux, N. Fish, W.B. Fourie, and G. Weibchen, Chemical characterization of territorial marking fluid of male Bengal tiger, *Panthera tigris*. J. Chem. Ecol. 34 (2008) 659-671.

[220] H.L. Monaco, The transthyretin-retinol-binding protein complex. Biochim. Biophys. Acta 1482 (2000) 65-72.
[221] M. Miyazaki, T. Yamashita, Y. Suzuki, Y. Saito, S. Soeta, H. Tiara, and A. Suzuki, Major urinary protein of the domestic cat regulates the production of felinine, a putative pheromone precursor. Chem. Biol. 13 (2006) 1071-1079.
[222] W.M. Brown, M. George, and A.C. Wilson, Rapid evolution of animal mitochondrial DNA. PNAS 76 (1979) 1967-1971.
[223] U. Brandt, Energy converting NADH:quinone oxidoreductase (complex I). Annu. Rev. Biochem. 75 (2006) 69-92.
[224] H. Song, J.E. Buhay, M.F. Whiting, and K.A. Crandall, Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. PNAS 105 (2008) 13486-13491.
[225] J.B.S. Haldane, Sex ratio and unisexual sterility in animal hybrids. J. Genet. 12 (1922) 101-109.
[226] B.U. Haq, J. Hardenbol, and P.R. Vail, Chronology of fluctuating sea levels since the triassic. Science 235 (1987) 1156-1166.

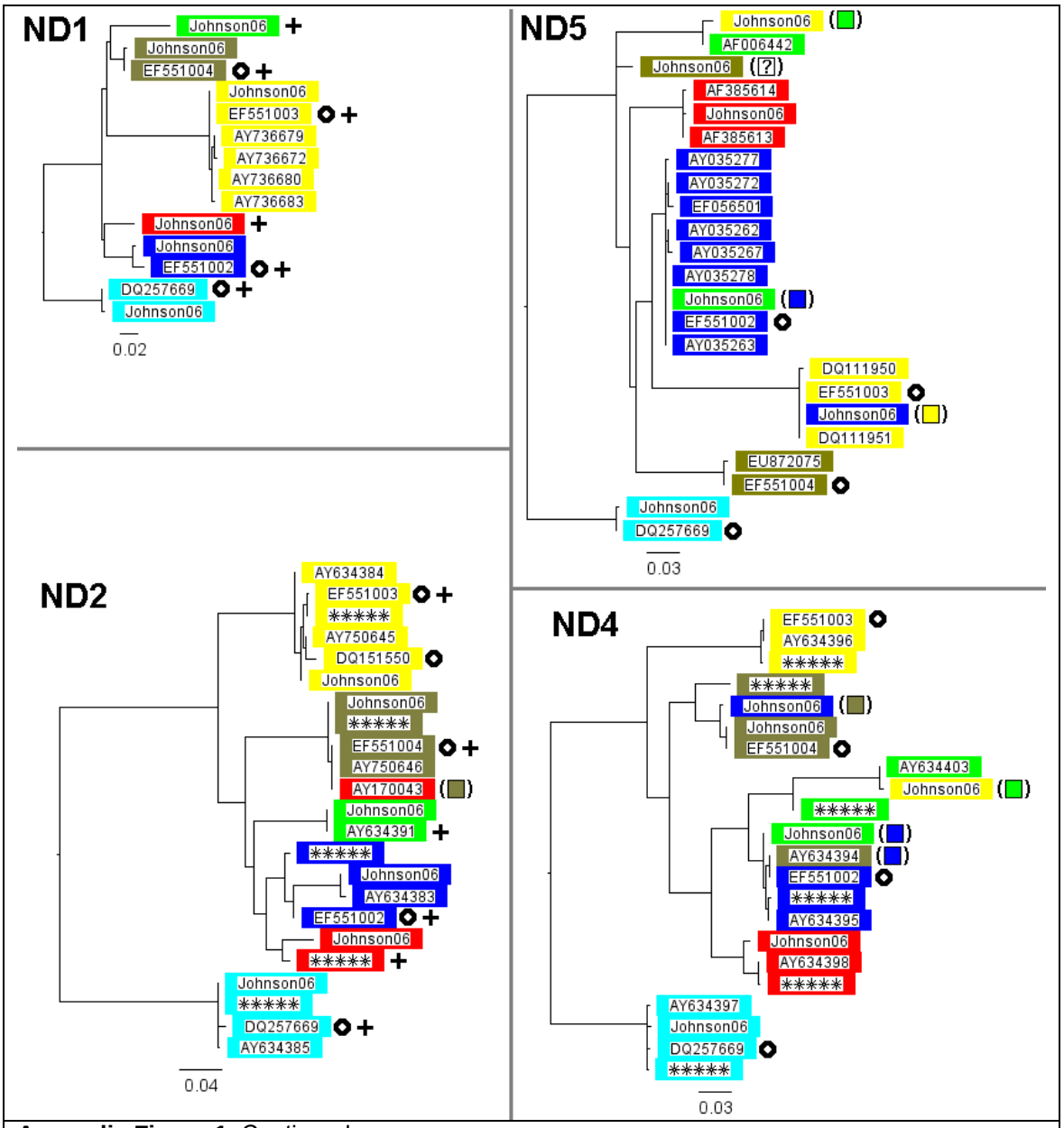
APPENDIX

Appendix Table 1. List of accession numbers for gene segments used in maximum likelihood mitochondrial analyses. Accession numbers not listed in GenBank are referenced by primary author and publication year. Some sequences were manually entered from within the body text of the Janczewski (1995) study and are denoted 'Jancz95'.

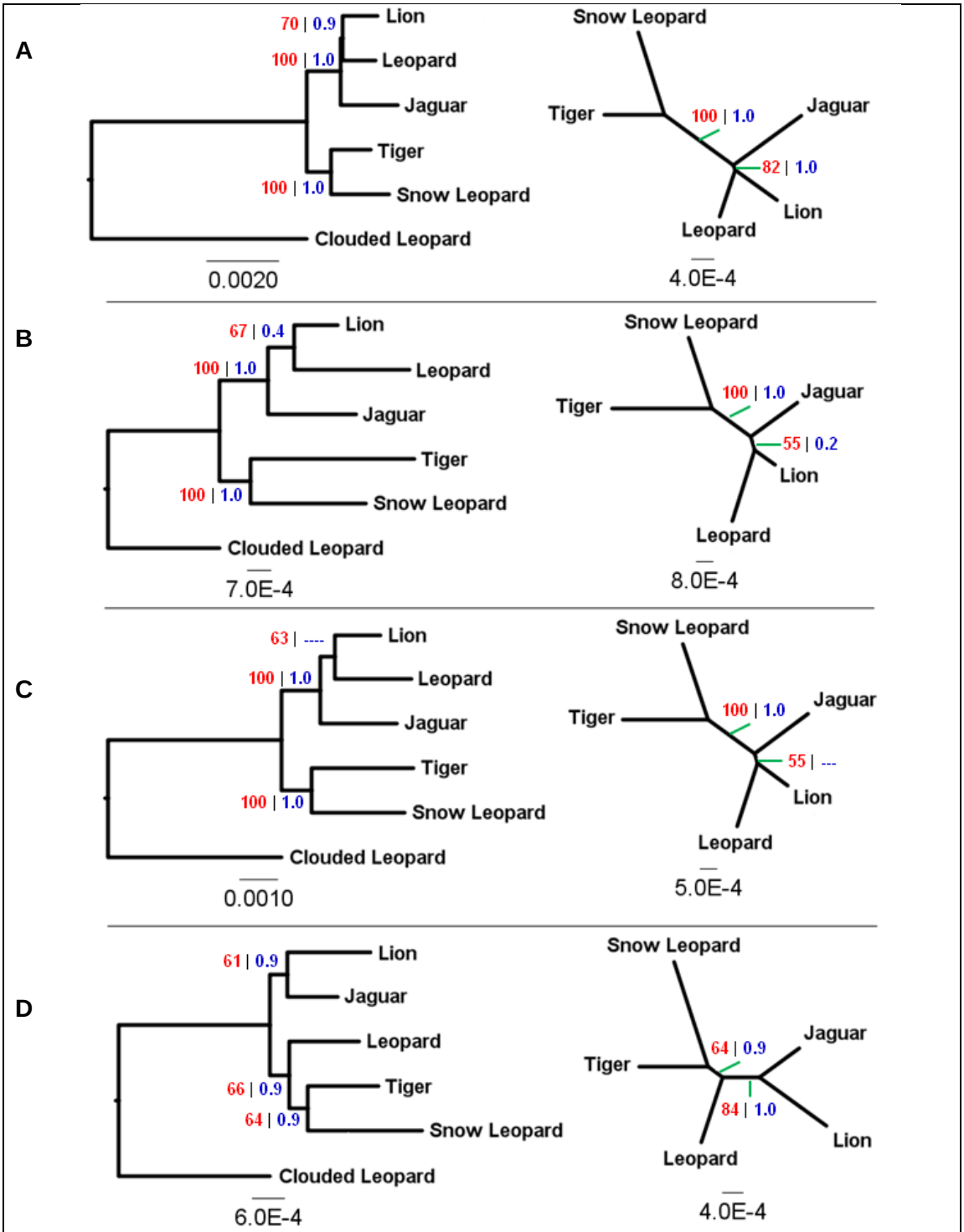




Appendix Figure 1. Maximum likelihood clustering trees for all mitochondrial segments after the removal of all putative numt sequences.



