学位論文(要約)

Study on the molecular evolution of insulin and seasonal control of metabolic regulators in the Japanese gecko

（ニホンヤモリにおけるインスリンの分子進化と代謝因子の季節制御に関する研究）

平成 29 年 12 月博士（理学）申請

東京大学大学院理学系研究科

生物科学専攻

山岸　弦記

学位論文(要約)

Study on the molecular evolution of insulin and seasonal control of metabolic regulators in the Japanese gecko

（ニホンヤモリにおけるインスリンの分子進化と代謝因子の季節制御に関する研究）

平成 29 年 12 月博士（理学）申請

東京大学大学院理学系研究科

生物科学専攻

山岸　弦記

Abstract

Among amniotes, ectothermic lineages are generally termed “reptiles” and distinguished from mammals and birds, two groups that have independently acquired endothermic system. Since reptiles do not generate body heat for thermoregulation, their body temperature is affected by ambient temperature fluctuation. Therefore, their activity is restricted by ambient climate. However, such disadvantage seems to have little effect on geographical distribution of one of reptilian groups called Squamata. Today, more than 9,000 squamate species are extant worldwide, from subarctic to tropical zones (Utez, 2015). Underlying this diversity seems to be their ability to acquire novel life-history traits in accordance with their habitat. Since every aspect of life requires energy (Andrade, 2016), regulatory mechanisms of energy metabolism is an interesting topic to understand prosperity of squamates. However, little is understood about the molecular mechanisms that regulate energy metabolism of squamates.

In this thesis, I aimed at establishing a basis for investigation of squamates’ metabolic regulation at the molecular level. Current understanding about evolution of organisms is that changes in coding sequences and those in transcriptional regulation of genes contribute to the acquisition of novel phenotypes. Therefore, I focused on 1) the variation of amino acid sequences of insulin among squamate lineages and 2) transcriptional regulation of key metabolic regulators.

In Chapter 1, I focused on the molecular evolution of insulin. Insulin is generally regarded as a central anabolic hormone for vertebrates. Reflecting its functional importance, amino acid sequence of insulin is generally well conserved (Conlon, 2001; Norris and Carr, 2013). However, insulin characterized from the Japanese gecko (*Gekko japonicus*) showed exceptional number of amino acid substitution. I predicted that such aberrant sequence reflects evolutionary background unique to squamates. In order to investigate the origin of aberrant sequence, I characterized insulin genes from multiple squamate lineages. Surprisingly, I found that the amino acid sequences of squamate insulin were more divergent than previously regarded. For the inference of evolutionary process acting on

squamate insulin, nonsynonymous-to-synonymous substitution ratio (dN/dS ratio) was calculated using PAML4 CODEML program (Yang, 2007). The values calculated suggested that squamate insulin had evolved under relaxed negative selection pressure. Furthermore, even positive selection pressure was detected in some lineages, implying that their insulin had undergone functional changes that increase their fitness. Since insulin alone is responsible for hypoglycemic regulation of vertebrates (Norris and Carr, 2013; Sherwood et al., 2013), these results suggest that aberrant insulin sequences increase fitness of squamates, possibly through unique metabolic regulation.

Interestingly, one of such lineages, the Japanese gecko, seems to have indeed undergone environmental adaptation. Geographical distribution and annual gametogenic cycle suggest that this species have achieved adaptation to temperate climate zone after divergence from a tropical ancestor (Ikeuchi, 2004; Rösler et al., 2011). While the energy balance of tropical species is relatively constant throughout the year, that of temperate species fluctuates as a result of winter dormancy and seasonal reproduction (Derickson, 1976; Dessauer, 1953; Di Maggio and Dessauer, 1963). Therefore, I predicted that the Japanese gecko had acquired regulatory mechanisms that optimize their energy metabolism for these phenomena. In chapter 2, I analyzed gene expression in the liver as a means to explore such regulatory mechanisms. I focused on the liver as the central organ for energy metabolism of the Japanese gecko because the Japanese gecko lacks abdominal adipose tissue (Ji and Wang, 1990), which is an important site for energy deposition in other squamates (Derickson, 1976), and seems to rely on the liver instead (Ji and Wang, 1990).

