LUND UNIVERSITY

The Molecular Evolution of Snakes as revealed by Mitogenomic Data

Douglas, Desiree

2008

Link to publication

Citation for published version (APA):
Douglas, D. (2008). *The Molecular Evolution of Snakes as revealed by Mitogenomic Data*. [Doctoral Thesis (compilation), Department of Biology].

Total number of authors:
1

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

The Molecular Evolution of Snakes as Revealed by Mitogenomic Data

DESIRÉE DOUGLAS

Department of Cell and Organism Biology
Division of Evolutionary Molecular Systematics
Lund University
2008

LUND UNIVERSITY

A doctoral thesis at a university in Sweden is produced as a monograph or as a collection of papers. In the latter case, the introductory part constitutes the formal thesis, which summarizes the accompanying papers. These have already been published or are manuscripts at various stages (in press, submitted or in preparation).

Cover: Rainbow Boa *Epicrates cenchria.* Photo by RainForest Adventures.

Reproduced with permission.

© Desirée Douglas 2008

The Molecular Evolution of Snakes as Revealed by Mitogenomic Data

ISBN: 978-91-85067-40-4

Printed in Sweden by Media-Tryck, Lund

THESIS SUMMARY

The snakes (Serpentes) are a diverse group of squamate reptiles that together with lizards (Lacertilia) and worm lizards (Amphisbaenia) belong to the reptilian order Squamata. In recent years there have been new and exciting fossil finds of snakes that show a mixture of primitive and advanced characters. This has led to a plethora of morphological studies addressing the placing of these fossils in a phylogenetic context and controversial hypotheses as to the origin of snakes. Despite this, no consensus has been reached either on the affinities of snakes or on the interrelationships of some snake families.

Inferring snake relationships has been very difficult because morphologically they are very derived with many unique characters and relatively few shared characters linking them with other groups of squamates (Estes et al., 1988). In addition, due to the presence of both primitive and derived characters, the phylogenetic position of the snake fossils relative to extant snakes has been contentious. Prior to the outset of this thesis there was very little molecular data on squamates and the few phylogenetic studies that had been published were limited to one or two mitochondrial (mt) or nuclear genes.

For this thesis, mt genomes of snakes and lizards were sequenced in order to yield a substantial body of data for phylogenetic analyses compared to the amount of data that had been used in previous studies. The focus was on basal snake families, as the phylogenetic relationships of these families were unresolved and their affinities would have implications for resolving snake origins. Only three squamate mt genomes had been described prior to the initiation of my PhD work: two lizards and one snake. Over the course of the project all coding regions of the mt genomes of ten squamates - eight snakes and two lizards - were sequenced. Besides phylogenetic analyses, the study also includes the examination of unusual features of snake genomes such as gene rearrangements and compositional biases in different lineages.

Two papers presented in the thesis, Papers I and III, deal with the affinities of snakes to other squamate reptiles and the root of the squamate

tree. While this work was being carried out interest in squamate mitogenomics had grown and additional squamate mt genomes became available. My results have supported a close relationship between snakes, amphisbaenians and lacertiform lizards. This agrees with a recent mitogenomic study but is in disagreement with nuclear gene analyses.

Paper II examines the properties of mt genomes and their composition. The results suggest that, from observing base composition, more than one replication mechanism may be present in snake mt genomes. In addition, the base composition of mt genes was found to be extremely divergent within snakes and there appeared to be a compositional bias towards adenine in one group, the Alethinophidia. Both of these peculiarities may have influenced the fast evolutionary rate of snake mt genes relative to other squamates.

A study was also carried out on the phylogenetic relationships of basal alethinophidian families, whose relationships hitherto have been unresolved. The results agree with previous molecular studies based on a limited number of genes in that they did not support the monophyly of the traditional higher taxonomic groups, Anilioidea and Booidea. Instead the results present two new hypotheses for the relationships of basal alethinophidians.

CONTENTS

INCLUDED PAPERS **6**

1. INTRODUCTION **7**

1.2 WHAT IS A SNAKE? 7
1.2 THE CLASSIFICATION OF SNAKES 8

2. THE ORIGIN OF SNAKES **11**

2.1 THE AFFINITIES OF SNAKES 12
2.1.1 The varanoid/mosasauroid hypothesis *13*
2.1.2 Criticism of the varanoid/mosasauroid hypothesis *14*
2.1.3 Snakes are sister-group to all other squamates *17*
2.1.4 The burrowing scincomorph/dibamid hypothesis *18*
2.1.5 The amphisbaenian hypothesis *19*
2.2 MOLECULAR STUDIES 20

3. RELATIONSHIPS OF BASAL ALETHINOPHIDIANS **23**

4. MY PHD PROJECT **25**

4.1 AIMS 25
4.2 MITOCHONDRIAL GENOMES 25
4.3 SEQUENCING MITOCHONDRIAL GENOMES 26
4.4 PHYLOGENETIC RECONSTRUCTION 27
4.5 USING MITOCHONDRIAL DATA FOR PHYLOGENETIC RECONSTRUCTION 28
PAPER I: THE PHYLOGENETIC POSITION OF SNAKES 31
Features of snake mitochondrial genomes *31*
The affinities of snakes *33*
PAPER II: COMPOSITION OF SNAKE MITOCHONDRIAL GENOMES 36
Strand bias in snake mt genomes *36*
Comparisons of base composition between snakes and other amniotes *38*
PAPER III: THE PHYLOGENY OF SQUAMATES 42
A STUDY ON THE RELATIONSHIPS OF BASAL ALETHINOPHIDIANS 47

CONCLUSIONS **50**

COPYRIGHT NOTICE **52**

REFERENCES 53
ACKNOWLEDGEMENTS 63
Appendix: Papers I-III 65

Included Papers

PAPER I

A mitogenomic study on the phylogenetic position of snakes

Douglas, D.A., Janke, A., Arnason, U. 2006 *Zoologica Scripta Vol. 35, Iss. 6 p545-558*

PAPER II

Base and amino acid composition in snake mitochondrial genomes: strand bias and compositional distinctions among different lineages

Douglas, D. (Submitted)

PAPER III

The mitochondrial genomes of *Dibamus novaeguineae* (Squamata: Dibamidae) and *Uromastyx aegyptia* (Squamata: Agamidae) and the phylogenetic tree of squamates

Douglas, D. (Submitted)

1. INTRODUCTION

1.2 What is a snake?

Everybody recognizes snakes for their elongate, limbless bodies and forked tongues. However, to define what a snake actually is and to pinpoint characters that distinguish snakes from their nearest relatives, lizards and amphisbaenians, is not a simple task. For example, many lizards and most amphisbaenians are also elongate and limbless. In fact, limb reduction, or total loss of limbs, has occurred many times among lizards including anguimorphs, skinks, dibamids and gekkotans. There are also lizards with forked tongues (Rieppel, 1988). Snakes are also known to have immovable eyelids and lack external ears (Greene, 1997). However, some geckos also have immovable eyelids (e.g. Underwood, 1970) and there are also lizards that lack external ears, such as the earless monitor lizard *Lanthanotus borneensis* and many skinks (Greer, 2002). What characters, then distinguish the snakes from lizards? There are in fact a number of morphological characters that separate these two groups, but they are mostly internal (e.g. Underwood, 1967; Estes et al., 1988). However, one conspicuous feature that is very exaggerated in advanced snakes compared to lizards and amphisbaenians is the degree of skull kinesis, that is, the flexibility of the skull.

Separating snakes from lizards scientifically has actually been a difficult process. The earliest classification schemes of snakes included caecilians (limbless amphibians), limbless lizards and amphisbaenians, as well as true snakes (cf. Rieppel, 1988). Subsequent classifications removed these other groups one by one, with the first classification scheme in which snakes were classified in the modern sense being produced by Wagler in 1830 (cf. Rieppel, 1988). There are approximately 3000 snake species currently known to science (Uetz et al., 2006) making them one of the most successful groups of reptiles. Snakes are also found on most continents and have colonized a variety of habitats, including desert, subterranean and open-ocean.

1.2 The classification of snakes

All extant snakes belong to the suborder Serpentes. However, snakes are sometimes referred to as belonging to the Ophidia, which also includes extinct forms not classified within the Serpentes. Below is a description of the main taxonomic divisions within the Serpentes that were taken from Rieppel (1988). The Serpentes is divided into two infraorders: the Scolecophidia (worm or blind snakes) and the Alethinophidia (true snakes - see Fig. 1). The Scolecophidia comprises three families of small to minute snakes specially adapted to burrowing that have reduced eyes. The Alethinophidia contains all other snakes and is further subdivided into Anilioidea and Macrostomata.

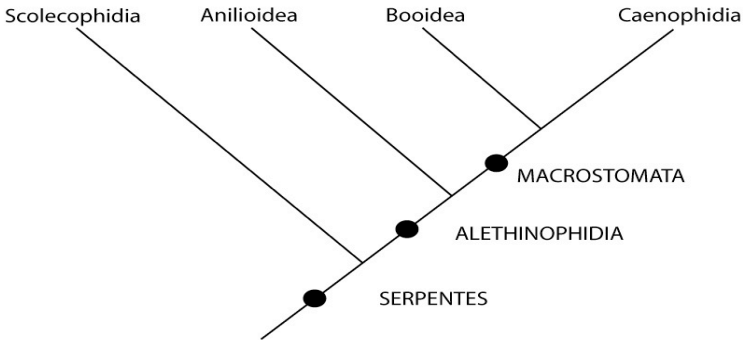


Fig. 1. A cladogram showing the relationships between the major lineages of snakes taken from Rieppel (1988). The photos above the cladogram show examples of a scolecophidian, the Texas Blind Snake *Leptotyphlops dulcis,* on the left and an alethinophidian, Rainbow Boa *Epicrates cenchria,* on the right. Photos taken by Gary Nafis and RainForest Adventures. Reproduced with permission.

The Anilioidea are comprised mostly of semi-fossorial (i.e. semi-burrowing) snakes. The Macrostomata, meaning "big mouth", is the largest group of snakes. Skull kinesis and flexibility of the jaws reaches their full extent in these snakes, which are able to increase the area in the mouth due to the lack of a hinge connecting the lower jaws (Fig. 2). This enables macrostomatan snakes to engulf prey with a diameter larger than the width of their heads. The Macrostomata is split into two groups, the Booidea and the Caenophidia. The Booidea contains the giants such as boas and pythons as well as a number of lesser-known families. The Caenophidia contains the large family Colubridae (e.g. milksnakes, kingsnakes and some venomous species) as well as cobras and vipers. Basal alethinophidians (that is, Anilioidea and Booidea) are commonly referred to as "Henophidia". Many of the more basal families of snakes - scolecophidians, several anilioids and booids - possess pelvic and hindlimb vestiges (Rage and Escuillié, 2003). However, no extant snakes are known to possess any vestiges of forelimbs.

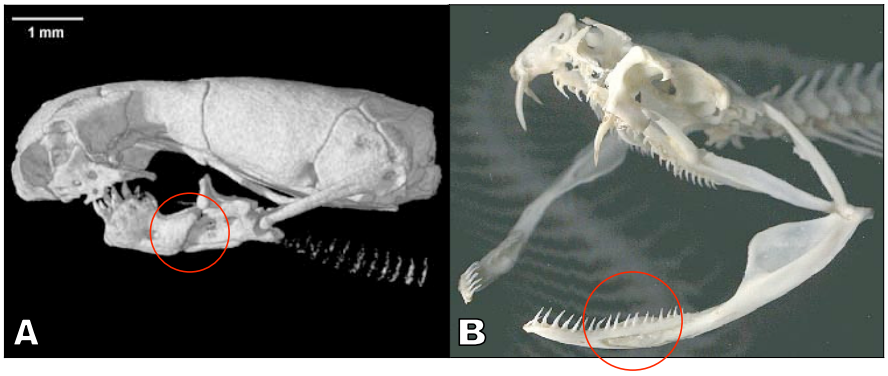


Figure 2. The differences in skull structure between non-macrostomatan (A) and macrostomatan (B) snakes. Note that in B the front of the mandibles (lower jaw) are not held together by a hinge. The intra-mandibular joints are circled. A: Texas blind snake *Leptotyphlops dulcis*, reproduced with permission from Digimorph.org. Source of specimen: Texas Memorial Museum. B: Cottonmouth (viper) *Agkistrodon piscivorus,* reproduced with permission from East Coast Natureworld Tasmania.

The higher taxonomic groups shown in Fig. 1 are those most frequently used in the literature. However, the number of families within certain groups

and their validity (especially those of basal alethinophidians) are uncertain due to the plasticity of the interfamilial relationships. This will be discussed in more detail in the section entitled "Relationships of basal alethinophidians". However, the biggest controversy regarding snake phylogenetics is the origin of snakes.

2. THE ORIGIN OF SNAKES

The snakes belong to a speciose group of reptiles known as the Squamata, or scaly reptiles. Besides squamation, one important feature of squamates is streptostyly (Vitt et al., 2003). Streptostyly is a condition whereby there is a special joint between the quadrate and squamosal bones in the skull, resulting in increased mobility of the quadrate in squamates compared to their sister-group the Sphenodontida (tuataras). The Squamata are made up of seven lineages: Iguania (iguanas, chameleons and kin), Gekkota (geckos and pygopodids), Scincomorpha (skinks and kin), Anguimorpha (monitor lizards, gila monster and kin), dibamids (blind skinks), Amphisbaenia (worm lizards) and Serpentes. The first five lineages are traditionally known as lizards.

A cladogram showing the traditional relationships of the Squamata is shown in Fig. 3. Phylogenetic analyses based on morphological data have divided the Squamata into Iguania and all remaining squamates into a group named Scleroglossa, so-called because the tongue is at least partly keratinized and is flattened compared to those of iguanians (Estes et al., 1988). Ecological shifts between different squamate groups appear to support this split. The Iguania are mainly ambush predators, using visual prey discrimination and capture their prey by lingual prehension, whereas the Scleroglossa are mainly active foragers that rely more on chemical cues to seek out prey (Cooper, 1995; Vitt et al., 2003; Vitt and Pianka, 2005). Scleroglossans also capture prey with their jaws instead of their tongue as iguanians do. The skulls of scleroglossans are thus less rigid (Vitt et al., 2003). However, some skull and visceral characters used for phylogeny may be associated with life history traits so concordance between these two attributes is not surprising. The Serpentes, Amphisbaenia and Dibamidae are recognized as scleroglossans, but their relationships are uncertain.