Appendix Figure 1. Continued



Appendix Figure 2. Maximum likelihood trees (rooted and unrooted) for the (A) Y chromosome (B) autosomal (C) nuclear (D) X chromosome partitions. ML bootstrap values (1000 replicates) in red, Bayesian posterior probabilities in blue.

Appendix Table 2. Maximum likelihood model parameters specified by Modeltest v3.4 [177]

Gene	Model	Maximum Likelihood Parameters
12S	[TrN+I]	Base=(0.3551 0.2384 0.1814) Nst=6 Rmat=(1.0000 31.0720 1.0000 1.0000 93.6756) Rates=equal Pinvar=0.8251
16S	[TIM+I)	Base=(0.3267 0.2209 0.2079) Nst=6 Rmat=(1.0000 37422698496.0 3924018688.0 3924018688.0 83880468480.0) Rates=equal Pinvar=0.7875
CYTB	[HKY+G]	Base=(0.2806 0.3057 0.1449) Nst=2 Tratio=18.6128 Rates=gamma Shape=0.2048 Pinvar=0
ND1	[HKY+G]	Base=(0.3176 0.3059 0.1089) Nst=2 Tratio=26.8145 Rates=gamma Shape=0.1514 Pinvar=0
ND2	[HKY+G]	Base=(0.3604 0.2955 0.0986) Nst=2 Tratio=21.2496 Rates=gamma Shape=0.1200 Pinvar=0
ND4	[K81uf+I]	Base=(0.3135 0.2901 0.1262) Nst=6 Rmat=(1.0000 29.7662 0.0001 0.0001 29.7662) Rates=equal Pinvar=0.6346
ND5	[HKY+G]	Base=(0.3142 0.2664 0.1282) Nst=2 Tratio=28.1177 Rates=gamma Shape=0.2455 Pinvar=0
DDX3Y	[HKY]	Base=(0.3095 0.1776 0.1739) Nst=2 Tratio=3.8930 Rates=equal Pinvar=0
EIF1AY	[TVM]	Base=(0.3290 0.1462 0.1540) Nst=6 Rmat=(0.0001 3.5554 0.2311 6.3969 3.5554) Rates=equal Pinvar=0
EIF2S3Y	[F81]	Base=(0.4137 0.1419 0.1274) Nst=1 Rates=equal Pinvar=0
SMCY	[HKY+I]	Base=(0.2839 0.1902 0.2261) Nst=2 Tratio=1.8734 Rates=equal Pinvar=0.8532
UBE1Y	[GTR]	Base=(0.2795 0.2242 0.1601) Nst=6 Rmat=(0.0000 7.2080 0.0000 2.9950 2.8408) Rates=equal Pinvar=0
USP9Y	[TVM+I]	Base=(0.3336 0.1595 0.1506) Nst=6 Rmat=(0.3694 2.1284 0.3171 1.6664 2.1284) Rates=equal Pinvar=0.7748
UTY	[HKY]	Base=(0.3777 0.1642 0.1372) Nst=2 Tratio=1.4398 Rates=equal Pinvar=0
SRY	[HKY+I]	Base=(0.2948 0.1852 0.2241) Nst=2 Tratio=2.1110 Rates=equal Pinvar=0.9198
ZFY	[HKY]	Base=(0.3266 0.1314 0.1770) Nst=2 Tratio=1.4370 Rates=equal Pinvar=0
FGB	[TVM+I]	Base=(0.2810 0.2031 0.1937) Nst=6 Rmat=(9.6652 14.3298 1.4122 5.3107 14.3298) Rates=equal Pinvar=0.8613
IRBP	[HKY]	Base=(0.1841 0.3147 0.3194) Nst=2 Tratio=3.3512 Rates=equal Pinvar=0
APP	[F81]	Base=(0.2658 0.2023 0.1776) Nst=1 Rates=equal Pinvar=0
CALB	[HKY]	Base=(0.3434 0.1676 0.1856) Nst=2 Tratio=459869221049005430000000000000000000.0000 Rates=equal Pinvar=0
CHRNA	[K80]	Base=equal Nst=2 Tratio=2.0093 Rates=equal Pinvar=0
CLU	[TIMef+I]	Base=equal Nst=6 Rmat=(1.0000 2.0046 0.0000 0.0000 6.1725) Rates=equal Pinvar=0.9214
CMA	[HKY+I]	Base=(0.2468 0.3137 0.1832) Nst=2 Tratio=3.6432 Rates=equal Pinvar=0.9074
DGKG2	[K81uf+I]	Base=(0.2600 0.2224 0.2104) Nst=6 Rmat=(1.0000 1.7658 0.1048 0.1048 1.7658) Rates=equal Pinvar=0.9370
FES	[K81uf+I]	Base=(0.1851 0.2730 0.3349) Nst=6 Rmat=(1.0000 1.3745 0.3449 0.3449 1.3745) Rates=equal Pinvar=0.9038
GATA	[K8luf]	Base=(0.3163 0.1838 0.2507) Nst=6 Rmat=(1.0000 1.4486 0.0000 0.0000 1.4486) Rates=equal Pinvar=0
GHR	[HKY]	Base=(0.3199 0.1935 0.1860) Nst=2 Tratio=4.0249 Rates=equal Pinvar=0
GNAZ	[HKY+I]	Base=(0.2224 0.3260 0.2809) Nst=2 Tratio=7.0577 Rates=equal Pinvar=0.9760
GNB	[HKY]	Base=(0.1949 0.2579 0.2953) Nst=2 Tratio=4.3585 Rates=equal Pinvar=0
HK1	[K8luf]	Base=(0.1878 0.2472 0.3015) Nst=6 Rmat=(1.0000 4.1561 0.0000 0.0000 4.1561) Rates=equal Pinvar=0
NCL	[HKY+I]	Base=(0.3041 0.1614 0.2277) Nst=2 Tratio=4.3783 Rates=equal Pinvar=0.9449
PNOC	[F81]	Base=(0.1442 0.3025 0.3059) Nst=1 Rates=equal Pinvar=0
RAG2	[HKY]	Base=(0.2684 0.1989 0.2325) Nst=2 Tratio=3.0117 Rates=equal Pinvar=0
RASA	[F81]	Base=(0.3330 0.1418 0.1904) Nst=1 Rates=equal Pinvar=0
SILV	[JC]	Base=equal Nst=1 Rates=equal Pinvar=0
TCP	[K8luf]	Base=(0.1833 0.2594 0.2732) Nst=6 Rmat=(1.0000 2.0302 0.0000 0.0000 2.0302) Rates=equal Pinvar=0
TTR	[K8luf]	Base=(0.2822 0.2128 0.1963) Nst=6 Rmat=(1.0000 2.5003 0.0000 0.0000 2.5003) Rates=equal Pinvar=0
CES7	[HKY+I]	Base=(0.2157 0.2583 0.2720) Nst=2 Tratio=2.5185 Rates=equal Pinvar=0.8856 clock=no
ALAS	[K80]	Base=equal Nst=2 Tratio=6.0296 Rates=equal Pinvar=0
ATP7A	[HKY]	Base=(0.2680 0.1873 0.2039) Nst=2 Tratio=5.0100 Rates=equal Pinvar=0
IL2RG	[JC]	Base=equal Nst=1 Rates=equal Pinvar=0
PLP	[HKY]	Base=(0.2209 0.2387 0.2265) Nst=2 Tratio=3.0163 Rates=equal Pinvar=0
ZFX	[HKY]	Base=(0.2969 0.1668 0.2207) Nst=2 Tratio=2.2587 Rates=equal Pinvar=0
Y Chromosome	[TVM+I]	Base=(0.3198 0.1725 0.1767) Nst=6 Rmat=(0.4527 2.7629 0.4131 1.6506 2.7629) Rates=equal Pinvar=0.7683
Mitochondrial	[TrN+I+G]	Base=(0.3233 0.2824 0.1345) Nst=6 Rmat=(1.0000 40.8743 1.0000 1.0000 48.8259) Rates=gamma Shape=0.3345 Pinvar=0.3018
Autosomal	[K81uf+I]	Base=(0.2457 0.2391 0.2479) Nst=6 Rmat=(1.0000 4.1593 0.6009 0.6009 4.1593) Rates=equal Pinvar=0.9371
Uniparental	[GTR+I+G]	Base=(0.3128 0.1996 0.1773) Nst=6 Rmat=(0.9048 10.2746 0.5547 1.4171 18.0246) Rates=gamma Shape=0.6017 Pinvar=0.8208
Nuclear	[TVM+I]	Base=(0.2938 0.1948 0.2010) Nst=6 Rmat=(0.5782 2.9995 0.3573 1.0300 2.9995) Rates=equal Pinvar=0.9045
Supermatrix	[GTR+I+G]	Base=(0.2868 0.2147 0.2018) Nst=6 Rmat=(0.8165 8.4354 0.4809 0.8814 13.7819) Rates=gamma Shape=0.4110 Pinvar=0.8078