During dormancy, squamates are anorexic and rely heavily on internal energy storage to sustain their life activities (Derickson, 1976; Pough, 2016). Focusing on this aspect, I investigated gene expression responses to a prolonged period of food deprivation. The gene expression patterns suggest that the Japanese gecko under prolonged food deprivation spares protein utilization. Instead, this species seems to rely on glycogen and TAG as energy sources. Since similar responses to food

deprivation have been reported from other ectothermic vertebrates (Qian et al., 2016; Wang et al., 2006; Zani et al., 2012), the Japanese gecko seems to employ ancestral regulatory mechanisms to tolerate food deprivation during dormancy.

In order to investigate the mechanisms that respond to increased energy demand for reproduction, I investigated seasonal gene expression patterns of free-living males and females. In addition, I investigated the effect of 17β-estradiol administration on the gene expression because estrogen has stimulatory role in oogenesis (Hara et al., 2016; Price, 2017; Wallace and Bergink, 1974). The gene expression patterns suggested reduction in hepatic glucose production and β-oxidation in females during the reproductive season. Instead, increase in VLDL synthesis was suggested. Therefore, females seem to concentrate energy investment on oogenesis at the cost of maternal activity. Since female Japanese geckos need to complete egg-laying within a short period (Ikeuchi, 2004), such concentrated effort on reproduction is likely to increase fitness in temperate environment. It seems that unique transcriptional regulation of metabolic regulators is the key to such trade-off regulation.

Unexpectedly, I found that the two glucose-6-phosphatase catalytic subunit 1 (*g6pc1*) duplicates had possibly undergone functional divergence. G6PC1 catalyzes the final step of hepatic glucose production and therefore critically important for systemic glucose homeostasis (Chou et al., 2010; Rui, 2014). While mammals and birds generally possess single *g6pc1* (Marandel et al., 2017), the Japanese gecko possesses two duplicates. In this research, I also found that the expression of one duplicate was downregulated by administration of 17β-estradiol, whereas that of the other was not affected. I further investigated evolutionary background of duplication and found that all ectothermic amniotes possessed 2 or 3 duplicates, whereas endotherms retained only single copy. Therefore, I speculated that the *g6pc1* duplicates play physiological roles advantageous only for ectotherms.

Overall, the results obtained in this research strongly suggest that unique coding sequences of insulin, together with transcriptional regulation of key metabolic regulators, increase fitness of the

Japanese gecko to temperate environment. Although further investigation in other species is necessary to understand metabolic regulation of squamates comprehensively, it seems that understanding about metabolic regulation at the molecular level is indispensable for the better understanding about evolutionary background of squamates' diversity and prosperity.

List of abbreviations

36B4: ribosomal protein lateral stalk subunit P0

aa: amino acid(s)

AARSD1: alanyl-tRNA synthetase domain containing 1

ACC1: acetyl-CoA carboxylase 1

AOC3: amine oxidase, copper containing 3

apoB: apolipoprotein B

ATGL: adipose triglyceride lipase

BEB: Bayes Empirical Bayes

bp: base pair(s)

cDNA: DNA complementary to RNA

CIDEB: cell death-inducing DFF45-like effector protein family B

CmC: Clade model C

DAG: diacylglycerol

dCTP: cytidine 5'-triphosphate

DGAT1: diacylglycerol acyltransferase 1

DGAT2: diacylglycerol acyltransferase 2

dN: number of nonsynonymous substitutions

DNL: *de novo* lipogenesis.

dNTP: deoxyribonucleotide 5’-triphosphate

dS: number of synonymous substitutions

E_2: 17β-estradiol

EF1A1: eukaryotic translation elongation factor 1 α 1

ESR1: estrogen receptor 1

F1,6P: fructose-1,6-bisphosphate

F6P: fluctose-6-phosphate

FA: fatty acid(s)