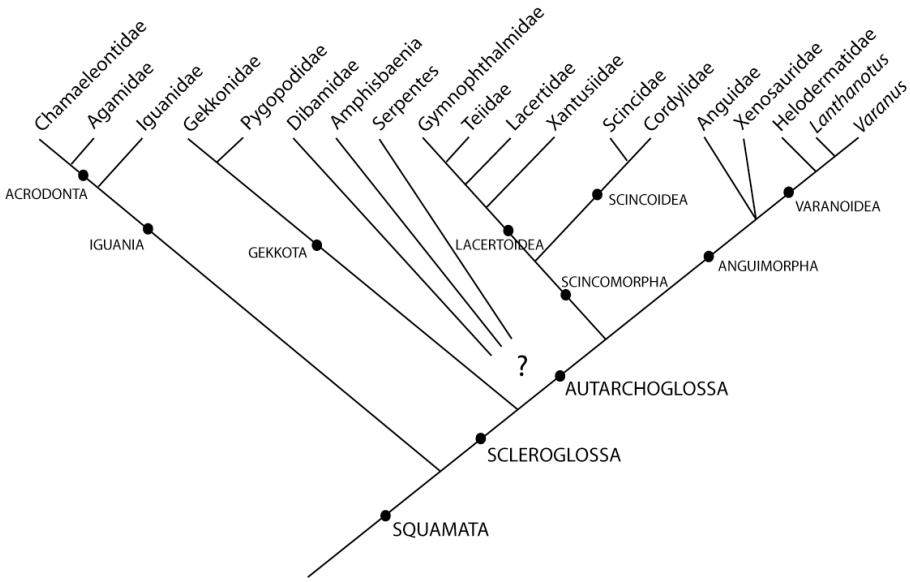


Fig. 3. Phylogenetic relationships of the Squamata based on morphological data, redrawn from Estes et al., 1988.

Of all debates regarding squamate phylogeny, the position of snakes has been the most contentious. The origin of snakes has been debated for 140 years. Snakes may have arisen from another squamate group or arisen independently from a "prolacertilian" (Fejérvary, 1918). Because many morphological characters of snakes and other squamates are associated with their life history, determining the habitat of ancestral snakes has become as much a key element to the debate as inferring the sister-group of snakes.

2.1 The affinities of snakes

There have been four main hypotheses as to the affinities of snakes (Rieppel, 1988):

1. The varanoid/mosasauroid hypothesis (Cope, 1869; Nopcsa, 1923)
2. Snakes arose independently of other squamates (Underwood, 1970)
3. The burrowing scincomorph/dibamid hypothesis (Senn and Northcutt, 1973)
4. The amphisbaenian hypothesis (Rage, 1982)

2.1.1 The varanoid/mosasauroid hypothesis

The oldest and most debated of the hypotheses has been that snakes are most closely related to varanoid lizards. Varanoids are a group of anguimorph lizards that include today's monitor lizards (Varanidae) and an extinct group of voracious marine predators known as mosasaurs (Fig. 4) that lived during the Cretaceous period (146 - 65 MYA). Similarities between snakes and mosasauroids, including the shape of the teeth, articulation and degree of flexibility in the lower jaw, position of certain skull bones and vertebrae, were first pointed out by Cope (1869, 1878). The strongest character linking snakes and mosasauroids is the intra-mandibular joint (see Fig. 2). Camp (1923) also allied snakes with varanoid lizards, but postulated that they evolved from grass-living ancestors rather than marine ones. However, the marine hypothesis has become the favoured scenario.

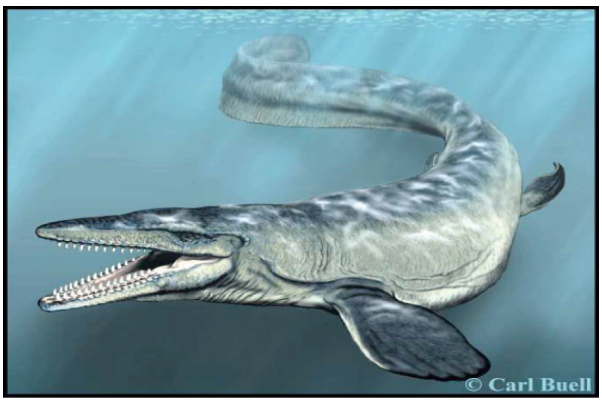


Fig. 4. Illustration of a mosasaur by Carl Buell. Reproduced with permission.

The marine hypothesis was later developed by Nopcsa (1923; 1925) who allied snakes with a group of more basal mosasauroids, the dolichosaurs, which lived in the mid-Cretaceous and more approached snakes in appearance than the more derived mosasauroids described by Cope (see Fig. 4). Nopcsa (1923) described *Pachyophis*, an incomplete fossil from the mid-Cretaceous of what was probably an aquatic snake (Rage and Escuillié, 2003). Nopcsa believed that *Pachyophis* is a missing link between snakes and

dolichosaurs. In a later paper (1925) he gives lines of evidence of a marine ancestry of snakes. These include strong median neck muscles, a slender skull and the strengthening of the vertebral column due to large processes known as zygosphenes. Other morphological works have also placed snakes with varanoids (McDowell and Bogert, 1954; McDowell, 1972; Schwenk, 1988).

In 1997 a fossil found in marine sediments dating from the mid-Cretaceous (~95 MYA) that was previously thought to be a varanoid lizard was re-described as a snake (Caldwell and Lee, 1997). This fossil, named *Pachyrhachis problematicus,* possessed not only a skull that had derived features of extant snakes (such as a macrostomatan skull) but also tiny, fully formed hindlimbs placed well back towards the end of the tail. In addition, the mid-dorsal vertebrae and ribs are thickened (pachyostosis), suggesting an aquatic lifestyle (Caldwell and Lee, 1997).

Since then other limbed snakes found in marine sediments of about the same age have been described: *Haasiophis terrasanctus* (Tchernov et al., 2000) and *Eupodophis descouensi* (Rage and Escuillié, 2000). Some authors have proposed that the aquatic species, because their hindlimbs are developed to a degree not seen in extant snakes, are basal to crown-group Serpentes (Caldwell and Lee, 1997; Lee and Caldwell, 2000; Scanlon and Lee, 2000; Rage and Escuillié, 2000; Fig. 5). These authors suggest that macrostomy is a primitive ophidian character that has been lost in scolecophidians and other non-macrostomatans. In addition to this, Palci and Caldwell (2007) recently described a dolichosaur, *Adriosaurus microbachis,* which shows extreme limb reduction in the pectoral region, claiming that this shows the transition from limbed to limbless that occurred in snakes.

2.1.2 Criticism of the varanoid/mosasauroid hypothesis

The characters purported to join snakes with mosasauroids were called into question. Rieppel and Zaher (2000) presented a critical assessment of characters associated with the intra-mandibular joint in varanoids, mosasaurs and snakes, whereby they identified certain structural differences and stated the possibility that this character complex could be convergent in these taxa.

An intra-mandibular joint has also been reported in fossil birds, *Hesperornis* and *Ichthyornis* (Gregory, 1952), and the forked tongue characteristic of snakes and varanids is also seen in another lizard *Tupinambis* (Scincomorpha: Teiidae - Rieppel, 1988). The forked tongue and the ability of *Tupinambis* to swallow large prey are also apparent in visual observations. This suggests that these characters have the potential to be homoplasious as they are associated with the ability to swallow large prey and may therefore not be reliable for phylogenetic analysis. Rieppel and Kearney (2001) and Kearney and Rieppel (2006) have argued against the structure of the teeth being a character uniting snakes and varanoids (Scanlon and Lee, 2000).

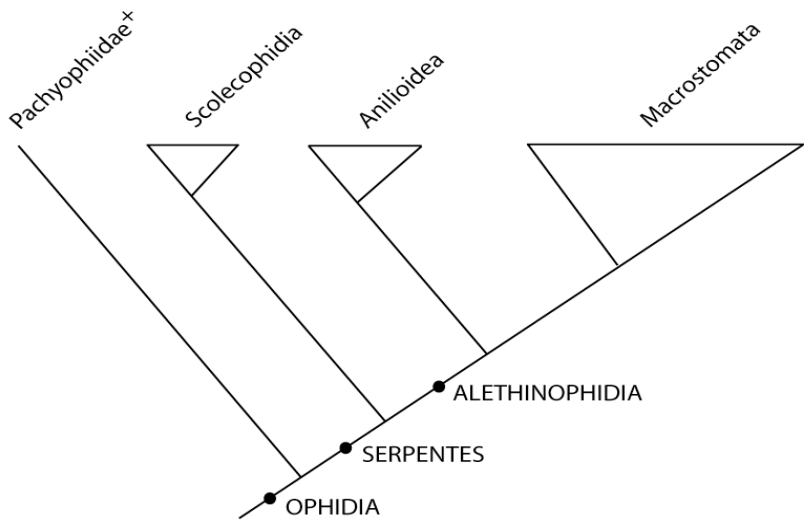


Fig. 5. A simplified cladogram of Ophidia showing the placement of fossil snakes – termed pachyophiids – as the sister-group to modern snakes (Serpentes). Redrawn from Scanlon and Lee (2000).

One character used to unite snakes and dolichosaurs has been the long, slender neck (Nopcsa, 1923). However, developmental studies have revealed that in snakes the neck is extremely short and expression of the trunk (thoracic) region expands anteriorly to just posterior to the head region (Cohn and Tickle, 1999). Studies on musculature in the cervical region also suggest

that the neck of snakes is very short (Tsuihiji et al., 2006). However Caldwell (2003) and Palci and Caldwell (2007) argued that anguimorph-like neck characters are found on many of the anterior vertebrae indicating that instead, the neck has expanded posteriorly.

Other authors posit that the fossil snakes are positioned within the Serpentes as opposed to being the sister-group of Serpentes (Zaher, 1998; Greene and Cundall, 2000; Rieppel and Zaher, 2000; Zaher and Rieppel, 2002; Fig. 6). Rage and Escuillié (2000) state that *Eupodophis* was found to have chevron bones on the caudal vertebrae, a primitive feature not found in other limbed snakes. However, this has been interpreted by other authors as a plesiomorphic trait retained in these snakes that would not affect a placing within the Serpentes (Zaher and Rieppel, 2002; Rieppel et al., 2003). In other words, this character has no bearing on the phylogenetic position of *Eupodophis.*

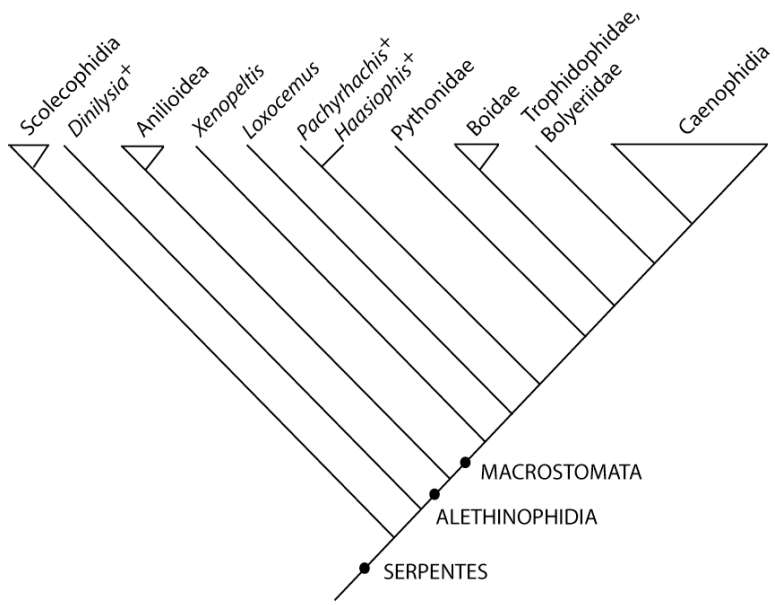


Fig. 6. A simplified cladogram showing the alternative placement of fossil snakes *Pachyrhachis* and *Haasiophis*. Redrawn from Tchernov et al. 2000. The cross + indicates extinct taxa.

Recently another limbed snake, *Najash rionegrina* from the Cenomanian-Turonian (100-89 MYA), has been described (Apesteguía and Zaher, 2006). Whereas *Pachyrhachis, Haasiophis* and *Eupodophis* were aquatic, *Najash* was terrestrial. This snake also showed adaptations to a subterranean way of life (Apesteguía and Zaher, 2006). In addition, it is unique among snakes, the authors say, in possessing a sacral region. The pelvis is also outside the ribcage whereas in all other snakes any pelvic remnants lie within the ribcage. This led the authors to suggest that this is the most primitive snake yet. In their phylogenetic analysis *Pachyrhachis, Haasiophis* and *Eupodophis* were placed as the sister-group to macrostomatans while *Najash* was positioned basal in relation to all other snakes.

2.1.3 Snakes are sister-group to all other squamates

Some authors suggest that the great number of unique characters in snakes suggest an origin independent of other squamates. As the hypothesis of a varanoid ancestry of snakes gained ground, Fejévary (1918) dismissed this, concluding that no fossil or extant lizards constitute snake ancestry. He instead proposed that snake ancestors were as yet unknown "prolacertilians". The varanoid/mosasauroid hypothesis was also rejected by other authors who believed snake ancestors were either burrowers like scolecophidians or burrowers with a more generalized body form seen in anilioid snakes (Mahendra 1938; Bellairs and Underwood, 1951). Studies by Walls (1940; 1942) on the eyes of squamates noted the many peculiarities within the eyes of snakes and the stark difference in accommodation technique compared to lizards. Walls (1942) proposed that snakes went through a burrowing phase early on in their evolution such that the eye degenerated then re-evolved new structures when above-ground living snakes evolved. This is consistent with scolecophidians being basal. Underwood (1957) describes various skeletal and soft characters that vary between lizards and snakes, which suggested an independent origin for snakes. In a later study on reptilian eyes Underwood (1970) also leaned towards this view but adapted Walls' hypothesis slightly, suggesting that the snake ancestor was both nocturnal and fossorial. It was

also noted that diurnal lizards have a simplex retina (cones only) whereas most snakes have a duplex (cones and rods) retina.

The issue of squamate eyes was readdressed by Caprette et al. (2004), who concluded that the eyes of snakes actually suggest more similarities to those adapted to an aquatic lifestyle than to any other. However, it may not be correct to say that an eye can be adapted to an aquatic lifestyle anymore than one can say it is 'adapted' to burrowing – the eye does not seem particularly 'adapted' to these environments in those animals that inhabit them.

Kochva's (1978; 1987) work on squamate oral glands suggests an independent origin of snakes. Kochva noted that snakes have labial glands in both the lower and upper jaws whereas most lizards, including anguimorphs, have glands in the lower jaw only. Iguanians, thought at this time to be the most basal of lizard lineages (see Fig. 3), are the only lizards to have labial glands in both the upper and lower jaws. This was thus taken to be the primitive condition for squamates. Kochva also reported that amphisbaenians, like snakes and iguanians, also have labial glands in both the lower and upper jaws.

2.1.4 The burrowing scincomorph/dibamid hypothesis

The dibamids, also known as blind skinks, are a poorly known, highly fossorial, family of limbless lizards that have been allied to burrowing limbless scincomorph lizards in some morphological studies (e.g. Camp, 1923). A study by Senn and Northcutt (1973) noted similarities in brain structure between some burrowing scincomorphs, dibamids and snakes. Of all the lizards examined, it was that of *Dibamus* that most closely approached snakes. It was thus proposed that this could represent a synapomorphy between dibamids and snakes. After reviewing characters that support the four main theories regarding snake origins, Rieppel (1988) considered a link between snakes and burrowing scincomorphs to be the least parsimonious hypothesis and rejected it.