Appendix Table 3. Bayesian model parameters specified by MrModeltest v2

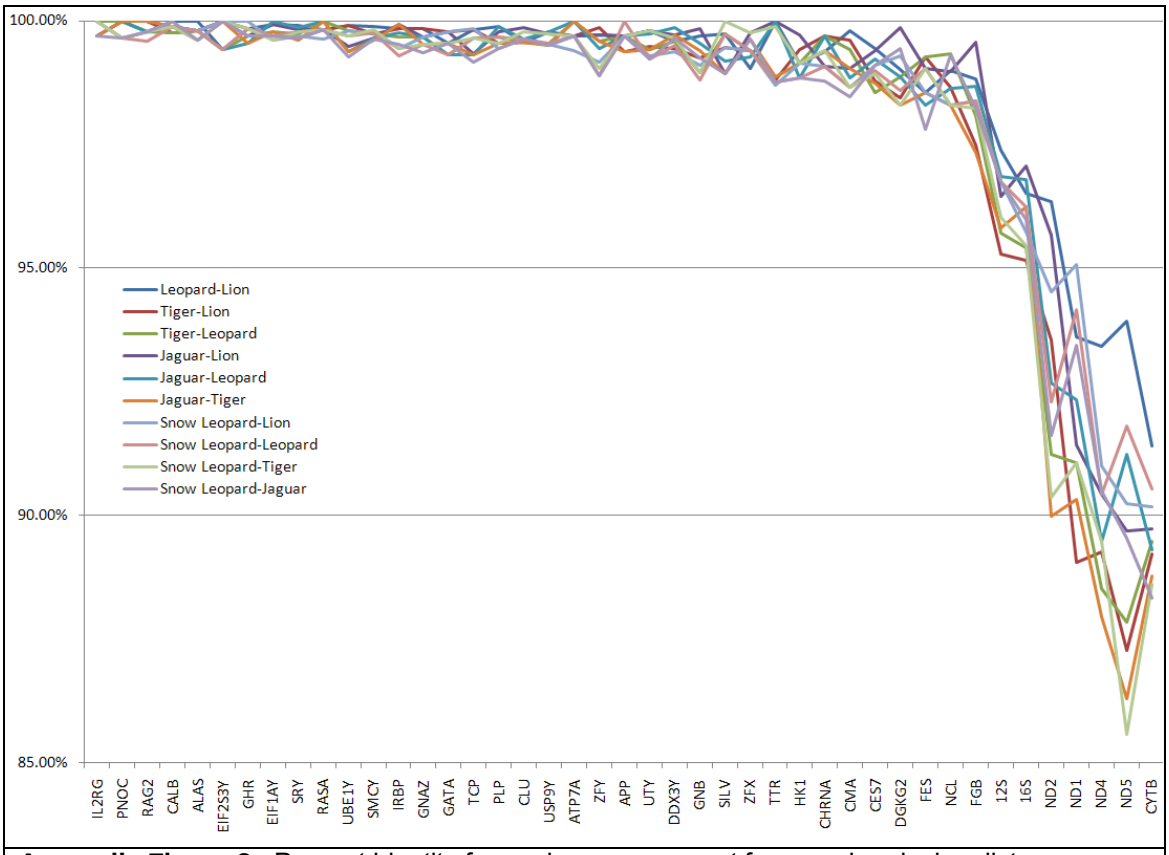
Gene	Model	Bayesian Parameters		
DDX3Y	[HKY]	nst=2	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
EIF1AY	[GTR]	nst=6	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
EIF2S3Y	[F81]	nst=1	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
SMCY	[HKY+I]	nst=2	rates=propinv;	statefreqpr=dirichlet(1,1,1,1);
UBE1Y	[GTR+I]	nst=6	rates=propinv;	statefreqpr=dirichlet(1,1,1,1);
USP9Y	[GTR]	nst=6	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
UTY	[HKY]	nst=2	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
SRY	[HKY+I]	nst=2	rates=propinv;	statefreqpr=dirichlet(1,1,1,1);
ZFY	[HKY]	nst=2	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
FGB	[GTR+I]	nst=6	rates=propinv;	statefreqpr=dirichlet(1,1,1,1);
IRBP	[HKY]	nst=2	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
APP	[F81]	nst=1	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
CALB	[HKY]	nst=2	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
CHRNA	[K80]	nst=2	rates=equal;	statefreqpr=fixed(equal);
CLU	[K80]	nst=2	rates=equal;	statefreqpr=fixed(equal);
CMA	[HKY+I]	nst=2	rates=propinv;	statefreqpr=dirichlet(1,1,1,1);
DGKG2	[HKY+G]	nst=2	rates=gamma;	statefreqpr=dirichlet(1,1,1,1);
FES	[HKY+I]	nst=2	rates=propinv;	statefreqpr=dirichlet(1,1,1,1);
GATA	[F81]	nst=1	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
GHR	[HKY]	nst=2	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
GNAZ	[HKY+I]	nst=2	rates=propinv;	statefreqpr=dirichlet(1,1,1,1);
GNB	[HKY]	nst=2	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
HK1	[HKY]	nst=2	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
NCL	[HKY+I]	nst=2	rates=propinv;	statefreqpr=dirichlet(1,1,1,1);
PNOC	[HKY]	nst=2	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
RAG2	[HKY]	nst=2	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
RASA	[F81]	nst=1	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
SILV	[JC]	nst=1	rates=equal;	statefreqpr=fixed(equal);
TCP	[F81]	nst=1	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
TTR	[HKY]	nst=2	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
CES7	[HKY+I]	nst=2	rates=propinv;	statefreqpr=dirichlet(1,1,1,1);
ALAS	[K80]	nst=2	rates=equal;	statefreqpr=fixed(equal);
ATP7A	[HKY]	nst=2	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
IL2RG	[JC]	nst=1	rates=equal;	statefreqpr=fixed(equal);
PLP	[HKY]	nst=2	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
ZFX	[HKY]	nst=2	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
12S	[GTR+I+G]	nst=6	rates=invgamma;	statefreqpr=dirichlet(1,1,1,1);
16S	[HKY+I]	nst=2	rates=propinv;	statefreqpr=dirichlet(1,1,1,1);
CYTB	[HKY+G]	nst=2	rates=gamma;	statefreqpr=dirichlet(1,1,1,1);
ND1	[HKY+G]	nst=2	rates=gamma;	statefreqpr=dirichlet(1,1,1,1);
ND2	[HKY+G]	nst=2	rates=gamma;	statefreqpr=dirichlet(1,1,1,1);
ND4	[GTR+I]	nst=6	rates=propinv;	statefreqpr=dirichlet(1,1,1,1);
ND5	[HKY+G]	nst=2	rates=gamma;	statefreqpr=dirichlet(1,1,1,1);
Y Chromosome	[GTR+I]	nst=6	rates=propinv;	statefreqpr=dirichlet(1,1,1,1);
Autosomes	[GTR+I]	nst=6	rates=propinv;	statefreqpr=dirichlet(1,1,1,1);
X Chromosome	[HKY]	nst=2	rates=equal;	statefreqpr=dirichlet(1,1,1,1);
Mitochondria	[HKY+G]	nst=2	rates=gamma;	statefreqpr=dirichlet(1,1,1,1);
Nuclear	[GTR+I]	nst=6	rates=propinv;	statefreqpr=dirichlet(1,1,1,1);
Uniparental	[GTR+I]	nst=6	rates=propinv;	statefreqpr=dirichlet(1,1,1,1);

Appendix Table 4. Molecular clock test results for individual gene segments. L_0 is the unconstrained –ln likelihood value and L_1 is the –ln likelihood value when the topology is constrained to clocklike behavior. The critical value for rejection of the null hypothesis is P < 0.05.

		-ln L	L_0-L_1	$2(L_0-L_{1)}$	P value (df=4)
12S	L_0	1853.20	0.71275	1.4255	0.83975
	L_1	1853.91			
16S	L_0	833.91	12.49956	24.99912	0.00005
	L_1	846.41			
CYTB	L_0	3220.38	5.10433	10.20866	0.03706
	L_1	3225.49			
ND1	L_0	1368.38	3.31873	6.63746	0.15633
	L_1	1371.70			
ND2	L_0	2647.47	30.01528	60.03056	0.00000
	L_1	2677.48			
ND4	L_0	3731.56	8.15281	16.30562	0.00264
	L_1	3739.71			
ND5	L_0	1988.32	15.2418	30.4836	0.00000
	L_1	2003.56			
DDX3Y	L_0	2133.94	14.80743	29.61486	0.00001
	L_1	2148.75			
EIF1AY	L_0	1854.14	12.92781	25.85562	0.00003
	L_1	1867.07			
EIF2S3Y	L_0	467.43	5.5103	11.0206	0.02633
	L_1	472.94			
SMCY	L_0	5652.02	25.21807	50.43614	0.00000
	L_1	5677.24			
UBE1Y	L_0	1412.98	4.93836	9.87672	0.04256
	L_1	1417.91			
USP9Y	L_0	8996.54	41.78574	83.57148	0.00000
	L_1	9038.32			
UTY	L_0	2296.76	8.97211	17.94422	0.00127
	L_1	2305.73			
SRY	L_0	3536.01	22.93268	45.86536	0.00000
	L_1	3558.94			
ZFY	L_0	1065.84	6.58924	13.17848	0.01044
	L_1	1072.42			
FGB	L_0	1157.25	1.83607	3.67214	0.45219
	L_1	1159.08			
IRBP	L_0	1787.63	3.86017	7.72034	0.10238
	L_1	1791.49			
APP	L_0	907.96	4.32244	8.64488	0.07062
	L_1	912.28			
CALB	L_0	1130.42	3.3025	6.605	0.15829
	L_1	1133.72			
CHRNA	L_0	513.84	3.6091	7.2182	0.12480
	L_1	517.45			
CLU	L_0	2013.16	0.48458	0.96916	0.91444
	L_1	2013.65			