FABP1: fatty acid binding protein 1 (L-FABP gene)

FASN: fatty acid synthase

FATP5: fatty acid transport protein 5

G1P: glucose-1-phosphate

G6P: glucose-6-phosphate

G6Pase: glucose-6-phosphatase

G6PC1: glucose-6-phosphatase catalytic subunit 1

G6PC2: glucose-6-phosphatase catalytic subunit 2

G6PC3: glucose-6-phosphatase catalytic subunit 3

GSD-Iα: glycogen storage disease type Iα

G6PT: glucose-6-phosphate translocase

GCK: glucokinase

GLUT2: glucose transporter 2

HGP: hepatic glucose production

HSL: hormone sensitive lipase

IGF1: insulin-like growth factor 1

IGF2: insulin-like growth factor 2

IIS: insulin and insulin-like signaling

ins: insulin gene

L-FABP: liver fatty acid binding protein

LPL: lipoprotein lipase

LRT: log likelihood ratio test

MAG: monoacylglycerol

ML: Maximum Likelihood

mRNA: messenger RNA

MTP: microsomal triglyceride transfer protein

N-J: Neighbor-Joining

ω: non-synonymous-to-synonymous substitution ratio

ORF: open reading frame

PC-1/3: prohormone convertase 1/3

PC-2: prohormone convertase 2

PCK1: phosphoenolpyruvate carboxykinase 1

PCK2: phosphoenolpyruvate carboxykinase 2

PCR: polymerase chain reaction

PNPLA2: patatin like phospholipase domain containing 2 (ATGL gene)

PPARG: peroxisome proliferator activated receptor gamma

PSS: positively selected site(s)

PYGL: glycogen phosphorylase, liver form

RACE: rapid amplification of cDNA ends

RPL8: ribosomal protein L8

RT: reverse transcription

SDHA: succinate dehydrogenase complex flavoprotein subunit A

SEM: standard error of mean

SVL: snout-to-vent length

TAG: triacylglycerol

VLDL: very-low-density lipoprotein

VTG2: vitellogenin-2

Table of contents

General introduction

Pros and cons of the ectothermic strategy of reptiles

During Late Carboniferous period, one group of vertebrate lineages, called Amniota (or amniotes), emerged. Unlike fish and amphibians, embryo of this group develops within water-containing space called amniotic sac. Advantage of this systems is enormous, because it liberates animals from aquatic stage, which exposes non-amniotic groups to the risk of mortality from droughts. Exploiting this advantage, amniotes have been dominating terrestrial environment to date (Benton, 2014).

However, there is one more obstacle for amniotes to overcome on land: temperature fluctuation. Because of a low specific heat of air, variation of temperature is larger in terrestrial environment than in aquatic environment. As a result, terrestrial animals must deal with a wider range of temperature fluctuation (Rastogi, 1971).

Among amniotes, two different strategies for thermoregulation exist: endothermy and ectothermy. Mammals and birds have independently acquired endothermy to cope with temperature fluctuation. By generating body heat constantly, they maintain their body temperature within narrow ranges. As a result, their activity is maintained regardless of ambient temperature. Owing to this regulation, endothermic amniotes are distributed all over the world, from tropical to polar zones. However, endothermy is an energetically costly strategy. It has been reported that endotherms invest about 98% of energy intake on heat production. As a result, only small portion of energy intake can be devoted to other activities, such as body maintenance, growth and reproduction (Pough, 2016).

On the other hand, polyphyletic group of ectothermic amniotes comprising four orders (Squamata, Crocodilia, Testudines and Sphenodontia), called reptiles, do not produce body heat (except a few exceptions) and their body temperature is exposed to ambient temperature fluctuation. Therefore, their activity is largely dominated by climate fluctuations. However, ectothermy is energetically less costly than endotherms. It has been reported that metabolic rate of reptiles is 7- to 10-fold lower than endotherms of the same body size (Pough, 2016). In addition, absence of energy expenditure for heat

production enables reptiles to invest more portion of energy intake on growth and reproduction (about 40~80% have been reported). As a result, reptiles can survive and reproduce in ecological niches with poor food availability (Pough, 2016).