Most studies that have suggested a close relationship between dibamids

and snakes also support a close relationship between these two groups and another highly fossorial group, amphisbaenians. Greer (1985) reported that dibamids shared more characters with amphisbaenians and snakes than burrowing scincomorphs, anguimorphs or gekkotans. He points out however that this could be due to convergence of characters correlated with the burrowing ecomorph, especially as the snakes he used in his comparisons were mostly scolecophidians, anilioids and *Dinilysia,* a fossil snake that has been allied with basal alethinophidians in some morphological phylogenetic analyses (e.g. Tchernov et al., 2000; Rieppel and Zaher, 2000). Two recent phylogenetic analyses recovered snakes to be the sister-group of a clade containing *Dibamus* and amphisbaenians (Rieppel and Zaher, 2000; Kearney, 2003). Both expressed reservations about these results however and hinted that they could be influenced by characters correlated with burrowing. This was despite Kearney (2003) removing all such characters, which resulted in "amphisbaenians in an unresolved polytomy with a snake-dibamid clade."

2.1.5 The amphisbaenian hypothesis

Rage (1982) hypothesized that snakes and amphisbaenians were each other's closest relatives based on certain characters including the presence of a retractor pterygoidei muscle (a muscle found in the skull), similarities in trunk musculature, the structure of inner ear hair cells and the loss of palatine glands. Rieppel (1988) reviewed all of Rage's (1982) characters and found some to be inconclusive. For example, one character regarding trunk musculature – the presence of a levator costae muscle – is also found in *Dibamus*.

Hallermann (1998) recovered a clade joining amphisbaenians, dibamids and snakes. He also states that certain characters that support this clade are not associated with burrowing. Estes et al. (1988) recovered snakes and amphisbaenians to be nested inside the Anguimorpha when performing a phylogenetic analysis on extant squamates. They did not accept this grouping, stating that most of the characters supporting a close relationship between snakes and amphisbaenians were losses correlated with fossorial

adaptations in both groups. Lee (1998), after obtaining a snake-amphisbaenian clade in his analysis, down-weighted all characters claimed to be correlated with the burrowing ecomorph. This resulted in snakes grouping with varanoid lizards instead.

Of all the hypotheses regarding snake origins, the varanoid/mosasauroid hypothesis appears to be the most widely accepted. In contrast, the burrowing scincomorph and amphisbaenian hypotheses have been questioned due to the great potential in phylogenetic analyses for snakes to cluster with these groups on account of characters correlated with burrowing. In my view support for a marine origin of snakes grew with the discovery of the marine limbed snakes, and with the intuition that snakes with hindlimbs as opposed to mere vestiges must be the most primitive of all snakes. However, terrestrial snakes of the same age have also been found, which means that the Serpentes had already diversified by 100 MYA and that the lineage is much older. Therefore we must wait for the discovery of older snake fossils, which are equally likely to be marine, terrestrial or fossorial.

2.2 *Molecular studies*

Up until recently there were very few molecular studies aimed at resolving snake or squamate relationships. One of the first was by Forstner et al (1995) whose phylogenetic analysis, based on one mt protein-coding gene (ND4) and three tRNA genes, recovered a clade joining snakes with *Varanus*, which would support the varanoid/mosasauroid hypothesis (Fig. 7A). However, the tRNA sequence alignment from this study was criticized by Macey and Verma (1997). The results of Macey and Verma's re-analysis were less conclusive than that of Forstner et al. (1995), with the *Varanus*-snake grouping receiving less support. A later phylogenetic analysis based on two mt protein-coding genes (ND1 + 2) and eight tRNA genes also supported a sister grouping between *Varanus* and snakes (Rest et al. 2003). However, a more recent study using all mt genes supported an independent origin of snakes (Kumazawa, 2004 – Fig. 7B), although amphisbaenians were not included in this analysis. Reasons for inconsistency in results between the

latter study and the ones mentioned previously were thought to be due to insufficient taxon and site sampling (Kumazawa, 2004).

Studies based on nuclear genes have given entirely different results. A number of studies used the genes c-mos and RAG-1 to infer squamate phylogeny (Saint et al., 1998; Harris, 2003; Vidal and Hedges, 2004; Townsend et al., 2004). All have supported a novel hypothesis, namely that snakes are recovered in a clade with anguimorphs and iguanians (see Fig. 7C). This appeared to be confirmed by a very recent study based on nine nuclear genes (including c-mos and RAG-1) that recovered the same result (Vidal and Hedges, 2005). A new taxonomic name was proposed for this group – Toxicofera – referring to the presence of venom (Vidal and Hedges, 2005) found only in these groups. Fry et al. (2006) reported the presence of venom in varanid lizards. Venom was thought only to occur in snakes, helodermatid lizards (gila monster) and bearded dragons (Agamidae, Iguania). Because nuclear studies have shown affinities between these three groups, a hypothesis arose whereby venom only evolved once, instead of three times in squamates. However, more studies would be needed on more species of these three groups to test this hypothesis further.

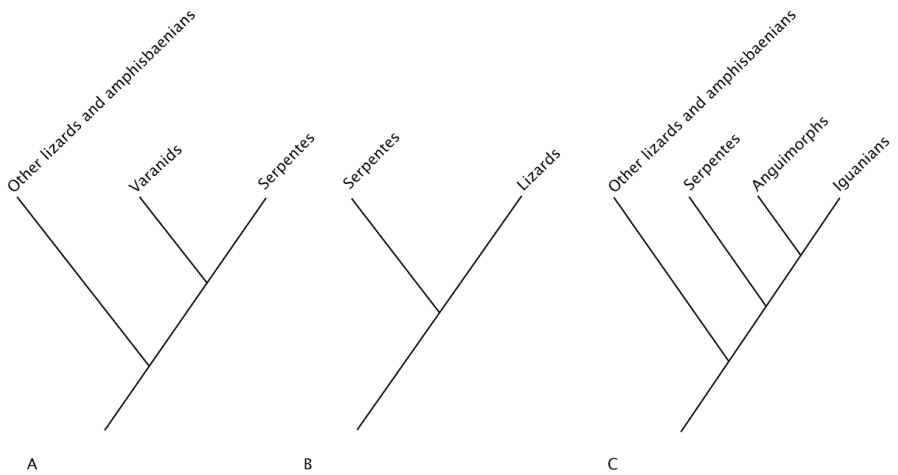


Fig. 7. Simplified cladograms of alternative hypotheses of snake affinities. A – Forstner et al. (1995). B – Kumazawa (2004). C – all nuclear gene studies.

A study on short interspersed nuclear elements (SINEs) found in squamate reptiles recovered the same result (Piskurek et al., 2006). However, the tree was rooted with lacertid lizards (instead of a proper outgroup taxon from outside the Squamata) with the only other squamates being iguanians, anguimorphs and snakes. So although it supports nuclear gene findings so far, the results are not conclusive. Furthermore, as SINEs are transposable elements, it is not known whether all SINEs included in this study were orthologous (i.e. from the same locus in the genome).

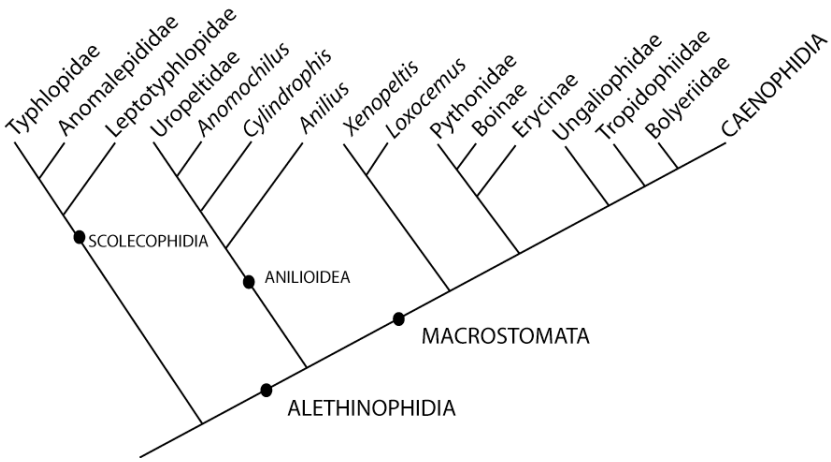
In general, one major difference between morphological and molecular studies has been that the latter have not supported the basal squamate split between Iguania and Scleroglossa. Iguania instead nests within Scleroglossa, which suggests that the morphological phylogenetic analyses may have reflected life histories of squamates rather than their phylogenetic relationships.

3. RELATIONSHIPS OF BASAL ALETHINOPHIDIANS

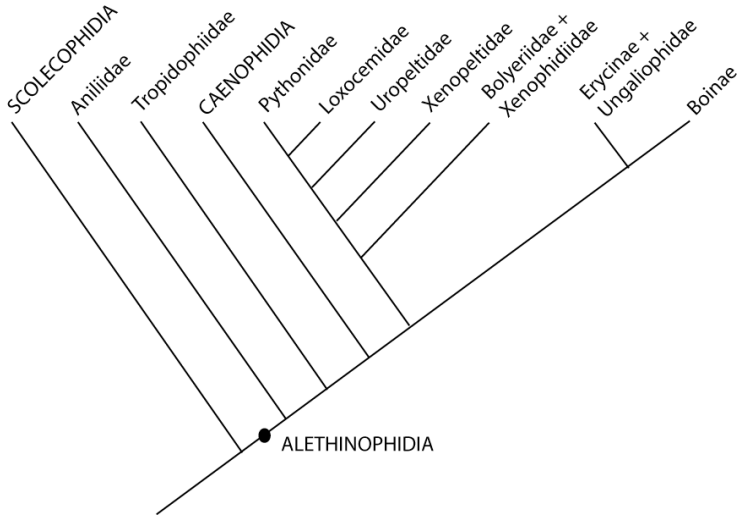
The traditional relationships of major snake lineages seen in Fig. 1 are based on increasing macrostomy from basal to derived lineages. As stated in the Introduction, relationships of basal alethinophidians are largely unresolved.

The monophyly of the Anilioidea and the Booidea are problematic. The Anilioidea typically contain the families Aniliidae, Anomochilidae, Cylindrophiidae and Uropeltidae (Scanlon and Lee, 2000; Lee and Scanlon, 2002). The Booidea (sensu Rieppel, 1988) contain the Boidae, Pythonidae (sometimes included within the Boidae), Bolyeridae, Loxocemidae, Tropidophiidae and Xenopeltidae. Some authors however refer to the Booidea simply as Boidae and Pythonidae (Lee and Scanlon, 2002).

The Anilioidea have been supported by some morphological studies (Tchernov et al., 2000; Scanlon and Lee, 2000; Scanlon, 2006) but not by others (Cundall et al., 1993). All morphological studies place anilioids at the base of the Alethinophidia (Fig. 8A). In contrast most molecular studies, based on one or two genes, place the Aniliidae basal to all other alethinophidians, making Anilioidea paraphyletic (Fig. 8B). The monophyly of Booidea has been at best tentative, supported by only two morphological characters (Rieppel, 1988). Morphological phylogenetic analyses have recovered this group as paraphyletic to Caenophidia, whereas in molecular studies uropeltids and cylindrophids, supposedly anilioids, tend to nest within the Booidea (e.g. Wilcox et al., 2002; Vidal and David, 2004; Lawson et al., 2004). Because of this last point, Macrostomata has been recovered paraphyletic in molecular analyses but monophyletic in morphological analyses.



A



B

Fig. 8. Alternative hypotheses on basal alethinophidian relationships. A – phylogenetic analysis based on morphological data redrawn from Scanlon (2006). B – molecular phylogenetic analysis redrawn from Lawson et al. (2004).

4. MY PHD PROJECT

4.1 Aims

The aims of my PhD project were to sequence mt genomes of snakes to:

- Obtain a reasonably large molecular dataset for phylogenetic analysis and to infer the affinities of snakes
- Examine features of snake mt genomes in order to find possible molecular markers and investigate certain properties such as nucleotide composition
- To infer relationships of basal alethinophidian snakes

4.2 Mitochondrial genomes

The mt genome of vertebrates is a circular, double-stranded molecule roughly 16-18 kilobases in length. It is compact compared to the nuclear genome in that intergenic regions are very short or non-existent and protein-coding genes do not contain introns. The mt genome typically contains a total of 37 genes: 13 protein-coding genes, 22 transfer RNAs (tRNA) genes and two ribosomal RNAs (rRNA) genes. The gene order was thought to be invariable within vertebrates until very recently when non-mammalian mt genomes were published containing several gene order rearrangements (e.g. Desjardins et al., 1990; Kumazawa and Nishida, 1995; Fonseca et al., 2006).

The protein coding genes are spread throughout the genome. They code for the various subunits of the cytochrome oxidase (CO), ATPase and NADH dehydrogenase (ND) enzymes. The two rRNA genes, 12S and 16S, represent the small subunit and large subunit respectively of the ribosome. Transfer RNA genes are approximately 60 to 75 base pairs in length and have highly conserved secondary structure. Each of the 20 amino acids has one tRNA gene attributed to them except leucine and serine, which each have two tRNA genes. In addition to the 37 genes, there is a large, non-coding region 1-1.5 kilobases in length known as the control region. This contains all the regulatory elements and promoter regions needed for replication and

transcription of genes. It also typically contains repeat sequences, which present a barrier for replication and transcription.

When embarking on my PhD project, non-avian reptiles were poorly represented by mitogenomic data. Only three complete squamate mitochondrial genomes were available: the Ryukyu snake *Dinodon semicarinatus*, a colubrid (Kumazawa et al., 1998) and two lizards, the green iguana *Iguana iguana* (Janke et al., 2001) and the mole skink *Plestiodon (Eumeces) egregius* (Kumazawa and Nishida, 1999). As I commenced my PhD, the complete mt genome of the tuatara (*Sphenodon punctatus*) was published (Rest et al., 2003). The tuatara genome was interesting as the largest mitochondrial protein-coding gene, ND5, was absent.

4.3 Sequencing mitochondrial genomes

Total DNA was extracted from muscle tissue using the organic extraction method, which involves the use of phenol and chloroform to separate DNA from proteins (Sambrook and Russell, 2001). Conserved snake primers were designed from aligned tRNA sequences. Secondary structure of the sequence was determined to avoid designing primers that would potentially self-anneal. PCR primers were also designed from conserved regions in protein coding genes such as COI or Cyt*b*. Other regions were amplified using published primers (Kumazawa and Endo, 2004). Once PCR products were obtained, and their ends sequenced, specific primers were designed to amplify interspersed regions of the genome. This also ensured that all fragments overlapped each other and that they could be aligned and contiged.