Appendix Table 4. Continued

		-ln L	L_0-L_1	2(L_0-L_1)	P value (df=4)
CMA	L_0	801.13	2.88277	5.76554	0.21736
	L_1	804.02			
DGKG2	L_0	1116.55	2.89857	5.79714	0.21482
	L_1	1119.45			
FES	L_0	683.35	2.57262	5.14524	0.27272
	L_1	685.92			
GATA	L_0	662.75	4.10261	8.20522	0.08434
	L_1	666.85			
GHR	L_0	979.27	3.97444	7.94888	0.09347
	L_1	983.25			
GNAZ	L_0	894.80	2.299	4.598	0.33108
	L_1	897.10			
GNB	L_0	1040.97	3.25236	6.50472	0.16449
	L_1	1044.22			
HK1	L_0	547.67	3.64109	7.28218	0.12171
	L_1	551.31			
NCL	L_0	467.76	3.1165	6.233	0.18241
	L_1	470.88			
PNOC	L_0	398.90	1.09943	2.19886	0.69924
	L_1	400.00			
RAG2	L_0	680.35	3.01262	6.02524	0.19727
	L_1	683.36			
RASA	L_0	768.70	4.1124	8.2248	0.08368
	L_1	772.81			
SILV	L_0	551.08	3.30395	6.6079	0.15812
	L_1	554.38			
TCP	L_0	883.43	2.77183	5.54366	0.23592
	L_1	886.20			
TTR	L_0	1308.81	1.4097	2.8194	0.58849
	L_1	1310.22			
CES7	L_0	9148.96	10.05674	20.11348	0.00047
	L_1	9159.02			
ALAS	L_0	759.85	4.40555	8.8111	0.06600
	L_1	764.25			
ATP7A	L_0	953.77	4.25245	8.5049	0.07474
	L_1	958.03			
IL2RG	L_0	456.56	2.19932	4.39864	0.35474
	L_1	458.76			
PLP	L_0	1351.02	7.39417	14.78834	0.00516
	L_1	1358.42			
ZFX	L_0	1217.00	2.9535	5.907	0.20620
	L_1	1219.96			
Concaten.	L_0	26809.79	11.51815	23.0363	0.00012
	L_1	26821.30			



Appendix Figure 3. Percent identity for each gene segment from each pairwise distance calculation.

Appendix Table 5. Pairwise identities and differences between all pairs of taxa. Total differences for each taxa pair indicated by the top value and total identities by the middle value for each gene segment. Percentage identity between taxa for each gene segment listed as the bottom number. The standard deviation of the percent identity for each gene segment and the total average dissimilarity is listed both with and without the most divergent species, the clouded leopard.

	DDX3Y	EIF1AY	EIF2S3Y	SMCY	UBE1Y	USP9Y	UTY	SRY	ZFY	FGB	IRBP	APP	CALB	CHRNA	CLU	CMA	DGKG2	FES	GATA	GHR	GNAZ	GNB
Partition Length:					19140																19124		
Gene Segment Length:	1428	1314	817	3793	977	6135	1575	2371	730	678	1254	642	817	326	1369	518	710	413	457	670	622	674
Leopard-Lion	6	1	2	4	1	14	3	2	2	8	2	2	0	2	6	1	7	6	2	1	1	2
	1422	1313	815	3789	976	6121	1572	2369	728	670	1252	640	817	324	1363	517	703	407	455	669	621	672
	99.58%	99.92%	99.75%	99.89%	99.90%	99.77%	99.81%	99.92%	99.73%	98.81%	99.84%	99.69%	100.00%	99.38%	99.56%	99.81%	99.00%	98.53%	99.56%	99.85%	99.84%	99.70%
Tiger-Lion	8	4	0	10	1	29	8	8	1	17	2	4	2	1	6	2	11	3	1	1	1	5
	1420	1310	817	3783	976	6106	1567	2363	729	661	1252	638	815	325	1363	516	699	410	456	669	621	669
	99.44%	99.69%	100.00%	99.74%	99.90%	99.53%	99.49%	99.66%	99.86%	97.43%	99.84%	99.37%	99.75%	99.69%	99.56%	99.61%	98.43%	99.27%	99.78%	99.85%	99.84%	99.25%
Tiger-Leopard	6	3	2	10	2	28	9	6	3	13	4	2	2	1	6	3	8	3	3	2	2	7
	1422	1311	815	3783	975	6107	1566	2365	727	665	1250	640	815	325	1363	515	702	410	454	668	620	667
	99.58%	99.77%	99.75%	99.74%	99.79%	99.54%	99.43%	99.75%	99.59%	98.05%	99.68%	99.69%	99.75%	99.69%	99.56%	99.42%	98.86%	99.27%	99.34%	99.70%	99.68%	98.95%
Jaguar-Lion	4	1	0	13	5	16	3	4	2	3	1	2	1	3	2	5	1	4	1	2	2	1
	1424	1313	817	3780	972	6119	1572	2367	728	675	1253	640	816	323	1367	513	709	409	456	668	620	673
	99.72%	99.92%	100.00%	99.66%	99.49%	99.74%	99.81%	99.83%	99.73%	99.56%	99.92%	99.69%	99.88%	99.07%	99.85%	99.03%	99.86%	99.02%	99.78%	99.70%	99.68%	99.85%
Jaguar-Leopard	2	0	2	13	6	14	4	3	4	9	3	2	1	1	5	6	8	7	3	3	2	3
	1426	1314	815	3780	971	6121	1571	2368	726	669	1251	640	816	325	1364	512	702	406	454	667	620	671
	99.86%	100.00%	99.75%	99.66%	99.38%	99.77%	99.75%	99.87%	99.45%	98.65%	99.76%	99.69%	99.88%	99.69%	99.63%	98.83%	98.86%	98.28%	99.34%	99.55%	99.68%	99.55%
Jaguar-Tiger	4	3	0	15	6	30	9	9	3	18	1	4	1	2	6	5	12	6	2	3	3	4
	1424	1311	817	3778	971	6105	1566	2362	727	660	1253	638	816	324	1363	513	698	407	455	667	619	670
	99.72%	99.77%	100.00%	99.60%	99.38%	99.51%	99.43%	99.62%	99.59%	97.27%	99.92%	99.37%	99.88%	99.38%	99.56%	99.03%	98.28%	98.53%	99.56%	99.55%	99.52%	99.40%
Snow Leopard-Lion	9	5	0	9	2	29	11	7	6	11	7	2	1	3	5	7	5	6	1	0	2	6
	1419	1309	817	3784	975	6106	1564	2364	724	667	1247	640	816	323	1364	511	705	407	456	670	620	668
	99.37%	99.62%	100.00%	99.76%	99.79%	99.53%	99.30%	99.70%	99.17%	98.35%	99.44%	99.69%	99.88%	99.07%	99.63%	98.63%	99.29%	98.53%	99.78%	100.00%	99.68%	99.10%
Snow Leopard-Leopard	7	4	2	9	3	28	12	5	8	11	9	0	1	3	5	7	10	4	3	1	3	8
	1421	1310	815	3784	974	6107	1563	2366	722	667	1245	642	816	323	1364	511	700	409	454	669	619	666
	99.51%	99.69%	99.75%	99.76%	99.69%	99.54%	99.23%	99.79%	98.89%	98.35%	99.28%	100.00%	99.88%	99.07%	99.63%	98.63%	98.57%	99.02%	99.34%	99.85%	99.52%	98.80%
Snow Leopard-Tiger	5	5	0	7	3	16	3	5	7	12	7	2	1	2	3	7	12	4	2	1	3	7
	1423	1309	817	3786	974	6119	1572	2366	723	666	1247	640	816	324	1366	511	698	409	455	669	619	667
	99.65%	99.62%	100.00%	99.82%	99.69%	99.74%	99.81%	99.79%	99.03%	98.20%	99.44%	99.69%	99.88%	99.38%	99.78%	98.63%	98.28%	99.02%	99.56%	99.85%	99.52%	98.95%
Snow Leopard-Jaguar	5	4	0	14	7	30	12	8	8	12	6	2	0	4	5	8	4	9	2	2	4	5
	1423	1310	817	3779	970	6105	1563	2363	722	666	1248	640	817	322	1364	510	706	404	455	668	618	669
	99.65%	99.69%	100.00%	99.63%	99.28%	99.51%	99.23%	99.66%	98.89%	98.20%	99.52%	99.69%	100.00%	98.76%	99.63%	98.43%	99.43%	97.77%	99.56%	99.70%	99.35%	99.25%
Clouded Leopard-Lion	26	15	3	41	4	98	23	32	6	17	4	2	4	4	7	4	5	6	1	7	1	8
	1402	1299	814	3752	973	6037	1552	2339	724	661	1250	640	813	322	1362	514	705	407	456	663	621	666
	98.15%	98.85%	99.63%	98.91%	99.59%	98.38%	98.52%	98.63%	99.17%	97.43%	99.68%	99.69%	99.51%	98.76%	99.49%	99.22%	99.29%	98.53%	99.78%	98.94%	99.84%	98.80%
Clouded Leopard-Leopard	24	14	5	41	5	96	24	31	8	13	6	0	4	5	5	5	8	5	3	8	2	10
	1404	1300	812	3752	972	6039	1551	2340	722	665	1248	642	813	321	1364	513	702	408	454	662	620	664
	98.29%	98.92%	99.38%	98.91%	99.49%	98.41%	98.45%	98.68%	98.89%	98.05%	99.52%	100.00%	99.51%	98.44%	99.63%	99.03%	98.86%	98.77%	99.34%	98.79%	99.68%	98.49%
Clouded Leopard-Tiger	24	17	3	43	5	92	23	31	6	18	4	2	4	4	7	4	10	5	2	8	2	9
	1404	1297	814	3750	972	6043	1552	2340	724	660	1250	640	813	322	1362	514	700	408	455	662	620	665
	98.29%	98.69%	99.63%	98.85%	99.49%	98.48%	98.52%	98.68%	99.17%	97.27%	99.68%	99.69%	99.51%	98.76%	99.49%	99.22%	98.57%	98.77%	99.56%	98.79%	99.68%	98.65%
Clouded Leopard-Jaguar	22	14	3	46	9	98	24	32	8	18	3	2	3	6	5	3	6	7	2	7	3	7
	1406	1300	814	3747	968	6037	1551	2339	722	660	1251	640	814	320	1364	515	704	406	455	663	619	667
	98.44%	98.92%	99.63%	98.77%	99.07%	98.38%	98.45%	98.63%	98.89%	97.27%	99.76%	99.69%	99.63%	98.13%	99.63%	99.42%	99.15%	98.28%	99.56%	98.94%	99.52%	98.95%
Clouded Leopard-Snow Leopard	23	18	3	42	6	91	26	30	12	14	7	0	3	6	6	6	7	8	0	7	3	8
	1405	1296	814	3751	971	6044	1549	2341	718	664	1247	642	814	320	1363	512	703	405	457	663	619	666
	98.36%	98.61%	99.63%	98.88%	99.38%	98.49%	98.32%	98.72%	98.33%	97.89%	99.44%	100.00%	99.63%	98.13%	99.56%	98.83%	99.00%	98.02%	100.00%	98.94%	99.52%	98.80%
Maximum	99.86%	100.00%	100.00%	99.89%	99.90%	99.77%	99.81%	99.92%	99.86%	99.56%	99.92%	100.00%	100.00%	99.69%	99.85%	99.81%	99.86%	99.27%	100.00%	100.00%	99.84%	99.85%
Minimum	98.15%	98.61%	99.38%	98.77%	99.07%	98.38%	98.32%	98.63%	98.33%	97.27%	99.28%	99.37%	99.51%	98.13%	99.49%	98.43%	98.28%	97.77%	99.34%	98.79%	99.35%	98.49%
Range	1.71%	1.39%	0.62%	1.12%	0.83%	1.39%	1.49%	1.28%	1.53%	2.28%	0.64%	0.63%	0.49%	1.57%	0.37%	1.38%	1.58%	1.50%	0.66%	1.21%	0.49%	1.36%
Average Dissimilarity	0.828%	0.553%	0.205%	0.562%	0.446%	0.780%	0.831%	0.605%	0.775%	1.949%	0.352%	0.292%	0.229%	0.973%	0.386%	0.950%	1.084%	1.360%	0.411%	0.532%	0.366%	0.900%
Average Dissimilarity (No Clouded)	0.394%	0.229%	0.098%	0.275%	0.370%	0.383%	0.473%	0.241%	0.608%	1.714%	0.337%	0.344%	0.123%	0.680%	0.359%	0.996%	1.113%	1.277%	0.440%	0.240%	0.371%	0.718%
Standard Deviation	0.648%	0.493%	0.195%	0.427%	0.240%	0.589%	0.561%	0.539%	0.424%	0.657%	0.199%	0.186%	0.173%	0.528%	0.098%	0.399%	0.447%	0.440%	0.202%	0.446%	0.143%	0.395%
Standard Deviation (No Clouded)	0.146%	0.135%	0.127%	0.090%	0.229%	0.120%	0.244%	0.098%	0.365%	0.655%	0.230%	0.178%	0.082%	0.321%	0.101%	0.467%	0.524%	0.481%	0.180%	0.145%	0.154%	0.346%