Prosperity of Squamata through the acquisition of novel life-histrory traits

Surprisingly, restriction of activity by ambient climate seems to have little impact on geographical distribution of reptiles. This is especially true to the order Squamata, a group comprising lizards and snakes. Currently, more than 9,000 squamate species are distributed worldwide, from subarctic to tropical zones (Herczeg et al., 2003; Utez, 2015; Vitt and Caldwell, 2013). This species diversity is superior to mammals (approximately 5,500 species), and is comparable to birds (approximately 10,000 species) (Baillie et al., 2010). This contrasts with the other orders of reptilian groups, from which less than 400 species have been reported (Utez, 2015). Also, squamates are ecologically diverse group. There are not only terrestrial “ground-crawlers”, but also arboreal, fossorial and even aquatic forms.

Underlying this success seems to be their capability to acquire novel life-history traits in accordance with environmental changes they encounter throughout their lives. At the interspecific level, the most significant example is the divergence of *Anolis* lizards inhabiting Caribbean islands. It has been reported that about 150 species have diverged from this genus as a result of selection of habitat use (Marnocha et al., 2011). Classic example of intraspecific difference is the selection of ecotypes between two populations of the western garter snake (*Thamnophis elegans*). The population inhabiting mountain meadows grows slowly, produces less offspring but lives longer. In contrast, the population inhabiting lakeshore grows faster, produces more offspring but shorter-lived. It has been suggested that in lakeshore environment, where more food is available but predation risk is higher, producing as many offspring as possible within short period by exploiting food-rich environment increases fitness of the population (Pough, 2016). Another example is reported from annual

reproductive activity of the common house gecko (*Hemidactylus frenatus*). Although this species is native to tropical environment of south-east Asia, it dispersed recently to subtropical zone presumably by human transport (Moritz et al., 1993). Interestingly, while tropical population produces offspring throughout the year, subtropical population ceases egg-laying in winter (Ota, 1994). Ikeuchi (Ikeuchi, 2004) suggested that this intraspecific difference is the result of adaptation to subtropical climate, because lower winter temperature in subtropical zone reduces survivorship of offspring.

Possible contribution of metabolic flexibility to environmental adaptation in squamates

Since every aspect of life needs to be supplied with energy (Andrade, 2016), acquisition of novel life-history traits must be accompanied by adjustment of metabolic regulation so as to provide new traits with sufficient energy (Figure I-1). Therefore, squamates as a whole seem to have great flexibility in regulation of energy metabolism. Understanding about such metabolic flexibility is likely to provide an important insight into the background of squamates' success. Unfortunately, too little is understood about energy metabolism of squamates. To date, patterns of energy utilization has been inferred from parameters such as body temperature, oxygen consumption, glycogen and lipid content, and enzymatic activities (Derickson, 1976; Pough, 2016; Vitt and Caldwell, 2013). However, little attention has been paid to the molecular mechanisms that regulate these patterns. This is because squamates have been unattractive to experimental biologists as they are of little importance for clinical, pharmacological and agricultural use. In addition, the difficulty in keeping and breeding seems to have made this group even less attractive to researchers (although recent booming of so-called 'herpetoculture', or habit of keeping reptiles as pets, largely solved this problem). As a result, genome data of squamate species has been largely lacking compared with other organisms. Considering the fact that metabolic pathways are regulated by interactions among multiple regulatory molecules, lack of knowledge at the molecular level is troublesome for the comprehensive understanding about metabolic regulation.

Aim of this thesis

In this research, I aimed at establishing basic knowledge about molecular mechanisms regulating squamates' energy metabolism for the inference of the mechanisms that bring flexibility to metabolic regulation of this group.