A few of the primers that were designed could be used for all snakes. More often however new primers had to be designed from other tRNAs or gene regions because snake mtDNA turned out to be more divergent than, for example, mammalian mtDNA (this is discussed in Paper II). It therefore took considerable effort to amplify and sequence snake genomes from both Scolecophidia and Alethinophidia. In total, eight snake mt genomes were sequenced: two scolecophidians, the Southern blind snake *Ramphotyphlops*

australis and Jan's blind snake *Typhlops mirus* (Papers I and II), and six alethinophidians: the false coral snake *Anilius scytale,* Peters' Philippine earth snake *Rhinophis philippinus,* the rosy boa *Charina trivirgata*, Columbian red-tailed boa *Boa constrictor,* the yellow anaconda *Eunectes notaeus* and the corn snake *Pantherophis (Elaphe) guttatus* (Paper I and an unpublished study - page 47). In addition, two lizard genomes were sequenced: the blind skink *Dibamus novaeguineae* and the agamid *Uromastyx aegyptia* (Paper III).

4.4 Phylogenetic reconstruction

Phylogenetic reconstruction, or cladistics, involves the use of morphological or molecular data to infer relationships of a group of related organisms. The founder of the cladistic method was German entomologist Willi Hennig. Phylogenetic reconstruction is aimed at finding monophyletic groups, that is, a group of organisms that all share common ancestry. Monophyletic groups are usually nested within one another. To take familiar examples, the monophyletic group Vertebrata is nested within Chordata, Amniota is nested within Vertebrata, and so on. Monophyletic groups are formed based on shared derived characters, or synapomorphies, which all taxa within the group possess. For morphological data, characters are chosen on this premise. For example, the limbs of all tetrapods are homologous. For molecular data, the chosen gene(s) must be homologous between species. In other words, they must have the same evolutionary history. The result of a phylogenetic analysis, a cladogram or a phylogenetic tree, can be produced using several methods: maximum parsimony, neighbor joining, maximum likelihood and Bayesian inference.

Most analyses that have aimed to reconstruct higher-level snake or squamate relationships have been based on morphological data. A caveat of morphological analyses, particularly evident in the case of snakes, is that the choice of characters and character weighting can be very subjective. For example, in his phylogenetic study of the affinities of snakes, Lee (1998) down-weighted characters he claimed were related to burrowing. Although he may have been correct in thinking that these characters are convergent,

there is no justification in assuming so *a priori* as the snake ancestor may have been a burrower. One advantage that morphological analyses have over molecular analyses is that in the former fossil data can be included. However, as discussed in the previous section, fossil snakes possess both primitive and derived characters that make inference of their relationships difficult with respect to extant snakes.

4.5 Using mitochondrial data for phylogenetic reconstruction

Mitogenomic data has been used for the phylogenetic inference of deep divergences such as the affinities of turtles and basal gnathostome divergences (Zardoya and Meyer, 1998; Arnason et al., 2004). It has also been used in estimating dates of divergence of taxa (e.g. Kumazawa, 2007; Roos et al., 2007). There are many advantages to using mtDNA for phylogenetic reconstruction. The mitochondria do not undergo Mendelian inheritance as does the nucleus. Instead only the mother's mitochondria, and hence her mtDNA, are inherited by the next generation as the father's mtDNA is either destroyed (Sutovsky et al., 1999) or undetected due to the vast surplus of maternal mtDNA. Thus, only one haplotype of mtDNA is present in every individual making it easier to determine the sequence of the molecule. Because of non-Mendelian inheritance, recombination is a rare, but not unknown, event (Ujarvi et al., 2007). However, recombination would only be a problem if one is working with very closely related species/individuals that are more likely to hybridize than individuals of different families. Most important however is the strict orthology of mt genes, which is brought about because there is only one copy of each gene (i.e. no duplicated genes or gene families). Furthermore the protein-coding genes are intronless, making sequencing more efficient. In addition, mtDNAs are very abundant in the cell. For example, >8,000 copies were estimated to exist in human HeLa cells (Bogenhagen and Clayton, 1974), so one can obtain a sufficient amount of DNA from small tissue samples. It is relatively easy to design conserved primers to be used for PCR because the gene order is for the most part conserved.

The disadvantage of using mtDNA for phylogenetic reconstruction is that it represents only one locus. All H strand protein-coding genes and rRNAs - the genes used in this work for phylogenetic analysis - are transcribed as a single polycistron (Fernández-Silva et al., 2003). In contrast, different nuclear genes represent sequences from different (independent) loci. In addition, mtDNA evolves much faster than nuclear DNA, which can be problematic for inferring relationships of taxa that had separated a long time ago as the sequences may become saturated with multiple substitutions. However, in the course of this study, and judging on evidence of previous single-gene studies, it became apparent that the lineages of squamates diversified rapidly, such that internal branches of the phylogenetic tree are short relative to external branches. Therefore the use of more slowly evolving genes will not necessarily be better at resolving higher-level relationships where short internal branches are present.

Prior to my studies, molecular datasets of squamates only contained the sequences from one or two protein-coding genes, limiting the amount of data to 2000 to 3000 nucleotides. The mt datasets used for Papers I and III were much greater, containing at least twelve genes (amounting to 9000 nucleotides or 3000 amino acids). The gene ND6 was excluded from the analysis as it was the only protein-coding gene encoded on the L strand and had a strikingly different base composition to all other genes and this may have been detrimental in model specification. Thus 12 protein-coding genes were used for phylogenetic analysis in Paper I. Whilst the first five genomes were being sequenced a mitogenomic study on the affinities of snakes was published that included all mt genes except ND6 (Kumazawa, 2004). Ribosomal RNA genes were not used in Paper I because these genes contain extensive secondary structure and there were no models available at that time that took non-independence of sites into account. In addition to secondary structure, tRNAs were very short and contained unalignable loop regions. Even if one could account for secondary structure, the amount of additional data obtained would be minimal. In Paper III both rRNA genes and protein-coding genes were used as models accounting for secondary structure became available

and were implemented in programs that allowed the combining of different types of molecular data. The ND6 gene was included in initial partitioned analyses for Paper III, but as this gene only contains at most 540 nucleotides (180 amino acids) and was too small for model parameters to be optimized efficiently.

PAPER I: THE PHYLOGENETIC POSITION OF SNAKES

The aims of this paper were firstly to test current hypotheses on the affinities of snakes and secondly to examine mitogenomic features of scolecophidians, basal alethinophidian and derived alethinophidian snakes. Snake mitogenomic features were also compared to those of lizard and amphisbaenian genomes in the search for molecular markers. Five mt genomes were sequenced in this study: two scolecophidians *Ramphotyphlops australis, Typhlops mirus,* two basal alethinophidians *Eunectes notaeus, Boa constrictor* and a derived alethinophidian (caenophidian) *Pantherophis guttatus*.

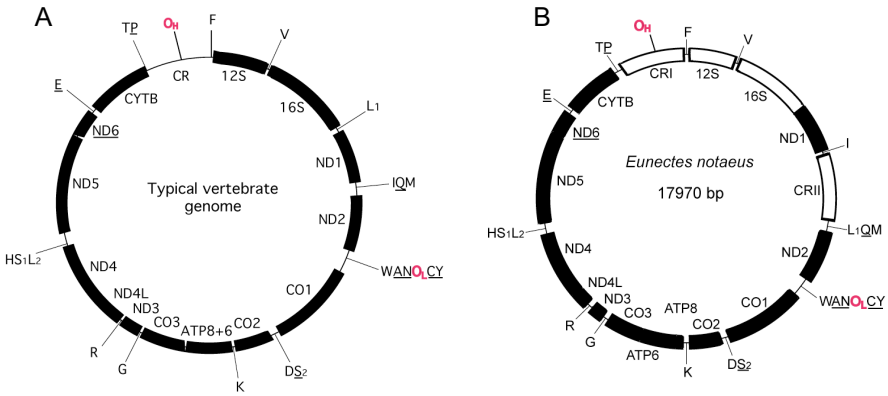


Fig. 9. Structure and gene organization of mt genomes. Letters outside the circle of each genome denote the one-letter code of the amino acid for each tRNA gene. The names of the protein-coding genes are found within the circle of each genome. Underlined genes denote genes encoded on the L strand. O_H and O_L: origin of H and L strand replication, respectively. A - typical gene organization seen in the majority of vertebrate genomes and scolecophidian snakes. B - gene organization of alethinophidian snakes represented by *Eunectes*.

Features of snake mitochondrial genomes

It was found that the alethinophidian snakes had a number of gene rearrangements, mostly involving tRNA genes. Fig. 9A shows a mitochondrial genome with the typical vertebrate gene order. Fig. 9B shows the mitochondrial genome of *Eunectes*, which was sequenced in its entirety. The mt genome of *Eunectes* is typical of alethinophidian genomes sequenced

to date in that it has a second control region within the IQM region. In addition, the tRNA-Leu UUR gene (L1 in Fig. 9) was translocated from its typical position between 16S rRNA and ND1 genes to downstream of the second control region sequence. This was in accordance with what had been found in a previous study where the IQM region was sequenced in a few snakes (Kumazawa et al., 1996), along with the regions flanking the first control region. In addition to these features, *Pantherophis* has a partial tRNA-Pro sequence downstream of the tRNA-Ile gene plus the functional gene upstream of the first control region. This is also observed in another colubrid snake, *Dinodon semicarinatus* (Kumazawa et al., 1998). These gene rearrangements are thought to be a result of an ancient duplication event where the region spanning the gene tRNA-Pro through to the tRNA-Leu UUR gene was duplicated with most of the duplicate genes being subsequently lost (Kumazawa et al., 1998).

Scolecophidian mt genomes sequenced in this study showed none of the rearrangements described for alethinophidians. However, the origin of L strand replication, a stem-loop structure typically situated between tRNA-Asn and tRNA-Cys genes, is absent in scolecophidians. In the only other scolecophidian genome sequenced, *Leptotyphlops dulcis,* the tRNA-Gln gene was translocated to the WANCY cluster of tRNAs (Kumazawa, 2004). This was not seen in the typhlopid mt genomes sequenced in this study, however.

The duplicate control regions of alethinophidian snakes are 95-100% identical. Furthermore, up to 600 bp of the control region 5′ ends was extremely conserved across alethinophidian snakes. This conservation was not observed in the control regions of lizard or amphisbaenian mt genomes. The mechanism responsible for this unexpected conservation and maintenance of both control regions in alethinophidian snakes is unknown.

As these events are expected to be rare, they can serve as phylogenetic markers (Kumazawa and Nishida, 1995). Comparisons between snake and other squamate mt genomes showed rearrangements to be unique to each lineage and so were inconclusive with respect to showing possible similarities between snakes and other squamates. Mitogenomic markers may only be

useful in determining genomic evolution at lower taxonomic ranks.

The affinities of snakes

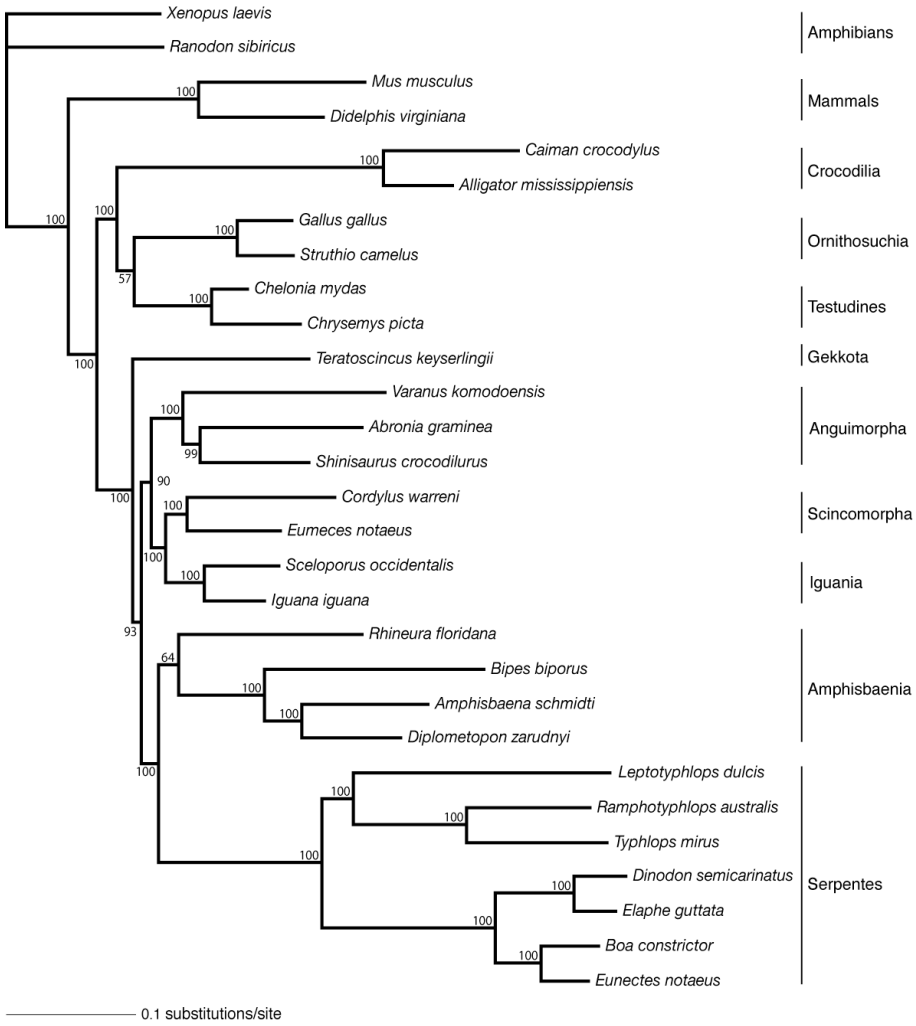


Fig. 10. The Bayesian tree (nucleotide data) based on 12 mt genes. This tree does not include *Sphenodon* as this taxon joined erroneously with snakes.

Seven snakes were included in the phylogenetic analysis: the five sequenced in this study in addition to *Dinodon* and *Leptotyphlops* (Kumazawa,

2004). Representatives of other major squamate lineages – Amphisbaenia, Iguania, Anguimorpha, Gekkota and Scincomorpha – were also included. It was therefore possible to test the hypotheses of snake affinities obtained in previous molecular studies (see Fig. 7). Outgroup taxa included *Sphenodon punctatus,* two birds and two turtles. Both nucleotide and amino acid data were analyzed, nucleotide data being analyzed using the GTR+I+Γ8 model of evolution (Lanave et al., 1984; Gu et al., 1995) and amino acid data analyzed with mtREV+I+Γ8 (Adachi and Hasegawa, 1996).

Contrary to previous studies, the phylogenetic analysis showed snakes to join with amphisbaenians (Fig. 10) instead of anguimorph or iguanian lizards, and lizards are shown to be paraphyletic. Initial results placed snakes in a clade with the iguanian lizard *Pogona vitticeps*, but this was shown to be a result of long-branch attraction (LBA). LBA is a phylogenetic reconstruction artifact that occurs if there are long branches in the tree that join together by chance (Page and Holmes, 1998). Both snakes and *Pogona* were at the end of long branches, and the latter evolved much faster than any other taxon in the dataset including the outgroups. Furthermore, *Pogona* was unstable in the tree when snakes were removed, joining with other fast evolving taxa such as crocodilians and never joined with other iguanians. This taxon was thus removed from subsequent analyses.