Appendix Table 5 Continued

	HK1	NCL	PNOC	RAG2	RASA	SILV	TCP	TTR	CES7	ALAS	ATP7A	IL2RG	PLP	ZFX	12S	16S	CYTB	ND1	ND2	ND4	ND5	TOTAL
Partition Length:												3223						6141				47628
Gene Segment Length:	350	297	290	473	551	371	596	885	6161	513	670	318	899	823	964	376	1140	548	1038	1368	707	
Leopard-Lion	3	3	0	1	1	1	1	0	32	0	2	0	1	8	25	13	98	35	38	90	43	470
	347	294	290	472	550	370	595	885	6129	513	668	318	898	815	939	363	1042	513	1000	1278	664	47158
	99.14%	98.98%	100.00%	99.79%	99.82%	99.73%	99.83%	100.00%	99.48%	100.00%	99.70%	100.00%	99.89%	99.02%	97.34%	96.42%	90.60%	93.18%	96.20%	92.96%	93.52%	99.00%
Tiger-Lion	2	4	0	0	1	2	1	10	69	1	2	0	4	5	45	18	123	60	67	147	90	776
	348	293	290	473	550	369	595	875	6092	512	668	318	895	818	919	358	1017	488	971	1221	617	46852
	99.43%	98.63%	100.00%	100.00%	99.82%	99.46%	99.83%	98.86%	98.87%	99.80%	99.70%	100.00%	99.55%	99.39%	95.10%	94.97%	87.91%	87.70%	93.10%	87.96%	85.41%	98.34%
Tiger-Leopard	3	2	0	1	0	1	2	10	83	1	0	0	3	5	41	17	120	49	91	157	86	797
	347	295	290	472	551	370	594	875	6078	512	670	318	896	818	923	359	1020	499	947	1211	621	46831
	99.14%	99.32%	100.00%	99.79%	100.00%	99.73%	99.66%	98.86%	98.63%	99.80%	100.00%	100.00%	99.67%	99.39%	95.56%	95.26%	88.24%	90.18%	90.39%	87.04%	86.15%	98.30%
Jaguar-Lion	1	3	0	0	1	4	4	0	34	1	2	1	2	2	34	11	117	47	45	131	73	589
	349	294	290	473	550	367	592	885	6127	512	668	317	897	821	930	365	1023	501	993	1237	634	47039
	99.71%	98.98%	100.00%	100.00%	99.82%	98.91%	99.32%	100.00%	99.45%	99.80%	99.70%	99.68%	99.78%	99.76%	96.34%	96.99%	88.56%	90.62%	95.47%	89.41%	88.49%	98.75%
Jaguar-Leopard	4	4	0	1	0	3	4	0	44	1	0	1	1	6	30	12	122	42	76	144	62	658
	346	293	290	472	551	368	592	885	6117	512	670	317	898	817	934	364	1018	506	962	1224	645	46970
	98.84%	98.63%	100.00%	99.79%	100.00%	99.18%	99.32%	100.00%	99.28%	99.80%	100.00%	99.68%	99.89%	99.27%	96.79%	96.70%	88.02%	91.70%	92.10%	88.24%	90.39%	98.60%
Jaguar-Tiger	3	5	0	0	0	4	4	10	72	2	0	1	4	3	40	14	128	53	104	165	97	855
	347	292	290	473	551	367	592	875	6089	511	670	317	895	820	924	362	1012	495	934	1203	610	46773
	99.14%	98.29%	100.00%	100.00%	100.00%	98.91%	99.32%	98.86%	98.82%	99.61%	100.00%	99.68%	99.55%	99.63%	95.67%	96.13%	87.35%	89.29%	88.87%	86.28%	84.10%	98.17%
Snow Leopard-Lion	3	5	1	1	2	2	1	11	51	1	4	0	5	5	31	16	112	27	57	123	69	661
	347	292	289	472	549	369	595	874	6110	512	666	318	894	818	933	360	1028	521	981	1245	638	46967
	99.14%	98.29%	99.65%	99.79%	99.64%	99.46%	99.83%	98.74%	99.17%	99.80%	99.40%	100.00%	99.44%	99.39%	96.68%	95.56%	89.11%	94.82%	94.19%	90.12%	89.18%	98.59%
Snow Leopard-Leopard	4	5	1	2	1	1	2	11	57	1	2	0	3	5	31	14	108	32	80	131	58	692
	346	292	289	471	550	370	594	874	6104	512	668	318	896	818	933	362	1032	516	958	1237	649	46936
	98.84%	98.29%	99.65%	99.58%	99.82%	99.73%	99.66%	98.74%	99.07%	99.80%	99.70%	100.00%	99.67%	99.39%	96.68%	96.13%	89.53%	93.80%	91.65%	89.41%	91.06%	98.53%
Snow Leopard-Tiger	3	5	1	1	1	0	2	1	61	2	2	0	4	2	38	17	130	49	100	144	102	779
	347	292	289	472	550	371	594	884	6100	511	668	318	895	821	926	359	1010	499	938	1224	605	46849
	99.14%	98.29%	99.65%	99.79%	99.82%	100.00%	99.66%	99.89%	99.00%	99.61%	99.70%	100.00%	99.55%	99.76%	95.90%	95.26%	87.13%	90.18%	89.34%	88.24%	83.14%	98.34%
Snow Leopard-Jaguar	4	2	1	1	1	4	5	11	51	2	2	1	5	3	31	15	133	36	87	130	74	750
	346	295	289	472	550	367	591	874	6110	511	668	317	894	820	933	361	1007	512	951	1238	633	46878
	98.84%	99.32%	99.65%	99.79%	99.82%	98.91%	99.15%	98.74%	99.17%	99.61%	99.70%	99.68%	99.44%	99.63%	96.68%	95.84%	86.79%	92.97%	90.85%	89.50%	88.31%	98.40%
Clouded Leopard-Lion	6	5	1	2	4	1	2	7	75	4	4	1	9	6	40	34	134	56	87	180	90	1066
	344	292	289	471	547	370	594	878	6086	509	666	317	890	817	924	342	1006	492	951	1188	617	46562
	98.26%	98.29%	99.65%	99.58%	99.27%	99.73%	99.66%	99.20%	98.77%	99.21%	99.40%	99.68%	98.99%	99.27%	95.67%	90.06%	86.68%	88.62%	90.85%	84.85%	85.41%	97.71%
Clouded Leopard-Leopard	7	5	1	3	3	0	3	7	81	4	2	1	7	8	34	30	145	54	131	175	93	1116
	343	292	289	470	548	371	593	878	6080	509	668	317	892	815	930	346	995	494	907	1193	614	46512
	97.96%	98.29%	99.65%	99.36%	99.45%	100.00%	99.49%	99.20%	98.67%	99.21%	99.70%	99.68%	99.22%	99.02%	96.34%	91.33%	85.43%	89.07%	85.56%	85.33%	84.85%	97.60%
Clouded Leopard-Tiger	6	5	1	2	3	1	3	9	92	5	2	1	10	5	41	30	135	62	151	182	116	1185
	344	292	289	471	548	370	593	876	6069	508	668	317	889	818	923	346	1005	486	887	1186	591	46443
	98.26%	98.29%	99.65%	99.58%	99.45%	99.73%	99.49%	98.97%	98.48%	99.02%	99.70%	99.68%	98.88%	99.39%	95.56%	91.33%	86.57%	87.24%	82.98%	84.65%	80.37%	97.45%
Clouded Leopard-Jaguar	7	2	1	2	3	3	5	7	69	5	2	2	9	4	43	31	149	56	151	184	87	1150
	343	295	289	471	548	368	591	878	6092	508	668	316	890	819	921	345	991	492	887	1184	620	46478
	97.96%	99.32%	99.65%	99.58%	99.45%	99.18%	99.15%	99.20%	98.87%	99.02%	99.70%	99.37%	98.99%	99.51%	95.33%	91.01%	84.96%	88.62%	82.98%	84.46%	85.97%	97.53%
Clouded Leopard-Snow Leopard	3	2	0	3	4	1	3	10	72	5	4	1	12	5	36	33	143	51	140	167	97	1118
	347	295	290	470	547	370	593	875	6089	508	666	317	887	818	928	343	997	497	898	1201	610	46510
	99.14%	99.32%	100.00%	99.36%	99.27%	99.73%	99.49%	98.86%	98.82%	99.02%	99.40%	99.68%	98.65%	99.39%	96.12%	90.38%	85.66%	89.74%	84.41%	86.09%	84.10%	97.60%
Maximum	99.71%	99.32%	100.00%	100.00%	100.00%	100.00%	99.83%	100.00%	99.48%	100.00%	100.00%	100.00%	99.89%	99.76%	97.34%	96.99%	90.60%	94.82%	96.20%	92.96%	93.52%	99.00%
Minimum	97.96%	98.29%	99.65%	99.36%	99.27%	98.91%	99.15%	98.74%	98.48%	99.02%	99.40%	99.37%	98.65%	99.02%	95.10%	90.06%	84.96%	87.24%	82.98%	84.46%	80.37%	97.45%
Range	1.75%	1.03%	0.35%	0.64%	0.73%	1.09%	0.68%	1.26%	0.99%	0.98%	0.60%	0.63%	1.24%	0.74%	2.23%	6.93%	5.63%	7.57%	13.22%	8.50%	13.15%	1.55%
Average Dissimilarity	1.139%	1.298%	0.185%	0.283%	0.304%	0.507%	0.473%	0.792%	1.032%	0.458%	0.300%	0.210%	0.591%	0.587%	3.883%	5.774%	12.498%	9.485%	10.072%	12.364%	13.302%	1.807%
Average Dissimilarity (No Clouded)	0.865%	1.298%	0.138%	0.170%	0.146%	0.598%	0.439%	0.732%	0.908%	0.235%	0.240%	0.126%	0.357%	0.538%	3.727%	4.072%	11.677%	8.556%	7.785%	11.085%	12.024%	1.498%
Standard Deviation	0.525%	0.455%	0.179%	0.208%	0.255%	0.384%	0.233%	0.503%	0.295%	0.347%	0.197%	0.195%	0.382%	0.224%	0.638%	2.566%	1.575%	2.300%	4.272%	2.416%	3.433%	0.495%
Standard Deviation (No Clouded)	0.273%	0.424%	0.179%	0.134%	0.115%	0.403%	0.255%	0.608%	0.272%	0.124%	0.190%	0.163%	0.165%	0.233%	0.689%	0.665%	1.168%	2.231%	2.498%	1.853%	3.271%	0.246%