Current understanding about the relationship between molecular evolution and acquisition of novel phenotypes is that changes in coding sequences and transcriptional regulation of genes are the determining factors for the acquisition of novel phenotypes (Little and Seebacher, 2016) . Therefore, as viewpoints for investigation, I focused on 1) the variation of amino acid sequences of insulin among squamates and 2) transcriptional regulation of key regulators for glucose and triacylglycerol (TAG) metabolism in the Japanese gecko.

Chapter 1. Molecular characterization of insulin from squamate reptiles and analyses on molecular evolution

In chapter 1, I focused on the molecular evolution of squamate insulin as a possible mechanism that contributed to environmental adaptation. Vertebrate insulin is generally released from pancreatic β-cells in response to food intake, and promotes anabolism of substrates such as carbohydrates, lipids and amino acids (Norris and Carr, 2013; Rhodes et al., 2005). This effect is so prominent for metabolic regulation that genetic variation of *ins* (insulin gene) results in the onset of diabetes in humans (Nishi and Nanjo, 2011). As a result, amino acid sequence of insulin is generally conserved among amniotes (Conlon, 2001).

Surprisingly, cDNA coding for insulin isolated from the Japanese gecko (*Gekko japonicus*) in our lab previously showed significantly lower sequence homology to those of other amniotes. Recent works on molecular evolution suggest involvement of insulin and insulin-like signaling (IIS) pathway

in the acquisition of unique life-history traits of squamates (McGaugh et al., 2015; Sparkman et al., 2012). Although insulin participates in IIS pathway, the significance of molecular evolution of this hormone seems to have been overlooked, because the existence of such aberrant insulin was unnoticed.

In this research, I aimed at evaluating the evolutionary significance of insulin for squamates by elucidating the evolutionary process that gave rise to the aberrant sequence of the Japanese gecko insulin by the following approaches. First, to find the lineage from which substitutions were introduced to insulin, I isolated cDNA coding for insulin from multiple lineages of squamates and compared deduced amino acid sequences. Second, to examine changes in the selection pressure imposed on the *ins* among squamate lineages, nonsynonymous-to-synonymous substitution ratio (dN/dS ratio) was calculated based on the nucleotide sequences isolated.

Chapter 2. Analyses on the expressional change of key regulators for hepatic glucose and lipid metabolism in the Japanese gecko

In chapter 2, I tried to elucidate the molecular mechanisms that regulate energy metabolism in accordance with life-history traits that are important for fitness of a squamate species, the Japanese gecko (*Gekko japonicus*).

Owing to recent advances in genome sequencing technologies, genome or transcriptome data of several squamate species are currently available (Castoe et al., 2013; Liu et al., 2015; Xiong et al., 2016). The Japanese gecko is one of such species. Interestingly, geographical distribution suggests that the Japanese gecko has undergone environmental adaption to temperate climate. Among the genus *Gekko*, this species and several close-relatives are grouped together to form a clade called "*Gekko japonicus* group" (Rösler et al., 2011). In contrast to the tropical habitat of most *Gekko* species, they are distributed in subtropical and/or temperate zones. Therefore, they are likely to have evolved from a tropical ancestor and subsequently adapted to climate in higher latitude.

In subtropical and temperate zones, seasonal fluctuations of temperature and food availability dominate annual activities of squamates (Lillywhite, 2016). For example, temperate squamates are generally dormant during winter. During this period, they are anorexic and completely dependent on internal energy storage (Pough, 2016). In addition, they are mostly seasonal breeders (Pough, 2016). Since reproduction is an energetically costly process, energy expenditure increases in the reproductive season. For example, daily energy expenditure of female green anoles (*Anolis carolinensis*) almost doubles in the reproductive season from that in the non-reproductive season (Orrell et al., 2004). As a result of these seasonal phenomena, temperate squamates undergo large fluctuation of metabolic status annually (Derickson, 1976; Di Maggio and Dessauer, 1963; Zani et al., 2012).

