Avian Raptor Evolution

A thesis presented in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

in

Genetics

Institute of Fundamental Sciences,
Massey University, New Zealand

Muhammad Tariq Mahmood

2015

ABSTRACT

Despite decades of research using a variety of data sources (such as morphological, paleontological, immunological, DNA hybridization and short DNA sequences) both the relationships between modern avian orders of raptors and their times of origin remain uncertain. This prior work is discussed in the opening introductory chapter. In order to address these issues, the second chapter reports a study that I undertook to develop a database that would have all sequence data from avian raptors (although it could easily be modified for other groups as well). Complete mitochondrial (mt) genomes have been used extensively to help study evolution, and I sequenced seven new bird mt genomes: from owls, secretary bird, falcons, and eagles in order to provide improved taxon sampling for the avian raptors. Adding three of these taxa to the avian mt genome dataset aids in resolving deep bird phylogeny and strongly supports the independent origin of the raptor life-style – so there is now agreement between nuclear and mitochondrial sequences. The final four newer genomes were then added in for a more detailed analysis of raptor relationships, and good progress is made on this. We had issues with one of the final four newer genomes where the sample was mis-identified. Therefore, it is emphasized that it is always important to keep good reference samples to check identifications. For the fifth chapter I undertook a population level study of spotted owlet (*Athene brama*) a nocturnal raptor from Pakistan, and so my study includes the use of sequence data to study both micro- and macro-evolution. The final chapter is a current overview of where we are at now in relation to avian raptor evolution.

ACKNOWLEDGEMENTS

This thesis is for my family including my mother, father, wife and my brave son (Ali). My sisters and brother have been a continuous source of unconditional support throughout this journey. Thank you for your support, your encouragement and, above all, your patience!

Sincere appreciation and a big deal of thanks to my supervisor David Penny! You are not only an amazing supervisor, but also a very good human being. I would always remember your ability to return the draft overnight. Hats off! I wish I could ever have as much wide ranging knowledge as you have.

I would like to express my gratitude to my co-supervisors: Dr Gillian Gibb (Institute of Agriculture and Environment), Professor Michael Wink (University of Heidelberg) and Dr Barbara Holland of the Mathematics from the University of Tasmania (Hobart).

A big thanks to my other collaborators, especially P. A. (Trish) McLenachan (Institute of Fundamental Sciences) for assistance with DNA extraction, sequencing, and most of all “It will be fine !”. Many thanks to Pete Lockhart for being around with a beautiful smile, and ready to provide a reference letter when I needed it. Olga Kardailsky deserves big thanks for her help in getting myself started as a wet lab worker, when I first came to Massey. My sincere thanks to Jing Wang and Ayesha Hakim, who helped me to start the database project, and Mauro Truglio for assistance with the database. No doubt, the Institute of Fundamental Sciences administration staff as a whole, and Ann Truter, Cynthia Cresswell deserve a sincere thanks for being as much helpful as they could.

Thank you Dr Lesley Collins, Dr Patrick Biggs and Dr Bennet McComish for your help at next generation sequencing assembly and bioinformatics support. Constructive feedback on my research proposal and overall project, by Dr Mary Morgan-Richards and Dr Steve Trewick was much helpful. Thanks you. I should not miss Dr Bojian Zhong, who has been keen to help me through evolutionary questions. Thank you!

Dr Juan-Carlos Garcia Ramirez, a special thanks for your sincere help in making sense out of population genetics. I hope it goes a long way! Many thanks to Ahmed Elwan for help with the maps.

My thanks to all the Boffin lounge inhabitants including Dr Ibrar Ahmed, Dr Simon Hills, Dr Robin Atherton and Dr Tim White for being in memories of my stay at Massey. Dr Barbara Schoenfeld and Dr Bennett McComish, I still remember good food and drinks at your place and no doubt your 'chickens!'

I would like to express my thanks to my friends Waheed Ahmed and Hina Cheema for their beautiful company and friendship. Abdul Jabbar and Munazza Saeed, your amazing formatting skills and sincerity are much appreciated.

The work carried out in this study was supported by a Higher Education Commission of Pakistan's scholarship – it has been very much appreciated, and to the Institute of Fundamental Sciences for an extension to that grant. I should not forget the generous travel grant from the Claude McCarthy scholarship committee. I appreciate your support for the research. I would also like to thank all my Pakistani friends at Massey, your help and friendship has been very much appreciated.

Dedicated To My Parents

TABLE OF CONTENTS

1 INTRODUCTION

1.1 Background

The avian raptors are a group of birds noted for the their capture and consumption of live prey, and they include falcons (Falconidae), eagles (Accipitridae), New World (Cathartidae) and Old World (Aegypiinae and Gypaetinae) vultures, as well as the owls (Strigiformes). Sharing the apex of the food chain with humans makes raptors invaluable for research on biomagnifications (when the concentration of a pollutant substance in an organism's body builds up from the same substance in its diet) (Wink, 2007). As ecologically susceptible predators, birds of prey are valuable indicators of environment quality (Lerner and Mindell, 2005). Moreover, birds of prey are often used as national and cultural mascots, and are of great interest to the community, especially related to bird watching and falconry. Many people simply want to know more on the subject of these 'cool' birds and are eager to buy books and watch videos on raptors. This public interest for knowledge about raptors provides researchers with a large audience. Certainly the enthrallment the public has for raptors has led to improved laws to protect and conserve them. These laws in turn have need of knowledge of raptors to guide the execution of conservation and management of programs (Wink, 2007).

Traditionally, the study of phylogeny and systematics of birds, as well as other organisms, is based on anatomical and morphological characters. Often behavioral, ecological and acoustical or geographical data are also included in the analysis. Since the main criterion is similarity, and because of the adaptive traits, convergence may sometimes obscure the picture (Wink et al. 2009). Diurnal raptors have traditionally been placed either in the order Falconiformes, or else in the infraorders Falconides and Ciconiides (Sibley and Monroe 1990) and grouped into five families; Accipitridae, Pandionidae, Sagittaridae, Falconidae, and Cathartidae. Many researchers (Wink 1995, Wink et al. 1998; Fain and Houde, 2004; and Ericson et al. 2006) have attempted to resolve whether or not the Falconiformes is a monophyletic group. However, it appears

from morphological and molecular data that Falconidae may not share direct ancestry with other diurnal raptors families (White et al. 1994)

Phylogenetic studies had a breakthrough with the advent of biochemical and molecular methods like DNA-DNA hybridization, protein electrophoresis, DNA restriction analysis (RFLP), or marker genes amplified through PCR (polymerase chain reaction) followed by DNA sequencing (Beebee & Rowe, 2004; Sibley and Ahlquist, 1990). Some of the important early molecular methods are shown in Table 1. Particularly the nuclear sequences analyzed by powerful computer programs such as PHYLIP (Felsenstein, 1993), PAUP* (Swofford, 2002) and MEGA (Tamura et al. 2007), has assisted the phylogenetic reconstruction in all kingdoms of life. The molecular approach is complementary, and it does not make the traditional analysis obsolete. The right evolutionary investigation can only be made if there is a solid framework based on acoustics, behavior, morphology and geography (Wink et al. 2009)

Table 1. Important methods of molecular biology that are useful in evolutionary and phylogeographic studies (adapted from Wink, 2007)

Method[a]	DNA Type	Adequate for Studying
Sequencing	mtDNA[a], nDNA[a]	Phylogeny, taxonomy, phylogeography
STR-Analysis[a]	Microsatellites	Population genetics, tracing of individuals, paternity, and pedigree
SNP-Analysis[a]	Point mutations in all genes	Population genetics, tracing of individuals, paternity, and pedigree
AFLP[a]	Nuclear genome	Population genetics, gene mapping
ISSR[a]	Nuclear genome	Phylogeny, population genetics, gene mapping, hybridization
DNA fingerprinting	Satellite DNA (VNTR, STR)	Paternity, tracing of individuals
Sexing	Sex chromosome	Molecular sexing

[a]STR = short tandem repeats; SNP = single nucleotide polymorphism; VNTR = variable number tandem repeats; mtDNA = mitochondrial DNA, nDNA = nuclear DNA; AFLP = amplified fragment length polymorphisms; ISSR = inter-simple sequence repeats.

Most of the recent molecular studies on birds have mitochondrial DNA (mtDNA) as their central focus (Mindell 1997), and the reason for this is that in birds mtDNA evolution is faster than nuclear DNA (nDNA) evolution. However, there is also a recent

emphasis on nuclear sequences as well (Suh et al. 2011; McCormack et al. 2013; Kimball et al. 2013). Many early studies used only the mitochondrial cytochrome b gene, because of the advantage that insertions, deletions or inversions are usually absent. So the sequence alignment is not problematic as compared to ribosomal genes, which are often used as markers. Wink et al. (2009) have recently studied the phylogenetics of owls using cytochrome b (Cyt *b*) and Recombination activating protein gene (RAG1).They are of the opinion that when using Cyt *b* as marker (also the other protein coding genes, for instance ND2, COI), resolution is lost when it comes to the divergences that happened more than 20 Mya (million years ago). The reason for this is that relatively quickly evolving mtDNA markers may have multiple nucleotide substitutions at the same position, which may result in undetected homoplasy. While mtDNA markers are good at resolving relationships in relatively recent timescales, more slowly evolving nuclear gene sequences are very helpful for resolution of deeper nodes. RAG1 is a nuclear gene that has been used in several vertebrate phylogenetic studies. So genes differ in the questions/time periods for which they are useful. Table 2 shows the relative rate of change for several mitochondrial genes used in phylogenetic studies, particularly in birds.

Table 2. Composition of mitochondrial DNA (adapted from Wink, 2007)

DNA	Number of Elements	Substitution Rate
16S rRNA	1	Low
12S rRNA	1	Low
tRNA	22	Low
Cytochrome b	1	Medium
Cytochrome oxidase (CO), subunits I-III	3	Medium
NADH dehydrogenase (ND), subunits I-VII	7	Medium
ATP synthase, subunits a, b	2	Medium
D-loop	1	High

1.2 DNA methods for Raptor systematics

One of the basic disciplines in biology is the classification of plants and animals, and the *Systema Naturae* by Linnaeus in 1753 proved to be a landmark in this field. Since that

time morphological and anatomical characters have been used by systematists to define species and subspecies. Behavior, vocalization, and biochemistry have also been used more recently. A broad set of genetic tools provided by molecular biology have immensely complemented existing methods and it is most likely soon that biologists will make a better taxonomy of the most of orders of the living world i.e., based on the relationships provided by phylogenetics and not solely on general similarity. Overall, the genetic characters are more numerous than the morphological ones and can help to clarify systematics. Darwin wrote to his friend T. H. Huxley in 1857,

> "In regard to classification, and all the endless disputes about the 'natural system' which no two authors define in the same way, I believe it ought, in accordance with my heterodox notions, to be simply genealogical. The time will come I believe, though I shall not live to see it, when we shall have fairly true genealogical trees of each kingdom of nature…"

Darwin's dream may now become a reality (reproduced from Wink, 2007).

Major DNA methods useful for studying Raptor systematics (as shown above in Table 1) are described briefly below. It would be useful to have a look at DNA as a logbook of life before starting on the applicability of DNA methods. The following paragraph on DNA is based on Avise, 1994; Hillis et al. 1996; Mindell 1997; Karp et al. 1998; Hall 2001; and Storch et al. 2001.

Similar to the DNA found in bacteria, mitochondrial DNA is a circular molecule derived from endosymbiotic bacteria taken up by the ancestral eukaryotic cell ≈1.4 billion years ago (Novikoff and Holtzman, 1976). In vertebrates, the mitochondrial genome consists of ≈16,000-19,000 base pairs (bp), containing 13 genes (that code for enzymes involved in respiratory chain), 22 genes for tRNA (transfer RNA) and 2 genes for rRNA (ribosomal RNA) (Table 2). The D-loop, or control region, is a short stretch of non-coding DNA in mitochondria and its variability is four to six times greater than protein coding genes, for instance Cyt *b*. Numbers of mitochondria in a typical animal cell ranges from 100 to 1000, each of which contains 5-10 copies of mtDNA (Novikoff and

Holtzman, 1976). Although it makes only 1% of all cellular DNA, this makes mitochondrial DNA an especially frequent molecule in cells. Therefore, mtDNA is an important source of genetic material for molecular studies, and it is inherited maternally, and so can be regarded as clonal in nature.

1.2.1 DNA Sequencing

A powerful method for phylogenetic reconstruction is the analysis of marker genes and their nucleotide sequences (Beebee and Rowe, 2004). The typical focus of the phylogenetic studies of raptors is on conserved marker genes' nucleotide sequences such as coding nuclear DNA (e.g. RAG1) and of protein coding mtDNA (Table 1) (Griffiths et al. 2004) or non-coding DNA (e.g. introns of protein coding genes, including ODC-6, and LDH). Sufficient divergence among geographically separated lineages is exhibited by the species that evolved several million years ago, and permit useful analysis (though the rate of change will determine how far back in time a gene is 'useful'). In identification of mitochondrial lineages of groups, or so-called haplotype, sequencing of mitochondrial genes (such as ND, Cyt *b* or CO) often proves helpful. Mitochondrial D-loop being more variable may provide an even higher resolution, but sometimes it can be difficult to amplify and sequence this region by PCR because of its high variability (Wink, 1998).

1.2.2 PCR Methods (Fingerprinting)

A high degree of resolution is required to analyze microevolutionary genetic differences within a species. At this intraspecific level, mtDNA sequences are sometimes less useful, in part because they are inherited maternally. An unambiguous allocation of individuals to populations, lineages and species can be masked by the normal processes of hybridization and introgression, because of the maternal inheritance of mtDNA (Karp et al. 1998). Also, in this respect DNA fingerprinting has been employed to trace individuals for pedigree studies. To overcome the above mentioned problems, more appropriate are nDNA molecular markers that have a higher degree of resolution and are

inherited by both sexes. In these methods polymorphic DNA markers are amplified by PCR and their separation involves high-resolution gel electrophoresis or capillary electrophoresis i.e. DNA sequence (Wink, 1998).

1.2.3 Microsatellite (STR) Analysis

Each raptor individual has two alleles for each nuclear locus: one derived from the mother and the other from the father. A high degree of length polymorphism is shown by the alleles of these STR loci. Several alleles exist for each polymorphic STR locus and they can often be distinguished by size because they differ in the number of tandem repeats (Wink, 2007).

As described by Nesje and Roed (2000) sequences used to amplify the STR loci can be identified through special efforts, as the sequences that might flank microsatellite loci vary between species. To generate species-specific STR sequences, several protocols have already been published. Wink (1998) selected various molecular techniques for studying raptors including their allele frequencies (used for STR) can be determined for characterizing populations. A 1/0 presence can be used to record the presence and absence of alleles, evaluated by cluster analysis, for instance UPGMA (Unweighted pair-group methods with arithmetic means) and neighbor-joining. Other programs used can be STRUCTURE (Pritchard et al. 2000) or GENELAND (Guillot et al. 2005).

1.2.4 Single Nucleotide Polymorphism (SNP)

A genetic map of populations can be built, so long as for each individual at least 30 loci are determined, by using single-nucleotide polymorphisms (SNPs) information. They are also analyzed via a 0/1 matrix (a similar way as STR data) and they have a similar resolution power (Lopez-Herraez et al. 2005). As suggested by Wink (2007), SNP marker systems are not available for raptors at present, because they have yet to be established for individual species. Most likely, this method would become an important tool in future, because DNA chips and mass spectrometry can automate SNP analysis.

1.2.5 Amplified Fragment Length Polymorphism (AFLP)

Including AFLP, genomic fingerprint methods provide an alternative if microsatellite PCR primer information is not available. The loci of AFLP are inherited co-dominantly (the allelic products coexist) and combining restriction-length analysis with PCR, AFLP makes it a powerful and convenient tool which results in a complex fingerprint detailed in a 0/1 matrix and it can be analyzed by cluster methods (Irwin et al. 2005). However, since the discovery of AFLP, little work has been done on wild species (including raptors). It is likely to be superior (in terms of relatively low cost and less start up time before data can be generated) to other more-established methods, such as microsatellites, SNP (single nucleotide polymorphism) analyses and multigene DNA sequencing (Bensch & Akesson, 2005).

1.2.6 ISSR (Inter-Simple Sequence Repeats)

The following paragraph on explanation and applicability of ISSR is based on Wink (1998, 2000, and 2002). In ISSR methods, the fingerprints produced are similar to AFLP. It is easier to carry out because it involves fewer experimental steps than AFLP. A single PCR primer is used, having its sequence identical to a common microsatellite motif, for instance, (CA) 10. Such loci are widely present in genomes and they occur in both orientations, and therefore only a single primer can be enough to amplify the stretches of DNA between adjacent microsatellite loci (i.e., between 10 and 80 loci) simultaneously. PCR products need to be analyzed by capillary electrophoresis or high resolution PAGE, as they differ in size. The ISSR loci provide information on the individual's genomic make up, since some of them are polymorphic. According to Wink (2007) these loci are inherited co-dominantly (the allelic products coexist). In practice, several ISSR are used, resulting in several hundred loci being available for analysis. In most plants and animals, the primers work universally, and that is the advantage of ISSR, PCR primers for any individual species are not needed to be defined. The results plotted in a 1/0 matrix, are evaluated by cluster analysis. In this way, individuals are

placed together on the basis of similarity of their band patterns in ISSR. For tracing back the individual birds to individual populations, ISSR can be useful to reveal population specific DNA bands (Wink et al. 2002).

1.2.7 Molecular Sexing

The following review of methodology and application of molecular sexing is based on Kahn et al. (1998); Morrison and Maltbie (1999); Hofle et al. (2000); Nesje and Roed (2000); Ristow and Wink (2004); Beckler and Wink (2003); Ristow et al. (2004); and Wink (2007).

Molecular sexing is useful molecular technique for raptor systematic work. This method allows the sexing of birds, which can otherwise be difficult in nestlings of monomorphic species, and also in many adult birds outside the breeding season. Compared with mammals sex chromosomes are in opposite order in birds: males are homogametic WW, whereas females have heterogametic ZW. On the sex chromosomes, the introns of the CHD gene are targeted by PCR methods. One PCR product can be obtained in males as opposed to two in females, because the alleles can differ in size. With the use of high resolution PAGE, this technique has been successful with almost all species of birds examined to date. For example, Hoefle et al. (2000) studied molecular sex determination in the Spanish Imperial Eagle (*Aquila adalberti*).

1.3 Raptor Evolution in General

Following is a short list of the main groups of raptors (based on del Hoyo et al. 1994a, b).

Order Falconiformes

1. Family Cathartidae (New World vultures)
2. Family Pandionidae (Osprey)
3. Family Accipitridae (Hawks & Eagles)
4. Family Sagittariidae (Secretarybird)

5. Family Falconidae (Falcons & Caracaras)

Order Strigiformes

1. Family Tytonidae (Barn owls)
2. Family Strigidae (Typical owls)

The relationship of Falconiformes with other birds is still not clear, and even among the Falconiformes, there is as yet no evidence from the fossil record that the different families have an immediate common ancestor, that is, they are monophyletic. So, it is still possible that the members of this order may be the result of convergent evolution between groups of a polyphyletic origin (del Hoyo et al. 1994a).

1.3.1 Order Falconiformes

According to DNA-DNA hybridization studies, the traditional Falconiformes, would be better placed in the infraorder Falconides within the larger order Ciconiiformes (storks). Falconides would not include the new world vultures (Cathartidae), which have closer affinities with other Ciconiiformes (Sibley and Ahlquist, 1990).

1.3.1.1 Family Cathartidae (New World vultures)

The following general account on evolution and classification of New World vultures is based on Houston (1994);

No. of genera	5	
No. of species	7	
No. of taxa	13	(including sub-species)

In superficial appearance, the seven species of New World Vultures are similar to the fifteen species of vultures found in Asia, Africa and Europe (i.e., Old World Vultures). However, the New World Vultures are most obviously different in not having a functional hind toe and also no internal separation of the nostrils. In general appearance, the two vulture groups are otherwise remarkably alike. Both of the vulture groups have

evolved large species which share many features, for instance, Condors (*Vultur* and *Gymnogyps*) from the New World have comparable size to the Griffon (*Gyps*) vultures in Old world. Similarly, medium to large sized species have also evolved in both of the vulture groups i.e. King vulture (*Sarcoramphus papa*) of the New world have the parallels in size to the Lappet-faced vultures (*Torgos tracheliotus*) and the White-headed vultures (*Trigonoceps occipitalis*) of the Old world vultures group. Both of the vulture groups have also evolved smaller species, of which *Coragyps* and *Cathartes* represent New world vultures, and the Egyptian vulture (*Neophron percnopterus*) and the Hooded vulture (*Necrosyrtes monachus*) represent the Old World vultures. The two groups of vultures (New World and Old World) are probably descended from different ancestors, and therefore these similarities between New World and Old World Vultures may have been derived independently, an example of convergent evolution. So they may not be closely related, and although they look alike this may be because they have adapted to a similar way of life.

Despite close similarities in feeding ecology and appearance of both the vulture groups, many aspects of their behavior and anatomy indicate that the two groups of vultures may not be closely related at all. Rather the features of cathartids are less close to those of eagles and hawks than to storks (Huxley 1876), so for more than 100 years it has been considered that New World vultures are more closely related to storks. Both New World vultures and storks squirt their legs with urine to maintain body temperature, and sometimes it looks as if their legs have been white-washed. Some modern stork species, particularly Marabou (*Leptoptilus crumeniferus*) obtains much of its food by scavenging dead fish, and it is found feeding alongside vultures at carcasses. One could imagine a stork species which was specialized for scavenging, and ultimately adapted to the lifestyle as Old World vultures.

But the new DNA results contradict the previous thinking. Recent nuclear DNA sequence-based phylogenetic analyses place the New World vultures in a sister relationship with eagles and hawks, and then the Old World vultures together (families

Sagittaridae, Pandionidae, and Accipitridae) (see Ericson et al. 2006; Hackett et al. 2008; Yuri et al. 2013; McCormack et al. 2013; Kimball et al. 2013).

The following description of cathartid fossils is based on Mayr (2009; p. 156-157, and references within). The fossil data on stem group representatives of New World birds of prey gives some glimpse of the possible evolutionary past of these taxa including New World vultures. They are grouped into Teratornithidae, which are now extinct. *Teratorns* closely resemble Cathartidae, due to many osteological features, however, their sister relationship is well supported by derived skull features. A fossil (*Taubatornis campbelli*) from the late Oligocene of Brazil is the only Paleogene fossil of the Teratornithidae. Olson and Alvarenga also support the hypothesis that *T. campbelli* is of South American origin. In contrast, and as pointed out by Mayr (2009; 2014) that Cathartidae had stem group representative fossils discovered from the paleogene of Europe. Another undescribed Cathartidae was mentioned from early Oligocene of Mongolia. The oldest known fossils of Cathartidae are from Paleogene. One of them *Phasmogyps patrius* is from the late Eocene while another one is *Brasilogyps faustoi* from the late Oligocene.
Traditionally, cathartid vultures have been placed in order Falconiformes, (Houston,. 1994). Brown and Amadon (1968) recognized that New World Vultures were different from the rest of the Falconiformes in many ways. But it has also been suggested that their present geographical isolation is geologically a comparatively recent event (Houston, 1994).

1.3.1.2 Family Pandionidae (Osprey)

The following description on evolution and classification of Osprey is based on Poole (1994).

No. of genera	1	
No. of species	1	
No. of taxa	4	(including sub-species)

Pandionidae is a monospecific family represented by Osprey (*Pandion haliaeetus*) and it is still matter of debate whether osprey is different enough from the rest of the raptors to warrant such a separate treatment. DNA-DNA hybridization studies (Sibley and Ahlquist, 1990) have suggested its sub-family status within the family Accipitridae. The lack of speciation is interesting among osprey and contrasts with the *Haliaeetus* sea-eagles. Ospreys are long distance migrants so there is more chance for individuals to interchange between populations, thus there are fewer chances for isolation of sub-species to become distinctive species.

The following description of pandionid fossils is based on Mayr (2009; p. 159-160, and references within). Although extant Ospreys are almost globally distributed, but all of their Paleogene fossils were discovered from Old World sites. The only identified fossil as an osprey, is the *Paleocircus cuvieri*. There are isolated fossil ungual phalanges discovered from the late Eocene of England and early Oligocene of Germany. Furthermore, a distal humerus and an incomplete carpometacarpus of a stem group of the Pandionidae was also reported from the early Oligocene of Egypt.

1.3.1.3 Family Accipitridae (Hawks & Eagles)

The given description on evolution and systematics of family Accipitridae is based on Thiollay (1994).

No. of genera	64	
No. of species	237	
No. of taxa	535	(including sub-species)

The Accipitridae have always been the most numerous of the families of Falconiformes. However, successive major classifications are rarely identical. Many of the taxa are converted from species to subspecies and vice versa. At the same time some of the species have been switched from one genus to another. Consequently, the number of species that have been recognized in recent times, range from about 212 to 240 (Thiollay, 1994).

Accipitridae is a highly diverse family and Thiollay (1994) suggests that it can usefully be divided into natural groups by bringing together species with high ecological or morphological convergence. In a systematic sense, any such arrangement may not reflect the evolution within the family, it illustrates the multiple radiations, occurred from many ancestral species to the supposedly derived and/or more evolved ones. Thiollay (1994) suggested that ten such natural groups within the larger group i.e., Accipitridae. These groups are explainded below.

- The kites, a heterogeneous assemblage of 33 species.
- The fish-eagles comprising of 10 species, considered to be derived from kites.
- 15 species of Old World Vultures, which share their scavenging habits.
- 15 assorted serpent-eagles or snake-eagles.
- The harriers, the harrier-hawks and also possibly the Crane Hawk.
- 58 species of goshawks and sparrow hawks, with rather uniform characteristics.
- 25 sub-buteonine hawks.
- 28 species of the genus Buteo.
- 4 large eagles of the America's tropical rain forests.
- 33 species of "booted" eagles.

The following description of accipitrid fossils is based on Mayr (2009; p. 159-161, and references within). According to the fossil record, the accipitrids and pandionids diverged before the Oligocene. The fossil record of the Accipitridae dates back to the late Eocene and early Oligocene. At least 18 species have been described so far. The oldest of the accipitrid fossils may be *Milvoides kempi* (Harrison and Walker 1979) from the late Eocene of England. Two further species of Accipitridae were described from Quercy fissure fillings: "*Aquila" ("Aquilavus") hypogaea* and "*A."corroyi.* The classification of these species in the same genus is conjectural, because the affinities between these two species cannot be established due to the lack of overlapping skeletal elements. Another species discovered from the late Oligocene of France, *Palaeohierax gervaisii*, was twice the size of the kite-sized *A. corroyi.* Three more species of putative

Accipitridae were described, from the middle Oligocene of Mongolia, as *Buteo circoides, Venerator dementjevi,* and *Gobihierax edax.* The description of another large accipitrid from the late Eocene of Egypt (Jebel Qatrani Formation) resembles to modern eagles of *Haliaeetus* taxon. Among the several species described from the Oligocene of North America, are *Buteo granger, B. Fluviaticus,* and *B. antecursor.* While another species *Palaeoplancus sternbergi,* roughly the size of Osprey seem to have some plesiomorphic feature which suggest its position outside crown group Accipitridae. Three species described from the late Oligocene of Nebraska, i.e. *Geranoaetus ales, Palaestur atavus,* and *Promilio efferus,* are assigned to Accipitridae as well. A couple of species described from the late Oligocene of Argentina are, *Climacarthrus incompletes* and *Cruschedula revola,* are assigned to Accipitridae based on, a badly preserved distal tarsometatarsus and the cranial extremity of scapula, respectively. But the phylogenetic affinities of these fragmentary bones cannot be reliably established until discovery of more complete specimens is made. From the late Oligocene of Australia, accipitrids are represented by *Pengana robertbolesi.*

1.3.1.4 Family Sagittariidae (Secretarybird)

The following systematic and general evolutionary account of Secretarybird is based on Kemp (1994).

No. of genera	1	
No. of species	1	
No. of taxa	1	(including sub-species)

Sagittariidae is a monospecific family with Falconiforms, and it is in its own sub-order, Sagittarii, but sometimes is even elevated to the rank of its own order, Sagittariiformes. It is placed adjacent to other diurnal birds of prey but the relationship with Cariamidae (Seriemas), Otididae (Bustards) or Gruidae (Cranes) in the order Gruiformes has also been proposed. To many people it looks like a “long legged marching eagle”. The skull and head also shares anatomy with eagles. Breeding habits are most similar to Storks in

order Ciconiiformes. And it is also notable that recent DNA-DNA hybridization studies (Sibley and Ahlquist, 1990) suggest close affinity between raptors and storks.

From the early and late Oligocene, *Pelargopappus schlosseri* is a stem group representative of Sagittariidae. Discovered from the Quercy fissure fillings, it is smaller in size than the single extant species *Sagittarius serpentarius* (Mayr 2009; p. 158-159, and references within).

1.3.1.5 Family Falconidae (Falcons & Caracaras)

The systematics and general evolution of family Falconidae, described below, is based on Kemp (1994).

No. of genera	10	
No. of species	61	
No. of taxa	149	(including sub-species)

Species of Falconidae have great diversity of life styles. They range from long-winged falcons (*Falco*) in open country, to ground dwelling scavenger caracaras (*Polyborinae*); and from almost chicken-like arboreal caracaras (*Daptrius*), to forest-falcons (*Micrastur*) of tropical forests that resemble the true hawks (*Accipiters*) with short wings There are also tiny falconets (*Microhierax*), being only the size of a thrush. However, they all are united taxonomically through six (or maybe more) distinctive morphological characteristics, for instance;

- Moult pattern of a characteristic flight feather (starting with primary 4).
- A bony tubercle developed to a degree in the nostril.
- Eggshells of the same chemical composition.
- The same *Mallophaga* feather lice.
- With the exception of pygmy-falcons having white eggs, their eggs are spotted reddish brown and are all blotched.

- The shells have reddish yellow translucence, rather than the internal greenish tint of accipitrid eggs.

Falconidae have been generally placed alongside the other diurnal raptors since the time of Linnaeus. This was supported by evidence from external morphological studies, internal anatomy, feather parasites, behavior, and moult patterns, as well as more recently the molecular evidences based on allozymes divergence (Randi et al. 1991) and DNA-DNA hybridization (Sibley and Ahlquist, 1990). However, detailed morphological studies of the families traditionally included in Falconiformes, the order was concluded to be polyphyletic (thus having more than one origin). It was also suggested that falconids were actually closer to plantain-eaters, owls, parrots and cuckoos. However, modern authorities regard the shared features of owls and falcons to be convergent (Wink et al. 2009) and most recent and comprehensive avian family classifications retain the Falconidae within the order Falconiformes (Lerner and Mindell, 2005; El-Sayed, 2007). Four main issues dominate the investigation concerning the systematics of the family.

- Within the family, how many subfamilies should be recognized, and also the composition of each of these.
- How different genera are related to each other and how should these relationships be reflected in a linear sequence.
- Number of species and their relationship within various genera.
- What should be the name of a particular species or genus.

Lack of agreement between different studies (refered to, in this section) creates more confusion. Therefore, it creates need for further investigations.