To test whether the snake-amphisbaenian relationship was due to LBA, separate analyses were run using different outgroups. *Sphenodon,* the closest outgroup to squamates, showed a tendency to group with snakes being as it is an isolated long branch, so analyses were also run with this taxon excluded. This did not affect the tree topology. Because snakes and amphisbaenians were shown to have faster rates of evolution than lizards, analyses were run with these taxa removed in succession to see if long branches of these taxa affected the topology. When snakes were removed, there was no change to the tree topology. When amphisbaenians were removed however, *Sphenodon* again grouped with snakes. Removal of *Sphenodon* resulted in snakes being the sister-group of lizards. It is not clear then whether this revealed that the snake-amphisbaenian clade was a result of LBA or whether this last result is

due to taxon sampling differences and attraction of the snake clade to the base of the ingroup. This would explain the inconsistency between results in this study and previous studies based on all mt genes (Kumazawa, 2004; Dong and Kumazawa, 2005).

Finally, we tested the snake-amphisbaenian clade against alternative hypotheses (Figs 7) using two paired sites tests, Kishino-Hasegawa and Shimodaira-Hasegawa tests (Kishino and Hasegawa, 1989; Shimodaira and Hasegawa, 1999), implemented in the program TREE-PUZZLE (Schmidt et al., 2002). Statistical tests showed the snake-amphisbaenian grouping to be preferred over alternative groupings. However, they did not reject the possibility of snakes being sister-group to all other squamates.

The results from Paper I thus preferred a sister-group relationship between snakes and amphisbaenians in contrast to nuclear gene studies that proposed a closer relationship between snakes, anguimorphs and/or iguanians. However, the present study did not reject the hypothesis proposed in other mitogenomic studies, namely that snakes are sister-group to lizards and amphisbaenians. While this paper was in press another mitogenomic study also placed snakes as the sister-group to other squamates, with amphisbaenians sister-group to gekkotans (Zhou et al., 2006).

Because of the extraordinarily fast evolutionary rates of snakes compared to other squamate lineages and the potential for LBA, these results were treated with caution. Even though most of the major squamate lineages (all except Dibamidae) were represented in this study, it was evident that greater taxon sampling would be needed from under-represented lineages in order to break up long branches. This issue is dealt with and discussed in Paper III.

PAPER II: COMPOSITION OF SNAKE MITOCHONDRIAL GENOMES

In light of the extraordinary rates of evolution observed in snake lineages, a comparative analysis of the base composition of snake mt genes was carried out. Base composition of snake mt genes was compared primarily to other squamates but also to other amniotes to get a better picture of mt gene evolution in snake lineages. Another aim of this paper was to examine mt genome strand bias in snakes in light of the duplicate control regions in alethinophidians. The type of strand bias exhibited has implications for the mode of replication of the genome.

Strand bias in snake mt genomes

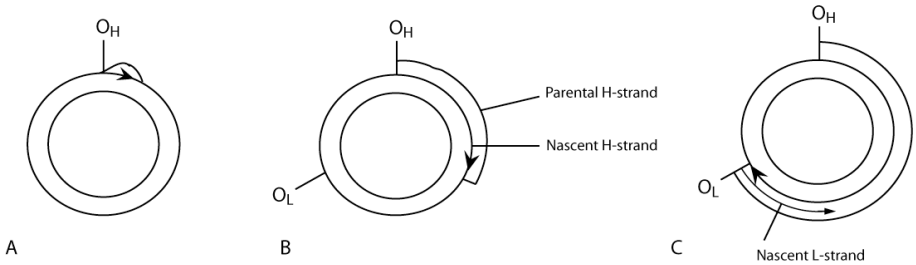


Fig. 11. Illustration of asynchronous strand replication (Clayton, 1982). Note that the mt genomes in this figure are drawn as mirror images of those in Fig. 9 to show replication proceeding clockwise. The three stages of replication shown here are explained in the text.

The most widely accepted mode of replication for mt genomes is the asynchronous replication, in which the two strands are replicated asynchronously. The mechanism behind this mode of replication has been reviewed in Clayton (1982) and Clayton (1991). The two strands of the mt genome are termed heavy (H) and light (L) strand due to their different buoyancies in CsCl gradients (Wolstenholme, 1992). Replication starts at the origin of H strand replication O_H, which is situated in the control region. A replication loop forms, in which the parental H and L strands are separated from each other and an RNA primer hybridizes to the L strand (Fig. 11A). H

strand synthesis proceeds clockwise, with the growing nascent H strand displacing the parental H strand (Fig. 11B). When the replication fork reaches the origin of L strand replication (O_L), L strand replication begins in the other direction using the parental H strand as a template (Fig. 11C). Replication in mt genomes has reported to take approximately 2 hours. This means that some regions on the H strand can spend a long time in the single-stranded state. Recently, however, evidence for other modes of replication have been presented whereby the time the H strand is spent single-stranded is reduced or absent (Holt et al., 2000; Reyes et al., 2005; Yasukawa et al. 2007).

Previous studies of mammalian genomes have shown that strand bias (when base frequencies on one strand are not at equilibrium – Sueoka, 1995) is significantly correlated with the time regions on the H strand were spent single-stranded during asynchronous replication, or D_{SSH} (Reyes et al., 1998). The longer a region or gene on one strand is spent single-stranded the more susceptible it is to oxidative damage. Oxidative deamination of cytosine to uracil and of adenine to hypoxanthin on the H strand results in guanine to adenine and thymine to cytosine substitutions, respectively, on the L strand (Reyes et al., 1998). L strand base composition thus becomes more skewed towards adenine and cytosine with increasing D_{SSH}.

Thirteen snakes were included in this analysis: three scolecophidians and ten alethinophidians. L strand mtDNA sequences were used in all cases. D_{SSH} was calculated for each of the 12 H strand encoded genes in addition to compositional skew (Perna and Kocher, 1995). D_{SSH} was calculated twice for alethinophidian snakes to take into account the two control regions as possible sites of replication initiation. Statistical tests were then performed to see whether compositional skew was correlated with D_{SSH}. It was found that for alethinophidians correlation of both AT and GC skew with D_{SSH} was significant, and whereas AT skew was positively correlated, GC skew was negatively correlated. However, no significant correlation was found in scolecophidians. This could be due to the much smaller sample size of scolecophidians compared to that of alethinophidians.

More statistical tests were done to see if C/T ratio and D_{SSH} were correlated in both groups. The proportions of cytosine to thymine were negatively correlated in accordance with the hypotheses put forth by Reyes et al. (1998). C/T ratio is a measure of compositional skew in which a high value is indicative of greater skew. C/T ratio has been shown to increase linearly with D_{SSH} (Faith and Pollock, 2003). Results again showed C/T ratio and D_{SSH} to be correlated in alethinophidians but not in scolecophidians. In another test the sample size of alethinophidians was reduced to three and an F-test for covariation was performed. The same results were found. Although sample sizes for alethinophidians and scolecophidians were the same in this last test, this should not be a factor responsible for the discrepancy between the two groups as correlation between C/T ratio and D_{SSH} should be apparent in individual genomes (Raina et al., 2005). This suggests that there may be another replication mechanism present in scolecophidians, especially since they do not have an O_L.

Comparisons of base composition between snakes and other amniotes

Nucleotide frequencies for ten snakes, lizards, crocodiles, birds, turtles and mammals were collected and analyzed using the statistical program SAS (SAS Institute Inc., 1990). The mt genes of seven amphisbaenians were also included. Within-group and between-group comparisons of composition were performed. Snakes were found to have the most divergent base composition (as exhibited by the high $\chi 2$-values) of any group at first and second codon positions (Table 1). This was in striking contrast to birds, which were found to be highly homogeneous. Between-group comparisons revealed the lineage specificity of nucleotide composition (Table 1). However, the highly significant result was mostly due to the aberrant composition in birds and mammals. Snakes along with turtles showed elevated amount of adenine at first and third codon positions compared to other lineages. None of the other squamate groups showed this trend.

Table 1. $\chi 2$-values and probability (p) values for tests of nucleotide composition within and between amniote groups at each codon position.

		SN	LI	AM	CR	BI	TU	MA	Between groups
1cp	$\chi 2$	123.31	119.67	34.68	52.18	38.29	46.62	95.69	827.18
	p	<0.0001	<0.0001	0.05	0.0025	0.07	0.05	<0.0001	<0.0001
2cp	$\chi 2$	24.15	12.48	16.77	21.39	7.54	3.14	10.38	186.54
	p	0.62	0.99	0.54	0.77	1.0	1.0	1.0	<0.0001
3cp	$\chi 2$	780.61	851.22	537.27	739.28	459.47	322.7	1052.46	3414.73
	p	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

[a] $\chi 2$ values represent the total $\chi 2$ for each test
[b] SN: snakes; LI: lizards; AM: amphisbaenians; CR: crocodilians; BI: birds; TU: turtles; MA: mammals

Because the elevated amount of adenine was mostly seen at 1st codon positions, it should also be reflected in the amino acid composition. A correspondence analysis was performed on squamate taxa to investigate how amino acids were being used in the 12 H-strand encoded genes in each species. Alethinophidian snakes were shown to cluster with amino acids methionine and lysine (Fig. 12). *Leptotyphlops* was shown to be closest to threonine, *Typhlops* to asparagine and isoleucine and *Ramphotyphlops* to serine and phenylalanine. The scolecophidians were not close to each other. Whereas snakes clustered in the left of the graph, lizards and amphisbaenians clustered on the right. This was with the exception of two acrodont lizards *Furcifer oustaleti* and *Xenagama taylori*, which clustered with alethinophidian snakes. *Furcifer* and *Xenagama*, together with *Pogona*, belong to a monophyletic group of iguanian lizards called Acrodonta. The similar base composition between acrodonts and snakes may explain why snakes and acrodont lizards cluster in phylogenetic analyses using mt data (Townsend et al., 2004; see Paper I).

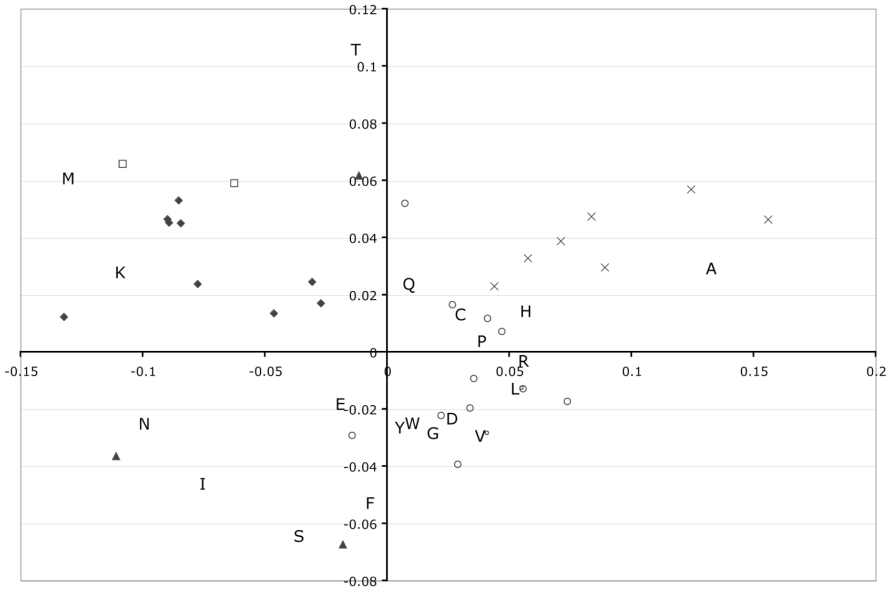


Fig. 12. Correspondence graph showing amino acid usage in squamates. Diamonds: alethinophidian snakes; triangles: scolecophidians; open squares: *Furcifer* and *Xenagama*; open circles: other lizards; crosses: amphisbaenians. Amino acids are denoted by their one-letter code.

With the exception of *Ramphotyphlops*, snakes clustered with amino acids that had adenine at the first codon position. Correlation analyses revealed that the proportion of these amino acids was significantly correlated with the frequency of adenine at both first and third codon positions. This was particularly the case for four-fold degenerate third codon positions at which selection was at a minimum. This ruled out Darwinian selection and suggested that mutation pressure was responsible for the bias towards adenine.

The results from this study show firstly that alethinophidian snake mt genomes show the same trends in strand bias as previous studies, which is evidence that the strand-asynchronous mode of replication (Clayton, 1982) is the predominant mode in these genomes. Alethinophidian mt genomes show a compositional bias towards adenine, which is probably the result of mutation pressure. Scolecophidian mt genomes do not exhibit any of the strand bias trends. This suggests that there may be another mode of

replication operating in scolecophidian genomes. More scolecophidian mt genomes would be needed to test this further.

Finally, base composition is extremely heterogeneous in snakes and particularly among scolecophidians. This may reflect the greater antiquity of the group or that the fossorial lifestyle of this group prevents a large amount of gene flow between populations.

PAPER III: THE PHYLOGENY OF SQUAMATES

It was understood from the first paper that the position of snakes could not be realized without greater taxon sampling of other squamates. In order to break up the long branches of the acrodont lizards and to reduce LBA, three acrodonts were to be sequenced. However, while this work was being carried out the mtDNA sequences of these same species or species closely related to them had been published (Macey et al., 2006; Kumazawa, 2007; Ujarvi et al., 2007). Nevertheless, the sequencing of one acrodont lizard, *Uromastyx aegyptia* that belongs to a subfamily of agamids whose species hitherto had not been sequenced (Uetz et al., 2006), was completed. A dibamid, *Dibamus novaeguineae*, was also sequenced. This taxon was important as it was found to be the most basal squamate in nuclear analyses (Townsend et al., 2004; Vidal and Hedges, 2005) and so would help root the mitochondrial tree, if that is its true position.

The dataset for this study included 41 squamate taxa: 13 snakes, five anguimorphs, five scincomorphs, seven iguanians, three gekkotans, seven amphisbaenians and *Dibamus*. Nucleotide data were partitioned into first codon positions, second codon positions and rRNA data. Third codon positions were excluded. Model parameters were thus optimized for each partition separately. Partitioning the data allows better fit of the model to the data. Amino acid data were left unpartitioned because initial phylogenetic analyses revealed that each of the gene partitions were too small for the parameters to be optimized efficiently.

Additional analyses were performed using models specifically designed to account for secondary structure within RNA molecules: mitoSLT+I+Γ (Smith et al., 2004) and the doublet model (Schöniger and von Haeseler, 1994). The data were analyzed using maximum likelihood and Bayesian inference methods, as these have been proven to be less susceptible to long-branch attraction than maximum parsimony and distance methods (Felsenstein, 1978). The model GTR+I+Γ8 (chosen by Modeltest - Posada and Crandall, 1998) and mtREV+I+Γ8 were used to analyze nucleotide and amino acid data,

respectively. In addition to these two categories (CAT) models (Lartillot and Philippe, 2004), CAT-GTR and CAT-Poisson, were also used to analyze nucleotide and amino acid data, respectively. CAT models are different from all others in that they partition the dataset according to their nucleotide or amino acid profiles. Profiling is based on the nucleotides or amino acids that are observed at each site, such that those sites with the same or very similar profiles would be grouped into one partition, or category. This is especially important for amino acids, as the model takes into account that not all 20 amino acids are likely to occur at all sites. Some amino acids will be frequent at some sites while others not at all. Because of this, CAT models should in theory be able to detect greater amounts of homoplasy in the dataset (Lartillot et al., 2007). The platform on which this model is implemented, PhyloBayes, uses Bayesian inference as the underlying criterion.