Winter dormancy and seasonal reproduction is reported from the Japanese gecko, too (Hisai, 1997; Ikeuchi, 2004). In contrast, tropical species of *Gekko* (*G. gecko* and *G. smithii*) are active and reproduce throughout the year (Goldberg, 2009). Therefore, the Japanese gecko seems to have acquired regulatory mechanisms that can deal with seasonal fluctuation of metabolic status. Investigation into such mechanisms may offer not only knowledge about the regulatory mechanisms themselves, but also implications for the relationship between metabolic regulation and environmental adaptation.

For the elucidation of such regulatory mechanisms, I focused on the gene expression changes of key regulators for major glucose and triacylglycerol (TAG) metabolic pathways in the liver. This is because glucose and TAG are the two most important energy substrates for animals. Glucose is indispensable for animals because neuronal cells depend almost exclusively on it as an energy source. On the other hand, TAG is best suited for energy reserve as it can store energy in the most concentrated form. While 9 kcal of energy is derived from 1 g of TAG, less than half this caloric value can be derived from the same amount of carbohydrates or amino acids (Berg et al., 2002). This feature makes TAG an ideal energy source for embryo development of oviparous species, because limited space is

available for substrate deposition within eggs. In addition, seasonal cycling of glycogen and lipid content implies the importance of these substrates for temperate squmates (Derickson, 1976; Dessauer, 1953; Di Maggio and Dessauer, 1963).

In this research, I focused on the liver as the central metabolic organ. Although multiple organs such as adipose tissue, kidney and skeletal muscle are involved in metabolism of glucose and TAG, the liver plays central roles in synthesis, storage and mobilization of these substrates. In addition, lipogenesis takes place in the liver, meaning that glucose and TAG metabolism is integrated in this organ (Rui, 2014). Furthermore, liver plays prominent roles in oogenesis. In response to estrogen stimulation, this organ synthesizes vitellogenin, a yolk protein precursor, and very-low-density lipoprotein (VLDL), TAG-rich particles coated with lipid layer (Deeley et al., 1975; Hara et al., 2016; Mullinix et al., 1976; Price, 2017; Wallace and Bergink, 1974; Walzem et al., 1999). These are released into circulation and are taken up by growing follicles to be catabolized to deposit yolk protein and TAG in the eggs (Price, 2017). Adding to these, liver seems to be an important site for energy deposition in the Japanese gecko. Temperate squamates generally possess abdominal adipose tissue to store lipids as an energy source during winter dormancy and for reproduction (Derickson, 1976). However, the Japanese gecko lacks this site. This feature is similar to tropical species that do not enter dormancy and reproduce year-round (Derickson, 1976), implying tropical origin of this species. Instead, the Japanese gecko seems to rely on lipid storage in the liver. Ji and Wang (Ji and Wang, 1990) reported significant reduction in hepatic lipid content during winter, implying mobilization of lipid as an energy source from this site.

The following are the details about the metabolic regulators and the pathways investigated in this research. Schematic representations of pathways are shown in Figures I-2 and I-3.

Regulation of glucose metabolism in the liver (Figure I-2)

1. Uptake and intracellular production of glucose by the liver

In the postprandial state, glucose absorbed in the small intestine flows into the liver through portal vein. In the liver, glucose enters hepatocytes mainly through glucose transporter 2 (GLUT2), a facilitative glucose transporter, and subsequently phosphorylated by glucokinase (GCK). This reaction lowers intracellular concentration of glucose and enables more glucose to flow in. Glucose-6-phosphate (G6P), the product of this reaction, is metabolized to generate ATP through glycolysis and TCA cycle. Excess amount of G6P is utilized for glycogen synthesis and stored as an energy reserve (Rui, 2014).