A poor Paleogene fossil record of falcons, includes no unambiguously identified specimens. Some recent molecular analyses support a clade that includes Falconidae, Cariamidae, Psittaciformes, and Passeriformes (Ericson et al. 2006; Hackett et al. 2008) which lead to temptation about possible affinities of falcons to the stem group

representatives of the Cariamidae i.e. the raptorial Phorusrhacidae. A couple of fossils described from the London Clay are, *Parvulivenator watteli* and *Stintonornis mitchelli.* They are tentatively assigned to the Falconidae, but are also too fragmentary to be reliably identified. Another fossil from the middle Eocene of Antarctica was assigned to the Falconidae, but again the specimen needs the additional skeletal elements to confirm the identification (Mayr 2009; p. 153-155, and references within).

1.3.2 Order *Strigiformes* (Owls)

The following classification of owls is based on Bruce (1994) and Marks et al. (1994).

	Family Strigidae	Family Tytonidae	
No. of genera	2	25	
No. of species	16	189	
No. of taxa	63	548	(including sub-species)

Taxonomists have been debating for over two centuries whether to consider owls (Strigiformes) as closely related to diurnal raptors (Falconiformes) or to nightjars including their allies (Caprimulgiformes). The hooked bill (to tear flesh) and the strong feet and claws (to capture prey) were significant to align owls to raptors, while the larger eyes, soft plumage and also other nocturnal hunting adaptations put them closer to Caprimulgiformes. The review of owl classification based on DNA-DNA hybridization work (Sibley and Ahlquist, 1990) shifted the opinion towards owls' closer affinities with Caprimulgiformes. This work was concluded by combining the owls with Caprimulgiformes in an expanded Strigiformes, though subsequent studies, including allozyme divergence (Randi et al. 1991), supported the traditional separation of Caprimulgiformes from Strigiformes. So it is now suggested that owls are closer to Caprimulgiformes and they are placed near this order, so far the evidence is not yet conclusive.

The fossil record provided a better understanding of the diversity and evolution of owls. Six families (two extant, and four extinct) are recognized at least, until late Paleocene.

The fossil record is quite extensive in the Eocene and Oligocene of the Northern Hemisphere. Notably, there is no Paleogene fossil record of Strigiforms from the Southern Hemisphere. The two early discoveries are *Berruornis orbisantiqui* and *B. Halbedeli* are known from the late Paleocene of France and Germany respectively. These two species were assigned to the family Sophiornithidae, a taxon which was originally erected for a *Bubo*-sized species *Sophiornis quercynus* from the Quercy fissure fillings. The second family Protostrigidae comprises of seven fossil species. Six species of this family from the Eocene of North America are *Eostrix mimica, E. martinelli* (early Eocene of Wyoming), *Minerva antiqua, M. leptosteus, M. saurodosis* (middle Eocene of Wyoming), and *M. californiensis* (late Eocene of California). The only species that represents Protostrigidae in Europe is *Oligostrix rupelensis* from the early Oligocene of Germany. The third family Ogygoptyngidae is represented by the single species *Ogygoptyngidae wetmorei* from the late paleocene of Colorado. Necrobyinae comprises of seven fossil species which were considered as stem group representatives of the extant Tytonidae. This group includes four taxa, *Necrobyas* (4 species: from late Eocene to late Oligocene), *Nocturnavis* (1 species: late Eocene), *Palaeobyas* (1 species: unknown locality of the Quercy fissure fillings), *Palaeotyto* (1 species: unknown locality of the Quercy fissure fillings). The fourth of the extinct owl families, the Palaeoglaucidae comprises of two species named *Palaeoglaux perrierensis* and *P. artophoron* which come from the late and middle Eocene respectively. The holotype of *Selenornis henrici* is the only known specimen that represents the monotypic taxon Selenornithinae. This specimen came from an unknown locality of the Quercy fissure fillings (Mayr 2009; p. 153-155, and references within).

The current division of living owls between the families Strigidae and Tytonidae, is further supported by the recent work on fossil material. Though forming a distinct subfamily, Phodilinae, the bay-owls (*Phodilus*) are probably best placed with Tytonidae. The subsequent studies on DNA-DNA hybridization (Sibley and Ahlquist, 1990) and allozyme divergence (Randi et al. 1991) support a Tytonidae family including *Phodilus*. The modern authorities regard the shared features of owls and falcons to be convergent (Wink et al. 2009).

1.4 Chronological survey of literature

The current emphasis of review should be the period since the publication of *Origin of Species* by Darwin in 1859. More complete accounts of early ornithologists who made a lasting impression on avian systematics can be followed in the introduction to Newton's *A Dictionary of Birds* (1896) and in Stresemann's *Ornithology from Aristotle to the present* (1975).

O'Hara (1988) gave a good review of the avian classification in the form of the *Diagrammatic Classifications of Birds* published between 1819 and 1901. Ornithologists were quick to embrace Darwinism. Perhaps one of the most respected among them, T.H. Huxley (1825-1895), was also an advocate of Darwin's ideas. Professor Huxley's 'Classification of Birds' (1867) has universally been recognized as an epoch-making memoir in the history of ornithology. The contributions of Alfred Newton (1829-1907) were less those of a participant than of an often acerbic critic. Reviews of the work of his predecessors and contemporaries can be seen in his historical critiques (1896). Sclater's work (1880) about classification, was based on some previous classifications, somewhat modified by his own work.

Moving to the modern era, there have been other classifications proposed during the past years, but mostly reflected the avian phylogenetic concepts inherited form Huxley, Furbringer and Gadow (Sibley and Ahlquist, 1990). Cracraft (1981) applied the formal principles of classification proposed by Willi Hennig (1966) for the analysis of morphological characters, a first attempt at phylogenetic classification of recent birds. Cracraft (1981) reintroduced the idea of including owls in the Falconiformes, because of derived tarsometatarsal and pelvic morphology as shared with Pandionidae and Accipitridae (including falcons). But, Olson (1982) vigorously criticized Cracraft's classification, showing that consensus had not yet been achieved in 1982. Olson (1982) criticized Cracraft's "division 3" (that contained Ciconiiformes and Falconiformes) vigorously as a category that contains both flamingos and owls as monophyletic, it will be in vain to search for a synapomorphy to define such a group, he claimed. The

classification by Amadon and Bull (1988) followed the traditional arrangement with all species in the Strigiformes and Strigidae, and the strigids divided into Tytoninae (Barn owls, Bay owls), and Striginae (Typical owls).

Among the molecular studies, nucleotide sequences of cytochrome b gene of the mitochondrion have been employed to study the systematics and evolution of diurnal raptors and owls (Heidrich and Wink 1994, 1998; Wink 1995, 1998, 2000; Wink et al. 1996, 1998; Griffiths 1997, Seibold and Helbig 1995a, b, 1996; Heidrich et al. 1995a, b; Mindell 1997; Matsuda et al. 1998, Wink & Heidrich 1999, 2000; Haring et al. 1999, 2001; Wink & Sauer-Gürth 2000, 2004; Groombridge et al. 2002, Olsen et al. 2002; Riesing et al. 2003, Hendrickson et al. 2003, Godoy et al. 2004, Griffiths et al. 2004, Kruckenhauser et al. 2004, Pearlstine 2004, Roques et al. 2004, Roulin & Wink 2004, Nittinger et al. 2005, Gamauf et al. 2005 and Proudfoot et al. 2006, 2007). Other valuable work on different genes includes that on Cyt *b*, RAG1, LDH int3, AK int5 studied in Aquiline eagles (Helbig et al. 2005); 12S, ATP8, ATP6, ND6 in Accipitridae (do Amaral et al. 2006); Cyt *b*, ND2, mt Control region and RAG1 in *Gyps* vultures (Arshad et al. 2008; Johnson et al. 2006); mitochondrial pseudo control region and ND6 in Buteonine raptors (Riesing et al. 2003).

Mitochondrial Cyt *b* gene was used to study the finer details of speciation and phylogeny of owls (Wink and Heidrich 1999, 2000; Wink et al. 2004). Michael Wink also reviewed the application of different DNA-markers to study the ecology and evolution and advances of DNA studies in both diurnal and nocturnal raptors (Wink, 1998; 2000). His work on the molecular systematics of African raptors' (Wink and Sauer-Gurth, 2000); Holarctic raptors (Wink et al 1998); Mediterranean raptors (Wink and Seibold, 1996) is remarkable. Helbig et al. (2005) studied the phylogenetic relationships among Falcon species using DNA sequence variation of Cyt *b* gene.

More recently Wink et al. (2009) included 120 taxa of the Strigidae and 23 taxa of Tytonidae (the dataset covers most of the genera) to study the phylogeny of owls (nocturnal raptors) based on Cyt *b* and nuclear markers (LDH b intron DNA, RAG1).

Their research provided insight into the phylogeny and evolution of owls. The phylogenetic tree inferred from sequences of the cytochrome b gene and nuclear RAG-1, was found to be generally in a good agreement with the classical taxonomy of owls (Sibley and Monroe 1990, Burton 1992, Hume 1991, König *et al.* 1999, Weick 2006). The genetic data agreed with the attribution of species to a given genus with exceptions evident in the polyphyletic genus *Otus* and the paraphyletic *Bubo* complex.

In the classification of birds by Cracraft (1981) in which the orders Ciconiiformes and Falconiformes were put together, he admittedly noted that placement in the same division may seem unwarranted. In a recent investigation on family Accipitridae (Lerner and Mindell, 2005) phylogenetic relationships using molecular sequence from two mitochondrial genes (ND2 and Cyt *b*) and one nuclear intron (b-fibrinogen intron 7) were inferred. Representatives of all 14 Accipitridae subfamilies and at least one representative from all genera were sampled, focusing on four subfamilies of eagles (booted eagles, sea eagles, harpy eagles, and snake eagles) and two subfamilies of Old World vultures (Gypaetinae and Aegypiinae) to include a majority of their species. Multiple well-supported relationships among accipitrids recognized with DNA sequences vary from those traditionally recognized based on morphology or life history characters. Monophyly of sea eagles (Haliaeetinae) and booted eagles (Aquilinae) was supported by this research; nevertheless, harpy eagles (Harpiinae), snake eagles (Circaetinae), and Old World vultures were inferred to be non-monophyletic. The Gymnogene (*Polyboroides typus*) and the Crane Hawk (*Geranospiza caerulescens*) were not found to be close relatives, presenting an illustration of convergent evolution for specialized limb morphology enabling predation on species nesting in cavities. Study of named subspecies within *Hieraaetus fasciatus* and *H. morphnoides* exposed significant genetic differentiation, or non-monophyly, supporting recognition of *H. spilogaster and H. weiskei* as individual species.

A recent study by El-Sayed (2007) provides valuable and additional information to understand the evolutionary relationships of birds of prey in the order *Falconiformes* generally. El-Sayed's (2007) study included almost half of the species recognized for the

Accipitridae (111 species) and for the Falconidae (35 species) that represents the largest molecular dataset for the Falconiformes studied so far. Intrageneric and intraspecific relationships of the Falconidae and the Accipitridae were of more focus of the above mentioned investigation. The analysis of combined genes having different evolutionary rates (Cyt *b* and RAG1) provided high resolution of both deep and shallow branches. The reconstructed topologies based on Cyt *b* were highly congruent with those based on RAG1. Moreover, results were in accordance with previous molecular studies and their Accipitridae tree showed a basal polytomy. New and Old World vultures represent two independent evolutionary lineages in this phylogenetic analysis. The New World vultures are not closely related to storks as suggested by Sibley and Ahlquist (1990), but they are closer to the Accipitridae. 'Raptors' may be a result of an evolutionary convergence between bird groups of polyphyletic origin. The traditional classification of Brown and Amadon (1968) which recognizes the order Falconiformes with the subgroups Cathartae, Accipitres, Sagittarii and Falcones is probably partly artificial and does not fully reflect phylogenetic descent. The results of El-Sayed (2007) did not support a common ancestor for this assemblage, and falcons form an independent lineage which is closer related to parrots than to Accipitridae.

The base of the Neoavian tree is suggested to be a hard polytomy (Poe and Chubb, 2004) but some recent phylogenetic analyses based on larger nuclear datasets appears to resolve some of these problematic interordinal relationships (Ericson et al. 2006; Hackett et al. 2008). They recovered a strongly supported group of 'higher land birds' which includes both nocturnal (Strigiformes) and diurnal (falcons and eagles). They recovered this previously unrecognized clade where falcons share a common ancestor with parrots (Psittaciformes) and song birds (Passeriformes). Apart from this node, eagles (including New World vultures) and owls were also recovered to have two separate origins. Although both of these studies were congruent on well-supported nodes, the study based on 19 nuclear loci (Hackett et al. 2008) was criticized for its data being predominantly introns based, which are known to have alignment issues (Morgan-Richards et al. 2008). The subsequent studies based on nuclear DNA sequences also re-evaluated the novel relationships reported by Hackett et al. (2008) and found the similar

relationships among falcons, owls and eagles (Wang et al. 2012; Yuri et al. 2013). The avian phylogenetic relationships inferred by the transposable elements (TE) insertions (Suh et al. 2011) and ultraconserved elements (UCEs) (McCormack et al. 2013) also corroborated Hackett et al. (2008) regarding the phylogenetic position of raptors. The section 1.5 (Figure 1) also shows a comparison of some of the previously published phylogenies of raptors. A more recent whole genome analysis also suggests that raptors are part of 'higher land birds' clade (Jarvis et al. 2014).

Morgans-Richards et al. (2008) found two different origins of falcons and eagles, in their analysis of phylogenetic relationships based on mitochondrial genomes, although with weaker support. In the subsequent mitogenomic phylogeny, Pratt et al. (2009) recovered a clade comprising of owls, eagles and osprey (excluding falcons). Later, a reasonably well sampled phylogenetic analysis of avian mitogenomes grouped eagles and falcons together, forming a monophyletic order Falconiformes (Pacheco et al. 2011). But more recently, we concluded (Mahmood et al. 2014: Chapter 3 of this thesis) that owls, eagles and falcons have different origins, that is, raptors are not a natural/ monophyletic group. In this case the eagles consisted of the true eagles, osprey, and the secretarybird.

1.5 Already published phylogenies/ trees of raptors

The following Figure 1 shows the several raptor phylogenies that have been published more recently.

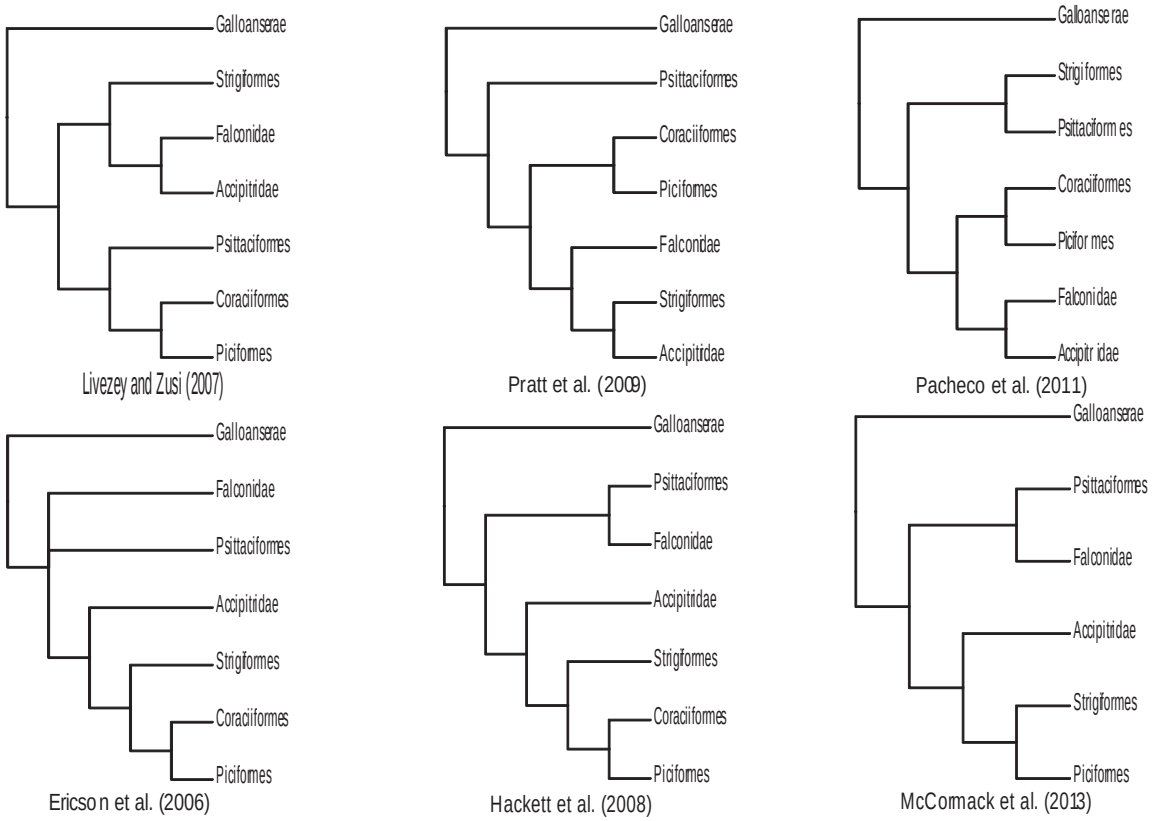


Figure 1. Some of the raptor phylogenies that have been published over the last decade.

1.6 Raptors of Pakistan

To have an idea about the diversity of birds of Pakistan, the enormous work of Tom J. Roberts, in the form of *Birds of Pakistan* (Roberts, 1991) can be regarded as the pioneer comprehensive account of avifauna of the region. Also, *The compact handbook of the birds of India & Pakistan* (Ali and Ripley, 1989) has been hailed as monumental work and is likely to remain an indispensable, standard publication for serious students of ornithology. Following Ali and Ripley (1989), a large number of species of diurnal raptors (Falconiformes) falling under families Accipitridae and Falconidae, can be found in the subcontinent. This book classifies owls (Strigiformes) by putting all owls of the region under family *Strigidae*, which further splits into subfamilies Tytoninae (Barn owls) and Striginae (other owls).

Almost no work was done on raptors' phylogeny and taxonomy in Pakistan until The World Conservation Union declared *Gyps* vultures as critically endangered (IUCN, 2006) following a steep decline in the population of > 95% (Prakash et al. 2003). Up to this point, what has been missing is consideration of their taxonomy and classification in detail. Responding to the issue, the systematics within *Gyps* vultures has been studied using mitochondrial Cyt *b*, ND2 and control region sequence data (Johnson et al. 2006). The phylogeny and phylogeography of the *Gyps* vultures of Pakistan has also been explored more recently, based on nuclear (RAG1) and mitochondrial (Cyt *b*) markers, which showed monophyly and no geographical partition within this group of vultures (Arshad et al. 2008). Similarly, more molecular phylogenetic work is needed for other raptors of Pakistan as well, to inform the conservation programs more explicitly.

1.7 Problems in resolution of the Neoavian tree

The reasons why there is conflict and difficulty in resolution of avian tree of life is worth discussing to help elaborate the scope of present investigation. In this concern the general issues are:

- Taxon sampling
- Number of genes (size of dataset)
- Compositional bias
- Substitution saturation
- Alignment issues

Early small sample studies got the tree partially incorrect because they did not sample enough e.g. some early mtDNA studies found passerines basal to Paleognaths (Mindell et al. 1997, 1999; Härlid and Arnason, 1999; Haring et al. 2001). Particularly these studies have suggested that Passerines are one of the most ancient of the extant avian taxa. The main reason for this placement was the faster evolution rate of Passerines which causes the long branch attraction problem (Hendy and Penny 1989) in a phylogenetic tree. Even though complete mitochondrial sequences have a large number of sites, the sequences analysed by Mindell et al. (1997, 1999) comprised only six avian

orders representing about 17% of the extant orders. Increased taxon sampling has reduced this problem. Braun and Kimball (2002) demonstrated that mitochondrial data is consistent with the passerines and neognaths monophyly. Their study also suggested that larger taxon samples generally perform better in phylogenetic analysis and increased the likelihood of recovering the correct topology. However, taxon addition is to be done carefully to avoid taxa that may create long branches which can further introduce biases in reconstruction (Mitchell et al. 2000; Poe, 2003; Hedtke et al. 2006). Similarly, phylogenetic inference improves by the addition of taxa that bisect long branches (Hillis, 1998; Poe and Swofford, 1999).

Increased sequence length is another factor that may reduce sampling bias and can be effective if sequence length does not go beyond a certain number of characters (Hillis et al. 2003). In other words, the sequence length is useful up to a certain limit. So the approach of Hillis et al. (2003) would depend on the situation to be examined. It contrasts to the approach of adding huge number of sites for fewer taxa (Rosenberg and Kumar 2002). Early studies of one or two genes may get a partially incorrect tree (or have little resolution). For instance, Fain and Houde (2004) suggested a split into neoaves tree based on intron 7 of the β-fibrinogen gene. They divided neoaves into Metaves and Coronaves. It was more recently that Morgan-Richards et al. (2008) rejected this division. Longer mitochondrial sequence data has proved helpful in stabilizing the avian tree (Gibb et al. 2007; Slack et al. 2007). Zwickl and Hillis (2002) and Pollock et al. (2002) demonstrated through extensive simulation studies that adding taxa to a phylogenetic analysis proved to be more effective than adding more characters. So, more genes are only part of the answer – we need a balance between more taxa and longer sequences!

The apparent conflict between the mitochondrial and nuclear genomes may be due to the difficulties in analysis of mitochondrial sequence data, particularly the 3rd codon positions may have biased base composition (Voelkers and Edwards, 1998). RY-coding is often used to reduce this bias at the 3rd codon position (Gibb et al. 2007). Greater nucleotide compositional bias in mtDNA is another difference between nuclear and

mitochondrial genes (Phillips et al. 2004). Making it further complicated each of two strands of mtDNA may show many different patterns in compositional bias, and these patterns may change due to gene rearrangements (Hassanin et al, 2005).

Faster evolving sites also more rapidly accumulate the 'non-historical biases' i.e. compositional non-stationarity (Lockhart et al. 1994). Compositional non-stationarity is a process characterized by the statistical properties that may vary over time. Mitochondrial genome has this problem and is most often discussed with reference to nucleotide bias in phylogenetics (Simon et al, 2006). If nucleotide bias pattern changes over lineages, models assuming stationary nucleotide distribution bias among taxa, may lead to systematic error in phylogenetic construction. For instance, Paton et al. (2002) used complete mitochondrial genome corrected for non-stationarity, to reject the idea that modern birds are descended from shorebirds or passerines.

Aligning non-coding sequences like introns can be difficult as they often have many gaps. The alignment of any intron having much indels and not any constant sites, may lead to artifacts, particularly when studying inter-relationship among orders having more than 60 million years of age (Morgan-Richards et al. 2008). Without the conserved sites, the problematic nucleotide sequence alignment is one of the sources for false signal for phylogeny (Kumar and Filipsky, 2007). Such erroneous alignment misleads the construction of tree by apparently giving phylogenetic signal among gaps (Lee, 2001).
Phylogenetic inferences are affected by combination of these factors. The combination of nuclear and mitochondrial data can be very important because the incongruence between the phylogenies of both may reveal important aspects of species history.

1.8 Questions for this thesis

1.8.1 What will the database of raptors achieve? (Chapter 2 of thesis)

At present it is difficult to find precisely what data is available, and to assemble all that data. A raptor specific database will enable us to more quickly, easily and accurately

identify which taxa require additional sequences to address questions of interest. Furthermore, the data available changes almost weekly, and it is difficult to update datasets. The database needs to output multiple sequence alignments and be updated regularly through GenBank/NCBI. We have already developed a database concentrating on raptors in a broad sense (including owls). It is extendable to other groups (avian and non-avian), and so we expect it will be adopted by others. New developments in this regard include PhyLoTa (Sanderson et al. 2008), and some more ideas will be imported from there.

1.8.2 Are raptors monophyletic? (Nocturnal, Diurnal and combined)

Nocturnal raptors: The representatives from the two owl families i.e. Strigidae and Tytonidae come together but quite weakly. Given that each of the two owl families are single sequences that form long branches, it may be that the long branches group together due to long branch attraction (Hendy and Penny, 1989) rather than shared ancestry. If we break these long branches by adding a new sequence for each family, we may be able to better resolve the deeper relationships with these complete mitochondrial sequences. Our aim is to break both long branches by adding two more genera, one for each family. For resolving deeper relationships we need complete mitochondrial DNA sequences. The complete mitochondrial genome sequencing for *Athene brama* (in the Strigidae) will be undertaken and a DNA sample for another owl from genus *Phodilus* (in the Tytonidae) will be imported from Germany and sequenced. The two species selected are widely separated from the existing mitochondrial genomes (*Strix* and *Tyto*), thus breaking up the long branches as much as possible.

Diurnal raptors: They have been described as having one, two or three origins (see Figure 1). Traditionally they were thought to be monophyletic, but Sibley and Ahlquist (1990) grouped the New World vultures as separate. More recently, Hackett et al (2008) have described Falconidae, Accipitridae and Cathartidae (New world vultures) as three different families. We will test these competing hypotheses more carefully. Again we will need complete mitochondrial genome sequences from a limited number of birds i.e.,

other diurnal raptors and their potential close relatives (reflected on previous studies; also refer to Figure 1), and that of the Secretary bird (*Sagittarius serpentarius*) will help. We expect the Secretary bird to be deepest on the combined Accipitridae/osprey lineage. Biogeographic questions are related and, for example, New World vultures appear to be particularly a South American radiation.

1.8.3 When and how often did raptor ecology evolve in birds—how long have raptors been raptors?

This is really associated with long-term niche stability. The first part of data collection is to continue with the mitochondrial genomes of the Secretary bird, and two owls including *Athene brama*. The general theme of our study will be the evolution of diurnal and nocturnal raptors. So, one aspect will be to interpret the results as a test of the apparent long term 'evolutionary-stable niche-discontinuity' hypothesis of Matt Phillips (see Poole et al. 2003). It seems likely that some modern raptor groups arose in the Cretaceous, and have been raptors ever since (Brown et al. 2008). This requires further testing.

1.8.4 What data do we need to test the questions?

We will sequence both mitochondrial as well as nuclear genes to go for phylogenetic analysis. Which nuclear genes will be the most useful! — will come from the database. Our aim would be to have an agreement between nuclear and mitochondrial data.

1.8.5 How would we analyze the data to test these questions?

Intron data as used particularly by Hackett et al. (2008) is expected to work well within an order or family, but this needs to be tested. But at the same time the alignment of the data may not be the same as a simultaneous alignment over all Neoaves. So a subsidiary goal is to check intron data within a small group, and compare the trees and alignment for deeper divergences. We predict a failure of intron data at deeper divergences but it is

expected to be very useful for more recent divergences, for example, when studying the raptors of Pakistan.

1.8.6 Current progress on Neoaves

A previous study by Pratt et al. (2009), attempted to resolve deeper level systematics among modern birds. The main approach in the above mentioned study was to improve signal for deeper divergences, presenting the formulae for calculation of probabilities to find predefined groupings in optimal tree and reporting nine new mitochondrial genomes that was a significant increase in data for each of the six main groups within Neoaves proposed by Cracraft (2001), thus reducing the number of long branches. In general this study was in agreement with Cracraft (2001) with two exceptions, the owls moved to the diurnal raptors, and shorebirds could be a separate group from "Conglomerati".

An important purpose of the present thesis will be to test further this apparent convergence of nocturnal and diurnal raptors. It will be of assistance in this respect to use improved down-weighting techniques for faster evolving sites (Phillips et al. 2010). This will help decide whether all sites support the convergence of nocturnal and diurnal raptors, or just the faster sites.

1.9 Thesis outline

My outline for the thesis is as follows:

Chapter 1. Introduction (this chapter).

Chapter 2. Database for raptors (see section 1.8.1.).

Chapter 3. Phylogenetic Position of Avian Nocturnal and Diurnal raptors (see sections 1.8.2. to 1.11.7.).

Chapter 4. Old and New world vultures (see sections 1.8.3.).

Chapter 5. Genetic diversity of Spotted owlet (*Athene brama*) from the Agro-Ecosystem of Punjab (Pakistan) (background for conservation).

Chapter 6. Thesis summation.

References

Ali S, Ripley S.1989. Compact handbook of the birds of India and Pakistan. Oxford University Press, Delhi.

Amadon D, Bull J.1988. Hawks and owls of the world. Proc. West. Found. Vert. Zool, 3:295-357.

Andrews, P. 1990. Owls, caves, and fossils: predation, preservation, and accumulation of small mammal bones in caves, with an analysis of the Pleistocene cave faunas from Westbury-sub-Mendip, Somerset, UK. University of Chicago Press

Arshad M, Gonzalez J, El-Sayed A, Osborne T, Wink M.2009. Phylogeny and phylogeography of critically endangered Gyps species based on nuclear and mitochondrial markers. J. Ornithol. 150:419-430.

Avise JC.1994. Molecular markers, natural history and evolution. Chapman & Hall, New York, London.

Becker P, Wink M. 2003. Influences of sex, sex composition of brood and hatching order on mass growth in common terns *Sterna hirundo*. Behav. Ecol. Sociobiol. 54:136-146.

Beebee T, Rowe G. 2004. An Introduction to Molecular Ecology. Oxford University Press, Oxford.

Bensch S, Akesson M. 2005. Ten years of AFLP in ecology and evolution: why so few animals? Mol. Ecol. 14:2899-2914.

Braun EL, Kimball RT. 2002. Examining basal avian divergences with mitochondrial sequences: Model complexity, taxon sampling, and sequence length. Syst. Biol. 51:614-625.

Brown JW, Rest JS, Garcia-Moreno J, Sorenson MD, Mindell DP. 2008. Strong mitochondrial DNA support for a Cretaceous origin of modern avian lineages. BMC Biol. 6:6

Brown LH, Amadon D. 1968. Eagles, hawks and falcons of the world, Vols. I & II. Hamlyn House, Feltham, Middlesex, U.K.

Bruce M. 1994. Family Tytonidae (barn-owls). Handbook of the Birds of the World. Lynx Edicions. Pp. 34–75.

Burton JA. 1992. Owls of the world: their evolution, structure and ecology. Peter Lowe, London.

Cracraft J. 1981. Toward a phylogenetic classification of the recent birds of the world (Class Aves). The Auk 98:681-714.

del Hoyo, J., A. Elliott, and J. Sargatal. 1999. Handbook of the birds of the world. Vol. 5, Barn-owls to Hummingbirds. Lynx edicions

del Hoyo J, Elliott A, Sargatal J, Cabot J. 1994. Handbook of the Birds of the World. Vol. 2, New world vultures to guineafowl. Lynx Edicions

Do Amaral F, Miller M, Silveira L, Bermingham E, Wajntal A. 2006. Polyphyly of the hawk genera *Leucopternis* and *Buteogallus* (Aves, Accipitridae): multiple habitat shifts during the Neotropical buteonine diversification. BMC Evol Biol. 6:10.

El-Sayed AA. 2007. Phylogenetic relationships of diurnal raptors based on nucleotide sequences of cytochrome b and RAG-1 genes. Institute of Pharmacy and Molecular Biotechnology. Ruperto-Carola University, Heidelberg. Pp. 90.

Ericson PGP et al. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. Biol. Letters 2:543-547.

Fain MG, Houde P. 2004. Parallel radiations in the primary clades of birds. Evolution 58:2558-2573.

Felsenstein J. 1993. PHYLIP, version 3.5c. Department of Genetics, University of Washington, Seattle.