Appendix Table 6. Calculations for the percentage of parsimony contribution of each gene segment based on the relative sequence size in the partition and in the supermatrix.

Gene	Length	% Total Length	% Partition Length	% Informative Sites (Total)	% Informative Sites (Partition)	% Total Sites / % Seq Length	% Partiton Sites / % Seq Length
DDX3Y	1,428	3.00%	7.46%	0.47%	8.33%	0.1561	1.1169
EIF1AY	1,314	2.76%	6.87%	0.16%	2.78%	0.0565	0.4046
EIF2S3Y	817	1.72%	4.27%	0.00%	0.00%	0.0000	0.0000
SMCY	3,793	7.96%	19.82%	0.78%	13.89%	0.0979	0.7009
UBE1Y	977	2.05%	5.10%	0.78%	13.89%	0.3803	2.7209
USP9Y	6,135	12.88%	32.05%	2.18%	38.89%	0.1696	1.2133
UTY	1,575	3.31%	8.23%	0.78%	13.89%	0.2359	1.6878
SRY	2,371	4.98%	12.39%	0.47%	8.33%	0.0940	0.6727
ZFY	730	1.53%	3.81%	0.00%	0.00%	0.0000	0.0000
FGB	678	1.42%	3.54%	1.72%	11.70%	1.2055	3.3066
IRBP	1,254	2.63%	6.55%	0.31%	2.13%	0.1185	0.3251
APP	642	1.35%	3.35%	0.16%	1.06%	0.1157	0.3175
CALB	817	1.72%	4.26%	0.16%	1.06%	0.0909	0.2495
CHRNA	326	0.68%	1.70%	0.16%	1.06%	0.2279	0.6252
CLU	1,369	2.87%	7.15%	0.62%	4.26%	0.2171	0.5955
CMA	518	1.09%	2.70%	0.47%	3.19%	0.4303	1.1804
DGKG2	710	1.49%	3.71%	1.09%	7.45%	0.7326	2.0094
FES	413	0.87%	2.16%	0.47%	3.19%	0.5397	1.4804
GATA	457	0.96%	2.39%	0.16%	1.06%	0.1626	0.4460
GHR	670	1.41%	3.50%	0.16%	1.06%	0.1109	0.3042
GNAZ	622	1.31%	3.25%	0.00%	0.00%	0.0000	0.0000
GNB	674	1.42%	3.52%	0.47%	3.19%	0.3307	0.9072
HK1	350	0.73%	1.83%	0.47%	3.19%	0.6369	1.7469
NCL	297	0.62%	1.55%	0.62%	4.26%	1.0007	2.7449
PNOC	290	0.61%	1.51%	0.16%	1.06%	0.2562	0.7028
RAG2	473	0.99%	2.47%	0.00%	0.00%	0.0000	0.0000
RASA	551	1.16%	2.88%	0.00%	0.00%	0.0000	0.0000
SILV	371	0.78%	1.94%	0.16%	1.06%	0.2003	0.5493
TCP	596	1.25%	3.11%	0.00%	0.00%	0.0000	0.0000
TTR	885	1.86%	4.62%	1.56%	10.64%	0.8396	2.3029
CES7	6,161	12.94%	32.16%	5.77%	39.36%	0.4462	1.2240
ALAS	513	1.08%	15.92%	0.00%	0.00%	0.0000	0.0000
ATP7A	670	1.41%	20.79%	0.00%	0.00%	0.0000	0.0000
IL2RG	318	0.67%	9.87%	0.00%	0.00%	0.0000	0.0000
PLP	899	1.89%	27.89%	0.16%	33.33%	0.0827	1.1950
ZFX	823	1.73%	25.54%	0.31%	66.67%	0.1806	2.6108
12S	964	2.02%	15.70%	5.15%	6.50%	2.5436	0.4138
16S	376	0.79%	6.12%	2.03%	2.56%	2.5690	0.4179
CYTB	1,140	2.39%	18.57%	18.41%	23.23%	7.6910	1.2511
ND1	548	1.15%	8.93%	6.71%	8.46%	5.8303	0.9484
ND2	1,038	2.18%	16.91%	12.95%	16.34%	5.9413	0.9665
ND4	1,368	2.87%	22.28%	22.62%	28.54%	7.8756	1.2811
ND5	707	1.48%	11.51%	11.39%	14.37%	7.6720	1.2480

Appendix Table 7. Gene jacknife bootstrap bipartition support for the complete supermatrix