Under food deprivation, the liver produces glucose and releases it into circulation. This process, called hepatic glucose production (HGP), is of critical importance for the survival of animals, because the brain is almost exclusively dependent on glucose as an energy source. At first, hepatic glycogen storage is broken down by glycogen phosphorylase (PYGL) to release G6P. This process, called glycogenolysis, is the primary source of glucose under short-term food deprivation (up to 12 hours in human). When food deprivation prolongs, gluconeogenesis takes over glycogenolysis as a source of G6P. This process generates G6P from non-carbohydrate sources such as amino acid, glycerol, lactate and pyruvate (Berg et al., 2002). In animals that depend on carbohydrate-rich diet (such as human and rodents), gluconeogenesis is activated only when food availability is low. However, carnivorous animals depend on gluconeogenesis as the primary source of glucose because their diet contains little carbohydrates. Therefore, activities of gluconeogenic enzymes are constitutively high in carnivores (S ørensen et al., 1995; Tanaka et al., 2005).

2. Gluconeogenesis and phosphoenolpyruvate carboxykinase (PCK)

Gluconeogenesis is a multi-step process which involves various enzymes. Among such enzymes, phosphoenolpyruvate carboxykinase (PEPCK or PCK) catalyzes the first irreversible step of

gluconeogenesis from amino acids, in which oxaloacetate is converted to phosphoenolpyruvate (Hanson and Garber, 1972; Rui, 2014; Xiong et al., 2011). There are two isoforms of PCK, one being cytosolic form (PEPCK-C or PCK1) and the other being mitochondrial form (PEPCK-M or PCK2). Tissue distribution of these isoforms differs among species. Mammals generally have equal amount of PCK1 and PCK2 in their liver. However, PCK1 is predominant isoform in rodents' liver. On the other hand, PCK2 is predominant in the liver of rabbit and avian species (Croniger et al., 2002). Although these two isoforms catalyze essentially the same reaction, physiological significance of PCK has been discussed based exclusively on studies on PCK1. However, recent studies revealed involvement of PCK2 in gluconeogenesis as a significant source of phosphoenolpyruvate (Stark and Kibbey, 2014). In this research, the expression of *pck1* was not detected in the liver of Japanese gecko (Figure I-4). Therefore, PCK2 seems to be the predominant isoform in the liver of this species.

3. Regulation of glucose release from hepatocytes by glucose-6-phosphatase catalytic subunit 1 (G6PC1)

G6P, the product of glycogenolysis and gluconeogenesis, cannot be transported out of hepatocyte because it cannot pass through glucose transporters (Rui, 2014). Therefore, dephosphorylation of G6P by glucose-6-phosphatase (G6Pase) is critically important for systemic glucose regulation. G6Pase is composed of glucose-6-phosphate translocase (G6PT) and glucose-6-phophatase catalytic subunit (G6PC), both of which resides on ER membrane. The former transports G6P into ER lumen, and the latter, whose catalytic domain is faced inside ER lumen, dephosphorylates G6P (Chou et al., 2010). There are three isoforms of G6PC, namely G6PC1 (more often called G6PC), G6PC2 and G6PC3. Among these, G6PC1 is expressed exclusively in the liver, kidney and small intestine and contributes to systemic glucose homeostasis (Hutton and O'Brien, 2009). Deficiencies in G6PC1 results in glycogen storage disease type Iα (GSD-Iα) (Chou and Mansfield, 2008), which is characterized by hypoglycemia, hyperlipidemia and hyperuricemia (Bali et al., 2006), all of which results from inability

to mobilize glucose from gluconeogenic organs.

Although mammalian G6PC1 is encoded by single gene, that of other vertebrates such as teleost fish, amphibians and reptiles are encoded by 2 to 3 duplicates (Marandel et al., 2017). This is also true to the Japanese gecko, which has two *g6pc1* genes. In this research, I designated these as *g6pc1-1* (NCBI accession: LOC107119344) and *g6pc1-2* (NCBI accession: LOC107119345), respectively.