Fowlie MK, Krüger O. 2003 The evolution of plumage polymorphism in birds of prey and owls: the apostatic selection hypothesis revisited. J. Evol. Biol. 16: 577-583

Frankham R, Ballou JD, Briscoe DA. 2002. Introduction to conservation genetics. Cambridge University Press, Cambridge, United Kingdom.

Gamauf A, Gjershaug JO, Rov N, Kvaloy K, Haring E. 2005. Species or subspecies? The dilemma of taxonomic ranking of some South-East Asian hawk-eagles (genus *Spizaetus*). Bird Conserv. Int. 15:99-117.

Gibb GC, Kardailsky O, Kimball RT, Braun EL, Penny D. 2007. Mitochondrial genomes and avian phylogeny: Complex characters and resolvability without explosive radiations. Mol. Biol. Evol. 24:269-280.

Godoy J, Negro J, Hiraldo F, Donazar J. 2004. Phylogeography, genetic structure and diversity in the endangered bearded vulture (*Gypaetus barbatus, L.*) as revealed by mitochondrial DNA. Mol. Ecol. 13:371-390.

Griffiths C. 1997. Correlation of functional domains and rates of nucleotide substitution in cytochromeb. Mol. Phyl. Evol. 7:352-365.

Griffiths CS, Barrowclough GF, Groth JG, Mertz LA. 2007. Phylogeny, diversity, and classification of the Accipitridae based on DNA sequences of the RAG-1 exon. J. Avian Biol. 38:587–602.

Griffiths C, Barrowclough G, Groth J, Mertz L. 2004. Phylogeny of the Falconidae (Aves): a comparison of the efficacy of morphological, mitochondrial, and nuclear data. Mol. Phyl. Evol. 32:101-109.

Groombridge J et al. 2002. A molecular phylogeny of African kestrels with reference to divergence across the Indian Ocean. Molecular Phylogenetics and Evolution 25:267-277.

Hackett SJ et al. 2008. A phylogenomic study of birds reveals their evolutionary history. Science 320:1763-1768.

Hall BG. 2001. Phylogenetic Trees made Easy. Sinauer Associates, Sunderland, Massachusets.

Haring E, Kruckenhauser L, Gamauf A, Riesing MJ, Pinsker W. 2001. The complete sequence of the mitochondrial genome of *Buteo buteo* (Aves, Accipitridae) indicates an early split in the phylogeny of raptors. Mol. Biol. Evol. 18:1892-1904.

Haring E, Riesing M, Pinsker W, Gamauf A. 1999. Evolution of a pseudo-control region in the mitochondrial genome of Palearctic buzzards (genus *Buteo*). J. Zool. Syst. Evol. Res. 37:185-194.

Harlid A, Arnason U. 1999. Analyses of mitochondrial DNA nest Ratite birds within the neognathae: supporting a neotenous origin of Ratite morphological characters. Proc. Royal Soc. London. Biol. Sci. 266:305-309.

Hassanin A, Leger N, Deutsch J. 2005. Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of Metazoa, and consequences for phylogenetic inferences. Syst. Biol. 54:277-298.

Hedtke S, Townsend T, Hillis D. 2006. Resolution of Phylogenetic Conflict in Large Data Sets by Increased Taxon Sampling. Syst. Biol. 55:522-529.

Heidrich P, Wink M. 1994. Tawny owl(*Strix aluco*) and Hume's tawny owl(*Strix butleri*) are distinct species: evidence from nucleotide sequences of the cytochrome b gene. Z. Naturforschung C. 49:230-234.

Heidrich P, Wink M. 1998. Phylogenetic relationships in holarctic owls (Order Strigiformes): evidence from nucleotide sequences of the mitochondrial cytochrome b gene. Holarctic Birds of Prey'.editors Chancellor RD, Meyburg BU, Ferrero JJ. pp:73–87.

Heidrich P, Konig C, Wink M. 1995a. Molecular phylogeny of South American Screech Owls of the *Otus atricapillus* complex(Aves: Strigidae) inferred from nucleotide sequences of the mitoch ondrial cytochrome b gene. Z. Naturforschung C. 50:294-302.

Heidrich P, König C, Wink M. 1995b. Bioakustik, Taxonomie und molekulare Systematik amerikanischer Sperlingskäuze (Strigidae: Glaucidium spp.). Stuttgarter Beitrage zür Naturkunde A.

Helbig A, Kocum A, Seibold I, Braun M. 2005. A multi-gene phylogeny of aquiline eagles (Aves: Accipitriformes) reveals extensive paraphyly at the genus level. Mol. Phyl. Evol. 35:147-164.

Hendrickson S, Bleiweiss R, Matheus J, de Matheus L, Jácome N, Pavez E. 2003. Low genetic variability in the geographically widespread Andean Condor. Condor 105:1-12.

Hendy MD, Penny D. 1989. A framework for the quantitative study of evolutionary trees. Syst. Zool. 38: 297–309.

Hennig W. 1966. Phylogenetic systematics. Univ. Illinois Press, Urbana.

Hillis DM. 1998. Taxonomic Sampling, Phylogenetic Accuracy, and Investigator Bias. Syst. Biol. 47:3-8.

Hillis DM, Moritz C. 1990. Molecular Systematics. Sinauer Publishers, Sunderland, Massachusetts.

Hillis DM, Moritz C, Mable BK. 1996. Molecular Systematics. Sinauer Associates, Sunderland, Massachusetts.

Hillis DM, Pollock D, McGuire J, Zwickl D. 2003. Is Sparse Taxon Sampling a Problem for Phylogenetic Inference? Syst. Biol. 52:124-126.

Hoefle U, Blanco J, Sauer-Guerth H, Wink M.. 2000. Molecular sex determination in Spanish imperial eagle (*Aquila adalberti*) nestlings and sex related variation in morphometric parameters. Raptor Biomedicine III. Zoological Education Network, Inc., Lake Worth, Florida:289–293.

Hoelzel AR. 1992. Molecular Genetic Analysis of Populations. IRL-Press, Oxford.

Holdaway RN. 1994. An exploratory phylogenetic analysis of the genera of the Accipitridae, with notes on the biogeography of the family. – In: Meyburg BU, Chancellor RD, editors. Raptor conservation today. World working group on birds of prey and owls, Berlin, pp. 601–649.

Houston DC. 1994. Family Cathartidae (New World Vultures). Handbook of the Birds of the World. Lynx Edicions. pp. 24-41.

Hume R. 1991. Owls of the world. Dragons World, Limpsfield.

Huxley T. 1867. On the classification of birds; and on the taxonomic value of the modifications of certain of the cranial bones observable in that class. Proc. Zool. Soc., London. 1867:415-472.

Irwin D, Bensch S, Irwin J, Price T. 2005. Speciation by distance in a ring species. Science. 307: 414-416.

IUCN. 2006. IUCN Red List of Threatened Species. [http://www.iucnredlist.org].

Jarvis ED et al. 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. Science. 346: 1320-1331.

Johnson J, Lerner H, Rasmussen P, Mindell D. 2006. Systematics within *Gyps* vultures: a clade at risk. BMC Evol. Biol. 6:1471-2148.

Kahn N, John J, Quinn T. 1998. Chromosome-specific intron size differences in the avian CHD gene provide an efficient method for sex identification in birds. The Auk 115:1074-1078.

Karp A, Isaac P, Ingram D. 1998. Molecular Tools for Screening Biodiversity. Chapman & Hall, London.

König C, Weick F, Becking J. 1999. Owls: a guide to the owls of the world. Pica Press, Robertsbridge.

Kruckenhauser L et al. 2004. Genetic vs. morphological differentiation of Old World buzzards (genus Buteo, Accipitridae). Zoologica Scripta 33:197-211.

Kumar S, Filipski A. 2007. Multiple sequence alignment: In pursuit of homologous DNA positions. Gen Res 17:127-135.

Lee MSY. 2001. Unalignable sequences and molecular evolution. Trends Ecol. Evol. 16:681-685.

Lerner H, Mindell D. 2005. Phylogeny of eagles, Old World vultures, and other Accipitridae based on nuclear and mitochondrial DNA. Mol. Phyl. Evol. 37:327-346.

Lopez-Herraez D, Schäfer H, Mosner J, Fries H, Wink M. 2005. Comparison of the exclusion power between microsatellite and single nucleotide polymorphism (SNP) markers in individual identification and parental analysis in a Galloway cattle population. Z. Naturforsch C. 60:637-643.

Mahmood MT, McLenachan PA, Gibb GC, Penny D. 2014. Phylogenetic position of avian nocturnal and diurnal raptors. Gen. Biol. Evol. 6:326–332.

Marks J, Cannings R, Mikkola H. 1994. Family Strigidae (typical owls). Handbook of the birds of the world. Lynx Edicions. pp. 76-242.

Masuda R et al. 1998. Genetic characterization of endangered Japanese Golden Eagles (*Aquila chrysaetos japonica*) based on mitochondrial DNA D-loop sequences and karyotypes. Zoo. Biol. 17:111-121.

Mayr G. 2009. Paleogene Fossil Birds. Springer Berlin Heidelberg. pp 156-157.

McCormack JE et al. 2013. A phylogeny of birds based on over 1,500 Loci collected by target enrichment and high-throughput sequencing. PloS One. 8: e54848.

Mindell D. 1997. Avian Molecular Evolution and Systematics. Academic Press, San Diego.

Mindell D et al. 1997. Phylogenetic relationships among and within select avian orders based on mitochondrial DNA. Avian Mol. Evol. Syst:213–247.

Mindell, D et al. 1999. Interordinal Relationships of Birds and Other Reptiles Based on Whole Mitochondrial Genomes. Syst. Biol. 48:138-152.

Mitchell A, Mitter C, Regier J. 2000. More Taxa or More Characters Revisited: Combining Data from Nuclear Protein-Encoding Genes for Phylogenetic Analyses of Noctuoidea (Insecta: Lepidoptera). Syst. Biol. 49:202-224.

Morgan-Richards M et al. 2008. Bird evolution: testing the Metaves clade with six new mitochondrial genomes. BMC Evol. Biol. 8:20.

Morrison J, Maltbie M. 1999. Methods for gender determination of Crested Caracaras. J. of Rapt. Res. 33:128-133.

Mossel E, Steel M. 2004. A phase transition for a random cluster model on phylogenetic trees. Math. Biosci. 187:189-203.

Nesje M, Roed K. 2000. Sex identification in falcons using microsatellite DNA markers. Hereditas 132:261-263.

Newton A. 1896. A dictionary of birds. Adam and Charles Black, London.

Nittenger F, Haring E, Pinsker W, Wink M, Gamauf A. 2005. Out of Africa: phylogenetic relationships between *Falco biarmkus* and the other Hierofalcons (Aves: Falconidae). J. Zool. Syst. Evol. Res 43:321-331.

Novikoff A, Holtzman E. 1976. Cells and organelles. Holt, Rinehart & Winston of Canada Ltd

O'Hara R. 1988. Diagrammatic classifications of birds, 1819-1901: Views of the natural system in 19th-century British ornithology. Acta XIX Congressus Internationalis Ornithologici (H. Ouellet, ed.). National Museum of Natural Science, Ottawa, Ontario, Canada:2746-2759.

Olsen J, Wink M, Sauer-Gurth H, Trost S. 2002. A new *Ninox* owl from Sumba, Indonesia. Emu 102:223–231.

Olson S. 1982. A critique of Cracraft's classification of birds. The Auk 99:733-739.

Pacheco MA et al. 2011. Evolution of modern birds revealed by mitogenomics: timing the radiation and origin of major orders. Mol. Biol. Evol. 28:1927–1942.

Paton T, Haddrath O, Baker A. 2002. Complete mitochondrial DNA genome sequences show that modern birds are not descended from transitional shorebirds. Proc. Royal Soc. B: Biol. Sci. 269:839.

Paulson D. 1973. Predator polymorphism and apostatic selection. Evolution 27: 269–277.

Pearlstine E. 2004. Variation in mitochondrial DNA of four species of migratory raptors. J. Rapt. Res. 38:250-255.

Peters JL. 1931. Check-list of the birds of the world. Vol. 1. Harvard Univ. Press, Cambridge.

Phillips MJ, Delsuc F, Penny D. 2004. Genome-scale phylogeny and the detection of systematic biases. Mol. Biol. Evol. 21:1455-1458.

Phillips MJ, Gibb GC, Crimp EA, Penny D. 2010. Tinamous and moa flock together: mitochondrial genome sequence analysis reveals independent losses of flight among ratites. Syst. Biol. 59: 90–107.

Poe S. 2003. Evaluation of the Strategy of Long-Branch Subdivision to Improve the Accuracy of Phylogenetic Methods. Syst. Biol. 52:423-428.

Poe S, Swofford D. 1999. Taxon sampling revisited. Nature 398:299-300.

Pollock D, Zwickl D, McGuire J, Hillis D. 2002. Increased taxon sampling is advantageous for phylogenetic inference. Syst. Biol.:664-671.

Poole A. 1994. Family Pandionidae (Osprey). Handbook of the Birds of the World. Lynx Edicions. pp. 42-50.

Poole AM, Phillips MJ, Penny D. 2003. Prokaryote and eukaryote evolvability. BioSystems. 69: 163-185

Prakash V et al. 2003. Catastrophic collapse of Indian white-backed *Gyps bengalensis* and long-billed *Gyps indicus* vulture populations. Biol. Conserv. 109:381-390.

Pratt RC et al. 2009. Towards resolving deep Neoaves phylogeny: data, signal enhancement and priors. Mol. Biol. Evol. 26:313–326.

Proudfoot G, Gehlbach F, Honeycutt R. 2007. Mitochondrial DNA Variation and Phylogeography of the Eastern and Western Screech-Owls. The Condor 109:617-628.

Proudfoot G, Honeycutt R, Slack R. 2006. Mitochondrial DNA variation and phylogeography of the ferruginous pygmy-owl (*Glaucidium brasilianum*). Conserv. Genet. 7:1-12.

Quinn T. 1997. Molecular evolution of the mitochondrial genome. *In:* Mindell DP, editor. Avian molecular evolution and systematics. Academic Press, San Diego. pp. 3-28

Randi E, Fusco G, Lorenzini R, Spina F. 1991. Allozyme divergence and phylogenetic relationships within the Strigiformes. Condor:295-301.

Riesing M, Kruckenhauser L, Gamauf A, Haring E. 2003. Molecular phylogeny of the genus *Buteo* (Aves: Accipitridae) based on mitochondrial marker sequences. Mol. Phyl. Evol. 27:328-342.

Ristow D, Wink M. 2004. Seasonal variation in sex ratio of fledging Eleonora's Falcon, *Falco eleonorae*. J. Rapt. Res. 38:320-325.

Ristow D, Witte L, Wink M. 2004. Sex Determination of Nestlings in Eleonora's Falcon *Falco eleonorae*: plumage characteristics and molecular sexing. In: Raptors Worldwide, editors. Chancelor RD, Meyburg BU.:459–466.

Roberts T. 1991. The birds of Pakistan. Oxford University Press, New York.

Roques S, Godoy J, Negro J, Hiraldo F. 2004. Organization and variation of the mitochondrial control region in two vulture species, *Gypaetus barbatus* and *Neophron percnopterus*. J. Heredit. 95:332-337.

Roulin A, Wink M. 2004. Predator-prey relationships and the evolution of genetic colour polymorphism: a comparative analysis in diurnal raptors. Biol. J. Linn. Soc. 81: 565–578.

Sanderson MJ, Boss D, Chen D, Cranston KA, Wehe A. 2008. The PhyLoTA browser: processing GenBank for molecular phylogenetics research. Syst. Biol. 57**:**335-346.

Sclater P. 1880. Remarks on the present State of the Systema Avium. Ibis 22:340-350, 399-411.

Seibold I, Helbig A, Meyburg B, Negro J, Wink M. 1995b. Genetic differentiation and molecular phylogeny of European Aquila eagles (Aves: Falconiformes) according to cytochrome b nucleotide sequences. Eagle Studies (World Working Group on Birds of Prey, Berlin)

Seibold I, Helbig A. 1995a. Evolutionary history of New and Old World vultures inferred from nucleotide sequences of the mitochondrial cytochrome b gene. Philosoph. Transac. Biol. Sci. 350:163-178.

Seibold I, Helbig A. 1996. Phylogenetic relationships of the sea eagles (genus *Haliaeetus*): reconstructions based on morphology, allozymes and mitochondrial DNA sequences. J. Zool. Syst. Evol. Res. 34:103-112.

Sibley CG, Monroe BL. 1990. Distribution and Taxonomy of Birds of the World. Yale University Press, New Haven, CT. U.S.A.

Sibley CG, Ahlquist JE. 1990. Phylogeny and classification of birds. Yale University Press, New Haven.

Simon C, Buckley TR, Frati F, Stewart JB, Beckenbach AT. 2006. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. Ann. Rev. Ecol. Evol. Syst. 37:545-579.

Slack KE, Delsuc F, Mclenachan PA, Arnason U, Penny D. 2007. Resolving the root of the avian mitogenomic tree by breaking up long branches. Mol. Phyl. Evol. 42:1-13.

Storch V, Welsch U, Wink M. 2001. Evolutionsbiologie. Springer Verlag, Heidelberg.

Stresemann E. 1975. Ornithology from Aristotle to the Present. Harvard University Press, Cambridge, Massachusetts.

Suh A et al. 2011. Mesozoic retroposons reveal parrots as the closest living relatives of passerine birds. Nat. Commun. 2: 443.

Swofford, DL. 2002. PAUP – Phylogenetic analysis using parsimony. Version PAUP 4.0b10.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA 4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24:1596-1599.

Thiollay J. 1994. Family Accipitridae (hawks and eagles). Handbook of the Birds of the World. Lynx Edicions. pp. 52–205.

Voelker G, Edwards SV. 1998. Can weighting improve bushy trees? Models of cytochrome b evolution and the molecular systematics of pipits and wagtails (Aves: Motacillidae). Syst. Biol. 47:589–603.

Wang N, Braun EL, Kimball RT. 2012. Testing hypotheses about the sister group of the passeriformes using an independent 30-locus data set. Mol. Biol. Evol. 29: 737–750.

Weick F. 2006. Owls (Strigiformes) Annotated and Illustrated Checklist. Springer Verlag, Heidelberg.

White C, Olsen V, Kiff L. 1994. Family Falconidae (falcons and caracaras). Handbook of the birds of the world 2:216-275.

White C, Olsen V, Kiff L. 1994. Family Falconidae (falcons and caracaras). Handbook of the birds of the world. Lynx Edicions. pp. 216-275.

Wink M. 1995. Phylogeny of Old and New World vultures (Aves: Accipitridae and Cathartidae) inferred from nucleotide sequences of the mitochondrial cytochrome b gene. Z. Naturforschung C. 50:868-882.

Wink M. 1998. Application of DNA-markers to study the ecology and evolution of raptors. In: Holarctic Birds of Prey, editors Chancellor RD, Meyburg BU. Ferrero JJ. ADENEX & WWGBP, Merida and Berlin. pp. 49–71.

Wink M. 2000. Advances in DNA studies of diurnal and nocturnal raptors. *in:* Chancellor RD, Meyburg BU, editors. Raptors at risk: WWGBP. Handcock House, Berlin. pp. 831–844

Wink M. 2007. Systematics. *In:* Bird DM, Bildstein KL, editors. Raptor research and management techniques. Hancock House Publishers, WA. pp. 57-73

Wink M, Sauer-Gürth H. 2000. Advances in the molecular systematics of African raptors. *In:* Chancellor RD, Meyburg BU, editors. Raptors at risk: WWGBP. Handcock House, Berlin. pp. 135–147

Wink M, Sauer-Gürth H. 2004. Phylogenetic relationships in diurnal raptors based on nucleotide sequences of mitochondrial and nuclear marker genes. *In:* Chancellor RD, Meyburg BU, editors. Raptors at risk: WWGBP. Handcock House, Berlin. pp. 483–498

Wink M, Seibold I. 1996. Molecular phylogeny of Mediterranean raptors (families Accipitridae and Falconidae). *In:* Muntaner JY, Mayol J, editors. Biología y conservación de las rapaces mediterráneas. pp. 335-344.

Wink M, Heidrich P. 1999. Molecular evolution and systematics of owls (Strigiformes). Owls: A Guide to Owls of the World. Yale University Press, New Haven, Connecticut. pp. 39–57.

Wink M, Heidrich P. 2000. Molecular systematics of owls (Strigiformes) based on DNA-sequences of the mitochondrial cytochrome b gene. *In:* Chancellor RD, Meyburg BU, editors. Raptors at risk: WWGBP. Handcock House, Berlin. pp. 819–828

Wink M, Sauer-Gürth H, Gwinner E. 2002. A molecular phylogeny of stonechats and related turdids inferred from mitochondrial DNA sequences and genomic fingerprinting by ISSR-PCR. Br. Birds 95:349–355.

Wink M, Heidrich P, Sauer-Gürth H, Elsayed AA, Gonzalez J. 2009. Molecular phylogeny and systematics of owls (Strigiformes). pp. 42-63 *In:*. Weick CKF, editor. Owls of the world. Yale University Press.

Wink M, Seibold I, Lotfikhah F, Bednarek W. 1998. Molecular systematics of holarctic raptors (Order Falconiformes). Holarctic Birds of Prey. Adenex, WWGBP, Calamonte. pp. 29–48.

Wink M, Heidrich P, Fentzloff C. 1996. A mtDNA phylogeny of sea eagles (genus *Haliaeetus*) based on nucleotide sequences of the cytochrome b-gene. Biochem. Syst. Ecol. 24:783-791.

Zwickl D, Hillis D. 2002. Increased Taxon Sampling Greatly Reduces Phylogenetic Error. Sys. Biol. 51:588-598.

2 *RAPTORBASE*: A DATABASE OF RAPTORS' NUCLEOTIDE SEQUENCES.

Muhammad Tariq Mahmood, Mauro Truglio and David Penny

GenBank is an indispensable source of biological data. To ask specific question related to raptor phylogenetics, it was felt that currently there is need of a database subset of GenBank which may provide a general overview of the available nucleotide data. *RaptorBase* is a database of raptor sequences which was created to address this need. GenBank has some shortcomings in its taxonomy, but we follow the same taxonomy, for consistency and a good flow of information. *RaptorBase* is a MySQL database of raptor sequences, it is a web application written in Python which queries the *Entrez* database of the NCBI, through its *eUtils* programming utilities. Since it is online, it is able to synchronize the nucleotide data with GenBank. It can display a matrix of available or otherwise of nucleotide information. The model of this database can be easily modified for other groups of organisms as well. The user gets the information in the form of a tabular display of gene names and species with ability to download the sequences. It will also help the users to get an idea of gaps in sequence availability. It is hoped that the database will help to direct the sequencing efforts related to raptors. Our aim is to link the current database with other biological databases with information relevant to the raptors.

MTM selected NCBI tools, conceived and designed the project. MT created the MySQL database, wrote the Python script to integrate the data pipeline with the local database. DP gave general oversight and supervision.

2.1 Introduction

Although computers have only been with us for a few decades, their impact has been profound, rivalling that of the telephone, automobile, or television. Their presence is felt by us all, in every field of life. In phylogenetics, the use of sequence data has increased with advances in sequencing technology and computing. It has, in turn, increased phylogeneticists' reliance on online GenBank and some of its associated tools. GenBank (Benson et al. 2013) is the largest widely accessible and easily queried database of biodiversity. But phylogenetic research was not the primary reason for the collection of most of its data. Here we discuss the problems involved in its processing and its optimization for utility in molecular phylogenetics. In particular, our main focus is the development of a web enabled subset of GenBank, tailored for nucleotide sequences related to the raptors/the birds of prey.

GenBank aims to have all the sequence data for every taxon, and lists taxa down to subspecies, and with information of the location where the sequence data comes from. The "taxonomy tree" for all taxa in GenBank is backed with sequences in the database, but it is known to have some shortcomings as a phylogeny. For instance, the new name of species proposed at the time of data submission, sometimes remains unchanged in GenBank, even if the name is revised later on. In that case, GenBank depends on the submitters discretion, who may or may not chose to update the submission (Federhen, 2003). Most of the times, these problems are observed by the researchers only when they come across their familiar clades. On the other hand, it provides a large and a more comprehensive set of candidate sequences for phylogenetic analysis.

Free and open access to, and deposition of, sequence data to the GenBank has encouraged a huge number of computational and empirical analyses of data across many disciplines, including phylogenetics. There are many instances of comprehensive phylogenetic studies based on bigger sample size of taxa or loci, which range from investigation on papilionoid legumes (McMahon and Sanderson, 2006) to the questions related to the resolution of the tree of life (Ciccarelli et al. 2006), and from the present day mammals (Bininda-Emonds et al. 2007) to the ray-finned fish (Li et al. 2007).

The needs and utilization of nucleotide sequences varies for a population biologist, a conservation biologist, and phylogeneticist. A phylogeneticist is often interested in the sequence data at a broad taxonomic level, while a population biologist or a conservation biologist might need a bigger sample set of certain loci but at a relatively lower level of taxonomy. Therefore, the ability to examine the entire matrix of available sequences at several levels of taxonomic specificity in GenBank is needed. This will not only allow efficient construction of multiple gene data sets, but will also identify the missing data or gaps in data sets. Of course, it will also direct sequencing efforts in future. In this study, we are mainly interested in building a database of raptor (birds of prey) sequences that are useful not only in phylogenetic but in population genetic analyses too. The example of the present effort could be followed, in future, to extend or build such a database across different levels of taxonomy among birds or of a different group of organisms (avian and non-avian), and so we expect it will be adopted by others. New developments in this regard include PhyLoTa (Sanderson et al. 2008). The PhyLoTa is an informatics processing pipeline. It presents clusters of homologous sequences at different levels of taxonomical hierarchy in the NCBI. We imported some ideas from there.

Raptors Database (‘*RaptorBase*’) is an online program which allows for a versatile management and retrieval of the data specific to avian raptors. The objective of the system is to run on the worldwide web and allow users easy access to the necessary relevant data. Currently family, genus, species and genes (mitochondrial and nuclear) sequenced for each of the species (together with their actual sequences) is the target. The information on locality of species would also be included. So, the major objectives can be elaborated through the following points:

- Keep data up to date with NCBI and GenBank.
- Select genes and species for multiple alignments, so that output/ export of sequences may be converted into ‘Nexus’ file format by using program Clustal W.
- User may search the system to get target information and view results immediately in form of printed/formatted reports.

General issues

The three most significant problems in a GenBank sequence record search are:

1- The taxon name

2- The gene name or a specific sequence region

3- Taxonomy

There are instances when errors have been noted in taxon names due to the original specimen misidentification (Vilgalys, 2003; McMahon and Sanderson, 2006), and we illustrate this later in chapter 4. Standardization of sequence names is a difficult, but necessary, task.

Since sequences often have different features (e.g. introns, exons, and genes), accepted scientific names by GenBank are not always helpful for easy retrieval of sequence data, particularly by sequence name or a text query. There are instances where such issues have been addressed through a list of synonyms and approximate matching algorithms (Bininda-Emonds, 2008).

The authors of the NCBI/GenBank taxonomy database accept that some parts of it are 'less' developed. They do not perform the phylogenetic analysis themselves, that may support their 'Taxonomy Database'. Therefore phylogenetic literature is regarded as the most reliable source of information (Federhen, 2003). One example of the above mentioned taxonomy problem of NCBI, is the inclusion of Osprey (*Pandion haliaetus*) under the family Accipitridae, rather than as a monotypic family of its own, i.e. Pandionidae.

2.2 Database Development

2.2.1 Raptors Data Manager – a desktop version

The initial version of the present '*RaptorBase'* is a local or desktop version. This is our first deliberation towards building and maintaining a database of raptor nucleotide sequences. The objective of the system is for an offline program that allows its users an easy access to the necessary relevant data. Initially, family, genus, species and genes (mitochondrial and nuclear) sequenced for each of the species - along with their actual

sequences - were the target. The information on locality of species is also included. So, the major objectives can be elaborated through following points;

A- Enable easy entry and retrieval of the data.

B- Keep data up-to-date with NCBI.

C- Select genes and species for multiple alignments, so that output of sequences may be converted into 'Nexus' file format by using the program Clustal W.

Raptors Data Manager has dual operation i.e. data entry and interactive searching. The user has the ability to enter data to each category (e.g. family, genus, species etc.) using a Graphical User Interface (GUI). The desktop (offline local database) version is based on the manual curation download, and then upload the data to a desktop computer. The update process depends on the regular email alerts from NCBI Entrez database. The user may search the system to get target information, and view results immediately in form of printed/formatted reports. The data can be used to make more informed decisions in designing research questions.

2.2.1.1 System Sections

2.2.1.1.1 General Section

General section has the following capabilities:

I. Add Family
II. Add Genus
III. Add Species
IV. Add Genes
V. Add Locality

2.2.1.1.2 Manager

Manager manages the hierarchal relationship between genus, species and genes. It provides the ability to add species information including all gene names, their accession numbers, sequences and annotation. It further provides following functionalities;

I. Add Species data
II. View whole data by selecting any genus or species.
III. Update records/ rename or reclassify e.g. if species A now becomes B.

2.2.1.1.3 Searching and Reports

This section shows the output of the system that enables the user to search data and view printable formatted reports. It provides the following functionalities:

I. Forward hierarchal search along with following reports:
   a. Detailed hierarchal report
   b. Sequence report
   c. Species locality report

II. Advanced search with the following options:
   a. Search detailed records on the basis of the gene name and display the report
   b. Search family and genus on the basis of species
   c. Search genes on the basis of gene's type

Appendices I and II show the Graphical User Interface (GUI) and its ability to search within the database respectively.

After maintaining the *Raptors Data Manager* for some time, we realized its reliance on manual curation and maintenance would not be a long-term sustainable task. The maintenance would also cost undue loss of time and effort. Particularly with the advancement of next generation sequencing technologies, there is exponential growth in the availability of DNA sequence data. We realized the need of an efficient and online database of raptor sequences, which would be equally useful for raptor phylogeny, as well as the population genetics research community.

2.2.2 RaptorBase – an online version

Before describing *RaptorBase*, it will be useful to know whether different online databases are standardized or not. The annotation, vocabulary and data structures used vary widely between the databases. The searching of the biological databases through crude text searches may not be efficient enough to pick up all the annotated information. In case of crude text searches, the annotated information can be a hit and miss. Therefore much of the current research is focused on standardizing the data structure

and vocabulary, to increase the querying abilities within and between them (Baker and Brass 1998).

There are good examples of specialist databases which present organized information from several sources. They use knowledge of experts to organize this information. The examples of such databases are, 'Ecocyc' – and encyclopedia of *E. coli* genes (Karp et al. 1997) and FlyBase – a comprehensive and useful database on *Drosophila* (The FlyBase Consortium 1997). Entrez is a collection of different databases including 'Nucleotide' database at NCBI server. It is a hypertext navigable database, which means it can be accessed through form-based single query system. Hypertext links enable the users to navigate between different member databases. Our aim is to build a comprehensive database pipeline of nucleotide data at NCBI, using relational architecture. Our present database, the *RaptorBase* - a cross-platform and cross-browser web application - is a subset of Entrez. It exploits the GenBank file format and its different fields to harvest the name of taxa, accession number, actual sequences, and their locality information. The GenBank data model is shown as Appendix III.