Supermatrix					
	(Lion-Leopard-Jaguar) (Tiger-Snow-Clouded)	Lion-Leopard	Tiger-Snow	Tiger-Clouded	Lion-Jaguar
ALL	100.0	99.9	93.7	6.0	0.0
DDX3Y	100.0	99.9	92.3	7.5	0.0
EIF1AY	100.0	99.7	92.2	7.8	0.0
EIF2S3Y	100.0	99.0	93.0	7.0	0.0
SMCY	100.0	99.1	88.0	12.0	5.1
UBE1Y	100.0	99.8	95.7	0.0	0.0
USP9Y	100.0	99.5	84.0	15.8	0.0
UTY	100.0	99.9	85.9	13.9	0.0
SRY	100.0	99.9	92.9	6.9	0.0
ZFY	100.0	99.6	94.4	5.5	0.0
FGB	100.0	99.9	90.0	9.9	0.0
IRBP	100.0	93.3	99.6	6.7	0.0
APP	100.0	99.7	94.9	0.0	0.0
CALB	100.0	98.9	92.8	7.1	0.0
CES7	100.0	97.5	89.7	10.3	0.0
CHRNA	100.0	99.6	93.7	6.1	0.0
CLU	100.0	99.9	94.1	5.7	0.0
CMA	100.0	99.7	94.5	5.4	0.0
DGKG2	100.0	99.9	96.4	0.0	0.0
FES	100.0	99.4	92.2	7.6	0.0
GATA	100.0	99.9	93.8	6.0	0.0
GHR	100.0	99.7	93.2	6.8	0.0
GNAZ	100.0	99.6	94.2	5.5	0.0
GNB	100.0	99.6	93.8	6.2	0.0
HK1	100.0	99.7	94.5	5.4	0.0
NCL	100.0	99.6	93.8	6.2	0.0
PNOC	100.0	99.4	94.4	5.5	0.0
RAG2	100.0	99.6	92.8	7.0	0.0
RASA	100.0	99.5	92.3	7.6	0.0
SILV	100.0	99.6	92.7	7.2	0.0
TCP	100.0	99.9	92.2	7.6	0.0
TTR	100.0	99.8	83.7	15.4	0.0
ALAS	100.0	99.7	93.2	6.8	0.0
ATP7A	100.0	99.8	94.2	5.8	0.0
IL2RG	100.0	99.6	93.8	6.1	0.0
PLP	100.0	99.8	94.2	7.3	0.0
ZFX	100.0	99.7	93.5	6.4	0.0
12S	100.0	99.4	97.0	0.0	5.5
16S	100.0	99.9	94.1	5.9	0.0
CYTB	100.0	99.6	98.0	0.0	0.0
ND1	100.0	99.5	93.6	6.3	0.0
ND2	100.0	100.0	95.3	0.0	0.0
ND4	100.0	96.2	92.2	7.8	37.7
ND5	100.0	99.4	95.4	0.0	7.4

Appendix Table 8. Gene jacknife bootstrap bipartition support for the autosomal partition

Autosomal					
	(Lion-Leopard-Jaguar) (Tiger-Snow-Clouded)	Lion-Leopard	Tiger-Snow	Lion-Jaguar	Snow-Clouded
ALL	99.8	66.6	93.4	33.4	6.1
FGB	100.0	83.7	96.3	16.3	0.0
IRBP	99.9	58.6	95.5	41.4	0.0
APP	100.0	76.3	95.5	23.7	0.0
CALB	99.8	58.2	92.5	41.8	7.4
CES7	90.3	5.1	69.5	94.9	30.0
CHRNA	99.9	69.7	95.3	30.3	0.0
CLU	99.9	73.8	93.4	26.1	6.5
CMA	99.7	56.2	95.8	43.8	0.0
DGKG2	100.0	80.4	95.4	19.6	0.0
FES	99.7	72.4	90.9	27.6	8.8
GATA	99.8	65.6	95.7	34.4	0.0
GHR	99.9	63.4	93.4	36.6	6.3
GNAZ	99.9	65.7	92.9	34.3	7.0
GNB	99.9	60.0	94.2	40.0	5.7
HK1	99.7	73.6	97.5	26.3	0.0
NCL	100.0	60.1	95.7	39.9	0.0
PNOC	99.8	67.8	95.1	32.2	0.0
RAG2	100.0	65.7	94.4	34.3	5.4
RASA	99.9	62.5	93.8	37.5	5.8
SILV	99.8	64.8	93.5	35.2	6.1
TCP	99.9	68.5	94.7	31.5	5.3
TTR	97.9	59.7	61.2	40.1	35.6
ALL	99.8	66.6	93.4	33.4	6.1

Appendix Table 9. Gene jacknife bootstrap bipartition support for the X chromosome partition

X Chromosome			
	(Tiger-Snow)	(Lion-Jaguar)	(Lion-Jaguar-Clouded)
ALL	64.3	61.1	66.0
ALAS	63.6	62.3	61.9
ATP7A	62.3	62.7	64.1
IL2RG	62.3	62.7	63.8
PLP	0.0	62.7	62.6
ZFX	63.1	0.0	0.0
ALL	64.3	61.1	66.0

Appendix Table 10. Gene jacknife bootstrap bipartition support for the Y chromosome partition

Y Chromosome	(Lion-Leopard-Jaguar) (Tiger-Snow-Clouded)	Lion-Leopard	Tiger-Snow
ALL	99.9	70.0	100.0
DDX3Y	100.0	68.5	99.9
EIF1AY	100.0	72.0	99.6
EIF2S3Y	100.0	71.2	100.0
SMCY	100.0	0.0	99.8
UBE1Y	99.9	72.5	100.0
USP9Y	99.4	77.3	97.3
UTY	99.9	75.0	99.7
SRY	100.0	75.4	99.9
ZFY	100.0	70.7	99.8

Appendix Table 11. Gene jacknife bootstrap bipartition support for the mitochondrial partition

Mitochondria							
	(Lion-Leopard-Jaguar) (Tiger-Snow-Clouded)	Lion-Leopard	Tiger-Snow	Snow-Clouded	Jaguar-Snow	(Lion-Snow)	(Lion-Leopard-Snow) (Tiger-Jaguar-Clouded)
ALL	71.2	99.8	94.3	5.0	12.9	0.0	6.5
12S	65.4	99.3	85.9	12.7	10.1	0.0	6.2
16S	71.1	99.6	93.4	6.5	11.7	0.0	5.5
CYTB	72.4	99.9	88.8	10.4	18.0	0.0	0.0
ND1	76.2	99.9	93.9	5.7	8.6	0.0	0.0
ND2	68.2	100.0	87.2	11.9	7.6	0.0	14.8
ND4	16.3	70.1	90.5	0.0	23.6	21.7	33.2
ND5	81.3	99.7	93.6	5.8	8.5	0.0	0.0

Appendix Table 12. Gene jacknife bootstrap bipartition support for the autosomal partition without CES7

Autosomal - NoCES7						
	(Lion-Leopard-Jaguar) (Tiger-Snow-Clouded)	Lion-Leopard	Tiger-Snow	Lion-Jaguar	Snow-Clouded	Leopard-Tiger
ALL	90.3	5.1	69.5	94.9	30.0	8.9
FGB	84.0	19.1	66.0	78.3	33.9	14.4
IRBP	90.1	0.0	72.9	97.1	26.2	8.8
TTR	44.2	0.0	8.3	96.0	86.2	54.4
APP	90.6	9.9	69.9	89.1	30.0	8.2
CALB	87.0	0.0	64.4	96.8	35.4	11.4
CHRNA	90.5	6.9	66.6	92.0	33.0	8.6
CLU	90.2	10.4	60.5	88.3	39.2	9.1
CMA	91.2	0.0	75.6	97.3	23.5	7.9
DGKG2	96.9	12.6	72.7	86.8	27.2	0.0
FES	94.7	7.4	63.2	92.2	36.6	0.0
GATA	90.7	5.0	72.1	94.6	27.2	7.8
GHR	91.1	5.5	67.5	94.4	31.9	7.8
GNAZ	90.1	6.4	69.0	93.0	30.7	8.7
GNB	85.4	0.0	68.6	96.5	30.5	12.8
HK1	92.4	8.7	78.3	90.7	20.6	6.8
NCL	94.3	0.0	78.7	96.1	20.9	0.0
PNOC	90.9	6.0	74.1	93.7	25.4	7.9
RAG2	90.2	7.7	67.8	92.0	31.6	8.9
RASA	90.7	5.3	65.3	94.4	34.5	8.8
SILV	89.0	5.2	58.1	94.2	41.2	10.0
TCP	90.8	7.0	66.8	92.8	32.3	8.0

Appendix Table 13. Parsimony informative sites for each species relationship within each gene segment for the Y chromosome.