Initial phylogenetic analyses using maximum likelihood joined acrodonts with snakes instead of with iguanids. The affinities of snakes and acrodonts were thus analyzed separately in subsequent analyses. Furthermore, the most fast-evolving taxa were removed by creating a distance matrix based on the mtREV+I+Γ8 model and compared distances between each squamate taxon and the turtles. Any squamate taxon that had a distance of 1.0 or more was removed. This left three snakes (all scolecophidians), five anguimorphs, four iguanians, four amphisbaenians, five scincomorphs, three gekkotans and *Dibamus*. The taxon sampling may have been compromised but it was only compromised within, not across, lineages. In addition, fast-evolving taxa only mask the true phylogenetic signal of the data and, as was evident, increase the likelihood of LBA (Baurain et al., 2007).

The CAT-GTR trees showing the affinities of squamate lineages with snakes excluded (Fig. 13A) and snakes included (Fig. 13B) are shown. Most of the relationships between lizard lineages are congruent with previous mitogenomic and nuclear studies. The Iguania joined with the Anguimorpha, Amphisbaenia was sister-group to lacertiform scincomorphs and *Dibamus* occupied a basal position in the tree. However, in all analyses except CAT, *Dibamus* joined with the scincoid lizard *Plestiodon egregius.*

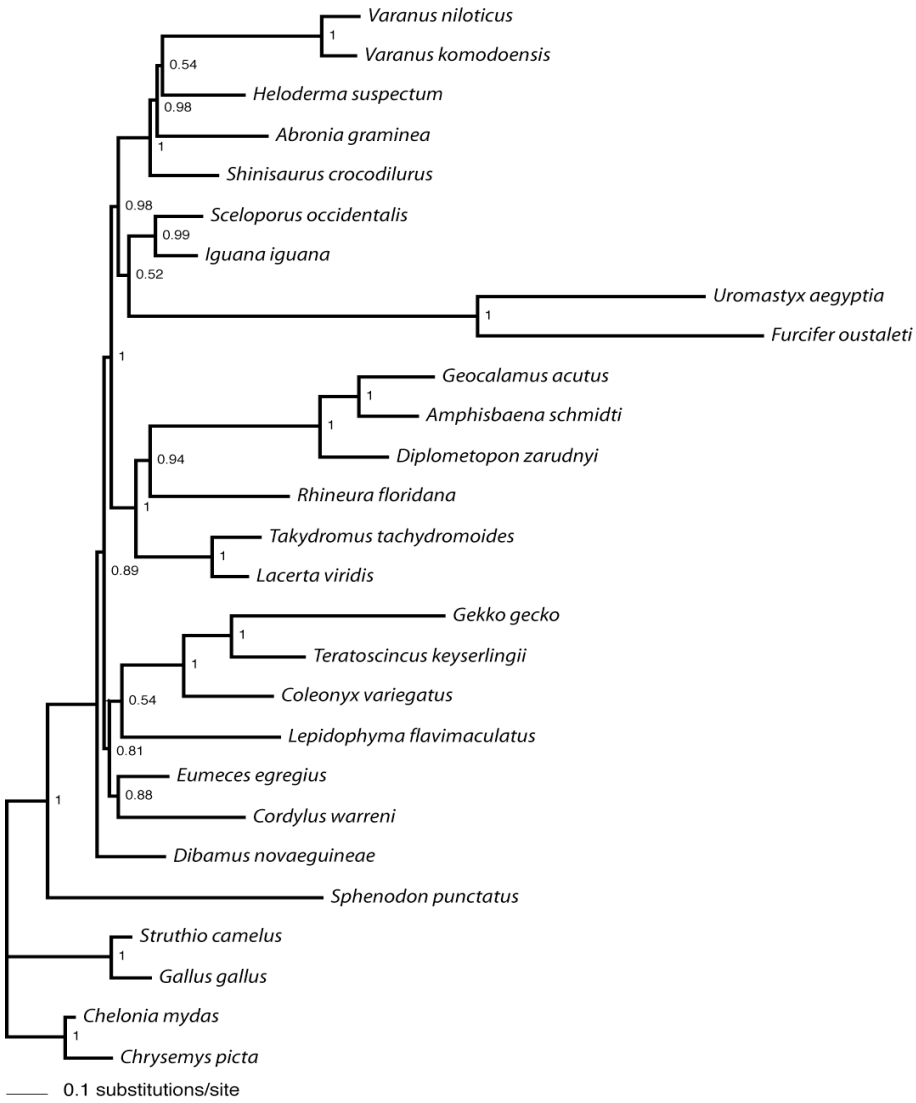


Fig. 13A. The CAT-GTR tree with snakes excluded. Note: *Eumeces egregius* = *Plestiodon egregius*.

In CAT analyses *Dibamus* alone was sister-group to the rest of Squamata. In most other analyses the other scincoids joined with gekkotans. Their positions, and that of *Dibamus*, were tested using six paired sites tests: Bootstrap Probability (Felsenstein, 1981), Expected Likelihood Weights (Strimmer and Rambaut, 2002) and the Approximately Unbiased test

(Shimodaira, 2002), in addition to Kishino-Hasegawa, Shimodaira-Hasegawa and Weighted Shimodaira-Hasegawa tests. These are implemented in the program Treefinder (Jobb, 2008). The preferred topology was that *Dibamus* and *Plestiodon* joined in a clade basal to all other squamates with other scincoids sister-group to gekkotans. However, no alternative topologies were rejected.

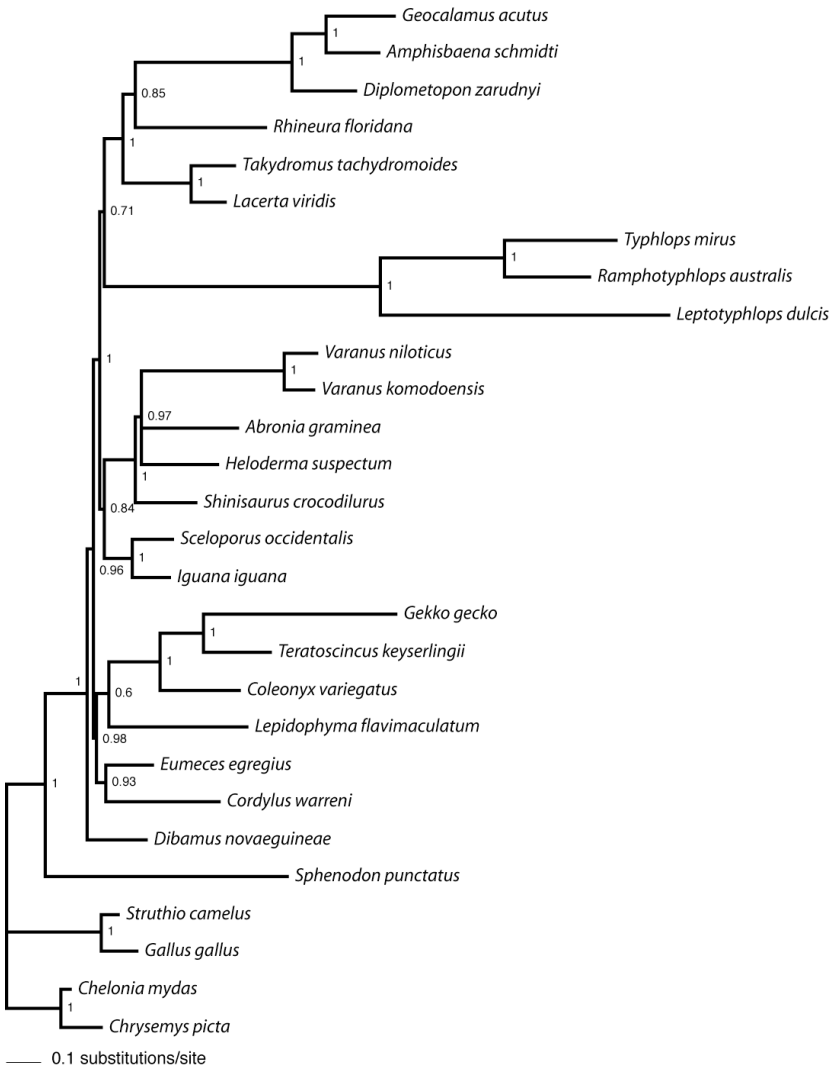


Fig. 13B. The CAT-GTR tree with snakes included (acrodont lizards excluded).

The snakes were recovered as sister-group to the Amphisbaenia/lacertiform clade, named the "Laterata" (Vidal and Hedges, 2005) in all analyses with lizard relationships unchanged (see Fig. 13B). As before, this result was compared to alternative positions for snakes using the paired sites tests mentioned above. Affinities of snakes to anguimorphs or iguanians were rejected by all paired sites tests except the Shimodaira-Hasegawa test.

These results are congruent with previous phylogenetic studies with respect to the affinities of lizard lineages. *Dibamus* was found to occupy a basal position, although its relationships with respect to scincoid lizards are still uncertain. These results support a recent mitogenomic study that also place snakes with the "Laterata" (Kumazawa, 2007) and congruent in this respect to paper I. These results are in contrast to nuclear gene studies that place snakes with anguimorphs and iguanids.

A STUDY ON THE RELATIONSHIPS OF BASAL ALETHINOPHIDIANS

The aim of this study was to infer the relationships of basal alethinophidians, which have hitherto been difficult to resolve. The mitochondrial genomes of the false coral snake *Anilius scytale*, Peters' Philippine earth snake *Rhinophis philippinus* and the rosy boa *Charina trivirgata* were sequenced.

Nucleotide and amino acid data from 12 H-strand encoded mitochondrial genes were analyzed using the models GTR+I+Γ8 and mtREV+I+Γ8, respectively. The CAT-GTR and CAT-Poisson models were also used. Nucleotide data was partitioned into first and second codon positions. Third codon positions were excluded. Maximum likelihood and Bayesian inference methods were used to infer relationships.

Fig. 14A shows a strict consensus tree of all ML and MB analyses. Fig. 14B show a strict consensus tree of the CAT analyses. ML and MB analyses showed different positions for the *Python/Xenopeltis* clade depending on whether nucleotides or amino acids were used.

However, both nucleotide and amino acid CAT analyses produced the same tree. In all analyses however, *A. scytale* is basal to all other alethinophidians, in agreement with other molecular studies (Wilcox et al., 2002; Lawson et al., 2004; Gower et al., 2005; Noonan and Chippendale, 2006). The monophyly of Anilioidea is thus not supported by mitogenomic data. The uropeltid snake *Rhinophis* invariably formed a clade with *Cylindrophis*, supporting the latter's inclusion in the Uropeltidae (contra Cundall et al., 1993). In accordance with Dong and Kumazawa (2005), *Python* and *Xenopeltis* form a strongly supported clade in all analyses. The Booidea, either sensu Rieppel (1988) or Lee and Scanlon (2002), was not supported by any analyses.

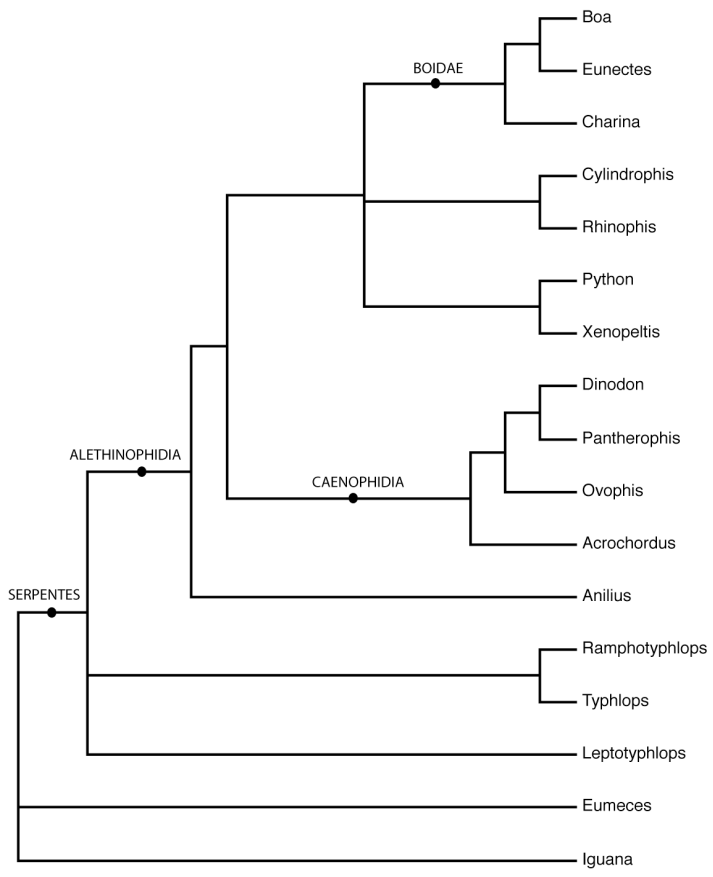


Fig. 14A. Strict consensus of nucleotide and amino acid maximum likelihood and Bayesian analyses

The positions of the other alethinophidians are less conclusive. In ML and MB analyses, all basal alethinophidians to the exclusion of *Anilius* formed a clade, but in the CAT analyses the Caenophidia was nested within this group with boids basal. The reason for the discrepancy between methods may be due to great rate heterogeneity between lineages and the attraction of Caenophidia towards the long branch of Alethinophidia. Paired sites tests, performed using maximum likelihood, preferred the topology of Fig. 14A over alternative topologies but did not reject the CAT tree (Fig. 14B). In all trees, Macrostomata was paraphyletic as anilioid taxa are nested within it. This suggests that macrostomy was lost in several snake families.