A summary on the sequences are reported, along with actual sequences. But we assume that most of the users want extensive manual control over this. At this stage, the user is able to see sequence and diversity summary information. Besides, this information also provides a link to the NCBI server, which in turns provides many links to other biodiversity databases of Entrez. The examples of these databases are PubMed, Protein, and Structure. Moreover, we use the NCBI classification as a convenient hierarchical framework to browse the sequence diversity. The important information at this stage is the user access to the sequence data matrix, at a particular node of the raptors tree. It may or may not be the complete matrix depending on the availability at NCBI. Although sequences are able to be downloaded at different levels of NCBI taxonomy/hierarchy, they may not necessarily reflect on the actual position of that particular node in the overall avian or raptor tree.

2.2.2.1 Methodology and system tools

2.2.2.1.1 NCBI eUtils and the customized data pipeline

The Entrez Programming Utilities, or eUtils, are a set of seven programs which run on the server-side. Each of them is a fixed URL syntax allowing for certain number of parameters to be retrieved from NCBI. The parameters, chosen by the user, are actually translated into certain values required by NCBI software. These softwares then search and retrieve the requested data depending on the values it receives through a particular eUtil or a set of them. In short, eUtils are the structured interface to the Entrez database system.

2.2.2.1.2 Computational language and the Database

The *RaptorBase* web-application is built using CherryPy (http://www.cherrypy.org), one of the most stable web application frameworks written using the Python language. Besides a rapid and easy-to-use interface, CherryPy can also work as a web server, which makes it an excellent all-in-one solution.

The workflow of *RaptorBase* consists of inputs, output and the core part. The user provides the Family or the lower level taxonomic information of interest as inputs (‘Terms’) to the web application through a Graphic User Interface (GUI) (Appendix IV). The request is sent to the database as a ‘MySQL query’ which returns a table of results in the form of a Python list. This is elaborated by an internal script that produces the output web page. The results page consists of a table that indicates, for each species or sub-species, the GenBank accession number(s) for each gene. The results can be downloaded at single or multiple species/sub-species level through a check box and a download link/button.

The core component of the *RaptorBase* is a MySQL database created from a subset of the NCBI’s *nucleotide* database (DB). A single python script, which can be manually executed by the administrator or set as a scheduled recurrent command, controls the creation of the *RaptorBase* DB and, in case it is already in place, its synchronization with the online NCBI version. This script uses Entrez library of the BioPython (Cock et al. 2009) to call the eUtils functions eSearch and eFetch, in order to query the NCBI

database and retrieve the results respectively. The results are received by the script as xml documents, which are in turn parsed and stored as MySQL records in *RaptorBase's* DB.

The DB is organized in four (4) related tables: families, genera, species and genes. The gene table in particular stores all the main features of the GenBank record as identifier, gene name, taxonomy and sequence (see Figure 1 for the details).

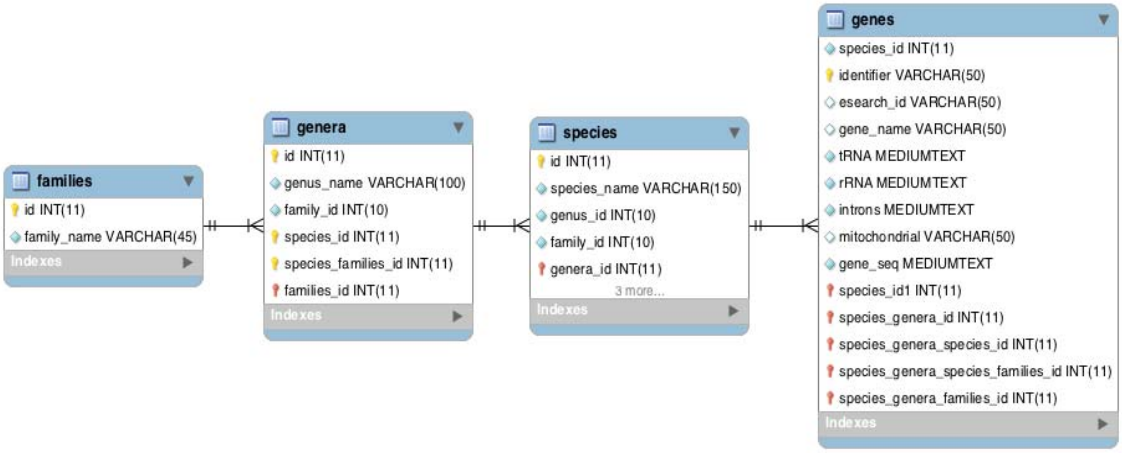


Figure 1. Four related tables of MySQL database (DB)

Figure 2 shows the workflow for the online database which is the secondary database or subset of a primary database. The source or the primary database is the Entrez nucleotide database, a subsection of NCBI.

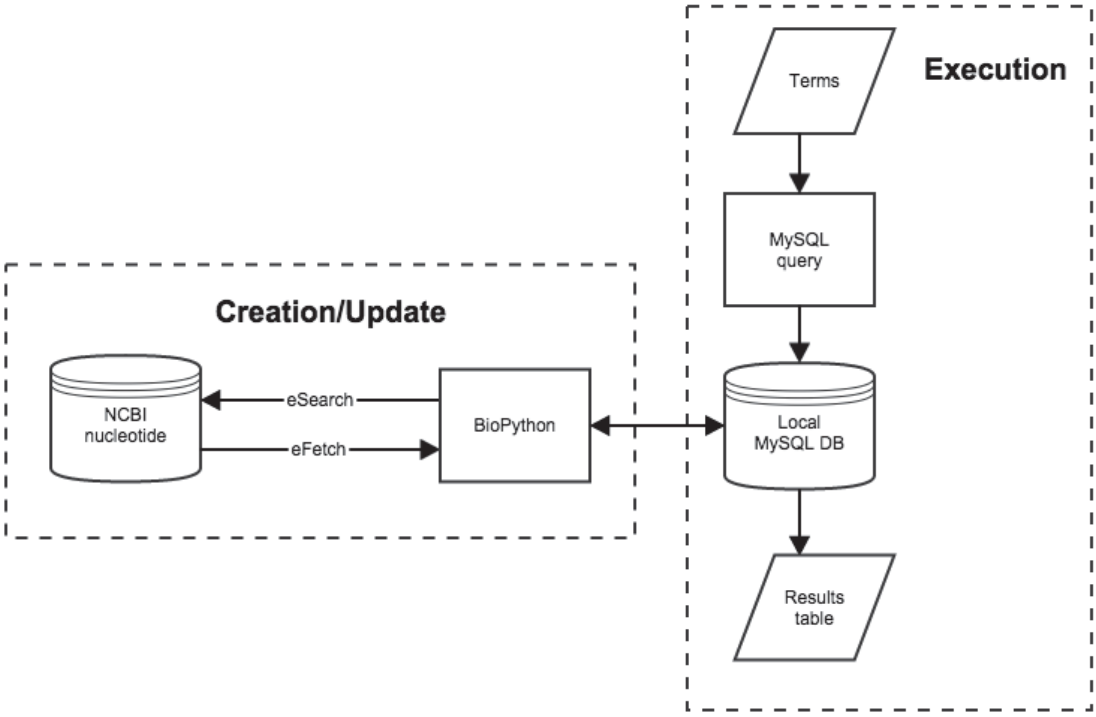


Figure 2. Workflow for the online database

2.3 Discussion

The request sent to the database as a 'MySQL query' returns a table that indicates, for each species or sub-species, the GenBank accession number(s) for each gene (Figure 3). The results can be downloaded at single or multiple species/sub-species level through a check box and a download link/button.

Species	EGR1	MUSK	PCBD1	cytb	COX1	...
Tyto glaucops	-	-	-	AJ003939.1	-	...
Tyto novaehollandiae calabyi	-	-	-	-	-	...
Phodilus badius	EU738838.1 EU739000.1	EU739856.1	EU738512.1	AJ004041.1 AJ004042.1 KJ456384.1	-	...
Tyto novaehollandiae	-	-	-	EU349009.1	-	...
Tyto sororcula sororcula	-	-	-	-	-	...
Tyto inexspectata	-	-	-	-	-	...
Tyto longimembris	-	-	-	EU349008.1	-	...
Tyto manusi	-	-	-	-	-	...
Tyto alba bargei	-	-	-	EU349000.1	FJ465378.1 FJ465379.1 FJ465380.1 FJ465381.1 KF432207.1	...
Tyto alba pratincola	-	-	-	AJ004073.1	-	...
Tyto furcata tuidara	-	-	-	EU349006.1	-	...
Tyto delicatula delicatula	-	-	-	EU349001.1	-	...
Tyto delicatula lulu	-	-	-	EU348996.1	-	...

Figure. 3. The output/ result table of species-sequences matrix.

The first column represents the taxonomic diversity, while the rest of the columns show the sequence diversity against the respective taxa. This view gives a quick impression of the evenness and density of the distribution of sequence data. Such information could be a useful guide for supertree construction (Sanderson et al. 2007; Sanderson and Driskell, 2003). The selection of the sequences can be retrieved as FASTA file format. The primary resource which the database provides is the ability to retrieve sequence information for the entire group of diurnal and nocturnal raptors, aside from their taxonomic diversity in GenBank.

The output of results in the present form could be very useful to identify the gaps in data availability. This, in turn, would help identify research priorities and direct the sequencing efforts. The present model of database could be equally useful for other groups of organisms, to gather the sequence information which has wide applications in research. But in that case, presence of sequence information for model organisms could skew one's perspective on sequence diversity. There could be an option introduced to either include or exclude the model organisms, if any, from the final table.

2.3.1 Link to other knowledge bases

We wish to gradually move towards an 'Object Oriented' model to integrate the related information from different databases, for instance, AviBase (http://avibase.bsc-eoc.org) which is enabled with different checklists and species distribution data for birds, and the IUCN red data list (IUCN 2014) which shows conservation and distribution status of different species. This will not only show the total number of species identified/ described so far, but also the completeness index matrix of their presence/ absence on NCBI/GenBank. Since we follow an accepted list of names in NCBI taxonomy, which is a small subset of names as compared to a huge list of names in taxonomic literature. It will be hard to work out the interoperability of *RaptorBase* and AviBase or IUCN. But it could be possible when the names of *RaptorBase* are mapped onto larger list of names in other databases mentioned above. But still, this interoperability will be through crude text searches. Unlike the above mentioned databases, there are some efforts which are interoperable with external databases, for example, the uBio system. It responds through XML queries which contain relevant taxonomic names. There is a strong need for object oriented biological databases, which could interoperate at the level of object rather than querying the detailed data representation. The existing databases also need to be re-engineered to standardize their data structure for easier querying within and between them (Baker and Brass 1998).

References

Baker PG, Brass A. 1998, Recent development in biological sequence databases. Curr. Opin. Biotech. 9: 54–58.

Benson, DA, Karsch-Mizrachi I, Lipman, J. Ostell, and D. L. Wheeler. 2013. GenBank. Nucleic Acids Res. 35:D21-D25.

Bininda-Emonds ORP. 2008. GenBankStrip.pl v. 2.0. Institut f ur Spezeille Zoologie und Evolutionsbiologie mit Phyletischem Museum. Friedrich-Schiller-Universitat. Computer program published by the author.

Bininda-Emonds ORP et al. 2007. The delayed rise of present-day mammals. Nature 446:507–512.

Ciccarelli FD et al. T. Doerks, C. von Mering, C. J. Creevey, B. Snel, and P. Bork. 2006. Toward automatic reconstruction of a highly resolved tree of life. Science 311:1283–1287.

Cock PJA et al. 2009. Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinf. 25:1422–1423

Consortium TF. 1997. FlyBase : a Drosophila database. Nucleic Acids Res. 25:63–66.

Federhen S. 2003. The taxonomy project. *In* The NCBI handbook http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=handbook. ch. 4.

IUCN .2014. IUCN Red List of Threatened Species. Version 2014.2. <www.iucnredlist.org>. Downloaded on 19 October 2014.

Karp PD, Riley M, Paley SM, Pellegrini-toole A, Krummenacker M. 1998. EcoCyc : Encyclopedia of Escherichia coli genes and metabolism. Nucleic Acids Res. 26:50–53.

Li CH, Orti G, Zhang G,. Lu GQ. 2007. A practical approach to phylogenomics: The phylogeny of ray-finned fish (Actinopterygii) as a case study. BMC Evol. Biol. 7:44.

McMahon M.M, Sanderson MJ. 2006. Phylogenetic supermatrix analysis of GenBank sequences from 2228 papilionoid legumes. Syst. Biol. 55:818–836.

Sanderson MJ, Boss D, Chen D, Cranston KA, Wehe A. 2008 The PhyLoTA browser: processing GenBank for molecular phylogenetics research. Systematic Biology, 57:335-346.

Sanderson MJ et al. 2007. Fragmentation of large data sets in phylogenetic analyses. In: Steel M, Gascuel O, editors. Reconstructing Evolution: New Mathematical and Computational Advances. Oxford University Press. p. 199-216

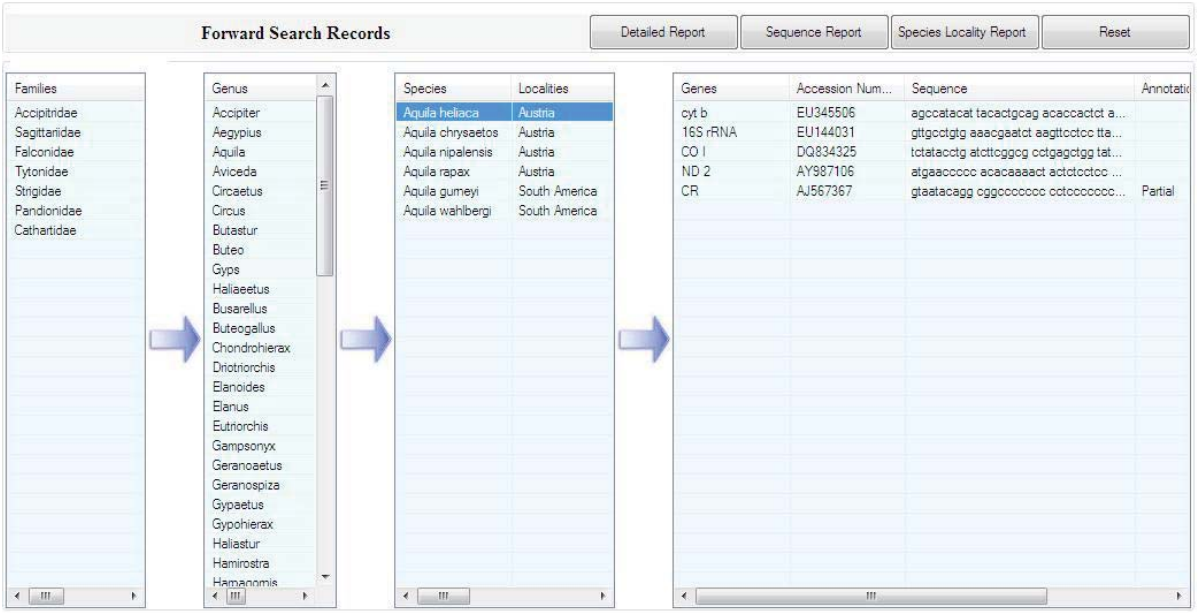
Sanderson MJ, Driskell AC. 2003. The challenge of constructing large phylogenetic trees. Trends Plant Sci. 8:374–9.

Vilgalys R. 2003. Taxonomic misidentification in public DNA databases. New Phytol. 160:4–5.

Appendices

Appendix I

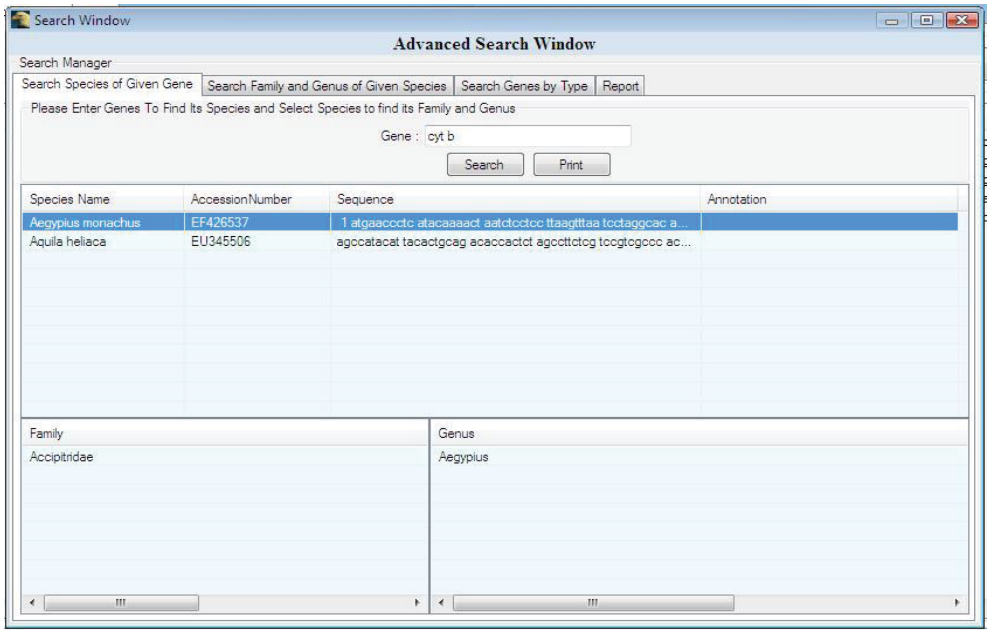
Search ability of *Raptors Data Manager* and data visualization.



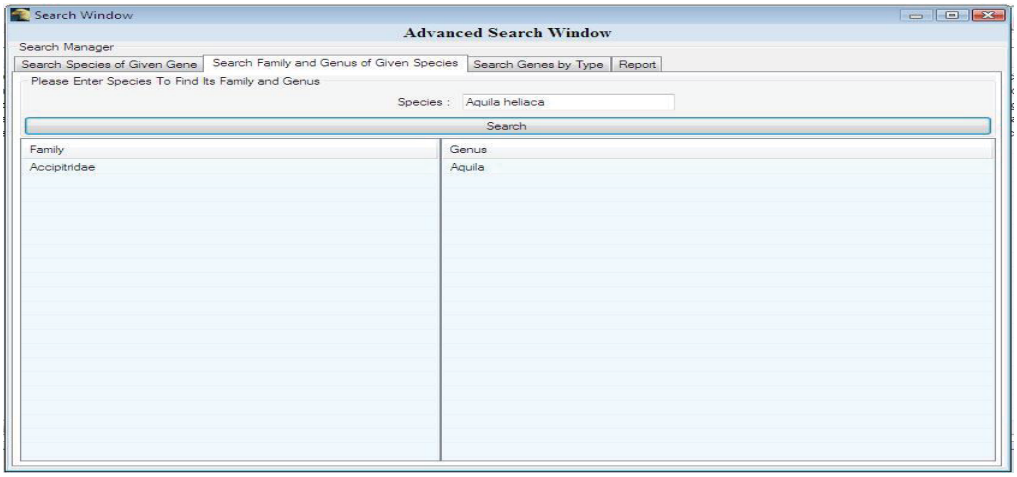
(i) Forward Search Window

Appendix II

Search ability of *Raptors Data Manager* by gene and species name.



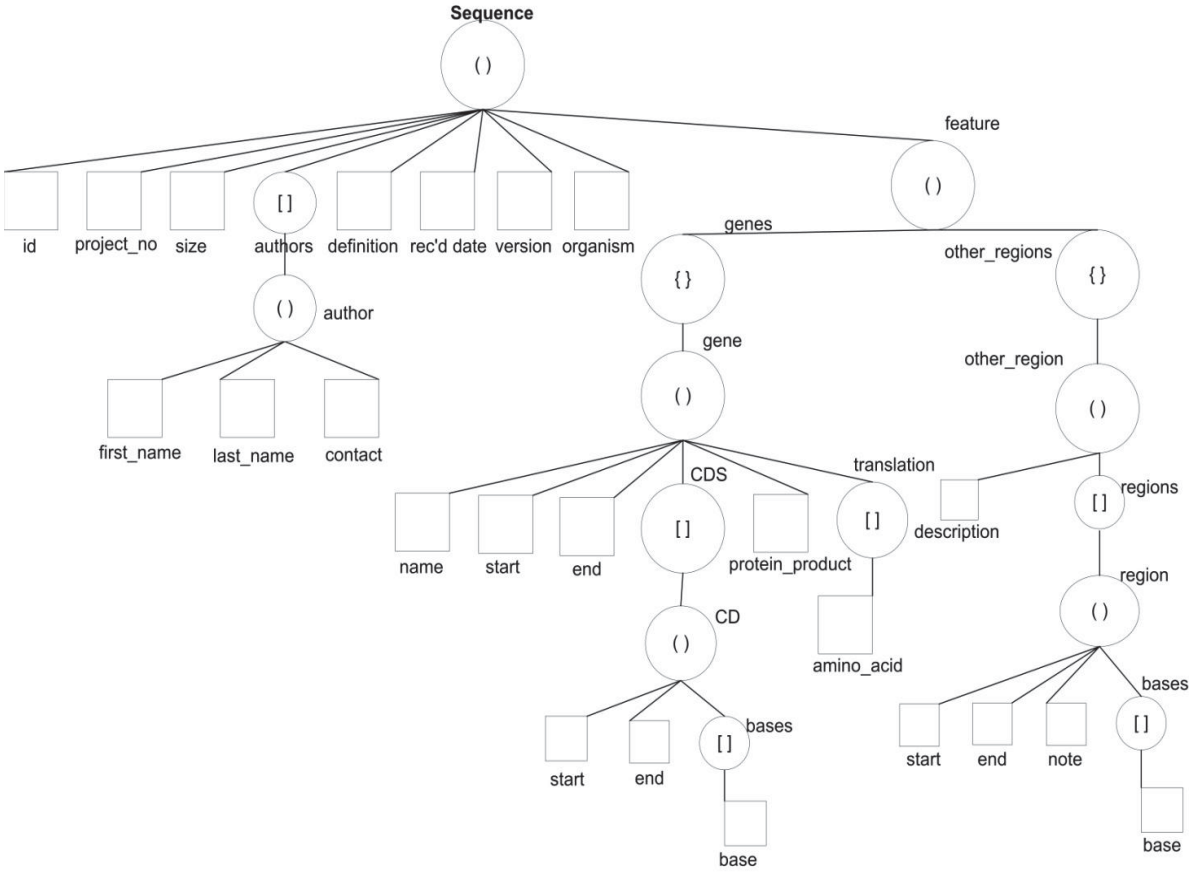
(ii) Search by Gene Name



(iii) Search by Species Name

Appendix III

GenBank data model. Each of the box/ circle represents the ‘identifier’ in GenBank file format. Every bit of the information needed from one of the fields mentioned below will be displayed through either a text query from the user or a drop down menu provided as part of its web interface.



Appendix IV

The Graphical User Interface (GUI) of the *RaptorBase.*

3 PHYLOGENETIC POSITION OF AVIAN NOCTURNAL AND DIURNAL RAPTORS

Muhammad Tariq Mahmood, Patricia A. McLenachan, Gillian C. Gibb and David Penny (2014) Phylogenetic position of avian nocturnal and diurnal raptors. *Genome Biology and Evolution*. 6: 326-332.

There are two classes of questions addressed in this chapter. The first class of question is a) whether the two families of owls (nocturnal raptors) do really come together phylogenetically, and b) whether the Secretarybird (*Sagittarius serpentarius*) is really the deepest divergence on the accipitrids (eagles) and the pandionid (Osprey) lineage. The second class of question concerns raptors as a group – whether the raptors are genuinely monophyletic. In this published chapter, three new avian mitochondrial genomes are reported, two from widely separated groups of owls [an Oriental Bay Owl (*Phodilus badius*) and the Spotted Owlet (*Athene brama*)], and then the Secretarybird. The phylogenetic analysis suggests that the two groups of owls do come together (it is not just long-branch attraction), and that the Secretarybird is the deepest divergence on the Accipitridae and the Pandionidae lineage. This is now agreed with both mitochondrial and nuclear sequences.

However, there is no evidence for the monophyly of the combined three groups of raptors (owls, eagles and falcons), and again this is agreed by nuclear and mitochondrial sequences. All three groups (owls, accipitrids (eagles) and falcons) do appear to be members of the 'higher land birds', and though there may not yet be full 'consilience' between mitochondrial and nuclear sequences for the precise order of divergences of the eagles, falcons and the owls, there is good progress on their relationships.

At each level of questions the conclusions are still accepted, and this is discussed again in the final summary chapter.

I did the DNA sequencing and primary analysis, PAM was responsible for important technical advice and laboratory supervision, GCG supplied the samples and was also a

co-supervised, GCG and DP gave general oversight and planning – all authors contributed to the final manuscript.

GENOME BIOLOGY AND EVOLUTION

SMBE

GBE

Phylogenetic Position of Avian Nocturnal and Diurnal Raptors

Muhammad Tariq Mahmood[1,*], Patricia A. McLenachan[1], Gillian C. Gibb[2], and David Penny[1]

[1]Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand

[2]Institute of Agriculture and the Environment, Massey University, Palmerston North, New Zealand

*Corresponding author: E-mail: t.mahmood@massey.ac.nz.

Accepted: January 14, 2014

Data deposition: This project has been deposited at GenBank under the accession numbers KF961184 (*Sagittarius serpentarius*), KF961183 (*Phodilus badius*), and KF961185 (*Athene brama*).

Abstract

We report three new avian mitochondrial genomes, two from widely separated groups of owls and a falcon relative (the Secretarybird). We then report additional progress in resolving Neoavian relationships in that the two groups of owls do come together (it is not just long-branch attraction), and the Secretarybird is the deepest divergence on the Accipitridae lineage. This is now agreed between mitochondrial and nuclear sequences. There is no evidence for the monophyly of the combined three groups of raptors (owls, eagles, and falcons), and again this is agreed by nuclear and mitochondrial sequences. All three groups (owls, accipitrids [eagles], and falcons) do appear to be members of the "higher land birds," and though there may not yet be full "consilience" between mitochondrial and nuclear sequences for the precise order of divergences of the eagles, falcons, and the owls, there is good progress on their relationships.

Key words: raptor evolution, mitochondrial genomes, owls, Secretarybird, Accipitridae, Strigiformes.

Introduction

There has been good progress in resolving issues about the deeper relationships of modern birds. But that still leaves the major group of birds as Neoaves, and here there is less certainty about basic divisions. However, there has been some good progress and, for example, both McCormack et al. (2013) with nuclear data and Gibb et al. (2013) with mitochondrial (mt) data have proposed a general group of "water carnivores" that includes Pelecaniformes, Ciconiiformes (including storks), and some related groups—but not the shorebirds (Charadriiformes). Similarly, a group often called the "higher land birds" has been proposed (Johansson et al. 2001; Ericson et al. 2006) that is quite distinct from the water carnivores. This higher land bird group includes groups such as the songbirds, parrots, owls, falcons, eagles, Piciformes, and Coraciiformes.

Here, we are particularly interested in the group of "raptors," both diurnal (e.g., falcons and eagles) and nocturnal (owls). Ideally, we would hope for "consilience" between nuclear sequences, mt sequences, and rare genomic changes (e.g., retroposons; Suh et al. 2011). The reasons why there is conflict and difficulty in resolution of the avian tree of life may be attributed to general issues such as taxon sampling, number of genes (size of data set), compositional bias, substitution saturation, and alignment issues (see Hernandez-Lopez et al. 2013; Kimball et al. 2013). Classical knowledge is often, but not always, right, for example, Cracraft (1981) included all raptors into Falconiformes, but more recently Hackett et al. (2008) proposed falconids as an order Falconiformes and they grouped accipitrids, cathartids, pandionid, and sagittarid into a separate order Accipitriformes. Similarly, the monophyletic status of some of the Neoavian orders remains uncertain (e.g., Gruiformes, Coraciiformes, Piciformes, and Falconiformes), and here again we would expect basic agreement for nuclear and mt data. Considerable progress has been made by using large data sets such as complete mtDNA genomes. It appears that the basal polytomy found in most early phylogenetic hypotheses proposed for Neoaves can be reduced by using complete mtDNA genomes and the phylogenetic signal can be improved by increasing the taxon sampling (Sorenson et al. 2003; Pereira and Baker 2006; Slack et al. 2007; Pratt et al. 2009; Gibb et al. 2013).

Cracraft (1981) reintroduced the idea of including owls in the Falconiformes, which was based on tarsometatarsal and pelvic morphology as shared with Pandionidae and Accipitridae (including falcons). His "division 3" included the

Ciconiiformes and Falconiformes (including Strigiformes). But, Olson (1982) vigorously criticized Cracraft's division that contains both flamingos and owls (combined) as monophyletic, because (according to Olson) it would be in vain to search for a morphological synapomorphy to define such a group within the class Aves. The classification by Amadon and Bull (1988) followed the traditional arrangement with all species of owls in the Strigiformes and Strigidae, which divided the strigids into Striginae (Typical owls) and Tytoninae (Barn owls, Bay owls).

More recently, Wink et al. (2009) included 120 taxa of the Strigidae and 23 taxa of the Tytonidae (the data set covers most of the genera) to study the phylogeny of owls (nocturnal raptors) based on cytochrome b and nuclear markers (LDH b intron DNA, RAG-1). This provides insight into the phylogeny and evolution of owls and the phylogenetic tree inferred from sequences of the cytochrome b gene, and nuclear RAG-1 was found to be generally in a good agreement with the classical taxonomy of owls (Sibley and Monroe 1990; Burton 1992; König et al. 1999; Weick 2006). The genetic data agreed with the attribution of species to a given genus with exceptions evident in the polyphyletic genus *Otus* and the paraphyletic *Bubo* complex.

It is expected that the two main owl lineages (barn owls and the ordinary owls) will come together. The two owl mt genomes currently available (*Tyto* and *Ninox*) do come together, but quite deeply, and could even be the result of "long-branch attraction (LBA)" (Hendy and Penny 1989). We support the suggestion (Pacheco et al. 2011) that additional complete mt genome sequences from deeply diverging Strigiformes are needed to discard the possibility of LBA; owls seem to have some of the highest rates of sequence evolution among Neoaves (Pratt et al. 2009). So, we expect that the availability of mt genomes for *Athene* and *Phodilus* will resolve fairly definitely that all the owls are monophyletic. Indeed, it would be surprising if the two groups of owls were not united. *Athene brama* (the Spotted Owlet) is expected to be deep on the same lineage as *Ninox novaeseelandiae* (an ordinary owl, see Wink et al. 2009), and *Athene* is expected to be the deepest divergence of common owls from *Ninox*. Similarly, *Phodilus badius* (the Oriental Bay Owl) should be about the deepest divergence with *Tyto alba* (the Barn Owl).

Regarding raptors, or birds of prey, there have been diverse opinions. One of our questions is whether the nocturnal raptors group (owls) joins together with either group of the diurnal raptors, together or combined (Falconidae and Accipitridae). There have been differences on this topic (see Hackett et al. 2008; Pacheco et al. 2011). Secretarybird (*Sagittarius serpentarius*) is predicted to lie deeper on the same combined lineage as Osprey (*Pandion haliaetus*) and the Accipitridae.