Gene	DDX3Y	EIF1AY	EIF2S3Y	SMCY	UBE1Y	USP9Y	UTY	SRY	ZFY
Length	1428	1314	817	3793	977	6135	1575	2371	730
% Total Length	3.00%	2.76%	1.72%	7.96%	2.05%	12.88%	3.31%	4.98%	1.53%
% Partition Length	7.46%	6.87%	4.27%	19.82%	5.10%	32.05%	8.23%	12.39%	3.81%
% Informative Sites (Total)	0.47%	0.16%	0.00%	0.78%	0.78%	2.18%	0.78%	0.47%	0.00%
% Informative Sites (Partition)	8.33%	2.78%	0.00%	13.89%	13.89%	38.89%	13.89%	8.33%	0.00%
% Total Sites / % Seq Length	0.16	0.06	0.00	0.10	0.38	0.17	0.24	0.09	0.00
% Partiton Sites / % Seq Length	1.12	0.40	0.00	0.70	2.72	1.21	1.69	0.67	0.00
(Lion, Leopard)				1					
(Jaguar, Tiger, Snow)				1					
(Tiger, Snow)	2	1		3	3	14	5	2	
(Lion, Leopard, Jaguar)	2	1		3	3	13	5	2	
(Lion, Jaguar)									
(Leopard, Tiger, Snow)								1	
(Leopard, Tiger)									
(Lion, Jaguar, Snow)									
(Leopard, Jaguar)									
(Lion, Tiger, Snow)									
(Jaguar, Snow)									
(Lion, Leopard, Tiger)									
(Lion, Tiger)									
(Leopard, Jaguar, Snow)									
(Leopard, Snow)									
(Lion, Tiger, Jaguar)									
(Lion, Leopard, Tiger, Jaguar)	1				1				
(Lion, Leopard, Tiger Snow)				1					
(Lion, Leopard, Jaguar, Snow)					1	1			
(Leopard, Tiger, Jaguar, Snow)									
(Lion, Tiger, Jaguar, Snow)									
(Tiger-Jaguar)									
(Lion, Leopard, Snow)									
Lion-Snow									
(Leopard-Tiger-Jaguar)									

Appendix Table 14. Parsimony informative sites for each species relationship within each gene segment for the autosomes.

Gene	FGB	IRBP	APP	CALB	CHRNA	CLU	CMA	DGKG2	FES	GATA	GHR
Length	678	1254	642	817	326	1369	518	710	413	457	670
% Total Length	1.42%	2.63%	1.35%	1.72%	0.68%	2.87%	1.09%	1.49%	0.87%	0.96%	1.41%
% Partition Length	3.54%	6.55%	3.35%	4.26%	1.70%	7.15%	2.70%	3.71%	2.16%	2.39%	3.50%
% Informative Sites (Total)	1.72%	0.31%	0.16%	0.16%	0.16%	0.62%	0.47%	1.09%	0.47%	0.16%	0.16%
% Informative Sites (Partition)	11.70%	2.13%	1.06%	1.06%	1.06%	4.26%	3.19%	7.45%	3.19%	1.06%	1.06%
% Total Sites / % Seq Length	1.21	0.12	0.12	0.09	0.23	0.22	0.43	0.73	0.54	0.16	0.11
% Partiton Sites / % Seq Length	3.31	0.33	0.32	0.25	0.63	0.60	1.18	2.01	1.48	0.45	0.30
(Lion, Leopard)		1		1			1				
(Jaguar, Tiger, Snow)				1			1				
(Tiger, Snow)	2					1					
(Lion, Leopard, Jaguar)	2					1		1			
(Lion, Jaguar)	1		1			2		1	1		
(Leopard, Tiger, Snow)	1		1			2		1	1		
(Leopard, Tiger)	1				1			4	1		
(Lion, Jaguar, Snow)								4	1		
(Leopard, Jaguar)					1				1		
(Lion, Tiger, Snow)						1			1		
(Jaguar, Snow)							1	1			
(Lion, Leopard, Tiger)							1	1			
(Lion, Tiger)											
(Leopard, Jaguar, Snow)											
(Leopard, Snow)											
(Lion, Tiger, Jaguar)											
(Lion, Leopard, Tiger, Jaguar)		1								1	
(Lion, Leopard, Tiger Snow)							1				1
(Lion, Leopard, Jaguar, Snow)											
(Leopard, Tiger, Jaguar, Snow)	1										
(Lion, Tiger, Jaguar, Snow)											
(Tiger-Jaguar)											
(Lion, Leopard, Snow)											
Lion-Snow											
(Leopard-Tiger-Jaguar)											

Appendix Table 14. Continued

Gene	GNAZ	GNB	HK1	NCL	PNOC	RAG2	RASA	SILV	TCP	TTR	CES7
Length	622	674	350	297	290	473	551	371	596	885	6161
% Total Length	1.31%	1.42%	0.73%	0.62%	0.61%	0.99%	1.16%	0.78%	1.25%	1.86%	12.94%
% Partition Length	3.25%	3.52%	1.83%	1.55%	1.51%	2.47%	2.88%	1.94%	3.11%	4.62%	32.16%
% Informative Sites (Total)	0.00%	0.47%	0.47%	0.62%	0.16%	0.00%	0.00%	0.16%	0.00%	1.56%	5.77%
% Informative Sites (Partition)	0.00%	3.19%	3.19%	4.26%	1.06%	0.00%	0.00%	1.06%	0.00%	10.64%	39.36%
% Total Sites / % Seq Length	0.00	0.33	0.64	1.00	0.26	0.00	0.00	0.20	0.00	0.84	0.45
% Partiton Sites / % Seq Length	0.00	0.91	1.75	2.74	0.70	0.00	0.00	0.55	0.00	2.30	1.22
(Lion, Leopard)		1									8
(Jaguar, Tiger, Snow)		1									8
(Tiger, Snow)		1		1				1		10	12
(Lion, Leopard, Jaguar)		1						1		10	12
(Lion, Jaguar)			1								2
(Leopard, Tiger, Snow)			1								2
(Leopard, Tiger)				1							
(Lion, Jaguar, Snow)				1							
(Leopard, Jaguar)				1							2
(Lion, Tiger, Snow)											2
(Jaguar, Snow)				2							
(Lion, Leopard, Tiger)				2							
(Lion, Tiger)											2
(Leopard, Jaguar, Snow)											2
(Leopard, Snow)											2
(Lion, Tiger, Jaguar)											2
(Lion, Leopard, Tiger, Jaguar)		1	2		1						2
(Lion, Leopard, Tiger Snow)											2
(Lion, Leopard, Jaguar, Snow)											3
(Leopard, Tiger, Jaguar, Snow)											1
(Lion, Tiger, Jaguar, Snow)											1
(Tiger-Jaguar)											
(Lion, Leopard, Snow)											
Lion-Snow											
(Leopard-Tiger-Jaguar)											

Appendix Table 15. Parsimony informative sites for each species relationship within each gene segment for the X chromosome and mitochondrial partitions.

Gene	ALAS	ATP7A	IL2RG	PLP	ZFX	12S	16S	CYTB	ND1	ND2	ND4	ND5
Length	513	670	318	899	823	964	376	1140	548	1038	1368	707
% Total Length	1.08%	1.41%	0.67%	1.89%	1.73%	2.02%	0.79%	2.39%	1.15%	2.18%	2.87%	1.48%
% Partition Length	15.92%	20.79%	9.87%	27.89%	25.54%	15.70%	6.12%	18.57%	8.93%	16.91%	22.28%	11.51%
% Informative Sites (Total)	0.00%	0.00%	0.00%	0.16%	0.31%	5.15%	2.03%	18.41%	6.71%	12.95%	22.62%	11.39%
% Informative Sites (Partition)	0.00%	0.00%	0.00%	33.33%	66.67%	6.50%	2.56%	23.23%	8.46%	16.34%	28.54%	14.37%
% Total Sites / % Seq Length	0.00	0.00	0.00	0.08	0.18	2.54	2.57	7.69	5.83	5.94	7.88	7.67
% Partiton Sites / % Seq Length	0.00	0.00	0.00	1.20	2.61	0.41	0.42	1.25	0.95	0.97	1.28	1.25
(Lion, Leopard)						5	1	12	3	4	20	4
(Jaguar, Tiger, Snow)						5	1	10	3	3	19	3
(Tiger, Snow)				1		5	2	7	2	12	13	5
(Lion, Leopard, Jaguar)				1		4	2	7	2	10	13	5
(Lion, Jaguar)					2	1	1	7	1	4	5	4
(Leopard, Tiger, Snow)					2		1	7	1	7	6	2
(Leopard, Tiger)						2	1	9	5	11	9	3
(Lion, Jaguar, Snow)						1	1	9	5	3	9	3
(Leopard, Jaguar)						1	1	9	2	2	9	5
(Lion, Tiger, Snow)						1	1	8	2		8	5
(Jaguar, Snow)						2		6	2	13	15	9
(Lion, Leopard, Tiger)						2		6	2	6	15	8
(Lion, Tiger)						1		7	1	2	6	3
(Leopard, Jaguar, Snow)						1		7	1	6	6	3
(Leopard, Snow)						2	1	12	1	4	6	2
(Lion, Tiger, Jaguar)						2	2	12	1	2	6	2
(Lion, Leopard, Tiger, Jaguar)								7		8	11	6
(Lion, Leopard, Tiger Snow)							1	10	2	1	6	6
(Lion, Leopard, Jaguar, Snow)						5	1	7	7	7	16	1
(Leopard, Tiger, Jaguar, Snow)						1		5	2	3	3	5
(Lion, Tiger, Jaguar, Snow)						2	1	5	1	5	9	2
(Tiger-Jaguar)						3		14	6	8	11	11
(Lion, Leopard, Snow)						3	2	13	6	6	11	12
Lion-Snow						3		6	5	4	3	
(Leopard-Tiger-Jaguar)						3		5	5	6	2	

VITA

Name:	Brian William Davis
Address:	Veterinary Integrative Biosciences Texas A&M University TAMU Mailstop 4458 College Station, TX 77843-4458
Email Address:	bdavis@cvm.tamu.edu
Education:	M.S., Biomedical Sciences Texas A&M University, College Station, TX, 2009 B.S., Genetics and Biochemistry Texas A&M University, College Station, TX, 2006 B.S., Molecular and Cellular Biology, Texas A&M University, College Station, TX, 2006