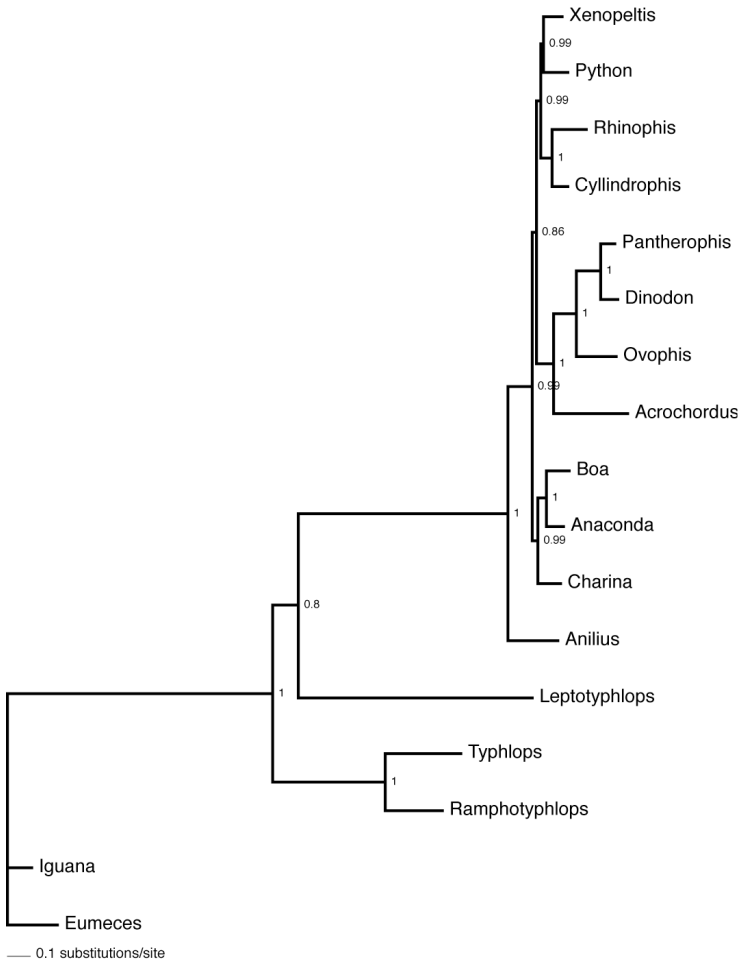


Fig. 14B. The CAT-Poisson tree. The CAT-GTR tree had an identical topology.

The novel position of boids in CAT analyses is intriguing, as this might support the 'macrostomy primitive' hypothesis proposed by proponents of the view that the fossil snakes that exhibit macrostomatan features are the most basal of all snakes. Furthermore, Cadle et al. (1990) found that boid albumin was more generalized than other alethinophidians, which would support a more basal position for Boidae.

CONCLUSIONS

This thesis work has shown that the mitogenomic evolution of snakes is very complex in that evolutionary rates, base composition and gene rearrangements differ markedly between lineages. Mitochondrial genes of all snakes have evolved at a much higher rate than in other lineages of squamates, which has made elucidating their phylogenetic affinities to other squamates difficult. However, with more advanced phylogenetic methods that ensure better fit of the model to the data, models that detect homoplasy more effectively and better taxon sampling, the phylogenetic signal of the data was increased as seen in the difference in the results of Paper I and Paper III.

The current results support a close relationship of snakes to amphisbaenians and lacertiform lizards. This is in contrast to nuclear genes that join snakes with anguimorphs and iguanians. This issue will only be resolved by the inclusion of more comprehensive data. Because the mitochondrial dataset contains a finite amount of data, elucidating the true position of snakes will ultimately be determined by the analysis of more nuclear genes, as initiated by the work of Vidal and Hedges (2005). Elucidating squamate relationships and the position of snakes will always be a challenge for molecular phylogenetics however, because of the fast evolutionary rates (also seen in nuclear genes - Hughes and Mouchiroud, 2001) and the short internal branches that signify rapid cladogenesis among squamates.

The striking differences in evolutionary rate between basal snakes are reflected by their heterogeneous base composition. Questions that could be asked are: did the duplication of the control region and flanking genes affect mutation pressure in alethinophidian mt genomes, and if so, how? Could the skew in amino acid composition towards lysine and methionine in alethinophidian genomes have functional significance? The picture appears to be even more complex, as in a recent study it was found that in snakes the

rate of evolution differs among mt genes in different snake lineages (Jiang et al., 2007).

Resolving the relationships of basal alethinophidian snake lineages has also been a complex issue. It is evident that the results from this work do not support the monophyly of the Anilioidea, Booidea or Macrostomata and that macrostomy may have been secondarily lost in some 'anilioid' lineages. This, taken with the macrostomatan features of fossil snakes, may suggest that macrostomy is a primitive ophidian character. However, current mitogenomic evidence does not support this. The mt genomes of all macrostomatans, and all alethinophidians, sequenced to date have two control regions. This is derived from the condition observed in scolecophidian mt genomes, which only have one control region. Although it is plausible to consider the anilioids as regressed macrostomatans, the same cannot be said of scolecophidians, whose morphology (Kley and Brainerd, 1999; Kley, 2001) is derived to a degree not seen in any other burrowing snake lineage. This suggests not only that scolecophidians are more ancient than the fossil snakes, but also that the true snake ancestor had a very different morphology to modern snakes.

It would be interesting to further explore the scolecophidians, their highly divergent base composition, population structure and whether or not they are a monophyletic group. Although the results from this work do not support scolecophidian monophyly, this may be due to the similarities in composition between *Leptotyphlops* and alethinophidians. Additional sampling would be needed to further elucidate the mitogenomic evolution of this group and to infer the phylogenetic relationships of scolecophidian taxa.

COPYRIGHT NOTICE

Permission to reproduce the following figures, images and articles was obtained from the publisher or copyright holder.

Cover photo and right-hand photo in Fig.1: The Rainbow Boa, *Epicrates cenchria* taken by RainForest Adventures.

Fig 1 left-hand photo: Texas Blind Snake *Leptotyphlops dulcis* taken by Gary Nafis, californiaherps.com.

Fig. 2A: Skull of *Leptotyphlops dulcis* taken from Digimorph.org.

Fig. 2B: Skull of the cottonmouth viper *Agkistrodon piscivorus*, taken by East Coast Natureworld Tasmania.

Fig. 4: Mosasaur illustration by Carl Buell.

Paper I reproduced with permission from Blackwell Publishing Ltd.

REFERENCES

Adachi J, Hasegawa M, 1996. Model of amino acid substitution in proteins encoded by mitochondrial DNA. *J. Mol. Evol.* **42**: 459-468

Apesteguía S, Zaher H, 2006. A Cretaceous terrestrial snake with robust hindlimbs and a sacrum. *Nature* **440**: 1037-1040

Arnason U, Gullberg A, Janke A, Joss J, Elmerot C, 2004. Mitogenomic analyses of deep gnathostome divergences: a fish is a fish. *Gene* **333**: 61-70

Baurain D, Brinkmann H, Philippe H, 2007. Lack of resolution in the animal phylogeny: closely spaced cladogenesis or systematic errors? *Mol. Biol. Evol.* **24(1)**: 6-9

Bellairs AD, Underwood U, 1951. The origin of snakes. *Biol. Rev.* **26**: 193-237

Bernhardt T. The Canadian Biodiversity Website http://canadianbiodiversity.ca

Bogenhagen D, Clayton DA, 1974. The number of mitochondrial deoxyribonucleic acid genomes in mouse L and human HeLa cells. Quantitative isolation of mitochondrial deoxyribonucleic acid. *J. Biol. Chem.* **249(24)**: 7991-5

Cadle JE, Dessauer HC, Gans C, Gartside DF, 1990. Phylogenetic relationships and molecular evolution in uropeltid snakes (Serpentes: Uropeltidae): allozymes and albumin immunology. *Biol. J. Linn. Soc.* **40**: 293-320

Caldwell MW, 2003. "Without a leg to stand on": on the evolution and development of axial elongation and limblessness in tetrapods. *Can. J. Earth Sci.* **40**: 573-588

Caldwell MW, Lee MSY, 1997. A snake with legs from the marine Cretaceous of the Middle East. *Nature* **386**: 705-709

Camp CL, 1923. Classification of the lizards. *Bull. Am. Mus. Nat. Hist.* **48**: 289-481

Caprette CL, Lee MSY, Shine R, Mokany A, Downhower JF, 2004. The origin of snakes as seen through eye anatomy. *Biol. J. Linn. Soc.* **81**: 469-482

Clayton DA, 1982. Replication of animal mitochondrial DNA. *Cell* **28**: 693-705

—, 1991. Replication and transcription of vertebrate mitochondrial DNA. *Ann. Rev. Cell Biol.* **7**: 453-478

Cohn MJ, Tickle C, 1999. Developmental basis of limblessness and axial patterning in snakes. *Nature* **399**: 474-479

Cooper Jr. WE, 1995. Foraging mode, prey chemical discrimination, and phylogeny in lizards. *Anim. Behav*. **50**: 973-985

Cope ED, 1869. On the reptilian order Pythonomorpha and Streptosauria. *Proc. Boston Soc. Nat. Hist.* **12**: 250-261

—, 1878. Professor Owen on the Pythonomorpha. *Bull. U.S. Geol. Geogr. Serv. Territories* **4**: 299-311

Cundall D, Wallach V, Rossman DA, 1993. The systematic relationships of the snake genus *Anomochilus. Zool. J. Linn. Soc*. **109**: 275-299

Desjardins P, Ramirez V, Morais R, 1990. Gene organization of the Peking duck mitochondrial genome. *Curr. Genet*. **17(6)**: 515-518

Dong S, Kumazawa Y, 2005. Complete mitochondrial DNA sequences of six snakes: phylogenetic relationships and molecular evolution of genomic features. *J. Mol. Evol.* **61(1)**: 12-22

Estes R, deQuieroz K, Gaulthier J, 1988. Phylogenetic relationships within Squamata in R. Estes and G. Pregill (eds) *Phylogenetic relationships of the lizard families* (pp 119-281). Stanford University Press, Stanford

Faith JJ, Pollock DD, 2003. Likelihood analysis of asymmetrical mutation bias gradients in vertebrate mitochondrial genomes. *Genetics* **165**: 735-745

Fejévary GJ, 1918. Contributions to a monography on the fossil Varanidae and Megalanidae. *Ann. Mus. Nat. Hung*. **16**: 341-467

Felsenstein J, 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool*. **27(4)**: 401-410

Felsenstein J, 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol*. **17**: 368-376

Fernández-Silva P, Enriquez JA, Montoya J, 2003. Replication and transcription of mammalian mitochondrial DNA. *Exp. Physiol.* **88(1):** 41-56

Fonseca MM, Froufe E, Harris DJ, 2006. Mitochondrial gene rearrangements and partial genome duplications detected by multigene asymmetric composition bias analysis. *J. Mol. Evol*. **63**: 654-661

Forstner MRJ, Davis SK, Arévalo E, 1995. Support for the hypothesis of anguimorph ancestry for the suborder Serpentes from phylogenetic analysis of mitochondrial DNA sequences. *Mol. Phylogenet. Evol.* **4(1)**: 93-102

Fry BG, Vidal N, Norman JA, Vonk, FJ, Scheib H, Ramjan SFR, Kuruppu S, Fung K, Hedges SB, Richardson, MK, Hodgson WC, Ignjatovic V, Summerhayes R, Kochva E, 2006. Early evolution of the venom system in lizards and snakes. *Nature* **439**: 584-588

Gower DJ, Vidal N, Spinks JN, McCarthy CJ, 2005. The phylogenetic position of Anomochilidae (Reptilia: Serpentes): first evidence from DNA sequences. *JZS* **43(4)**: 315-320

Greene H, 1997. *Snakes, the evolution of mystery in nature.* Berkeley and Los Angeles, California: University of California Press

Greene H, Cundall D, 2000. Limbless tetrapods and snakes with legs. *Science* **287**: 1939-1941

Greer AE, 1985. The relationships of the lizard genera *Anelytropsis* and *Dibamus*. *J. Herp*. **19(1)**: 116-156

Greer AE, 2002. The loss of the external ear opening in scincid lizards. *J. Herp*. **36(4)**: 544-555

Gregory JT, 1952. The jaws of the Cretaceous toothed birds, *Ichthyornis* and *Hesperornis*. *The Condor* **54(2):** 73-88

Gu X, Fu Y, Li W, 1995. Maximum likelihood estimation of heterogeneity of substitution rate among nucleotide sites. *Mol. Biol. Evol.* **12(4)**: 546-557

Hallermann J, 1998. The ethmoidal region of *Dibamus taylori* (Squamata: Dibamidae), with a phylogenetic hypothesis on dibamid relationships within Squamata. *Zool. J. Linn. Soc*. **122**: 385-426

Harris DJ, 2003. Codon bias variation in *c-mos* between squamate families might distort phylogenetic inferences. *Mol. Phylogenet. Evol.* **27**: 540-544

Holt IJ, Lorimer HE, Jacobs HT, 2000. Coupled leading- and lagging- strand synthesis of mammalian mitochondrial DNA. *Cell* **100(5)**: 515-524

Hughes S, Mouchiroud D, 2001. The evolutionary rates in nuclear genes of squamates. *J. Mol. Evol.* **53:** 70-76

Janke A, Erpenbeck D, Nielsson M, Arnason U, 2001. The mitochondrial genomes of a lizard, *Iguana iguana*, and the caiman, *Caiman crocodylus*: implications for amniote phylogeny. *Proc. R. Soc. Lond. B* **268**: 623-631

Jiang ZJ, Castoe TA, Austin CC, Burbrink FT, Herron MD, McGuire JA, Parkinson CL, Pollock DD, 2007. Comparative mitochondrial genomics of

snakes: extraordinary substitution rate dynamics and functionality of the duplicate control region. *BMC Evol. Biol.* **7**: 123

Jobb G, 2008. *TREEFINDER, version of January 2008*. [Computer software and manual] Munich, Germany. Distributed by the author at www.treefinder.de

Kearney M, 2003. Systematics of the Amphisbaenia (Lepidosauria: Squamata) based on morphological evidence from recent and fossil forms. *Herpetological Monographs* **17**: 1-74

Kearney M, Rieppel O, 2006. An investigation into the occurrence of plicidentine in the teeth of squamate reptiles. *Copeia* **2006(3)**: 337-350

Kishino H, Hasegawa M, 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data and the branching order in Hominoidea. *J. Mol. Evol.* **29**: 170-179

Kley NJ, 2001. Prey transport mechanisms in blindsnakes and the evolution of unilateral feeding systems in snakes. *Amer. Zool.* **41:** 1321-1337

Kley NJ, Brainerd EL, 1999. Feeding by mandibular raking in a snake. *Nature* **402**: 369-370

Kochva E, 1978. Oral glands of the reptiles in C. Gans and K.A. Gans (eds) *Biology of the Reptilia* **vol. 8** (pp 43-161) Academic Press, London

—, 1987. The origin of snakes and evolution of the venom apparatus. *Toxicon* **25(1)**: 65-106

Kumazawa Y, 2004. Mitochondrial DNA sequences of five squamates: phylogenetic affiliation of snakes. *DNA Res.* **11**: 137-144

Kumazawa Y, 2007. Mitocondrial genomes from major lizard families suggest their phylogenetic relationships and ancient radiations. *Gene* **388**: 19-26

Kumazawa Y, Endo H, 2004. Mitochondrial genome of the Komodo dragon, efficient sequencing method with reptile-oriented primers and novel gene rearrangements. *DNA Res.* **11**: 115-125

Kumazawa Y, Nishida M, 1995. Variations in mitochondrial tRNA gene organization of reptiles as phylogenetic markers. *Mol. Biol. Evol.* **12(5)**: 759-772

Kumazawa Y, Ota H, Nishida M, Ozawa T, 1996. Gene rearrangements in snake mitochondrial genomes, highly concerted evolution of control region-like sequences duplicated and inserted into a tRNA gene cluster. *Mol. Biol. Evol.* **13(9):** 1242-1254