Recently, there have been an important attempt to integrate the phylogeny of Neoaves with biogeography (Ericson 2012), and this is paralleled by our recent attempt to integrate phylogeny and macroecology (Gibb et al. 2013). Without prejudging that the proposed phylogenetic groups are correct biogeographically, accipitrids, woodpeckers, and owls belong to Afroaves of Ericson (2012), whereas falconids belong to Australavis (though all are within higher land birds). So, an example of convergence in ecological adaptations would be the parallel evolution of diurnal predators in the two clades, "Accipitriformes" in Afroaves and "Falconiformes" in Australavis. This would predict that the falconids and accipitrids do not come together as a monophyletic group. Similarly, most striking are the parallels in lifestyle and behavior between the Secretarybird in Afroaves and the seriemas in Australavis.

There are several unresolved questions that we address; apart from the prediction that the two owls are monophyletic, as well as the Secretarybird grouping with Pandionidae and Accipitridae. We do not know yet whether the three raptor groups are monophyletic, though perhaps the consensus is now against their being so. So, we find that the owls are a natural group, and that the Secretarybird is deepest in Accipitridae. However, we find no good evidence that all raptors are monophyletic (unless there is reversion to nonraptor behavior in several groups).

Results and Discussion

The three new mt genomes are sequenced and deposited in GenBank. The genomes are Secretarybird (*S. serpentarius*), GenBank accession number KF961184, 16,773 bp (complete); Oriental Bay Owl (*P. badius*), GenBank accession number KF961183, 17,086 bp (gap in control region [CR]); and the Spotted Owlet (*A. brama*), GenBank accession number KF961185, 16,194 bp (CR incomplete).

Oriental Bay Owl (*P. badius*) and Spotted Owlet (*A. brama*) follow the standard avian gene order which was first described in chicken (Desjardins and Morais 1990) and referred to as "ancestral avian" by Gibb et al. (2007). This gene order is consistent within Strigiformes. In contrast to other eagles and hawks (pandionid and accipitrids), Secretarybird (*S. serpentarius*) also has the ancestral avian gene order. Although it has been pointed out by Mindell et al. (1998) following all major avian phylogenies that avian gene order may have evolved independently several times, it will be interesting to see how gene order of seriemas compares with the falcons. This new information could be helpful to investigate parallel lifestyle and behavior between the Secretarybird and the seriemas (placed in Afroaves and in Australavis, respectively; see Ericson 2012).

Our main approach to improve the raptor tree was the inclusion of additional taxa. Our main phylogenetic result is shown in figure 1 and is from a maximum likelihood (ML) analysis using GTR + gamma + I model using RAxML. As is our usual practice, the third codon position was RY coded. We have improved sampling within Strigiformes in order to

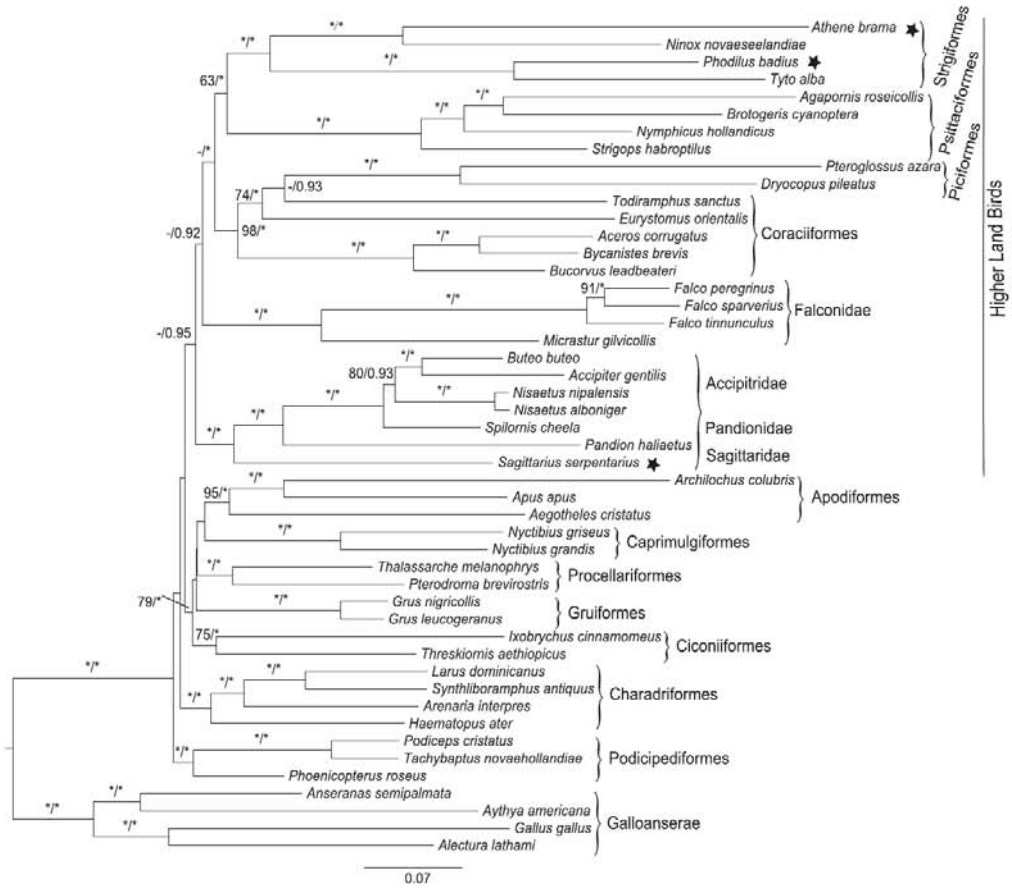


FIG. 1.—Rooted phylogram recovered from RAxML (using GTR + gamma + I). Five data partitions were used (third base position was RY-coded). New genomes reported in present investigation are marked with a star. Bootstrap and PP values are indicated for each node, with an * equaling 100% bootstrap support or a PP of 1.0 and – equaling less than 60% bootstrap or <0.9 PP. Values are not shown where both PP and bootstrap are less than 60%.

avoid the possibility of LBA (the question raised by both Pratt et al. 2009 and Pacheco et al. 2011). This was a possibility, given the owls' apparent high rate of evolution within Neoaves (Pratt et al. 2009). The two new mt genomes, of Oriental Bay Owl (*P. badius*) and Spotted Owlet (*A. brama*), represent Tytonidae and Strigidae, respectively. All four owls, the two common owls and the two barn owls, do come together on the tree; there appears to be no problem of LBA in this case. Presumably, this nocturnal group of raptors only evolved once within birds. This result is in agreement with the result of Hackett et al. (2008) based on nuclear sequences.

The next result was also as predicted; the Secretarybird is deepest on the Accipitridae lineage. There is a robust support for a clade that includes the families Accipitridae, Pandionidae, and Sagittaridae; thus, the Secretarybird (*S. serpentarius*) is deepest on the branch with Osprey (*Pandion haliaeetus*) and Accipitridae. This finding is in congruence with Lerner and Mindell (2005) and Hackett et al. (2008), although Wink and Sauer-Gürth (2004) (based on relatively short sequences) placed Secretarybird (*S. serpentarius*) with storks (Ciconiidae).

Although a morphological study (Livezey and Zusi 2007) recovered a monophyletic order Falconiformes consisting of five traditional families (Falconidae, Accipitridae, Pandionidae, Sagittaridae, and Cathartidae), none of the molecular studies that have included all five groups has found this relationship (Cracraft et al. 2004; Ericson et al. 2006; Hackett et al. 2008; Pacheco et al. 2011). We could not recover a sister relationship between Strigiformes and either or both of Falconidae and Accipitridae. This is consistent with previous molecular studies (Sibley and Ahlquist 1990; Gibb et al. 2007; Hackett et al. 2008; Pratt et al. 2009; Pacheco et al. 2011; McCormack et al. 2013), suggesting that such a relationship is not correct.

Hackett et al. (2008) found that Falconidae was closely related to a clade of Passeriformes and Psittaciformes (parrots). A similar relationship of Falconidae and Passeriformes–Psittaciformes clade was recovered by some studies primarily based on nuclear introns (Wang et al. 2012) and retroposons (Suh et al. 2011). We did not find a Falconidae/Psittaciformes relationship in our analyses, neither with Bayesian (fig. 2) nor ML (fig. 1) analyses.

The phylogenetic analysis returned a relatively well-supported clade (posterior probability or PP = 1.0, ML = 63%) between Strigiformes (owls) and Psittaciformes (parrots), which had a good (PP = 1.0, ML = 58%) sister relationship with another well-supported clade (PP = 1.0, ML = 98%) between Piciformes (woodpeckers) and Coraciiformes (kingfishers). The addition of trogon (*Trogon viridis*) in the analysis slightly lowered the resolution. That could be partly because it forms an isolated long branch; therefore, it was excluded from most of our analyses (data not shown). At least one more deeply diverging mt genome within Trogoniformes would be helpful to improve the resolution of that particular node. However, we find that the bootstrap results are just small local changes in the underlying tree, and so the tree is "locally stable" in the sense of Cooper and Penny (1997). Our results, that Strigiformes and Psittaciformes are related groups, are in agreement with previous studies which recovered the same relationship (Sorenson et al. 2003; Harrison et al. 2004; Gibb et al. 2007; Brown et al. 2008; Wright et al. 2008; Pacheco et al. 2011). We present a resolved relationship of Strigiformes–Psittaciformes clade with Piciformes–Coraciiformes (referred to as SPPC henceforth) with support values PP = 1.0 and ML = 58%, which improved upon the support values (PP = 0.60–0.64, ML = 20%) reported by Pacheco et al. (2011) who found the same relationship. The inclusion of Passeriform taxa did not change the grouping of owls with parrots (data not shown). Nevertheless, and it is certainly significant that we also find a higher land group of birds, and that this includes all three groups of raptors.

Falconidae (falcons) appear to have shared a common ancestor with SPPC (PP = 0.92). Our ML analysis also found this relationship, although with poor support (ML = 25%). Accipitridae (hawks) were sister to the group comprising Falconidae and SPPC (PP = 0.95, ML = 26%) (fig. 1). We never recovered a direct sister relationship between accipitrids

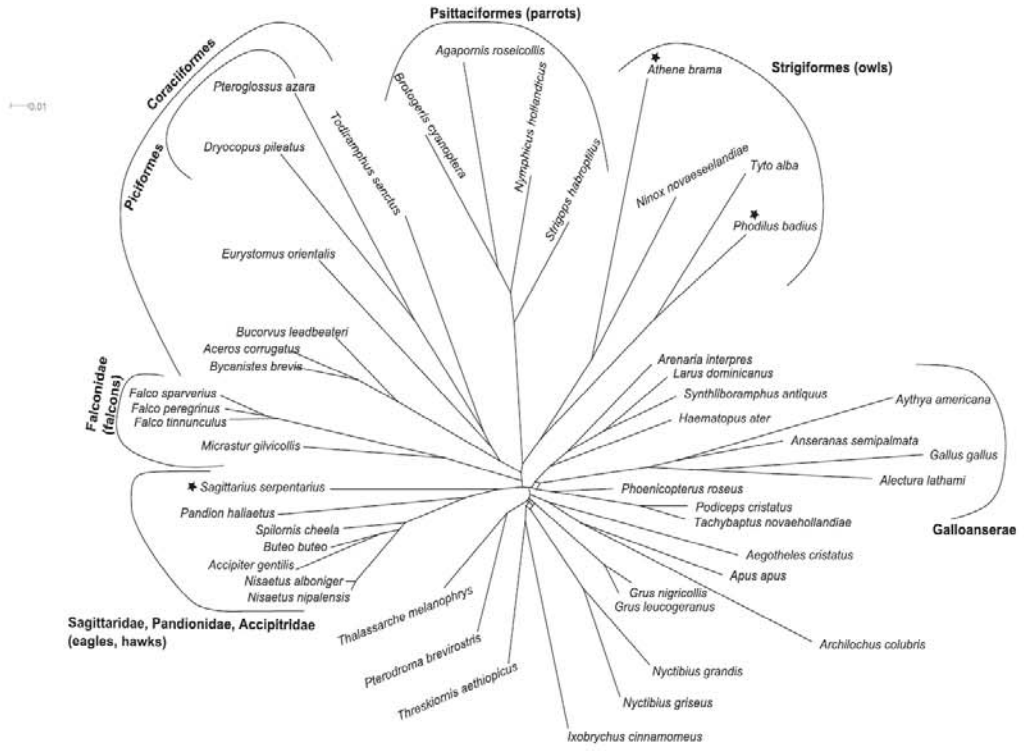


FIG. 2.—Consensus network of 48 Neoavian species based on analysis using MrBayes. New genomes reported in present investigation are marked with a star. Trees were sampled by Bayesian MCMC. Threshold = 0.2.

and falconids, which was in agreement with previous studies (Gibb et al. 2007; Slack et al. 2007). This finding (fig. 1) contrasts with Pacheco et al. (2011) who found a monophyletic relationship of diurnal raptors, which could be due to the choice of partitioning scheme (see Powell et al. 2013). Because Turkey Vulture (*Cathartes aura*) generated a long branch, it was removed after preliminary analysis (see Gibb et al. 2007; Slack et al. 2007). This could be partly because Cathartidae is possibly not a genuine raptor and its resemblance to Old World vultures is an example of convergent evolution (Tagliarini et al. 2009). Here again, additional sequences, including nuclear data, will be helpful.

Of the four nodes under investigation, the ML bootstrap values are quite low, that is (accipitrids and falconids–SPPC), (falconids and SPPC), (piciforms–coraciiforms and strigiforms–psittaciforms), and (strigiforms and psittaciforms). There are at least four possible explanations for the low resolution of ML tree. First, a "star tree paradox" (Steel and Matsen 2007) might also affect Bayesian methods when three or more lineages with high PP values (apparently resolved) diverge more or less simultaneously (Lewis et al. 2005; Kolaczkowski and Thornton 2006). Second, LBA might affect the Bayesian methods more than ML ones (Kolaczkowski and Thornton 2009). Third, Bayesian estimation of tree often gives quite high support values, which sometimes could be deceptively high (Douady et al. 2003). Finally, as compared with Bayesian methods, ML bootstrapping might underestimate the support values (Erixon et al. 2003). In a recent phylogenetic study on birds, McCormack et al. (2013) also observed weak ML support and high PP for some nodes when they used 416 locus data set, but ML support increased when they increased the size of their data set to 1,541 loci.

A consensus network of our Bayesian analysis (fig. 2) also suggested no close relationship between falconids and accipitrids. Further resolution of relationships between raptor groups would probably require additional complete mt genomes. In the present data set, Forest-Falcon (*Micrastur gilvicollis*) is the only representative of subfamily Polyborinae, among falconids. It appears to be a long branch. More representative mt genome sequence of this subfamily would be possibly helpful to discard the possibility of LBA problem. At least one more complete mt genome sequence from each of elanid kites (Elaninae), Old World vultures (Aegypiinae), and New World vultures (Cathartidae) would help resolve the relationship of cathartids, accipitrids, falconids, and SPPC within Neoaves.

Following the present tree, we find no good evidence that all raptors are monophyletic (unless there is reversion to nonraptor behavior in several groups). A recent description of a Middle Eocene skeleton (Mayr 2011) of a stem parrot (Pan-Psittaciformes) fossil *Messelastur gratulator* (Messelasturidae) may support such an argument, though it was previously considered to have closer affinities to either falconiform or strigiform birds (Mayr 2005). If future data on this fossil provide more affinities toward Psittaciformes, it would further support the idea that stem group parrots were predatory birds, but at present we cannot really support the idea.

Conclusions

We are able to conclude here that owls (Strigiformes) are monophyletic, Secretarybird (*S. serpentarius*) forms a group with Accipitridae and Pandionidae, higher land birds are a natural group, and raptors are not a natural (monophyletic) group.

Materials and Methods

Taxon Sampling

The Secretarybird (*S. serpentarius*) was provided by Donna Dittman (Lousiana), sample number LSMUZ B-2458; the Oriental Bay Owl (*P. badius*) was supplied by Michael Wink from the Institute of Pharmacy and Molecular Biotechnology, Heidelberg, Germany, sample number 28304; and the Spotted Owlet (*A. brama*) was sampled by M.T.M. and was from Multan (south-west of Lahore) in Pakistan.

Molecular Methods

Extractions of genomic DNA from each of the birds were performed at the Institute of Fundamental Sciences from 25 to 50 mg of muscle tissue using the High Pure PCR Template Preparation Kit (Roche Applied Science, Mannheim, Germany) according to the manufacturer's instructions. To minimize the chance of obtaining nuclear copies of mt genes (NUMTs), 2–4 overlapping long-range polymerase chain reaction (PCR) fragments (3.5–12 kb in length) were first amplified using the Expand Long Template PCR System (Roche Applied Science). The products were excised from agarose gels and purified using a QIAquick Gel extraction kit (Qiagen GmbH, Hilden, Germany) as per the manufacturer's instructions. These long-range products were subsequently used as template DNA for short-range PCRs to generate overlapping fragments 0.5–3 kb in length. Short-range primer combinations were found using our laboratory database as described in Slack et al. (2006), and any new primers required were designed using Geneious 5.5.7 (Drummond et al. 2011). Sequencing was performed using BigDye Terminator Cycle Sequencing reagents according to the manufacturer's instructions (Applied Biosystems, Foster City, CA), and the reactions were run on an ABI 3730 automated sequencer (Applied Biosystems) by Massey Genome Service. Sequences were aligned using Sequencher 4.7 (Gene Codes Corp., Ann Arbor, MI) and then manually edited and checked for complete concurrence between overlapping sequences.

Where necessary (e.g., with length heteroplasmy in CRs from microsatellite repeats), PCR products were cloned using the TOPO TA cloning kit for sequencing (Invitrogen,

Carlsbad, CA). For each region, at least three clones were sequenced to safeguard against PCR errors. In all cases, overlaps between sequences were sufficient to ensure synonymy and sequence identity was confirmed through Blast searches (http://www.ncbi.nlm.nih.gov/blast/), confirmation of amino acid translation in coding regions, and alignment with other species.

In addition to the 3 new bird mt genomes reported in this article, 47 other complete avian mt genomes from NCBI GenBank were included in the analyses: 43 Neoaves and 4 Galloanserae. Paleognath taxa were not included in this data set because their overall placement is now well established (Gibb et al. 2007; Slack et al. 2007). Instead, we rooted our Neoaves trees with the Galloanserae sequences (Gibb et al. 2007). We also repeated our analyses with six passerines included. The passerines do fall within the higher land birds, and their inclusion did not affect our conclusions of nonmonophyly of raptors (data not shown). The full data set is available from the authors on request.

Phylogenetic Analysis

Sequences were aligned in Geneious 5.5.7 (Drummond et al. 2011) at the amino acid level for protein-coding genes and based on secondary structure for RNA genes (see Gibb et al. 2013). The data set has 12 protein-coding genes, 2 ribosomal RNAs (rRNA), and 22 transfer RNAs (tRNA). Gaps, ambiguous sites adjacent to gaps, NADH6 (light-strand encoded), and stop codons (often incomplete in the DNA sequence) were excluded from the alignment. The 12 protein-coding genes were separated into first-, second-, and third-codon positions (the third-codon position was RY-coded as explained by Gibb et al. 2007), whereas rRNA and tRNA genes were partitioned into stems (S) and loops (L), thus we use five data partitions (see Harrison et al. 2004). Protein-coding genes were checked for NUMTs by translating into amino acids.

A combined total of 13,430 nucleotides (excluding gaps) were used for analyses. We ran analyses in RAxML (Stamatakis et al. 2008) to carry out bootstrap replicates on the data sets where bootstrapping automatically stopped using the "majority rule" criterion. In addition to ML support, Bayesian posterior probabilities were also estimated. Bayesian analyses were carried out in MrBayes (Huelsenbeck and Ronquist 2001). Bayesian analyses were run for 10,000,000 generations with a burn-in value of 10%. Both RAxML and MrBayes were run using CIPRES Science Gateway (Miller et al. 2010). Trees were visualized in FigTree v1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/ [accessed April 2013]). Consensus networks were implemented in SplitsTree version 4 (Huson and Bryant 2006).

Acknowledgments

The authors thank the Higher Education Commission, Pakistan (HEC) for supporting this research by a PhD scholarship to M.T.M. They also thank Prof. M. Wink (Heidelberg, Germany, for the Phodilus sample) and Dr D. Dittman (Louisiana, USA, for the Secretarybird sample).

Literature Cited

Amadon D, Bull J. 1988. Hawks and owls of the world: an annotated list of species. Proc West Found Vert Zool. 3:297–330.

Brown JW, Rest JS, García-Moreno J, Sorenson MD, Mindell DP. 2008. Strong mitochondrial DNA support for a Cretaceous origin of modern avian lineages. BMC Biol. 6:6.

Burton JA. 1992. Owls of the world: their evolution, structure and ecology. London: Peter Lowe.

Cooper A, Penny D. 1997. Mass survival of birds across the Cretaceous/Tertiary boundary. Science 275:1109–1113.

Cracraft J. 1981. Toward a phylogenetic classification of the recent birds of the world (Class Aves). Auk 98:681–714.

Cracraft J, et al. 2004. Phylogenetic relationships among modern birds (Neornitlies). In: Cracraft J, Donaghue MJ, editors. Assembling the tree of life. New York: Oxford University Press. p. 468–489.

Desjardins P, Morais R. 1990. Sequence and gene organization of the chicken mitochondrial genome—a novel gene order in higher vertebrates. J Mol Biol. 212:599–634.

Drummond AJ, et al 2011. Geneious v5.5. Auckland: Biomatters Ltd.

Douady CJ, Delsuc F, Boucher Y, Doolittle WF, Douzery EJP. 2003. Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. Mol Biol Evol. 20:248–254.

Ericson PGP. 2012. Evolution of terrestrial birds in three continents: biogeography and parallel radiations. J Biogeogr. 39:813–824.

Ericson PGP, et al. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. Biol Lett. 2:543–547.

Erixon P, Svennblad B, Britton T, Oxelman B. 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. Syst Biol. 52:665–673.

Gibb GC, Kardailsky O, Kimball RT, Braun EL, Penny D. 2007. Mitochondrial genomes and avian phylogeny: complex characters and resolvability without explosive radiations. Mol Biol Evol. 24:269–280.

Gibb GC, Kennedy M, Penny D. 2013. Beyond phylogeny: pelecaniform and ciconiiform birds, and long-term niche stability. Mol Phylogenet Evol. 68:229–238.

Hackett SJ, et al 2008. A phylogenomic study of birds reveals their evolutionary history. Science 320:1763–1768.

Harrison GL, et al. 2004. Four new avian mitochondrial genomes help get to basic evolutionary questions in the late Cretaceous. Mol Biol Evol. 21:974–983.

Hendy MD, Penny D. 1989. A framework for quantitative study of evolutionary trees. Syst Biol. 38:297–309.

Hernandez-Lopez A, Raoultb D, Pontarotti P. 2013. Birds perching on bushes: networks to visualize conflicting phylogenetic signals during early avian radiation. C R Palevol. 12:333–337.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol. 23:254–267.

Johansson U, Parsons T, Irestedt M, Ericson P. 2001. Clades within the higher land birds', evaluated by nuclear DNA sequences. J Zool Syst Evol Res. 39:37–51.

Kimball RT, Wang N, Heimer-McGinn V, Ferguson C, Braun EL. 2013. Identifying localized biases in large datasets: a case study using the avian tree of life. Mol Phylogenet Evol. 69:1021–1032.

Kolaczkowski B, Thornton JW. 2006. Is there a star tree paradox? Mol Biol Evol. 23:1819–1823.

Kolaczkowski B, Thornton W. 2009. Long-branch attraction bias and inconsistency in Bayesian phylogenetics. PLoS One 4:e7891.

König C, Weick F, Becking JH. 1999. Owls: a guide to the owls of the world. Robertsbridge: Pica Press.
Lerner HRL, Mindell DP. 2005. Phylogeny of eagles, Old World vultures, and other Accipitridae based on nuclear and mitochondrial DNA. Mol Phylogenet Evol. 37:327–346.
Lewis PO, Holder MT, Holsinger KE. 2005. Polytomies and Bayesian phylogenetic inference. Syst Biol. 54:241–253.
Livezey BC, Zusi RL. 2007. Higher-order phylogeny of modern birds (Theropoda, Aves: Neornithes) based on comparative anatomy. II. Analysis and discussion. Zool J Linn Soc. 149:1–95.
Mayr G. 2005. The postcranial osteology and phylogenetic position of the Middle Eocene *Messelastur gratulator* PETERS, 1994—a morphological link between owls (strigiformes) and falconiform birds? J Vert Paleontol. 25:635–645.
Mayr G. 2011. Well-preserved new skeleton of the Middle Eocene *Messelastur* substantiates sister group relationship between Messelasturidae and Halcyornithidae (Aves,? Pan-Psittaciformes). J Syst Palaeontol. 9:159–171.
McCormack JE, et al. 2013. A phylogeny of birds based on over 1,500 Loci collected by target enrichment and high-throughput sequencing. PLoS One 8:e54848.
Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Proceedings of the Gateway Computing Environments Workshop (GCE); 2010 November 14; New Orleans, LA. p. 1–8.
Mindell DP, Sorenson MD, Dimcheff DE. 1998. Multiple independent origins of mitochondrial gene order in birds. Proc Natl Acad Sci U S A. 95:10693–10697.
Olson SL. 1982. A critique of Cracraft's classification of birds. Auk 99:733–739.
Pacheco MA, et al. 2011. Evolution of modern birds revealed by mitogenomics: timing the radiation and origin of major orders. Mol Biol Evol. 28:1927–1942.
Pereira SL, Baker AJ. 2006. A mitogenomic timescale for birds detects variable phylogenetic rates of molecular evolution and refutes the standard molecular clock. Mol Biol Evol. 23:1731–1740.
Powell AFL, Barker FK, Lanyon SM. 2013. Empirical evaluation of partitioning schemes for phylogenetic analyses of mitogenomic data: an avian case study. Mol Phylogenet Evol. 66:69–79.
Pratt RC, et al. 2009. Toward resolving deep Neoaves phylogeny: data, signal enhancement, and priors. Mol Biol Evol. 26:313–326.
Sibley CG, Ahlquist JE. 1990. Phylogeny and classification of birds. New Haven (CT): Yale University Press.
Sibley CG, Monroe BL. 1990. Distribution and taxonomy of birds of the world. New Haven (CT): Yale University Press.
Slack KE, Delsuc F, McLenachan PA, Arnason U, Penny D. 2007. Resolving the root of the avian mitogenomic tree by breaking up long branches. Mol Phylogenet Evol. 42:1–13.
Slack KE, et al. 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. Mol Biol Evol. 23:1144–1155.
Sorenson MD, Oneal E, Garcia-Moreno J, Mindell DP. 2003. More taxa, more characters: the hoatzin problem is still unresolved. Mol Biol Evol. 20:1484–1498.
Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML Web servers. Syst Biol. 57:758–771.
Steel M, Matsen FA. 2007. The Bayesian "star paradox" persists for long finite sequences. Mol Biol Evol. 24:1075–1079.
Suh A, et al. 2011. Mesozoic retroposons reveal parrots as the closest living relatives of passerine birds. Nat Commun. 2:443.
Tagliarini MM, Pieczarka JC, Nagamachi CY, Rissino J, De Oliveira EHC. 2009. Chromosomal analysis in Cathartidae: distribution of heterochromatic blocks and rDNA, and phylogenetic considerations. Genetica 135:299–304.
Wang N, Braun EL, Kimball RT. 2012. Testing hypotheses about the sister group of the passeriformes using an independent 30-locus data set. Mol Biol Evol. 29:737–750.
Weick F. 2006. Owls (Strigiformes) annotated and illustrated checklist. Heidelberg: Springer Verlag.
Wink M, Heidrich P, Sauer-Gürth H, Elsayed A-A, Gonzalez J. 2009. Molecular phylogeny and systematics of owls (Strigiformes). In: Weick F, Konig C, editors. Owls of the world. New Haven (CT): Yale University Press. p. 42–63.
Wink M, Sauer-Gürth H. 2004. Phylogenetic relationships in diumal raptors based on nucleotide sequences of mitochondrial and nuclear marker genes. In: Chancellor R, Meyburg BU, editors. Raptors worldwide. Berlin: World Working Group on Birds of Prey and Owls. p. 483–498.
Wright TF, et al. 2008. A multilocus molecular phylogeny of the parrots (Psittaciformes): support for a Gondwanan origin during the Cretaceous. Mol Biol Evol. 25:2141–2156.

Associate editor: Bill Martin

4 FOUR NEW MITOCHONDRIAL GENOMES SUPPORT THE DISTINCTIVENESS OF THREE SEPARATE GROUPS OF RAPTORS

Muhammad Tariq Mahmood, McLenachan P.A. Zhong B., Wink M. and Penny, D. (2015). Four new mitochondrial genomes support the distinctiveness of three separate groups of raptors. *Genome Biology and Evolution* (to be submitted).

The four new avian mitochondrial (mt) genomes reported here. Three of them are members of the family Accipitridae, and one is from the family Falconidae. These are: an Elanid kite (*Elanus caeruleus*), a Griffon vulture (*Gyps fulvus*), a Cinereous vulture (*Aegypius monachus*) and a Striated caracara (*Phalcoboenus australis*). These four taxa were chosen because they would allow a better understanding of the evolution of raptors generally. This study was a follow up on the previous chapter that showed that diurnal and nocturnal raptors had evolved at least three, and probably four, times – owls, eagles, falcons and New World (NW) vultures.

Previously we had excluded the Turkey vulture sequence after the preliminary analysis (Mahmood et al. 2014) because it was an isolated long branch and caused an LBA (Long branch attraction) problem. A previous mitogenomic study also excluded it from their dataset because of the same LBA problem (Pacheco et al. 2011). To eliminate this possibility, we also intended to include another NW vulture (not an OW vulture) in addition to the already available Turkey vulture (*Cathartes aura*) genome. The DNA sample of the Black vulture (*Coragyps atratus*) was processed for the next generation sequencing pipeline. But after it was already sequenced and analysed it turned out that the sample we had received had been mis-identified, and it was really an OW vulture namely, Cinereous vulture (*Aegypius monachus*). This was quickly re-identified because there was a good specimen available in Heidelberg to check the identification. This emphasizes the importance of good samples being retained for any further study; the problem was very quickly identified and solved. As mentioned above, we removed the Turkey vulture from the dataset after the preliminary analysis, since it disrupts the overall tree topology and phylogenetic resolution (Figure 1).

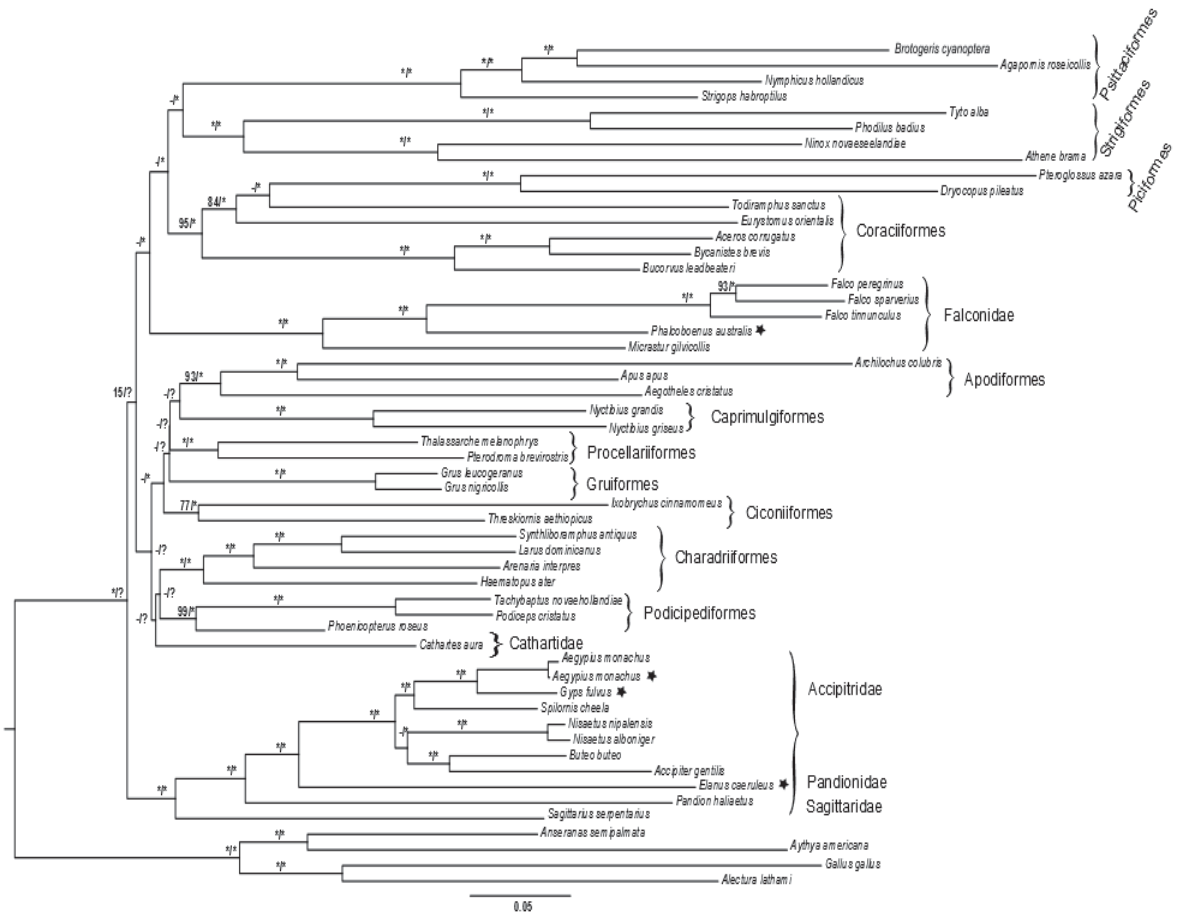


Figure 1. Rooted phylogram (including Turkey vulture; *Cathartes aura*) recovered from RAxML (using GTRGAMMA). In this tree 5 data partitions were used and the third base position was RY-coded. A star ★ highlights the new genomes reported in present investigation. Bootstrap and posterior probability (PP) values are indicated for each node, where an asterisk * represents 100% bootstrap support/PP of 1.0, and '-' represents a less than 60% bootstrap support or < 0.9 PP. Values for both bootstrap and PP are not shown where they are less than 60% or 0.6.

We subsequently obtained a DNA sample of King vulture (*Sarcoramphus papa*), another NW vulture and it has been sequenced using the Illumina MiSeq platform for next generation sequencing. However, the process of the mitochondrial genome assembly has been relatively slow because that sample has turned out to mostly give nuclear sequences.

For future phylogenetic analysis, inclusion of yet another NW vulture to the existing dataset would help test whether OW and NW vultures have morphological and

ecological convergence, or whether vulture ecology has evolved twice on evolutionary time scale.

As pointed out by Mayr (2009; 2014) that Cathartidae had stem group representative fossils discovered from the Paleogene of Europe. Another undescribed Cathartidae was mentioned (Olson 1985) from early Oligocene of Mongolia. Falcons, seriemas, and NW vultures now have a South American radiation and may have diversified there further. Another possibility could be that falcons and NW vultures migrated from eastern hemisphere (possibly Europe or North Africa?) before becoming extinct there, to western hemisphere where their diversification occurred. Yet another perspective on biogeography also suggests that avian diversification in Australia and Africa followed a pattern where ancient radiations filled the ecological space in their region of origin (Ericson 2012).

The results of the analysis show that the Accipitridae, and the falcons had separately become raptors (and captured and ate birds and other animals). The suggestion had been made that the bird tree might prove susceptible to 'long branch attraction' problems, and the likelihood of this is much reduced if other taxa are available that breakup long branches on the tree. So the taxa we selected continued the analysis of raptor evolution, and showed very good agreement with earlier suggestions as to the placement of vultures and falcons. However, there is still a need for another NW vulture genome to support that the NW and OW vultures are genuinely separate origins of the raptor life-style.

I did the DNA sequencing and primary analysis, PAM was responsible for important technical advice and laboratory supervision, BZ helped with advice for deleting the fastest evolving sites, MW supplied the samples and gave co-supervision, MW and DP gave general oversight and planning – all authors contributed to the final manuscript.

Four new mitochondrial genomes support the distinctiveness of three separate groups of raptors.

Mahmood, M.T.[1], McLenachan. P.A.[1], Zhong B.[1,2]., Wink, M.[3] and Penny, D.[1]

[1]Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand.

[2] Present address: College of Life Sciences, Nanjing Normal University, Nanjing, China

[3] Institute of Pharmacy and Molecular Biotechnology, University of Heidelberg, Heidelberg, Germany.

Corresponding author: E-mail t.mahmood@massey.ac.nz

Data deposition: The sequences from this project have been deposited at GenBank under the accession numbers xxxx *Elanus caeruleus*, xxxx *Phalcoboenus australis*, xxxx *Gyps fulvus*, and xxxx *Aegypius monachus*.

Abstract:

Four mitochondrial genomes of raptors are reported here and analysed. There are three Accipitridae and one falcon. The birds include an Elanid kite (*Elanus caeruleus*), a caracara (*Phalcoboenus australis*), Griffon vulture (*Gyps fulvus*) and an Old World (OW) vulture (*Aegypius monachus*), Although we had intended to sequence a New World (NW) vulture that sample turned out to have been mislabelled, and it really was *Aegypius monachus*, an OW, and not a NW, vulture. Thus it is very important to always have good reference samples in order to check any sequence data against – we think that there are problems in the literature where samples have been mis-identified, or a nuclear copy of a mitochondrial gene has been sequences (a 'numt'), or even hybrids of more than one genome may be reported. The results generally support the hypothesis that the falcon and eagle families have separate origins. Given also our earlier results with owls, it appears that the 'raptor' lifestyle evolved on at least three, and probably four, occasions (namely owls, eagles, falcons and probably NW vultures).

Key words. Higher land birds, raptors, *Elanus*, *Phalcoboenus*, *Gyps*, *Aegypius*.

4.1 Introduction

The avian tree is rapidly being resolved, and there is agreement from both nuclear and mitochondrial data to the effect that paleognaths (ratites and tinamous) are an ancient divergence (Mitchell et al. 2014). Then the Neognaths split into two groups, the combined chicken and duck lineages (Galloanseriformes) and then the large major group of all other birds – the Neoaves. This Neoavian group has been harder to resolve, but good progress is being made (see Jarvis et al. 2014). Here we are only interested in one subgroup of Neoaves, the diurnal raptors, because our recent paper (Mahmood et al. 2014) agreed with earlier work that the nocturnal raptors (owls) and diurnal raptors (for example, eagles and falcons) were not monophyletic.

Here we are interested in one aspect of Neoavian relationships - raptors (birds of prey). It is well documented that there are localised biases in both avian nuclear (Kimball et al. 2013) and mitochondrial datasets (Meiklejohn et al. 2014). Phylogenetic relationships and attribution of different genera within family Falconidae, based on short nuclear sequences (Griffiths et al. 2004), and within the family Accipitridae, based on osteological characters (Holdaway 1994), short mitochondrial (Wink and Sauer-Gurth 2004), short nuclear (Griffiths et al. 2007) as well as combined both short nuclear and mitochondrial sequences (Lerner and Mindell 2005) may not be fully reliable. Particularly there have been conflicting results regarding the monophyly (Holdaway 1994) or polyphyly (Wink and Sauer-Gurth 2004; Lerner and Mindell 2005) of OW vultures within the family Accipitridae. One study based on short mitochondrial sequences (Wink and Sauer-Gurth 2004) placed the Secretarybird (*Sagittarius serpentarius*) with storks (Ciconiidae), but the full mitochondrial sequence of the Secretarybird did not support that relationship (Mahmood et al. 2014). Although it is known that some subset of the mitochondrial genome can serve as a proxy for the whole mitogenome across the Metazoa, but the identity and the exact number of these genes is lineage specific (Havird and Santos 2014). Also, it is now established that in birds whole mitochondrial genome perform better than a small subset as a proxy for the whole genome (Powell et al. 2013; Meiklejohn et al. 2014) as well as in mammals (Duchene et al. 2011). Since incongruence among mitochondrial regions (Cox et al. 2007; Meiklejohn et al. 2014) and more potential of mitogenomes is known (Duchene et al. 2011; Meiklejohn et al. 2014; also see above), the placement of the taxa based on

either limited nuclear or mitochondrial regions is likely to be less resolved than the mitogenomic species tree.

It is important to revisit previously unresolved relationships within the tree of raptors. It will be helpful to see whether breaking some long branches (increased taxon sampling), better fitting models, and noise reduction (either by RY coding or removal of faster-evolving sites), would improve the overall resolution. Reflecting on the previous mitogenomic studies (Pacheco et al. 2011; Mahmood et al. 2014) we chose *Elanus caeruleus, Pahalcoboenus australis, Gyps fulvus, and Aegypius monachus* for the following reasons:

The Elanid kite

The lack of clear fossil record attributable to the ancient group of Elanid kites, such as *Elanus caeruleus*, may explain why they have been overlooked when studying evolutionary relationships. The phylogenies based on mitochondrial Cyt *b* (Roulin and Wink 2004) and nuclear RAG plus Cyt *b* sequences (Lerner and Mindell 2005) place them basal to the largest raptor family i.e. Accipitridae, and they are only quite distantly related to Falconidae. A few studies have indicated their distinctiveness in osteology (Holdaway 1994), genetics (Wink et al. 1998), and karyotype (Bed'Hom et al. 2003) but without questioning their placement in the family Accipitridae. Negro et al. (2006) pointed out six phenotypical affinities between owls (Strigiformes) and the Elanid kites that give them something of a hybrid appearance between hawks and owls. Such affinities are absent in the rest of the raptors. The ecological, physiological, and morphological affinities between *Elanus* and the owls might have resulted from convergent evolution (Roulin and Wink 2004; Lerner and Mindell 2005). We also suggested (Mahmood et al. 2014) that inclusion of *Elanus caeruleus* in the mitogenomic phylogeny of raptors will improve the overall stability of the tree.

The Striated caracara

In the previous mitogenomic attempts at resolving phylogenetic relationships among birds in general (Gibb et al. 2007; Pacheco et al. 2011) and raptors in particular (Mahmood et al. 2014), the Forest-Falcon (*Micrastur gilvicollis*) was a deep branch on the tree of available mitochondrial genomes. In our previous study (Mahmood et al. 2014), we noted that Forest-Falcon being the only representative of sub-family

Polyborinae may likely cause long branch attraction (LBA) problems. We therefore suggested that inclusion of one of the Caracaras (i.e. *Phalcoboenus australis*) would help to stabilise the raptor tree.

Griffon vulture and Cinereous vulture - the OW vultures

In a recent phylogenetic study on raptors, we noticed the lack of vulture representation within the family Accipitridae (Mahmood et al. 2014). We have decided to include the mitochondrial genome sequence of Griffon vulture (*Gyps fulvus*) and Cinereous vulture (*Aegypius monachus*) in our present study. During the process of assembling the present dataset, another complete mitogenome of Cinereous vulture (*Aegypius monachus*) was published (Li et al. 2014). We included both vulture genomes in the present analysis.

4.2 Results and Discussion

The four new mt genomes are sequenced and deposited in GenBank. The mt genomes are as follows. Elanid kite (*Elanus caeruleus*), GenBank number xxxx; a caracara (*Phalcoboenus australis*), GenBank number xxxx; Griffon vulture (*Gyps fulvus*), GenBank number xxxx and another Old World vulture, Cinereous vulture (*Aegypius monachus*), GenBank number xxxx.

The DNA sample from IPMB (Heidelberg, Germany) with the Institute number: 16571 upon phylogenetic analysis turned out not to be a *Coragyps atratus* and it was rapidly confirmed by IPMB (Heidelberg) that the sample has been mislabelled. Instead, it was an OW vulture (*Aegypius monachus*) collected from Mongolia. This species already has a mitochondrial genome sequenced (Li et al, 2014). Its gene sequence is very similar (almost identical) to the existing genome, but that bird was from a different location, namely Hefei, from China.

None of the new genomes reported here follow the standard avian gene order, first described in chicken (Desjardins and Morais 1990). Based on the nomenclature of gene order from Gibb et al. (2007), we found that all of the four new genomes had a remnant CR(2) gene order which was first described in a falcon (Mindell et al. 1998). This is the same gene order that was known previously for falcons and Accipitridae (see Fig 1E of Gibb et al. 2007).

Intra-familial relationships

We tested four strategies for phylogenetic analysis. We divided our dataset into 'full data without RY coding', 'full data with RY coding of the 3rd positions of proteins', 'full data with fast evolving sites removed', and 'RY coded full data with fast evolving sites removed'. Thus we could eliminate the fastest evolving sites, on the grounds that they may show more marked divergences and have a poor fit to the substitution models (Zhong et al. 2011). Inclusion of additional taxa and breaking some long branches was one main approach to improve the overall resolution of the raptor tree.

The results from each of the four datasets suggest that Elanid kite (*Elanus caeruleus)* is basal in relation to the rest of the family Accipitridae (see Figs 2 and 3). This finding is in agreement with the previous studies (Lerner and Mindell 2005; Griffiths et al. 2007) based on shorter sequences. The ecological, physiological, and morphological affinities between *Elanus* and the owls must therefore be the result of convergent evolution (Roulin and Wink 2004; Lerner and Mindell 2005). Because relatively shorter sequences sometimes obscure the true phylogenetic placement of taxa, the inclusion of Elanid kite is valuable for the tree of families Accipitridae, Pandionidae, and Sagittaridae.

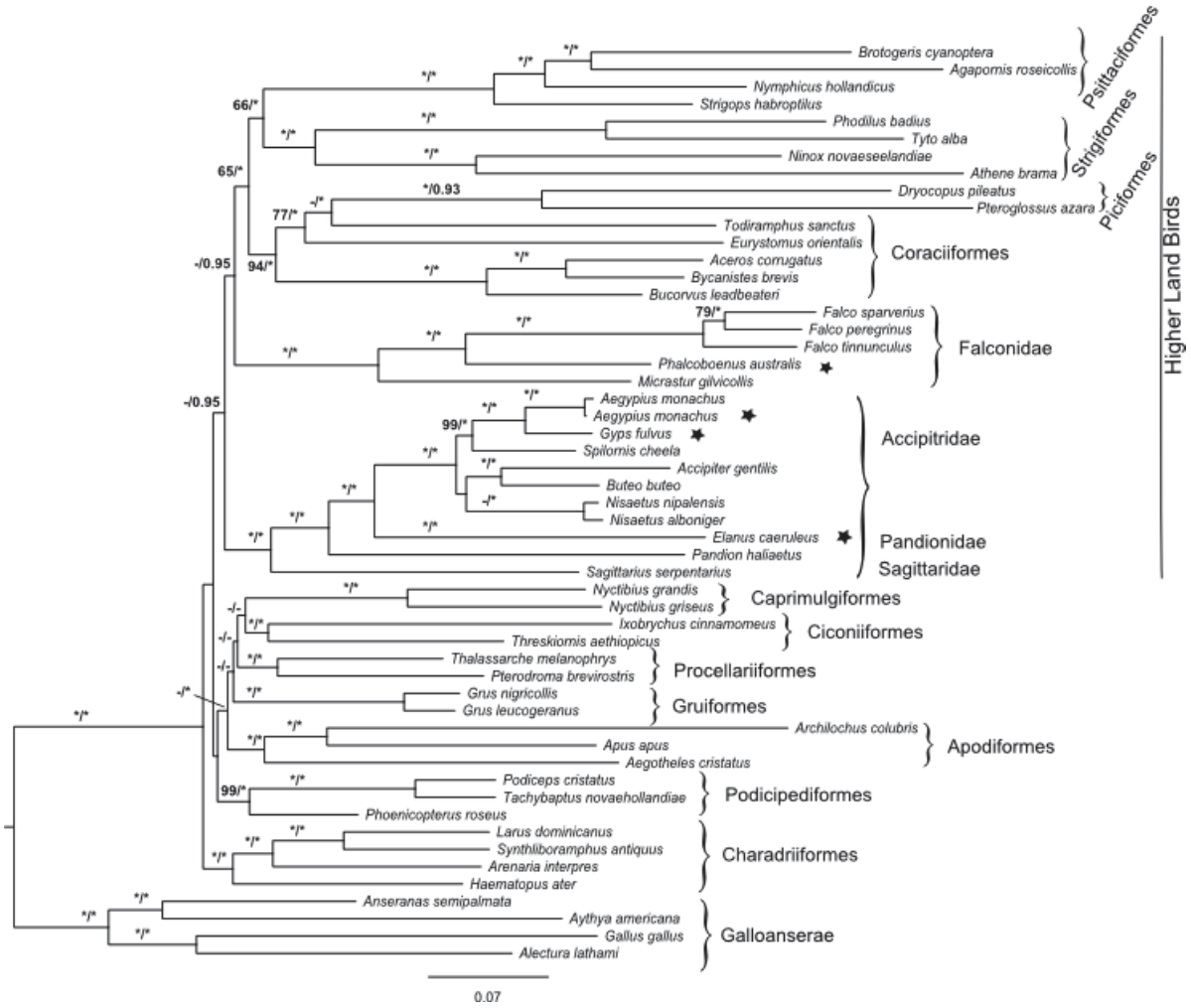


Figure 2. Rooted phylogram (excluding Turkey vulture; *Cathartes aura*) recovered from RAxML (using GTRGAMMA). In this tree 5 data partitions were used and the third base position was RY-coded. A star ★ highlights the new genomes reported in present investigation. Bootstrap and posterior probability (PP) values are indicated for each node, where an asterisk * represents 100% bootstrap support/PP of 1.0, and '-' represents a less than 60% bootstrap support or < 0.9 PP. Values for both bootstrap and PP are not shown where they are less than 60% or 0.6.

All analyses placed the Cinereous vulture (*Aegypius monachus*) and the Griffon vulture (*Gyps fulvus*), together. The group of the two vultures was found within the family Accipitridae with strong support values (ML=100, PP=1.0). Such a placement of Old World vultures within family Accipitridae is in agreement with the previous studies (Wink and Sauer-Gurth 2004; Lerner and Mindell 2005; Griffiths et al. 2007).

Because it was noted that forest falcon (*Micrastur gilvicollis*) was a long branch in the tree of family Falconidae, and it might cause a long-branch attraction (LBA) problem (Mahmood et al. 2014). We sequenced a caracara (*Phalcoboenus australis*) mitochondrial genome to break that long branch. We recovered a tree of family

falconidae with caracara placed between forest falcon and the other falcons from the genus *Falco.* The phylogenetic relationship of caracara with other falcons is in congruence with the previous studies (Griffiths 1999; Griffiths et al. 2004). All four types of analyses recovered a monophyletic family falconidae with strong support (ML=100, PP=1.0).

An important conclusion is that the placement of each of the 4 raptors, and as shown in Figs 1 and 2, is 'as expected' on current classification. The caracara (*P. australis)* came in its expected position as deep on the lineage to the falcons. Similarly, the Elanid kite (*E. caeruleus*), and the OW vultures (*G. fulvus* and the *A. monachus)* sequences all came within the Accipitridae, with the *Elanid* kite being currently the deepest lineage on that group. Furthermore, it does not disrupt the position of the Osprey (*P. haliaetus)* or Secretarybird (*S. serpentarius)* sequences on the trees, which come just outside the Accipitrid group.

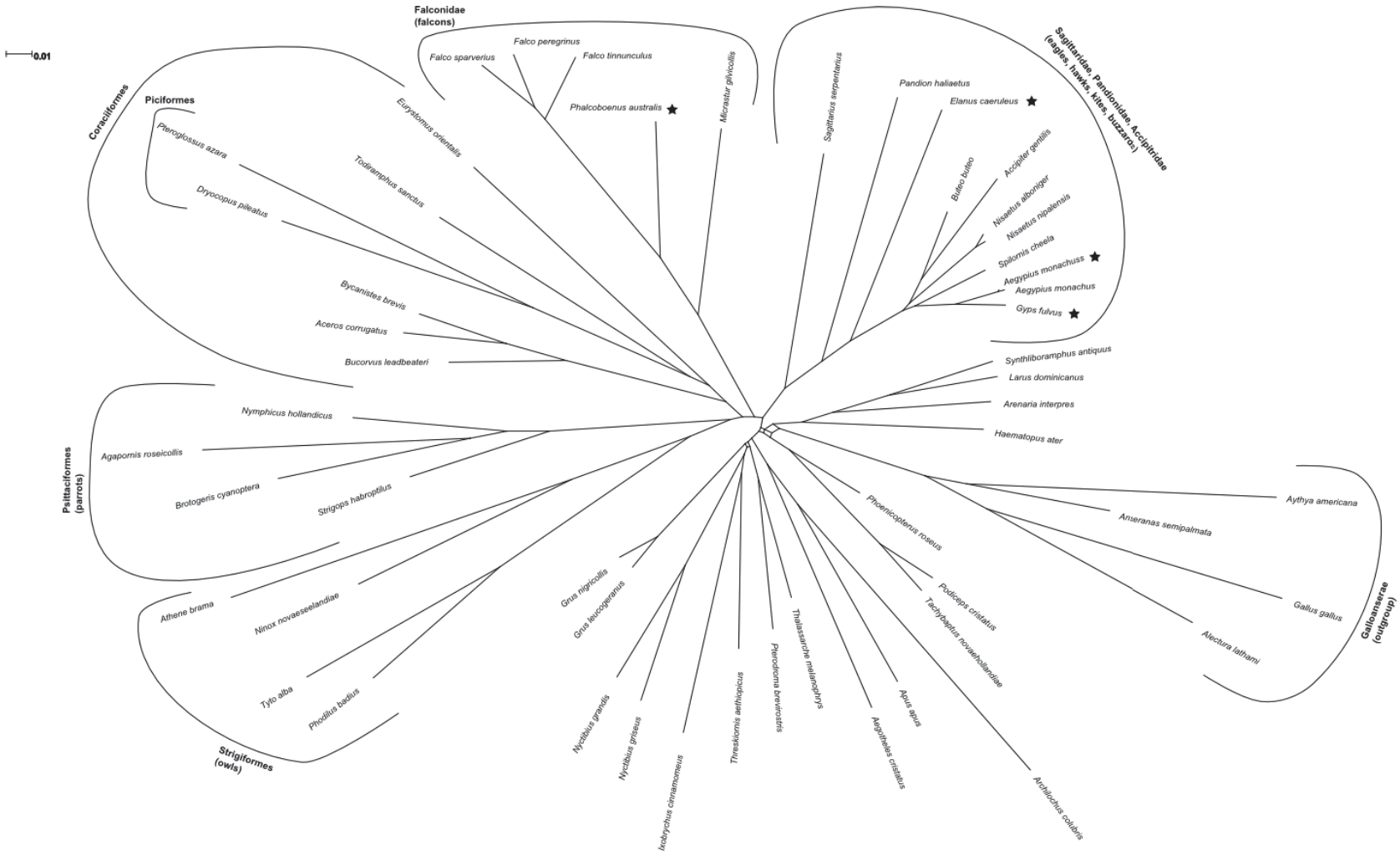


Figure 3. Consensus network of 52 Neoavian species based on analysis using MrBayes. A star highlights the new genomes reported in present investigation. Trees were sampled by Bayesian Markov Chain Monte Carlo (MCMC). Threshold = 0.2.

Inter-ordinal/ Inter-familial relationships

We carried out an analysis that examined the effect of sites with multiple changes. The purpose of dividing the dataset into four (4) strategies was to determine if there was any further resolution within Neoaves. There are interesting phenomena to follow up here. Of these four datasets, full data with RY coding performed best, and our conclusions regarding inter-ordinal/inter-familial relationships are based on that. However, we also discuss the other three datasets.

Our results are consistent with the common ancestry of owls (Strigiformes) and parrots (Psittaciformes) (ML=66%, PP=1.0). Tree topology supports that Strigiformes-Psittaciformes group forms a clade with Piciformes-Coraciiformes which is in congruence with the previous studies (Pacheco et al. 2011; Mahmood et al. 2014, referred to as SPPC) with a modest increase in nodal support values. There was high support in our Bayesian analysis (PP=1.0) while the maximum likelihood bootstrap support values for SPPC clade was low (ML=65%). Initially we included Trogon (*Trogon viridis*) in our preliminary analysis, but we decided to exclude it from final analysis (because it forms an isolated long branch) since it slightly lowered the overall resolution of tree (results not shown). Resolution of that particular node could be increased through a mitochondrial genome of at least one more deeply diverging representative of Trogoniformes. A slight increase in ML support was also noted on a node where Falconidae seems to share a common ancestor with SPPC clade (PP=0.95, ML=38%) which supports the previous mitogenomic study on raptors (Mahmood et al. 2014). Although Pacheco et al. (2011) also recovered such a relationship but in that case Falconidae was found to group with the Accipitridae-Pandionidae-Sagittaridae lineage, which was placed relatively basal to SPPC.

In all four strategies mentioned above we found a strong monophyletic relationship (ML=100, PP=1.0) between the three families Accipitridae, Pandionidae and Sagittaridae, which is congruent with previous studies (Pacheco et al. 2011; Gibb et al. 2013; Mahmood et al. 2014). Turkey Vulture (*Cathartes aur*a) was removed after preliminary analysis, because it generated a long branch (see Gibb et al. 2007; Slack et al. 2007; Mahmood et al. 2014; also see introductory section of this chapter). We suggest that inclusion of at least one more mitochondrial genome of another NW

vulture would be very helpful to improve the resolution of that particular node. OW vultures have been reported to have polyphyletic origins, based on shorter nuclear and mitochondrial sequences (Wink and Sauer-Gurth 2004; Lerner and Mindell 2005; Griffiths et al. 2007). So, the OW vultures sub families Aegypinae and Gypaetina appear to have two different origins. In the present study, only OW vultures from the sub family Aegypina are represented. That still leaves the question of polyphyly or monophyly of OW vultures. Therefore a mitochondrial genome sequence of OW vultures group, which is in sister to pernine kites (Lerner and Mindell 2005; Griffiths et al. 2007), along with two NW and the existing OW vulture genomes, would help to resolve the question of polyphyly, or monophyly, of OW vultures and their relationship with NW vultures. Overall, it would improve the resolution of the raptor tree.

We recovered a topology with Accipitridae, Pandionidae, and Saggitaridae clade (eagles, hawks, osprey and secretarybird) in a relatively basal position to Falconidae-SPPC clade as part of the 'higher land birds' group. That is in congruence with previous studies based on both nuclear data (Hackett et al. 2008; Ericson 2012; McCormack et al. 2013) and mitogenomic data (Pacheco et al. 2011; Mahmood et al. 2014). We recovered eagles, hawks, osprey and secretarybird as part of higher land birds in our ML analysis, although with low support, ML=32%. Our Bayesian analysis also recovered such a relationship but with high support (PP=0.95).

We did not recover a monophyletic order Falconiformes (comprising of families Falconidae, Accipitridae, Pandionidae, Sagittaridae, and Cathartidae), as reported by a morphological study (Livezey and Zusi 2007). None of the molecular phylogenetic studies have found a monophyletic relationship placing these five families within a single order (Cracraft 2004; Ericson 2006; Hackett et al. 2008; Pacheco et al. 2011; Mahmood et al. 2014).

Removal of potentially fast evolving sites

With our full data without RY coding, we recovered the 'higher land birds' group topology in ML analysis, though with low support (SPPC, ML=37%; Falconidae-SPPC, ML=19%; Eagles/hawks/osprey/secretarybird-falconidae-SPPC, 13%). In this case, Bayesian analysis did not recover this topology, with only SPPC coming

together with good support (PP=0.92). All the ML and Bayesian support values are shown in Table 1. There were cases when the topology changed to an alternative topology, upon exclusion of the potentially fast evolving/misleading sites (results not shown). As it was expected that increased taxon sampling and breaking some possible long branches would increase the overall resolution of the tree, there was a modest increase in ML support (Bayesian support is already high) particularly at deep internal nodes. This increase in ML support has improved on the previous studies (Pacheco et al. 2011; Mahmood et al. 2014; see Table 1) which had relatively fewer taxa than the present one. Overall, we recovered a topology where raptors are found as part of 'higher land birds' clade which is congruent with the studies based on, nuclear (Ericson 2012; Hackett et al. 2008) including a more recent whole genome analysis (Jarvis et al. 2014) and mitochondrial datasets (Pacheco et al. 2011; Mahmood et al. 2014).

Table 1. Comparison of nodal support values across different avian mitochondrial phylogenomic studies.

	Pacheco et al. (2011)	Mahmood et al. (2014)	Present study			
Nodes			Full data (RY coding)	Full data (without RY coding)	Reduced data (RY coding)	Reduced data (without RY coding)
Strigiformes-Psittaciformes (SP)	-/0.92a	63/1.0	66/1.0	58/0.94	52/0.68	-/1.0
Piciformes-Coraciiformes (PC)	-/1.0a,b	98/1.0	94/1.0	93/1.0	82/1.0	17/-
SP-PC	20/0.60-0.64b	58/1.0	65/1.0	37/0.92	41/0.75**	-/1.0
Falconidae (F)-SPPC	-/-	25/0.92	38/0.96	19/-	-/0.71	-/-
Accipitridae-F-SPPC	na/1.0 a, *	26/0.95	32/0.95	13/-	-/0.45	-/0.67**

bootstrap/posterior probability, '-'= node not supported, * = Accipitridae-F-SPPC come together as a group, ** = alternative topology, a = rooted tree, b = unrooted tree, na = data not available.

We agree with the suggestions of Meiklejohn et al. (2014) regarding the need of more robust evolutionary models to analyse mitogenomic data. In our case, the stripping/removal of fast evolving sites, was not helpful in increasing the overall resolution of the raptor tree, rather it decreased resolution. One possibility could be the accumulation of more substitutions, or more heterogeneous base composition, in certain taxa than the rest represented in the present dataset. In that case, most of the evolutionary models, which assume compositional homogeneity, might not be sufficient. Any possibility of presence of such instances would be difficult to model. In such a situation, a more robust evolutionary model would be needed. The presence of relatively more heterogeneous taxa could be measured individually or tested through removal of potentially problematic taxa by applying a jack-knifing approach (Lanyon 1985), adopted by Slack et al. (2007), where taxa are deleted sequentially to measure the effect (Penny and Hendy, 1985). In the sense of Cooper and Penny (1997) our tree is locally stable (that is, there are no major changes to the tree) and that the changes in bootstrap values are small local changes in the underlying tree. Possibly, rapid adaptive radiation of Neoaves around the K-Pg (Cretaceous-Paleogene) boundary could be a reason of short internal branches on Neoavian tree. But smaller dinosaurs, leading to present day class Neoaves, did have more space for innovation due to their smaller size. In turn that led those lineages to accumulate more diversity in birds (Moen and Morlon 2014).