Kumazawa Y, Ota H, Nishida M, Ozawa T, 1998. The complete nucleotide sequence of a snake (*Dinodon semicarinatus*) mitochondrial genome with two identical control regions. *Genetics* **150**: 313-329

Lanave C, Preparata G, Saccone C, Serio G, 1984. A new method for calculating evolutionary substitution rates. *J. Mol. Evol.* **20(1)**: 86-93

Lartillot N, Philippe H, 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Mol. Biol. Evol.* **21(6)**: 1095-1109

Lartillot N, Brinkmann H, Philippe H, 2007. Suppression of long-branch attraction artifacts in the animal phylogeny using a site-heterogeneous model. *BMC Evol. Biol.* **7 (suppl. 1)** S4

Lawson R, Slowinski JB, Burbrink FT, 2004. A molecular approach to discerning the phylogenetic placement of the enigmatic snake *Xenophidion schaeferi* among the Alethinophidia. *J. Zool. Lond.* **263**: 285-294

Lee MSY, 1998. Convergent evolution and character correlation in burrowing reptiles: towards a resolution of squamate relationships. *Biol. J. Linn. Soc.* **65**: 369-453

Lee MSY, Caldwell MW, 2000. *Adriosaurus* and the affinities of mosasaurs, dolichosaurs and snakes. *J. Palaeont.* **74(5)**: 915-937

Lee MSY, Scanlon JD, 2002. Snake phylogeny based on osteology, soft anatomy and ecology. *Biol. Rev.* **77**: 333-401

Macey JR, Verma A, 1997. Re: homology in phylogenetic analysis: alignment of transfer RNA genes and phylogenetic position of snakes. *Mol. Phylogenet. Evol.* **7**: 272-279

Macey JR, Schulte II JA, Fong JJ, Das I, Papenfuss TJ, 2006. The complete mitochondrial genome of an agamid lizard from the Afro-Asian subfamily Agaminae and the phylogeneti position of *Bufoniceps* and *Xenagama*. *Mol. Phylogenet. Evol.* **39**: 881-886

Mahendra BC, 1938. Some remarks on the phylogeny of the Ophidia. *Anat. Anz.* **86**: 347-356

McDowell SB, 1972. The evolution of the tongue of snakes and its bearing on snake origins, in T. Dobzhansky, M.K. Hecht and W.C. Steere (eds) *Evolutionary Biology* **Vol. 6** (pp 191-273) Appleton-Century-Crofts, New York

McDowell SB, Bogert CM, 1954. The systematic position of *Lanthanotus* and the affinities of anguimorphan lizards. *Bull. Am. Mus. Nat. Hist.* **105**: 1-142

Noonan BP, Chippindale PT, 2006. Dispersal and vicariance: the complex evolutionary history of boid snakes. *Mol. Phylogenet. Evol*. **40(2)**: 347-358

Nopcsa F, 1923. *Eidolosaurus* und *Pachyopohis*. Zwei neue Neocom-Reptilien. *Palaeontographica* **65**: 97-154

Nopcsa F, 1925. Ergebnisse der Forschungsreisen Prof. E. Stromers in den Wüsten Ägyptens, II. Wirbeltier-Reste der Baharrîje-Stufe (unterstes Caenoman), 5, Die *Symoliophis*-Reste. *Abhandlungen der Bayerischen Akademie der Wissenschaften, mathematisch-naturwissenschaftliche Abeitlung* **30**: 1-27

Page RDM, Holmes EC, 1998. *Molecular Evolution, a phylogenetic approach.* Oxford: Blackwell Scientific

Palci A, Caldwell MW, 2007. Vestigial forelimbs and axial elongation in a 95 million-year-old non-snake squamate. *J. Vert. Palaeont*. **27(1)**: 1-7

Perna NT, Kocher TD, 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J. Mol. Evol*. **41**: 353-358

Piskurek O, Austin CC, Okada N, 2006. Sauria SINES: novel short interspersed retroposable elements that are widespread in reptile genomes. *J. Mol. Evol*. **62**: 630-644

Posada D, Crandall KA, 1998. Modeltest, testing the model of DNA substitution. *Bioinf.* **14(9)**: 817-818

Rage J-C, 1982. La phylogénie des Lépidosauriens (Reptilia): Une approche cladistique. *CR Acad. Sci. Paris* **294**: 563-566

Rage J-C, Escuillié F, 2000. Un nouveau serpent bipède du Cénomanien (Crétacé). Implications phylétiques. *CR Acad. Sci. Paris Sciences de la Terre et des Planètes* **330**: 513-520

Rage J-C, Escuillié F. 2003. The Cenomanian: stage of hindllimbed snakes. *Carnets de Géologie Maintenon* Article 2003/01

Raina SZ, Faith JJ, Disotell TR, Seligmann H, Stewart C, Pollock DD, 2005. Evolution of base-substitution gradients in primate mitochondrial genomes. *Genome Res*. **15**: 665-673

Rest JS, Ast JC, Austin CC, Waddell PJ, Tibbetts EA, Hay JM, Mindell DP, 2003. Molecular systematics of primary reptilian lineages and the tuatara mitochondrial genome. *Mol. Phylogenet. Evol.* **29(2)**: 289-297

Reyes A, Gissi C, Pesole G, Saccone C, 1998. Asymmetrical directional mutation pressure in the mitochondrial genome of mammals. *Mol. Biol. Evol.* **15**: 957-966

Reyes A, Yang MY, Bowmaker M, Holt IJ, 2005. Bidirectional replication initiates at sites throughout the mitochondrial genome of birds. *J. Biol. Chem.* **280(5)**: 3242-3250

Rieppel O, 1988. A review of the origin of snakes. *Evol. Biol.* **22**: 37-130

Rieppel O, Kearney M, 2001. The origin of snakes: limits of a scientific debate. *Biologist* **48(3)**: 110-114

Rieppel O, Zaher H, 2000. The intramandibular joint in squamates and the phylogenetic relationships of the fossil snake *Pachyrachis problematicus* Haas. *Fieldiana (Geology)* **43**: 1-69

Rieppel O, Zaher H, Tchernov E, Polcyn MJ, 2003. The anatomy and relationships of *Haasiophis terrasanctus*, a fossil snake with well-developed hind limbs from the mid-Cretaceous of the Middle East. *J. Palaeont.* **77(3)**: 536-558

Roos J, Aggarwal RK, Janke A, 2007. Extended mitogenomic phylogenetic analyses yield new insight into crocodylian evolution and their survival of the Cretaceous-Tertiary boundary. *Mol. Phylogenet. Evol.* **45**: 663-673

Saint KM, Austin CC, Donnellan SC, Hutchinson MN, 1998. C-mos, a nuclear marker useful for squamate phylogenetic analysis. *Mol. Phylogenet. Evol.* **10(2)**: 259-263

Sambrook J, Russell DW, 2001. *Molecular Cloning, a laboratory manual.* Cold Spring Harbor, NY: Cold Spring Harbor Press

SAS Institute Inc., 1990. *SAS user's guide, 5th ed.* SAS Institute Inc., Cary, N.C.

Scanlon JD, 2006. Skull of the large non-macrostomatan snake *Yurlunggur* from the Australian Oligo-Miocene. *Nature* **439**: 839-842

Scanlon JD, Lee MSY, 2000. The Pleistocene serpent *Wonambi* and the early evolution of snakes. *Nature* **403**: 416-420

Schmidt HA, Strimmer K, Vingron M, von Haeseler A, 2002. TREE-PUZZLE, a maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinf.* **18**: 502-504

Schwenk K, 1988. Comparative morphology of the lepidosaur tongue and its relevance to squamate phylogeny, in R. Estes and G. Pregill (eds), *Phylogenetic relationships of the lizard families.* Stanford University Press, Stanford

Schöniger M, von Haeseler A, 1994. A stochastic model for the evolution of autocorrelated DNA sequences. *Mol. Phylogenet. Evol.* **3(3)**: 240-247

Senn DG, Northcutt RG, 1973. The forebrain and midbrain of some squamates and their bearing on the origin snakes. J. Morph. 140: 135-152

Shimodaira H, 2002. An approximately unbiased test of phylogenetic tree selection. *Syst. Biol.* **51(3)**: 492-508

Shimodaira H, Hasegawa M, 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol*. **16**: 1114-1116

Smith AD, Lui TWH, Tillier ERM, 2004. Empirical models for substitution in ribosomal RNA. *Mol. Biol. Evol*. **21(3)**: 419-427

Strimmer K, Rambaut A, 2002. Inferring confidence sets of possibly misspecified gene trees. *Proc. R. Soc. Lond. B* **269**: 137-142

Sueoka N, 1995. Intrastrand Parity Rules of DNA base composition and usage biases of synonymous codons. *J. Mol. Evol*. **40**: 318-325

Sutovsky P, Moreno RD, Ramalho-Santos J, Dominko T, Simerly C, Schatten G, 1999. Ubiquitin tag for sperm mitochondria. *Nature* **402**: 371-372

Tchernov E, Rieppel O, Zaher H, Polcyn MJ, Jacobs LL, 2000. A fossil snake with limbs. *Science* **287**: 2010-2012

Townsend TM, Larson A, Louis E, Macey JR, 2004. Molecular phylogenetics of Squamata: the position of snakes, amphisbaenians, dibamids, and the root of the squamate tree. *Syst. Biol*. **53(3)**: 735-757

Tsuihiji T, Kearney M, Rieppel O, 2006. First report of a pectoral girdle muscle in snakes, with comments on the snake cervico-dorsal boundary. *Copeia* **2006(2)**: 206-215

Uetz P et al., 2006. The Reptile Database, http://www.reptile-database.org

Ujvari B, Dowton M, Madsen T, 2007. Mitochondrial DNA recombination in a free-ranging Australian lizard. *Biol. Lett.* **3**: 189-192

Underwood G, 1957. *Lanthanotus* and the anguimorph lizards: a critical review. *Copeia* **1957**: 20-30

—, 1967. *A contribution to the classification of snakes*. British Museum (Natural History), London

—, 1970. The eye. In C. Gans and TS Parsons (eds) *Biology of the Reptilia* (pp 1-97). London: Academic Press.

Vidal N, David P, 2004. New insights into the early history of snakes inferred from two nuclear genes. *Mol. Phylogenet. Evol*. **31**: 783-787

Vidal N, Hedges SB, 2004. Molecular evidence for a terrestrial origin of snakes. *Proc. R. Soc. Lond. B. (suppl. 4)* **271**: 226-229

Vidal N, Hedges SB, 2005. The phylogeny of squamate reptiles (lizards, snakes and amphisbaenians) inferred from nine nuclear protein-coding genes. *CR Biol*. **328**: 1000-1008

Vitt LJ, Pianka ER, Cooper Jr. WE, Schwenk K, 2003. History and the global ecology of squamate reptiles. *Am. Nat*. **162(1)**: 44-60

Vitt LJ, Pianka ER, 2005. Deep history impacts present-day ecology and biodiversity. *PNAS* **102(22)**: 7877-7881

Walls G, 1940. Ophthalmological implications for the early history of snakes. *Copeia* **1940**: 1-8

Walls G, 1942. The vertebrate eye and its adaptive radiation. The Cranbrook Institute of Science, Bloomfield MI pp 785

Wilcox TP, Zwickl DJ, Heath TA, Hillis DM, 2002. Phylogenetic relationships of the dwarf boas and a comparison of Bayesian and bootstrap measures of phylogenetic support. *Mol. Phylogenet. Evol*. **25**: 361-371

Wolstenholme DR, 1992. Animal mitochondrial DNA: structure and evolution, in DR Wolstenholme and KW Jeon (eds) *Mitochondrial genomes.* International Review of Cytology **vol. 141** (pp 173-216) Academic Press, Inc.

Yasukawa T, Reyes A, Cluett TJ, Yang M, Bowmaker M, Jacobs HT, Holt IJ, 2006. Replication of vertebrate mitochondrial DNA entails transient ribonucleotide incorporation throughout the lagging strand. *EMBO* **25**: 5358-5371

Zaher H, 1998. The phylogenetic position of *Pachyhrachis* within snakes (Squamata, Lepidosauria). *J. Vert. Palaeont*. **18(1)**: 1-3

Zaher H, Rieppel O, 2002. On the phylogenetic relationships of the Cretaceous snakes with legs, with special reference to *Pachyhrachis problematicus* (Squamata: Serpentes) *J. Vert. Palaeont*. **22(1)**: 104-109

Zardoya R, Meyer A, 1998. Complete mitochondrial genome suggests diapsid affinities of turtles. *PNAS* **95(24)**: 14226-31

Zhou K, Hongdan L, Han D, Bauer AM, Feng J, 2006. The complete mitochondrial genome of *Gekko gecko* (Reptilia: Gekkonidae) and support for

the monophyly of Sauria including Amphisbaenia. *Mol. Phylogenet. Evol*. **40**: 887-892

ACKNOWLEDGEMENTS

I would like to thank:

Ulfur and Axel for giving me the opportunity to work on snake mt genome evolution.

Morgan and Björn for being good corridor mates and helping me out with any computer stuff, running phylogenetic analyses and interesting discussions.

Torbjörn for helping me out with statistical tests for Paper II and discussions about statistics.

Ulfur, Axel and Paul for giving helpful comments on the thesis.

The late Anette Gullberg for her help in the lab and on the teaching course.

Ulfur, Axel, Maria and Dave Gower for giving helpful comments on my papers.

Dave Gower, Jakob Hillman-Szabo, Steve Donnellan, Frank Madsen and all those who donated their squamate specimens.

Kristina and Pernilla for your friendship and support and the numerous Friday lunches to celebrate the end of each grueling week!

Anna Härlid for some good chats, running the *Broloppet* with Anders and I and for giving me the lovely Gubbe!

Former exams students Rasa and Jonas - good luck for the future.

Tina, Christine, Lena, Bengt Olle and everybody at Genetikhuset – thanks for Tuesday *fika* and *genetikkaffe*. It's been nice knowing you all!

My good friends Danni, Darren, Lotta, Mattias and Elena – I will miss all the pool nights, parties, movie nights, pub dinners and all our other social get-togethers.

The Sparta party animals and great corridor-mates: Karin, Elizabeth, Anders, Tomas, Andrew, Eric, Robert – thanks for all the great *korridorfester* and other fun times.

Alex, one-time classmate on the Swedish course in which we spent most of the time trying unsuccessfully to control our fits of laughter. We need you in

Fagans – it's been a while since we've won the quiz! I enjoyed all the little holiday breaks and hope for many more to come.
Jonas – my badminton partner, and the first person I got to know when I came to Lund. Thanks for being a good friend and for helping me to get over the culture shock (remember the *fil mjölk* incident?).
Megan - good luck with the rest of your projects and look forward to enjoying some more coffee time-outs.
Mattias, Björn and all other guys in Skånes Herpetologiska Föreningen – thanks for introducing me to your society and inviting me to give talks at meetings.
Ramiro, Anders and Guille (R.I.P. my friend), Luis and Christine and everyone else I got to know during my time in Lund - take care and good luck with everything.

Appendix: Papers I-III