4.3 Conclusions

Our results are in agreement with the current classification of falcons and Accipitridae, and this is a positive finding. They also corroborate the findings of our recent study (Mahmood et al. 2014) and previous studies (Gibb et al. 2007; Hackett et al. 2008) that combined raptors are not a natural/monophyletic group; eagles, falcons, owls, and NW vultures are distinct groups. Particularly NW vultures are currently a South American lineage. Now, that ‘higher land birds’ group is regarded as a natural group (Johansson et al. 2001; Ericson et al. 2006; Mahmood et al. 2014), our analysis also supports that relationship (Fig. 1).

4.4 Materials and Methods

Taxon sampling

All taxa were supplied from the collection at the Institute of Pharmacy and Molecular Biotechnology, University of Heidelberg. The sample numbers are as follows. Elanid kite (*Elanus caeruleus*), 9965; a caracara (*Phalcoboenus australis*), 7281; Griffon vulture (*Gyps fulvus*), 3677 and (*Aegypius monachus*), 16571.

Molecular Methods

Extractions of genomic DNA from each of the birds were performed at the Institute of Pharmacy and Molecular Biotechnology from 25 to 50 mg of muscle tissue using the High Pure PCR Template Preparation Kit (Roche Applied Science, Mannheim, Germany) according to the manufacturer's instructions. Preparation for Illumina sequencing was done at Massey University.

Paired end short reads of 100 base pair length were generated from an Illumina-Solexa pipeline, at Beijing Genomics International (BGI). De-novo assembly of the short reads was performed using a CLC Genomics Workbench (version 5.1.5, CLC Bio), and the reference genome, closest to the respective sample, was used to identify mitochondrial contigs through the built-in local BLAST tool. Contig assemblies were aligned to the annotated reference genome using Geneious 6.1.6 (http://www.geneious.com). Genomes had small gaps which were filled by amplifying the relevant regions through Polymerase Chain Reaction (PCR) using previously published protocols and primers. Where needed, new primers were designed using Geneious 6.1.6 (http://www.geneious.com). The process of choice of primers, PCR amplification, sequencing protocol, and confirmation of sequence identity is described in detail in Mahmood et al. (2014) and references therein.

In addition to the four new bird mt genomes reported in this paper, 51 other complete avian mt genomes from GenBank were included in the analyses: 46 Neoaves and 4 Galloanserae (as an outgroup). Since, overall placement of Paleognath taxa is now well established (Gibb et al. 2007; Slack et al. 2007; Phillips et al. 2010; Mitchell et al.

2014), they were omitted from this data set. Galloanserae sequences were used to root our Neoaves trees (Gibb et al. 2007). The full data set is available from the authors on request.

Phylogenetic Analysis

Sequences were aligned in Geneious 6.1.6 (Drummond et al. 2011) at the amino acid level for protein-coding genes and based on secondary structure for RNA genes (see Gibb et al. 2013). Gaps, ambiguous sites adjacent to gaps, NADH6 (light-strand encoded), and stop codons (often incomplete in the DNA sequence) were excluded from the alignment. Protein-coding genes were checked for NUMTs by translating into amino acids. The 12 protein-coding genes were separated into first-, second-, and third-codon positions (the third-codon position may be RY-coded as explained by Gibb et al. 2007), whereas rRNA and tRNA genes were partitioned into stems (S) and loops (L), thus we use five data partitions (see Harrison et al. 2004) as first strategy for phylogenetic analysis. The data was not RY coded in the second strategy. In both cases data was also re-ordered/partitioned based on the rate of evolution. Fast-evolving sites were determined using a tree independent method implemented in the TIGER program (Cummins and McInerney, 2011). Sites were re-ordered and assigned into 30 rate bins, and the most fast-evolving rate bin (bin30) is excluded for further phylogenetic analyses. This generated four types of datasets. PartitionFinder (Lanfear et al. 2014) was used to find the substitution models for all four types of datasets based on the respective rate bins/partitions. A combined total of 13,445 nucleotides (excluding gaps) were used for analyses. We ran analyses in RAxML (Stamatakis et al. 2008) to carry out 500 bootstrap replicates on the data sets, using a general time reversible model with gamma distribution (GTRGAMMA). Bayesian posterior probabilities were also estimated in MrBayes (Huelsenbeck and Ronquist 2001). Bayesian analyses were run for 40,000,000 generations with a burn-in value of 10%. Both RAxML and MrBayes were run using CIPRES Science Gateway (Miller et al. 2010). Trees were visualized in FigTree v1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/ [accessed February 2014]). Consensus networks were implemented in SplitsTree version 4 (Huson and Bryant 2006).

Acknowledgments

The authors thank the Higher Education Commission, Pakistan (HEC) for supporting this research by a PhD scholarship to M.T.M. We also thank Gillian C. Gibb, Bennet McComish, and Lesley J. Collins for general advice.

References

Armstrong MH, Braun EL, and Kimball RT. 2001. Phylogenetic utility of avian ovomucoid intron G: A comparison of nuclear and mitochondrial phylogenies in Galliformes. The Auk 118:799-804.

Bed'Hom B et al. 2003. Characterization of the atypical karyotype of the black-winged kite *Elanus caeruleus* (Falconiformes: Accipitridae) by means of classical and molecular cytogenetic techniques. Chromos. Res. 11:335–343.

Berlin S, Smith NGC, Ellegren H. 2004. Do avian mitochondria recombine? J. Mol. Evol. 58:163–167.

Boore JL. 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27:1767–1780.

Cooper A and Penny D. 1997. Mass survival of birds across the Cretaceous- Tertiary boundary: molecular evidence. Science. (275):1109–1113

Cox WA, Kimball RT, Braun EL. 2007. Phylogenetic position of the New World quail (Odontophoridae): Eight nuclear loci and three mitochondrial regions contradict morphology and the Sibley-Ahlquist tapestry. The Auk 124:71–84

Desjardins P, Morais R. 1990. Sequence and gene organization of the chicken mitochondrial genome - a novel gene order in higher vertebrates. J. Mol. Biol. 212:599–634.

Drummond AJ, et al 2011. Geneious v5.5. Auckland: Biomatters Ltd.

Duchêne S, Archer FL, Vilstrup J, Caballero S, Morin PA. 2011. Mitogenome phylogenetics: the impact of using single regions and partitioning schemes on topology, substitution rate and divergence time estimation. PLoS One. 6:e27138.

Ericson PGP et al. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. Biol. Lett. 2:543–547.

Ericson PGP. 2012. Evolution of terrestrial birds in three continents: biogeography and parallel radiations. J. Biogeogr. 39: 813–824.

Gibb GC, Kardailsky O, Kimball RT, Braun EL, Penny D. 2007. Mitochondrial genomes and avian phylogeny: Complex characters and resolvability without explosive radiations. Mol. Biol. Evol. 24:269–280.

Griffiths CS, Barrowclough GF, Groth JG, Mertz L. 2004. Phylogeny of the Falconidae (Aves): a comparison of the efficacy of morphological, mitochondrial, and nuclear data. Mol. Phylogenet. Evol. 32:101–109.

Hackett SJ et al. 2008. A phylogenomic study of birds reveals their evolutionary history. Science. 320:1763–1768.

Havird JC, Santos SR. 2014. Performance of single and concatenated sets of mitochondrial genes at inferring metazoan relationships relative to full mitogenome data. PLoS One. 9: e84080

Holdaway, RN. 1994. An exploratory phylogenetic analysis of the genera of the Accipitridae, with notes on the biogeography of the family. In: Meyburg BU,

Chancellor RD, editors. Raptor Conservation Today.WGBP, Pica Press, Berlin, Mountfield. p. 601–649.

Jarvis ED et al. 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. Science. 346: 1320-1331.

Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO. 2012. The global diversity of birds in space and time. Nature. 491:444–448.

Johansson U, Parsons T, Irestedt M, Ericson P. 2001. Clades within the higher land birds', evaluated by nuclear DNA sequences. J. Zool. Syst. Evol. Res. 39: 37–51.

Kimball RT, Wang N, Heimer-McGinn V, Ferguson C, Braun EL. 2013. Identifying localized biases in large datasets: a case study using the avian tree of life. Mol. Phylogenet. Evol. 69:1021–32.

Lanfear R, Calcott B, Kainer D, Mayer C, Stamatakis A. 2014. Selecting optimal partitioning schemes for phylogenomic datasets. BMC Evol. Biol. 14:82

Lanyon SM. 1985. Detecting internal inconsistencies. Syst. Zool, 34:397-403.

Lerner HRL, Mindell DP. 2005. Phylogeny of eagles, Old World vultures, and other Accipitridae based on nuclear and mitochondrial DNA. Mol. Phylog. Evol. 37:327–346.

Li B, Liu G, Zhou L, Gu C. 2014. Complete mitochondrial genome of Cinereous vulture *Aegypius monachus* (Falconiformes: Accipitridae). Mitochondrial DNA. 1736:1–2.

Mahmood MT, McLenachan PA, Gibb GC, Penny D. 2014. Phylogenetic position of avian nocturnal and diurnal raptors. Genome Biol Evol. 6:326–332.

Mayr G. 2009. Paleogene Fossil Birds. Springer Berlin Heidelberg. Pp 156-157.

Mayr G. 2014. The origins of crown group birds: molecules and fossils. Palaeontology. 57:231–242.

Meiklejohn KA et al. 2014. Incongruence among different mitochondrial regions: A case study using complete mitogenomes. Mol. Phylogenet. Evol. 78:314–323

Mindell DP, Sorenson MD, Dimcheff DE. 1998. Multiple independent origins of mitochondrial gene order in birds. Proc. Natl. Acad. Sci. USA. 95:10693–10697.

Mitchell KJ et al. 2014. Ancient DNA reveals elephant birds and kiwi are sister taxa and clarifies ratite bird evolution. Science. 344:898—900.

Moen D, Morlon H. 2014. From dinosaurs to modern bird diversity: extending the time scale of adaptive radiation. PLoS Biol. 12:e1001854.

Negro JJ et al. 2006. Convergent evolution of Elanus kites and the owls. J. Raptor Res. 40:222–225.

Olson SL. 1985. The fossil record of birds . In: Farner DS, King JR, Parkes KC, editors, Avian biology, vol 8. Academic Press , New York, pp 79 – 238.

Pacheco MA et al. 2011. Evolution of modern birds revealed by mitogenomics: timing the radiation and origin of major orders. Mol. Biol. Evol. 28:1927–1942.

Penny D, Hendy MD. 1985. The use of tree comparison metrics. Syst. Zool. 34:75-82

Phillips MJ, Gibb GC, Crimp EA, Penny D. 2010. Tinamous and moa flock together: mitochondrial genome sequence analysis reveals independent losses of flight among ratites. Syst. Biol. 59: 90–107.

Powell AFL a, Barker FK, Lanyon SM. 2013. Empirical evaluation of partitioning schemes for phylogenetic analyses of mitogenomic data: an avian case study. Mol. Phylogenet. Evol. 66:69–79.

Roulin A, Wink M. 2004. Predator-prey polymorphism: relationships and the evolution of colour a comparative analysis in diurnal raptors. Biol. J. Linn. Soc. 81:565–578.

Slack K E et al. 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. Mol. Biol. Evol. 23:1144-1155.

Slack KE, Delsuc F, McLenachan PA, Arnason U, Penny D. 2007. Resolving the root of the avian mitogenomic tree by breaking up long branches. Mol. Phylogenet. Evol. 42: 1–13.

Wink M, Sauer-Gürth H. 2004. Phylogenetic relationships in diurnal raptors based on nucleotide sequences of mitochondrial and nuclear marker genes. In: Raptors Worldwide. Chancellor and Meyburg, B. -U., RD, editor. WWGBP: Berlin pp. 483–498.

Wink M, Seibold I, Lotfikhah F, Bednarek W. 1998. Molecular systematics of holarctic raptors (Order Falconiformes). Holarctic Birds of Prey. Adenex, WWGBP, Calamonte. P. 29–48.

Zhong B, Deusch O, Goremykin VV, Penny D, Biggs PJ, Atherton RA, Nikiforova SV, Lockhart PJ. 2011. Systematic error in seed plant phylogenomics. Genome Biol Evol. 3:1340–1348.

5 GENETIC DIVERSITY OF SPOTTED OWLET (*ATHENE BRAMA*) FROM THE AGRO-ECOSYSTEM OF PUNJAB (PAKISTAN)

Muhammad Tariq Mahmood, Patricia A. McLenachan, Gillian C. Gibb, and David Penny

Spotted owlet (*Athene brama*) is a small owl which breeds from Pakistan, India, and Southeast Asia (except Sri Lanka). In Pakistan, it is the resident species of the agro-ecosystem of Punjab. The analysis was carried out to test whether the population of the agro-ecosystem, which is heavily affected with pesticides, shows any trends of decline in diversity. Our ultimate goal is to discover the effect of excessive use of pesticide on the ecology of the spotted owlet. As a first step, we investigate the current genetic diversity of the spotted owlet within Punjab. A total of 22 blood samples were collected from the areas of Pirowal forest reserve, Multan, Muzaffargarh, and Choti Zareen areas of Punjab. Five of them were misidentified by non-expert contributors of samples, and upon sequencing of their DNA, turned out to be Collared scops owl (*Otus bakkamoena*). The birds from forest reserve and agricultural lands were initially considered two populations. The amplification of partial ND5 gene and partial control region (CR1) of ten (10) birds was successful and they were subsequently sequenced. The genetic analysis suggested six haplotypes (h = 6) in two local populations (one from a reserve forest, and another from agricultural lands). The analysis found gene flow between the two populations which suggested they were not fully separate populations. Despite the fact that we could not get samples from the areas where the agro-ecosystem was absent, the number of haplotypes suggest a good genetic diversity for a smaller sample set. Apparently, the population from the agro-ecosystem has more genetic diversity than the one from a forest reserve. Possibly, the population from heavily pesticide sprayed areas may not be affected because of the nocturnal feeding and activity. But we wish to test the same hypothesis with a larger sample size, and from the areas having different levels of pesticides use.

MTM did the sequencing and primary analysis, Patricia A. McLenachan was responsible for important technical advice and laboratory supervision, Gillian C. Gibb and David Penny gave general oversight and planning.

5.1 Introduction

The spotted owlet (*Athene brama*) is one of the small owls which breeds from Pakistan and India to Southeast Asia (tropical Asia), although the species is absent from Sri Lanka. Although it has adapted to live closer to human habitations in the cities, the farmlands are also its common habitat. They nest in a hole in a building or a tree and roost in small groups in cavities in buildings or rocks or the hollows of trees. The number of eggs it lays ranges from 3–5. Higher breeding success of nests near human habitations may be due to increased abundance in rodents for feeding the young (Pande et al. 2007). The very similar little owl (*Athene noctua*) and the spotted owlet (*Athene brama*) form a superspecies that show great clinal variation in size. Spotted owlet is stocky and small bird (21 cm), with the grey-brown upper parts which are heavily spotted with white. The white under parts are streaked with brown. The facial disc is pale with yellow irises. The supercilium and the neckband is white. The flight pattern is deeply undulating, and the sexes are similar (Rasmussen and Anderton 2005).

As mentioned above, the spotted owlet and the little owl (*Athene noctua*) tend to form a superspecies complex, and the spotted owlet has sometimes been treated as a subspecies of the little owl. Of the several subspecies of *A. brama* which have been described so far, four or five are widely accepted. The five subspecies which are widely recognized, are; *A. b. albida* of western Asia (Iran and Pakistan) (Koelz, 1950); *A. b. indica* of northern India (Franklin, 1831); *A. b. brama* of southern India (Temminck, 1821); *A. b. ultra,* which is not always recognized, of northeastern India (Ripley, 1948; and *A. b. pulchra* of Southeast Asia (Hume, 1873). However, the race which was named *poikila* (Yang and Li 1989) is considered invalid and is actually revised to a different species *Aegolius funereus* (Sun et al. 2003). Somewhat similarly, *Athene brama fryi* of southern India and *A. b. mayri* from northern Thailand (Baker 1920; Deignan 1941) are usually considered synonymous with *A. b. brama* and *A. b. pulchra* respectively. There is no natural dividing boundary between the northern and southern Indian populations, and they are known to intergrade. The size of the species decreases from north to south (Ali and Ripley 1981; Baker 1919).

Although nocturnal, sometimes the spotted owlet is seen in the day time. They stare at intruders and bob their head when disturbed from their daytime nest or roost site (Ali 1996). Sometimes when the species is perched in a tree, it can be located and mobbed by the small birds. The species preys on a variety of small vertebrates and insects. The data from Pakistan suggests that they mostly take insect prey (Shah and Beg 2001; Shah et al. 2004; Shah and Beg 2001; Beg et al. 1990). In the arid region of Jodhpur (India), they have been found to prey more on rodents, and especially they picked the genus *Mus* and tended to avoid other rodents, for instance *Tatera*), particularly prior to the breeding season (Jain and Advani 1983). In some other studies they have been noted to catch small snakes such as *Ramphotyphlops braminus*, toads, bats, scorpions, and molluscs (Jadhav and Parasharya 2003; Pande et al. 2004). Their call is a loud and harsh churring and chuckling chirurr-chirurr-chirurr that ends with a chirwak-chirwak. Early dawn or just after sunset, is the usual time when they mainly call (Rasmussen and Anderton 2005; Brahmachary et al. 1972).

In their breeding season, they are found to be allopreening, bill grasping, and ritual feeding as part of courtship behaviour. The female may deflect its tail in invitation, call with the male, and bob its head (Hassan 2008). Their breeding season starts from November and lasts until April (Rasmussen and Anderton 2005). More than one male may copulate with a female. How the family groups are organized is not yet clear. The attempt of pseudocopulation by females (Pravin and Kasambe), may possibly be a kind of displacement behaviour (Watson 1957; Kasambe 2004). Because they nest in cavities, they are often found competing with other hole-nesters, for instance, mynas. They are also found to nest in holes in vertical embankments (Pande et al. 2006). The nest may be lined with feathers and leaves. They may also use the existing nest from a prior occupant with its original lining. The clutch size ranges from three to five spherical white eggs. The incubation begins after the first egg is laid and it leads to the chicks with a wide variation in their sizes. Initially, the chicks are fed on insects (e.g. cockroaches) and later fed on small vertebrate prey (e.g. mice). The chicks are noted to gain weight during the early stages, but before fledging they are known to lose weight (Pande et al. 2011). The number of fledglings may be only one or two and they have been observed to leave the nest at the age of about 20–28 days (Jadhav and Parasharya 2003).

Their brain has a pineal gland, although this was thought to be absent in the Strigiforms (the owls) (Haldar and Guchhait 2000). Because a low level of melatonin is associated with high foraging activity and alertness and a high level is associated with sleep, other birds also show variation in the levels of melatonin between day and night. However, spotted owlets show only a slightly decreased melatonin level at night while it slightly increases in the early afternoon. Other owls also show little day-night variation in melatonin, for instance the barn owl (Guchhait and Haldar 1999; Martin et al. 2005). Environmental factors such as humidity and temperature are known to be associated with the seasonal changes in glandular activity (Haldar et al. 2009)

These birds have been very familiar to humans, and particularly their loud calling is associated with bad omens (Rose 1910). The species name *brama* is given by way of homage to the Hindu supreme spirit, Brahma. In Hindu mythology, particularly, the owl is a 'mode of transport' of Lakshmi i.e. the goddess of wealth (Pittie 2004).

In Pakistan

The spotted owlet *(Athene brama*) is one of the common resident nocturnal raptors found in Pakistan (Roberts 1991). The species has a strong association with the agricultural landscapes of Punjab (Mahmood-ul-Hassan et al. 2007). The little Owl (*Athene noctua*), which is known to avoid densely forested areas, is considered ecological counterpart of spotted owlet in Pakistan (Osieck and Shawyer 1997). Until now, no data on population density or size of the spotted owlet are available. However, considerable data is available on its European ecological counterpart Little Owl (*Athene noctua*). There is no data available on genetic diversity of the Spotted Owlet. The availability of the sequences of Cyt *b* and RAG1 genes (Wink et al. 2009) has helped to understand its phylogenetic relationship within the Order Strigiformes. Also, we have reported a nearly complete mitochondrial genome of this species (Mahmood et al. 2014; also see chapter 2) which has helped to clarify the monophyly of the order Strigiformes, and contributed to the understanding of their evolutionary relationship with the diurnal raptors.

As ecologically susceptible predators, birds of prey are valuable indicators of environment quality (Lerner and Mindell, 2005). The World Conservation Union declared *Gyps* vultures as critically endangered (IUCN, 2014) following a steep decline in the population of > 95% (Prakash et al. 2003) which was later known to be caused by renal failure due to bioaccumulation of diclofenac sodium (Oaks et al. 2004). A recent report discovers that *Aquila* eagles of South Asia are also being affected by the above mentioned drug (Sharma et al. 2014). Another recent study has reported higher levels of heavy metals in birds, particularly raptors including owls (Abbasi et al. 2014). The persistence of heavy metal contamination poses elevated risk for the studied species of birds in Pakistan. Such a scenario necessitates the evaluation and remediation of ecological damage (Movalli 2000; Naccari et al. 2009). It is also needed to prioritise the study the genetic potential of the resident species of raptors which may be directly vulnerable to the deterioration of local environment (Movalli 2000). Similarly, more molecular phylogenetic work is needed for other raptors of Pakistan as well, to inform the conservation programs more explicitly.

The agricultural use of pesticides is on the rise in Pakistan. There has been a tremendous increase in pesticide consumption, in a period just over two decades, which was just 12,530 metric tonnes in 1985 and reached 117,513 metric tonnes in 2005-06 (Khan et al. 2011). This huge increase in pesticides has increased resistance in pests, which in turn has disturbed the delicate balance between predators and pests (Hasnain 1999). Another investigation suggests that birds are either killed due to direct exposure to the pesticides or indirectly through consuming the contaminated food/prey (Pimentel 2005). Based on above mentioned reasons, we chose Spotted Owlet (*Athene brama*), to investigate whether its population shows some trends.

This is a pilot study to investigate the popuation diveristy of *A. brama* wihtin the Punjab region (Pakistan). We plan to collect avian blood samples from the areas of Pirowal (PP), Multan (BM), Choti Zareen (DG), Muzaffargarh (NB, NW, BJ, HP). Our choice of the species and research area is based on the following;

i) It is the resident species of the agro-ecosystem of Punjab (Mahmood-ul-Hassan et al. 2007).
ii) The study on prey/ food preference of spotted owlet (*Athene brama*) in Punjab, has showed that it constituted insects (47%), small rodents (28%), birds

(12%), and plant tissues (11%). This is an unusual component (Shah et al. 2004). Bird population would be under the effect of direct poisoning if they consumed pesticide infested plant tissues.

iii) The areas in Punjab are under direct effect of massive pesticide use while the sampling localities in other provinces are not. It will give us a comparison of demographic trends (whether affected by overuse of pesticides or not?) and genetic diversity.

5.2 Materials and Methods

Sampling

The samples were collected from all of the localities mentioned (see Introduction). The whole birds were collected from the game birds hunters who would get them as a by-kill. The detail of samples with the information on their locality is given in the Table 1 below.

Table 1. The details of samples collected from Punjab (Pakistan).

Species	Locality/ code	No. of samples
Otus bakkamoena	Nonari basti/ NB	1
	Kacha/ K1-2	2
	Nai wala/ NW2	1
	Choti Zareen/ DG3	1
Athene brama	BZ university, Multan/ BM1-3	3
	Pirowal plantation/ PP1-9	9
	Nai wala/ NW1	1
	Choti Zareen/ DG1-2	2
	Basti Jalal/ BJ1	1
	Hassan Pur/ HP1-2	2

Fresh tissue samples were not collected before the expert identification of the bird. So the birds were stored by our collaborators, in a freezer which was unattended during a huge power crisis. Unfortunately the samples were decomposed and degraded. In a

second sampling attempt, the primary author (MTM) undertook the sampling while some of the samples were contributed by some acquaintances. But only samples from Punjab were possible. Visits to Kohat, Nagar parkar, and Muzaffarabad could not be made due to security issues at that particular time. Five, out of the 23 samples in total, were misidentified by the samplers and they were later identified as *Otus bakkamoena* upon sequencing. Figure 1 (below) shows the sampling sites in Pakistan.

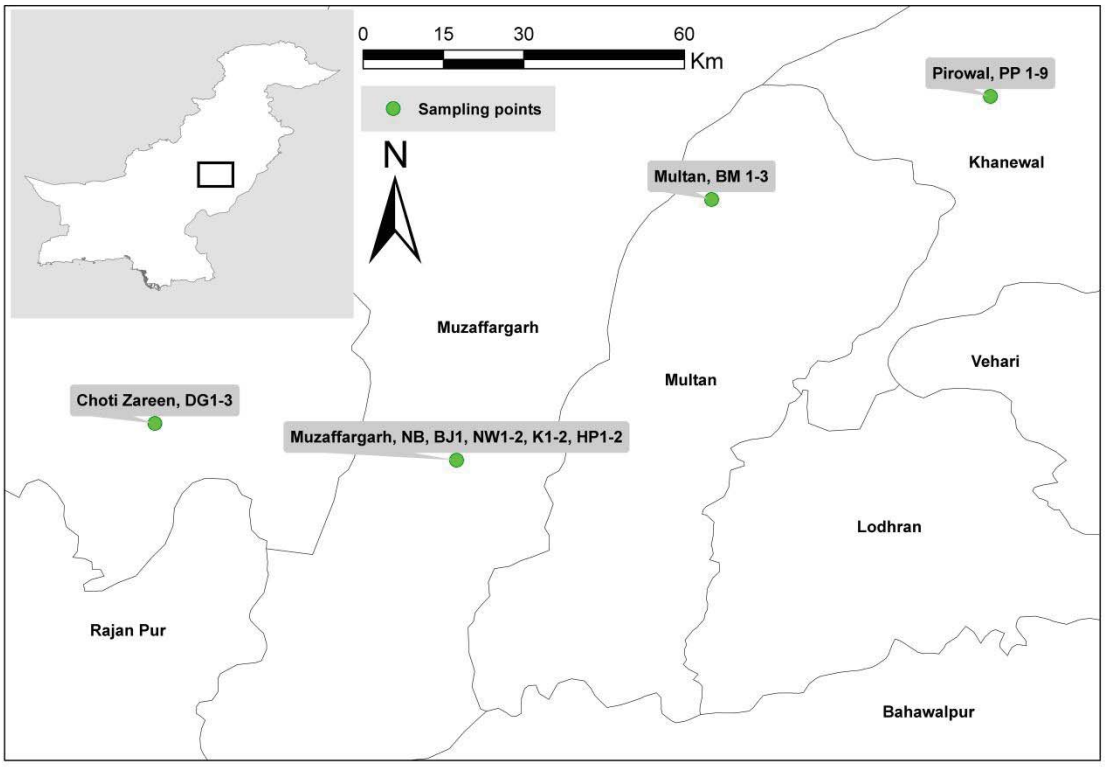


Figure 1. Sampling localities for Spotted owlet (*Athene brama*) Collared Scops Owl (*Otus bakkamoena*) from Punjab (Pakistan).

Molecular Methods

Following description of the molecular methods is adopted from Mahmood et al. (2014). Extractions of genomic DNA from each of the birds were performed at the Institute of Fundamental Sciences from the liver tissue using the High Pure PCR Template Preparation Kit (Roche Applied Science, Mannheim, Germany) according to the manufacturer's instructions. The extracted genomic DNA samples were subsequently used as template DNA for short-range PCRs to generate overlapping fragments 0.5–3 kb in length. Short-range primer combinations were found using our laboratory database as described in Slack et al. (2006), and any new primers required were designed using Geneious Pro (http://www.geneious.com). The primers used in the present study, with their nucleotide sequences are listed in Table 2. Sequencing was

performed using BigDye Terminator Cycle Sequencing reagents according to the manufacturer's instructions (Applied Biosystems, Foster City, CA) and the reactions were run on an ABI 3730 automated sequencer (Applied Biosystems) by Massey Genome Service. In all cases, overlaps between sequences were sufficient to ensure synonymy and sequence identity was confirmed through BLAST searches (http://www.ncbi.nlm.nih.gov/blast/), confirmation of amino acid translation in coding regions, and alignment with other species. Sequences were aligned using Sequencher 4.7 (Gene Codes Corp., Ann Arbor, MI) and then manually edited and checked for complete concurrence between overlapping sequences.

Table 2. Primer pairs used for the PCR amplification of the avian mtDNA regions used in the present study.

	Primer Sequence (5' to 3')	
Primer Code	Forward	Reverse
Av1753F12S-Av2150R12S	AAACTGGGATTAGATACCCCACTAT	GAGGGTGACGGGCGGTRTGTAC
Av438FDloopB-Av907DloopR	TCACGTGAAATSAGCAACCC	TGTTTAGAAAGTTAGAGGAARTG
Bat12769ND5F-Av16137tProR	TGCTCAGGATCAATTATTCA	ARAATRCCAGCTTTGGGAGTTGG

Analysis

Sequences were edited, assembled, and aligned in Geneious 6.1 (Biomatters Ltd., Auckland, New Zealand http://www.geneious.com/) at the amino acid level for protein-coding genes and verified by eye. The Table 3 shows the detail of species, their locality codes, and the names of amplified genetic markers with length of each marker in base pairs (bp). If any marker was not successfully amplified for any of the samples, it is shown as 'No'. Genetic variation within each population was calculated by using DnaSP 5.0 (Librado and Rozas 2009). Following summary statistics were calculated: nucleotide diversity per site (π) (Nei and Li 1979), haplotype diversity (Hd), Watterson's per-site population mutation rate estimator (θ_W) (Watterson 1975), and the number of polymorphic (segregating) sites (S). Sites with missing data or indels were eliminated in pairwise sequence comparisons. A combined total of 1,074 nucleotides,

concatenated from 603 bp (base pair) of ND5 and 471 bp of Control Region (CR1), were used for analyses.

Table 3. The detail of the sequenced markers with their respective species and locality codes.

Species	code	12S rRNA	ND5	Control region Part I (CRI)	Control region Part II (CRII)
Otus bakkamoena	NB1	Yes (373 bp)	Yes (325 bp)	No	No
	K1	Yes (373 bp)	No	No	No
	K2	Yes (373 bp)	No	No	No
	NW2	Yes (373 bp)	Yes (603 bp)	No	No
	DG3	Yes (373 bp)	Yes (603 bp)	No	No
Athene brama	BM1	Yes (366 bp)	No	Yes (471 bp)	No
	BM2	Yes (366 bp)	Yes (603 bp)	Yes (471 bp)	No
	BM3	Yes (366 bp)	Yes (603 bp)	Yes (471 bp)	Yes (876 bp)
	PP1,3	Yes (366 bp)	No	Yes (471 bp)	Yes (876 bp)
	PP2, 4-6	Yes (366 bp)	Yes (603 bp)	Yes (471 bp)	Yes (876 bp)
	PP7,8	Yes (366 bp)	Yes (603 bp)	Yes (471 bp)	No
	PP9	Yes (366 bp)	No	Yes (471 bp)	No
	NW1	Yes (366 bp)	Yes (603 bp)	Yes (471 bp)	No
	DG1	Yes (366 bp)	No	Yes (471 bp)	No
	DG2	Yes (366 bp)	No	Yes (471 bp)	Yes (589 bp)
	BJ1	Yes (366 bp)	Yes (603 bp)	Yes (471 bp)	No
	HP1-2	Yes (366 bp)	No	No	No

5.3 Results and Discussion

All of the 23 samples were amplified for 12S rRNA and sequenced. Both the 16S and 12S ribosomal RNA (rRNA) mitochondrial genes are suggested as reliable markers to taxonomically classify the animal tissues (Yang et al. 2014). Five of the sequences, in a 366 base pairs alignment, were distinct from the rest. They showed a 94% identity to the Japanese Scops Owl (*Otus semitorques*), when compared with the available nucleotide data through the BLAST searches. After a careful consideration these sequences were assigned to the Collared Scops Owl (*Otus bakkamoena*) because that is the only species of the genus *Otus* reported in the sampling area (Grimmett et al. 1999). BLAST searches were also performed for the rest of the sequences and they showed 99-100% identity with the Spotted Owlet (*Athene brama*). This study presents the first 12S rRNA sequence data for both of these species. The single nucleotide polymorphism (SNPs), as it is highlighted with red rectangle (Figure 2), could indicate the different morphs or sub species of the respective species. The sequence data of more individuals of both species with their possible morphological traits, can further confirm this. Overall, this confirms that 12S rRNA gene can be a potential marker for the tissue identification of owls.

Figure 2. An alignment (366 base pairs) of 12S rRNA sequences of the Collared Scops Owl (*Otus bakkamoena*; sequences 1-5) and the Spotted Owlet (*Athene brama*; sequences 6-23).

The Pirowal (PP) is a forest reserve, and nine samples were obtained. However three of the nine samples coming from this locality (PP1, PP3 and PP9) were not successfully amplified for *ND5* region. All other localities belong to the agro-ecosystem. Amplification failed for both of the samples from Choti Zareen (DG1-2) which is 160 km (kilometres) apart from Pirowal. Two samples from Muzaffargarh (HP1-2) and one sample from Multan (BM1) also failed to amplify for the ND5 region. The possible reason for failure of PCR (Polymerase Chain Reaction) could be the presence of an inhibitor.

The distance of Pirowal from Muzaffargarh is about 70 km, and from Multan it is even less than that (see Figure 1). For that reason all of the samples may be considered to belong to a single population, although there is river 'Chenab' between Pirowal, Multan and Muzaffargarh (but the river should not be a barrier for strong fliers like owls). Furthermore, calculations for the 'Index for genetic differentiation between populations' (F_{ST}) shows no differentiation between individuals samples from two sides of the above mentioned river. That suggests a gene flow between individuals belonging to two localities. Only the longest possible sequence regions of the available data with more variation were selected for analysis. In the present case they are ND5 and CR1. The rest of the samples/ individuals either failed to amplify or they had only shorter sequence regions available without any meaningful variation and were not alignable to the sequences selected for analysis. The concatenated sequence alignment for a total of 10 individuals (N = 10) were analysed for genetic diversity/ polymorphism (Table 4).

Table 4. Genetic polymorphism data for mitochondrial partial *ND5 and CR1* sequences from the spotted owlet (*Athene brama*) population(s) from Punjab (Pakistan).

Population	N (number of sequences)	h (number of haplotypes)	S (segregating sites)	θ_W (population mutation rate per site)	π (nucleotide diversity per site)
PP2,4-8, BM1,3, BJ1, NW1	10	6	6	0.00197	0.00157

We are aware of the range of sample sites and the low number of samples. We found six segregating sites (S = 6) and six haplotypes (h = 6) among all the individuals analysed. There are two haplotypes from Pirowal (PP), and Multan (BM2) each with one

segregating site for each of the haplotypes. Two haplotypes from Muzaffargarh i.e. NW1 and BJ1, both were differentiated from the rest of the conspecifics by one and two segregating sites respectively. One of the two haplotypes from Muzaffargarh (BJ1) with greater differentiation as compared to their conspecifics from Pirowal (a forest reserve) could be a random variation. For the whole set of individuals, the population mutation rate per site is 0.00197, while the nucleotide diversity per site is 0.00157. Moreover, 12S rRNA data does not add lot more information but it is consistent with ND5 and CR results The following Figure (3) shows a Neighbour-Joining tree of spotted owlet (*Athene brama*) haplotypes.

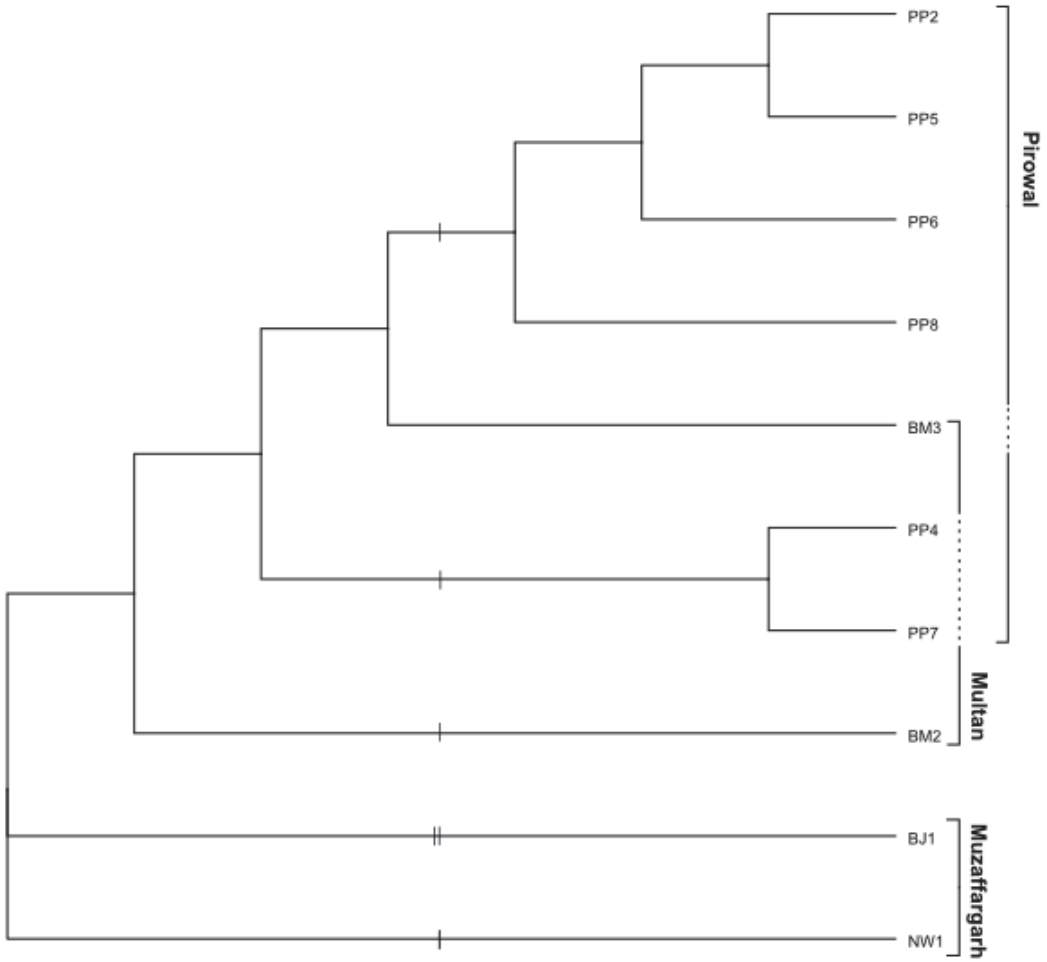


Figure 3. Neighbour-Joining tree of spotted owlet (*Athene brama*) haplotypes using ND5 and CR sequences. The number of bars at different nodes represent number of segregating sites. The sequences from the Multan and Pirowal localities overlap.

Two owl species i.e. eagle owl (*Bubo bubo*) and little owl (*Athene noctua*) are reported to be more susceptible to organohalogen pollutants than a diurnal kestrel (*Falco*

tinnunculus) because of their prey preference. i.e. brown rat (*Rattus norvegicus*) for both the owl species, and Eurasian tree sparrow (*Passer montanus*) for the kestrel. The kestrel food chain (sparrow-kestrel) was found with lower level of organohallogent pollutants than owl food chain (rat-owl) (Yu et al. 2013). Similarly, our hypothesis that the excessive use of pesticides in agro-ecosystem could be a bottle-neck for the population, may stand true. But with the present sample set it appears that spotted owlet (*Athene brama*) populations are thriving more successfully in the agro-ecosystem which is in agreement with previous observations (Shah et al. 2004; Mahmood-ul-Hassan et al. 2007). Since, little owl (*Athene noctua*) is considered the ecological counterpart of the spotted owlet (*Athene brama*), therefore an expanded sample set throughout its species range including other subspecies may show some demographic trends between agro-ecosystem and any alternative habitat. The alternative explanation for the divergent haplotypes could be migrants from a nearby unsampled locality.

To compare the present study to the little owl from Europe could be helpful in deciding future sampling strategies. A total of 146 haplotypes of the little owl (*Athene noctua*) across the whole of Europe were recovered based on mitochondrial control region (CR) sequences, and the phylogenetic analysis was in congruence with the four subspecies described for that species (Pellegrino et al. 2014). It was noted that the Alps were the real barrier in their dispersal and diversification which could be a reason for the diversification of different populations into subspecies. It should be noted that the CR sequences from Pellegrino et al (2014) do not overlap with the pieces sequenced for the present study. The Suleman ranges in Pakistan are of comparable height to the Alps, and it will be of interest to get samples from the both sides of the mountain range to see if the sequence data overlapping with that of Pellegrino et al (2014) shows any demographic changes in the spotted owlet populations. Also, the choice of localities from the foothills of Himalayas (Muzaffarabad), to the dry desert-like area (Nagar Parkar) would provide a comparative data on the genetic diversity of the spotted owlet.

In the second part of the control region sequence of the spotted owlet, a 109 base pair repeat sequence was found at least 5 times but the actual number of repeats could be variable in different populations. The BM3 sample had a transition from T to C at position 16 of the repeat. Note that BM3 was sampled 5 years before the rest of the sample set and has been reported as complete mitochondrial genome (Accession:

KF961185). Overall, the above mentioned repeat sequence could be a good population marker for future investigations, but again a larger sample set will help to see if this transition is linked to any particular subspecies.

To the best of our knowledge, there is no home range data available for this species. According to the IUCN red data list (IUCN 2014), *Athene brama* is declared 'least concern' having extremely large range and stable population trend although its population size has not been quantified yet. But it is not believed that this species could approach 'vulnerable' status under population size criterion.

References

Aasheesh P. 2004. A dictionary of scientific bird names originating from the Indian region. *Buceros* 9: 1–31

Abbasi NA, Jaspers VLB, Chaudhry MJI, Ali S, Malik RN. 2014. Influence of taxa, trophic level, and location on bioaccumulation of toxic metals in bird's feathers: A preliminary biomonitoring study using multiple bird species from Pakistan. Chemosphere. 120C:527–537.

Ali S, Ripley SD. 1981. Handbook of the birds of India and Pakistan. Oxford University Press. pp. 299–302.

Ali S. 1996. The Book of Indian Birds. (12 ed.). Oxford University Press.

Anika J, Parasharya BM. 2003. Some observations on the nesting behaviour and food of the spotted owlet *Athene brama*. Zoos' Print J. 18: 1163–1165.

Baker ECS. 1919. Descriptions of subspecies of *Carine brama*. Bulletin BOC. 40: 60–61.

Baker, ECS. 1920. Notes on a collection of bird-skins formed by Mr. E.G.Herbert, C.M.Z.S, M.B.O.U". J. Nat. Hist. Soc. Siam 4: 25–43.

Beg MA, Maqbool M, Mushtaq–ul–Hassan M. 1990. Food habits of spotted owlet, *Athene brama*. Pak J. Agri. Sci. 27: 127–131.

BirdLife International. 2012. *Athene brama*. IUCN Red List of Threatened Species. Version 2014.2. International Union for Conservation of Nature. Retrieved 19 October 2014.

Brahmachary RL, Basu TK, Sengupta AJ. 1972. On the daily screeching time of a colony of spotted owls *Athene brama* (Temminck). J. Bombay Nat. Hist. Soc. 69: 649–651.

Charde P, Kasambe R. 2007. Study of the mounting behaviour of Spotted Owlets *Athene brama* in Maharashtra, India. Abst. World Owl conf.

Deignan HG. 1941. New birds from the Indo-Chinese sub-region. The Auk. 58: 396–398.

Geneious version 6 created by Biomatters. Available from http://www.geneious.com/

Grimmett R, Inskipp C, Inskipp T. 1999. Pocket guide to the birds of the Indian Subcontinent. Oxford University Press, New Delhi, India.

Guchhait P, Haldar C. 1999. Circadian rhythms of melatonin and sex steroids in a nocturnal bird, Indian spotted owlet *Athene brama* during reproductively active and inactive phases. Biol. Rhythm Res. 30: 508–516.

Haldar C, Guchhait P. 2000. Pineal gland of a nocturnal bird, Indian spotted owlet, *Athene brama*: Morphological and endocrine observations. J. Exp. Zool. 287: 145–150.

Haldar C, Kumari S, Guchhait P. 2002. Seasonal Adreno-Cortical Cycle of a Nocturnal Bird, Indian Spotted Owlet *Athene brama*: Biochemical and Morphological Observations. Biol. Rhythm Res. 33: 53–63.

Hasnain T. 1999. Pesticide Use and Its Impact on Crop Ecologies: Issues and Options. Sustainable Development Policy Institute (SDPI), Islamabad. Working Paper Series.

Jain AP, Advani R. 1983. Winter food of spotted owlet, *Athene brama indica*. J. Bom. Nat. Hist. Soc. 80: 415–416.

Kasambe R. 2004. Unusual mounting behaviour of a female Spotted Owlet (*Athene brama*). *Newsletter for Birdwatchers* 44: 63–64.

Khan M, Akram N, Husnain MIU, Padda IU H, Qureshi SA. 2011. Poverty-Environment Nexus: Use of Pesticide in Cotton Zone of Punjab, Pakistan. J. Sustain. Dev. 4:163–173.

Lan Y, Li. 1989. A New Subspecies of The *Athene brama* (Spotted Little Owl) — *A. b. poikila* (Belly-mottled Little Owl). Zool. Res. 10: 303–308.

Lerner HRL, Mindell DP. 2005. Phylogeny of eagles, Old World vultures, and other Accipitridae based on nuclear and mitochondrial DNA. Mol. Phylogenet. Evol. 37: 327–346.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 25:1451–2.

Mahmood MT, McLenachan PA, Gibb GC, Penny D. 2014. Phylogenetic position of avian nocturnal and diurnal raptors. Genome Biol. Evol. 6:326–332.

Mahmood-ul-hassan M, Beg MA, Mushtaq-ul-Hassan M, Rana SA. 2007. Nesting and Breeding Habits of the Spotted Owlet (Athene brama) in Punjab , Pakistan Your use of this PDF , the BioOne Web site , and all posted and associated content. J. Raptor Res. 41:50–52.

Mehmood-ul-Hassan M. 2008. Some observations on behaviour of Spotted Owlet (*Athene brama*) during its breeding season. J. Anim. Pl. Sci. 18: 47–49.

Movalli P. 2000. Heavy metal and other residues in feathers of laggar falcon Falco biarmicus jugger from six districts of Pakistan. Environ. Pollut. 109:267–275.

Naccari C, Cristani M, Cimino F, Arcoraci T, Trombetta D. 2009. Common buzzards (Buteo buteo) bio-indicators of heavy metals pollution in Sicily (Italy). Environ. Int. 35:594–8.

Nei M, Li W-H. 1979. Mathematical model for studying variation in terms of restriction endonucleases. Proc. Nat. Acad. Sci. USA. 76: 5269–5273.

Oaks JL et al. 2004. Diclofenac residues as the cause of vulture population decline in Pakistan. Nature. 427:2002–2005.

Osieck E, Shawyer C. 1997. Barn Owl, *Tyto alba* . In: Hagemeijer EJM, Blair MJ, editors, The EBCC atlas of European breeding birds: their distribution and abundance. T. and Poyser AD, London, UK. pp 398–399.

Pande S, Amit P, Mahajan MN, Joglekar C, Mahabal A. 2007. Effect of Food and Habitat on Breeding Success in Spotted Owlets (*Athene brama*) Nesting in Villages and Rural Landscapes in India. J Rapt Res. 41: 26–34.

Pande S, Pawashe A, Bastawade DB, Kulkarni PP. 2004. Scorpions and molluscs: some new dietary records for Spotted Owlet *Athene brama* in India. Newsletter Ornith. 1: 68–70.

Pande S, Pawashe A, Mahajan MB, Mahabal A, Yosef R, Dahanukar N. 2011. Biometry based ageing of nestling Indian Spotted Owlets (*Athene brama brama*) . Zookeys 132: 75–88.

Pande S, Pawashe A, Mahajan MN, Mahabal A. 2006. Changing nest site preference for holes in earth cuttings in Spotted Owlet *Athene brama*. *Indian Birds* 2: 7–8.

Pellegrino I et al. 2014. Phylogeography and Pleistocene refugia of the Little Owl *Athene noctua* inferred from mtDNA sequence data. Ibis. 156:639–657.

Pimentel D. 2005. Environmental and Economic Costs of the Application of Pesticides Primarily in the United States. Environ. Dev. Sustain. 7:229–252.

Prakash V et al. 2003. Catastrophic collapse of Indian white-backed *Gyps bengalensis* and long-billed *Gyps indicus* vulture populations. Biol. Conserv. 109:381–390.

Rasmussen PC, Anderton JC. 2005. Birds of South Asia: The Ripley Guide. 2. Smithsonian Institution & Lynx Edicions. pp. 246–247.

Roberts TJ. 1991. Birds of Pakistan. Vol. 1. Oxford Univ.Press, Oxford, U.K.

Rose HA. 1910. Panjab Folklore Notes. Folklore 21: 216–217.

Santhanakrishnan R; Ali AHMS; Anbarasan U. 2011. Food habits and prey spectrum of Spotted Owlet (*Athene brama*) in Madurai District, Tamil Nadu, southern India. Chinese Birds 2: 193–199.

Shah ZA, Beg MA, Khan AA. 2004. Prey Preferences of the Spotted Little Owl (*Athene brama*) in the Croplands Near Faisalabad – Pakistan. Int. J. Agric. Biol. 6:278–280.

Shah ZA, Beg MA, Khan AA. 2004. Prey Preferences of the Spotted Little Owl (*Athene brama*) in the Croplands Near Faisalabad–Pakistan. Int. J. Agri. Biol. 6: 278–280.

Shah ZA, Beg MA. 2001. Food of the Spotted Little Owl (*Athene brama*) at a place where a cropland and a sandy wasteland met. Pak J. Zool. 33: 53–56.

Sharma AK et al. 2014. Diclofenac is toxic to the Steppe Eagle *Aquila nipalensis*: widening the diversity of raptors threatened by NSAID misuse in South Asia. Bird Conserv. Int. 24:282–286.

Slack K E et al. 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. Mol. Biol. Evol. 23:1144-1155.

Sun YH, Lin BZ, Scherzinger W. 2003. Belly-mottled little owl *Athene brama poikila* should be boreal owl (*Aegolius funereus beickianus*) . Current Zoology 49: 389–392.

Watson A. 1957. The behaviour, breeding, and food-ecology of the snowy owl *Nyctea scandiaca*. Ibis. 99: 419–462

Watterson GA. 1975. On the number of segregating sites in genetic models without recombination. Theoret. Popul. Biol. 7: 256–276.

Wikelski M, Tarlow EM, Eising CM, Groothuis TGG, Gwinner, E. 2005. Do night-active birds lack daily melatonin rhythms? A case study comparing a diurnal and a nocturnal-foraging gull species. J. Ornithol. 147: 107–111.

Yang L et al. 2014. Species identification through mitochondrial rRNA genetic analysis. Sci. Rep. 4:4089.

Yu L et al. 2013. Occurrence and biomagnification of organohalogen pollutants in two terrestrial predatory food chains. Chemosphere. 93:506–511.

6 THESIS SUMMATION

This chapter presents an overview of this thesis and the possible future questions for avian raptor evolution.

The first chapter (Chapter 1) gives an overview of various attempts at inferring relationships of raptors, with other Neoavian orders in general, and within the broad category of raptors themselves, i.e. both diurnal and nocturnal raptors. The attempt has been made to portray an array of techniques and various types of datasets which could be employed to study birds of prey, both at the level of macroevolution as well microevolution. The description about each of the raptor families includes the information related to their diversity that involves both extant and fossil taxa, as well as a glimpse of their ecology. Despite a variety of data sources and decades of research both the relationships between modern avian orders of raptors, and their times of origin are still uncertain. The review of prior work generally consists of studies based on morphological (Livezey and Zusi 2007), paleontological (Mayr 2009; 2014), DNA-DNA hybridization (Sibley and Ahlquist 1990) and short DNA sequences (Wink et al. 1998; Lerner and Mindell 2005). This chapter also presents a general overview of the problems in resolution of the avian tree of life.

The review of previous published/available data is crucial to identify new possibilities and strategies to resolve the problem. In addition to the different datasets which have been used to study avian raptor evolution, it was felt that currently there is need of a tool which could give a general overview of the available DNA sequence data. That is where *RaptorBase* (a database of raptor sequences) comes in. The creation, maintenance and workflow for this database is discussed in Chapter 2 of this thesis. Although GenBank is an indispensable source of biological information, it has some shortcomings in its taxonomy (Federhen, 2003). But we follow the same taxonomy, for a good flow of information to build a MySQL based online database of raptor sequences. *RaptorBase* is able to update itself with every new submission of nucleotide data to GenBank. It can display a matrix of availability or otherwise of nucleotide information, by species as well as by gene. This database could be easily modified for other groups as well. There are examples of some comprehensive phylogenetic studies based super matrix of taxa or

loci (McMahon and Sanderson, 2006; Bininda-Emonds et al., 2007) and the same could be followed to build a supertree phylogeny of raptors, with the help of *RaptorBase*. PhyLoTa (Sanderson et al. 2008) is a good example of harvesting GenBank information for phylogenetics, but it does not give the information in the form of a matrix of gene names and species. *RaptorBase* provides such a matrix with these properties, which gives phylogeneticists more manual control to design a project.

Chapter 3 has already been published (Mahmood et al. 2014). Our choice of complete mitochondrial genomes is due to some specific reasons: they reduce 'stochastic error' associated with short datasets, are small and easier to isolate as compared to nuclear DNA, provide an independent dataset which is complementary to nuclear dataset, have no recombination due to their maternal inheritance, and have higher mutation rates than the nuclear DNA (see Kvist 2000 and references therein). In this study we ask some specific questions. The core of these questions is to test whether owls (nocturnal raptors) are monophyletic and whether raptors (nocturnal and diurnal combined) are monophyletic or not? According to previous studies owls appear to be monophyletic on the overall avian tree but that could be due to their faster rate of evolution (Pratt et al. 2009; Pacheco et al. 2011). In that case they created long branches on the tree. Long-branch attraction is known to mislead the phylogenetic inference (Penny and Hendy 1985) which could lead to a misleading conclusion. We have reported two new almost complete mitochondrial genomes of owls (spotted owlet, *Athene brama*; and oriental bay owl, *Phodilus baadius*) to break two existing long branches. It improved the taxon sampling on the tree of raptors. In addition to that, there was no representation of Secretarybird (*Sagittarius serpentarius*) on the tree of raptors, which left its position as unclear on the tree. We have reported a complete mitochondrial genome of the Secretarybird too. After the inclusion of these three genomes in the dataset of raptors mitochondrial genomes, we conducted the phylogenetic analysis. This analysis leads us to conclude that owls are monophyletic and Secretarybird has a basal position in relation to the lineage of both Pandionidae and Accipitridae. Here we find the agreement between nuclear and mitochondrial datasets.

Chapter 4 is built on findings of the Chapter 3 and the suggestions made there. We have improved further the taxon sampling through breaking some potential long branches. As we noticed, there was a lack of representation of Elanid kites and Old World vultures

within Accipitridae, and caracaras within Falconidae. We also intended to include another New World vulture, but during the sequencing process we noticed it was a misidentified/ mislabelled sample of an Old World vulture instead. Therefore, we place emphasis on keeping good voucher specimens. We have reported four more nearly complete mitochondrial genomes to improve the taxon sampling on the tree of raptors. These genomes include *Elanus caeruleus* (and Elanid kite), *Gyps fulvus* (Griffon vulture) *Aegypius monachus* (Cinereous vulture), and *Phalcoboenus australis* (common caracara). These final four genomes were then added in for a more detailed analysis of raptor relationships. The unique aspect of methodology of this chapter from the Chapter 3 is removal of the fast evolving sites from the final dataset. We followed a two pronged strategy to test whether removal of fast evolving sites help to improve the overall resolution of the tree. We noticed that removal of the faster evolving sites decreased rather than increased the resolution of the raptor tree. We can confirm the placement of caracara within Falconidae, and Elanid kite and Old World vultures within Accipitridae. Particularly, the Elanid kite is at the basal position in Accipitridae. A moderate increase in nodal support values for raptors tree was noticed, in comparison to the previous mitogenomic phylogenetics efforts. There was another consistent finding in both the Chapters 3 and 4, that both diurnal and nocturnal groups of raptors are part of 'Higher Land birds' group, which is congruent with previous studies (Ericson et al. 2006; Hackett et al. 2008).

As the last part of this thesis, a population study (Chapter 5) was undertaken on spotted owlet (*Athene brama*). The analysis was carried out as a first step to investigate the current genetic diversity of the potted owlet within the agro-ecosystem of the Punjab province (Pakistan). Our ultimate goal is to discover whether the excessive use of pesticides has any effect on the ecology of this species. In the present study we establish that 12S rRNA gene is a useful marker for species identification within owls, although its utility to the level of sub species can be further confirmed with a larger sample set aided with morphological data. The genetic analysis suggests six haplotypes (h = 6) among the individuals from four different localities. Despite the fact that we could not get samples from the areas where the agro-ecosystem was absent, the number of haplotypes suggest a good genetic diversity for a smaller sample set. But we wish to test the above mentioned hypothesis about the effect of excessive pesticide use, with a larger sample size. So, ranging from testing inter-ordinal relationships of raptors to the

population genetics of one of the owls, my study makes use of sequence data to study both the micro-evolution and the macro-evolution of these beautiful birds.

We wish to extend the phylogenetic investigations on raptors in future with more mitochondrial genome and nuclear sequences. Currently, only one of the Old World groups of vultures is represented on the tree, and therefore inclusion of a representative from both groups would help to test whether Old World vultures have one or two origins. Also the presence of at least two New World vultures would help to stabilise the tree, since the inclusion of only one available New World vulture disrupts the overall topology of the tree. Moreover, the forest falcon (*Micrastur gilvicollis*) still appears to be a long branch which might be misleading the overall raptor phylogeny. Therefore, at least one more representative mitochondrial genome from the sub-family Polyborinae would help to discard the possibility of long-branch attraction problem for the forest falcon. Inclusion of these additional taxa will help to improve the overall resolution of the raptor tree.

A recent genome scale investigation places the New World vultures with the eagles (Jarvis et al. 2014). As mentioned above, long-branch attraction problem affects the mitogenomic phylogeny and that may or may not be true in case of the whole genome datasets. We need to confirm whether the phylogenetic placement of New World vultures in the whole genome phylogeny reflects their actual position in the tree of raptors or it is also a result of the long-branch attraction problem.

It is known that some subset of mitochondrial genome can be used as a proxy for the whole mitogenome to recover the historic phylogenetic signal across the Metazoa (Havird and Santos 2014), but the identity and the exact number of these genes changes for a particular lineage in question. But in case of avian mitogenomic phylogenetics, whole mitochondrial genome performs better than a small subset as a proxy for the whole genome (Powell et al. 2013; Meiklejohn et al. 2014) as well as in mammals (Duchene et al. 2011). We want to explore the possibility of improving the resolution of the raptor tree through investigating different partitioning strategies and RY coding to see if this improves the overall phylogenetic resolution.

References

Bininda-Emonds ORP et al. 2007. The delayed rise of present-day mammals. Nature 446:507–512.

Duchêne S, Archer FL, Vilstrup J, Caballero S, Morin PA. 2011. Mitogenome phylogenetics: the impact of using single regions and partitioning schemes on topology, substitution rate and divergence time estimation. PLoS One. 6:e27138.

Ericson PGP et al. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. Biol. Lett. 2: 543–547.

Federhen S. 2003. The taxonomy project. In: The NCBI handbook http:// www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=handbook. ch. 4.

Hackett SJ et al. 2008. A phylogenomic study of birds reveals their evolutionary history. Science. 320: 1763–1768.

Havird JC, Santos SR. 2014. Performance of single and concatenated sets of mitochondrial genes at inferring metazoan relationships relative to full mitogenome data. PLoS One. 9: e84080

Jarvis ED et al. 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. Science. 346:1320-1331

Kvist L. 2000. Mitochondrial DNA. *In:* Phylogeny and phylogeography of European parids. Oulu University Library. Pp. 15-17.

Lerner HRL, Mindell DP. 2005. Phylogeny of eagles, Old World vultures, and other Accipitridae based on nuclear and mitochondrial DNA. Mol. Phylogenet. Evol. 37: 327–346.

Livezey BC, Zusi RL. 2007. Higher-order phylogeny of modern birds (Theropoda, Aves: Neornithes) based on comparative anatomy. II. Analysis and discussion. Zool. J. Linn. Soc. 149: 1–95.

Mahmood MT, McLenachan PA, Gibb GC, Penny D. 2014. Phylogenetic position of avian nocturnal and diurnal raptors. Genome Biol. Evol. 6:326–332.

Mayr G. 2009. Paleogene Fossil Birds. Springer Berlin Heidelberg. pp 156-157.

Mayr G. 2014. The origins of crown group birds: molecules and fossils. Palaeontology. 57: 231–242.

McMahon, MM, Sanderson MJ. 2006. Phylogenetic supermatrix analysis of GenBank sequences from 2228 papilionoid legumes. Syst. Biol. 55:818–836.

Meiklejohn KA et al. 2014. Incongruence among different mitochondrial regions: A case study using complete mitogenomes. Mol. Phylogenet. Evol. 78C:314–323

Pacheco, MA et al. 2011. Evolution of modern birds revealed by mitogenomics: timing the radiation and origin of major orders. Mol. Biol. Evol. 28:1927–1942.

Penny D, Hendy MD. 1985. The use of tree comparison metrics Syst. Zool., 34:75-82

Powell AFLA, Barker FK, Lanyon SM. 2013. Empirical evaluation of partitioning schemes for phylogenetic analyses of mitogenomic data: an avian case study. Mol. Phylogenet. Evol. 66:69–79.

Pratt RC et al. 2009. Toward resolving deep Neoaves phylogeny: data, signal enhancement, and priors. Mol. Biol. Evol. 26: 313–326.

Sanderson MJ, Boss D, Chen D, Cranston KA, Wehe A. 2008 The PhyLoTA browser: processing GenBank for molecular phylogenetics research. Systematic Biology, 57:335-346.

Sibley CG, Ahlquist J E. 1990. Phylogeny and classification of birds. Yale University Press, New Haven, CT. USA.

Wink M, Seibold I, Lotfikhah F, Bednarek W. 1998. Molecular systematics of holarctic raptors (Order Falconiformes). Holarctic Birds of Prey. Adenex, WWGBP, Calamonte. pp. 29–48.