Regulation of TAG metabolism (Figure I-3)

1. Synthesis of triacylglycerol (TAG) in the liver

Triacylglycerol (TAG) is composed of three fatty acids esterified with a glycerol. In the liver, two pathways synthesize TAG. One is *de novo* synthesis, in which fatty acids derived from *de novo* lipogenesis (DNL: synthesis of fatty acids from intermediates of carbohydrate metabolism) are sequentially esterified with nascent glycerol-3-phosphate. The other is re-esterification of diacylglycerol (DAG) and monoacylglycerol (MAG), products of partial hydrolysis of TAG, with exogenous fatty acids. The former results in net increase of TAG storage in the liver, while the latter replenishes TAG storage following its partial mobilization. Although final step of both pathways, esterification of DAG with a fatty acid, is essentially the same reaction, the enzyme involved is specific to each pathway. Diacylglycerol acyltransferase 2 (DGAT2) is involved in *de novo* synthesis, as it selectively utilizes fatty acids derived from DNL as substrates. This enzyme is mainly responsible for net increase of TAG in the liver. In contrast, DGAT1 favors exogenous fatty acids taken up by transporters such as FATP5 as substrates and is mainly involved in replenishment of TAG. Not only such substrate selectivity but also intracellular distribution and membrane topology differentiates physiological function of these isozymes. DGAT2 is located on ER membrane and cytosolic droplet with its catalytic site always exposed to cytosol. As a result, DGAT2 contributes exclusively to cytosolic TAG storage. In contrast, DGAT1 is located on ER membrane with dual topology. In other

words, each DGAT1 molecule can expose its catalytic site either to cytosol or to ER lumen. Therefore, DGAT1 is involved in development of TAG storage both within cytosol and within ER lumen (Mashek et al., 2015; Zammit, 2013).

2. Mobilization of TAG from the liver

There are two pathways to mobilize hepatic TAG storage; β-oxidation and synthesis of very-low-density lipoprotein (VLDL).

2-1. β-oxidation

β-oxidation is a process that generates ATP to fuel metabolism in the liver utilizing fatty acids as substrates. Sequential hydrolysis of TAG by adipose triglyceride lipase (ATGL), hormone sensitive lipase (HSL), and monoglyceride lipase (MGL) provides this process with fatty acids. It has been reported that hydrolysis of TAG mediated by ATGL selectively partitions fatty acids to β-oxidation pathways (Ong et al., 2011). The fatty acids released from TAG are transported into mitochondria, where β-oxidation takes place. Liver fatty acid binding protein (L-FABP), which has facilitative effect on β-oxidation (Atshaves et al., 2010), has been suggested to mediate this transport, although its contribution is under debate (Ong et al., 2014). β-oxidation is also important for systemic metabolism under food deprivation, because ketone bodies, byproduct of this process, can substitute glucose as an energy substrate for brain. In human, more than half the energy consumed by the brain is fueled by ketone bodies after several weeks of food deprivation (Cahill, 2006).

2-2. Synthesis of VLDL

VLDL is a class of lipoprotein that abundantly contains TAG in its core. It is synthesized in the liver and secreted into circulation to provide other organs such as skeletal muscle, adipose tissue and growing follicles with TAG. Uptake of TAG content in the VLDL by target organs generally requires lipoprotein lipase (LPL) that resides on vascular endothelial cells. This lipase captures circulating VLDL and subsequently hydrolyzes TAG content. The product of this reaction, fatty acids and

monoacylglycerol, are subsequently taken up by the target organs (Wang and Eckel, 2009). However, ovarian follicles of birds do not need LPL, as it incorporates VLDL directly through basal lamina of the follicle granulosa (Price, 2017).

Synthesis of VLDL commences with formation of a precursor particle in the ER lumen by apolipoprotein B (apoB), a component of VLDL which resides on its surface. ApoB is co-translationally inserted into ER lumen. In this process, it cooperates with microsomal triglyceride transfer protein (MTP) to transfer TAG from within inter-leaflet site (between lipid bilayer) of ER membrane to ER lumen (Shelness and Sellers, 2001). This process produces VLDL precursor, which is small and contains less TAG. Maturation of this precursor occurs through its fusion with apoB-free TAG particles that resides within ER lumen. Cell death-inducing DFF45-like effector protein family B (CIDEB) is involved in this process and subsequent budding from ER (Mashek, 2013; Tiwari et al., 2013).

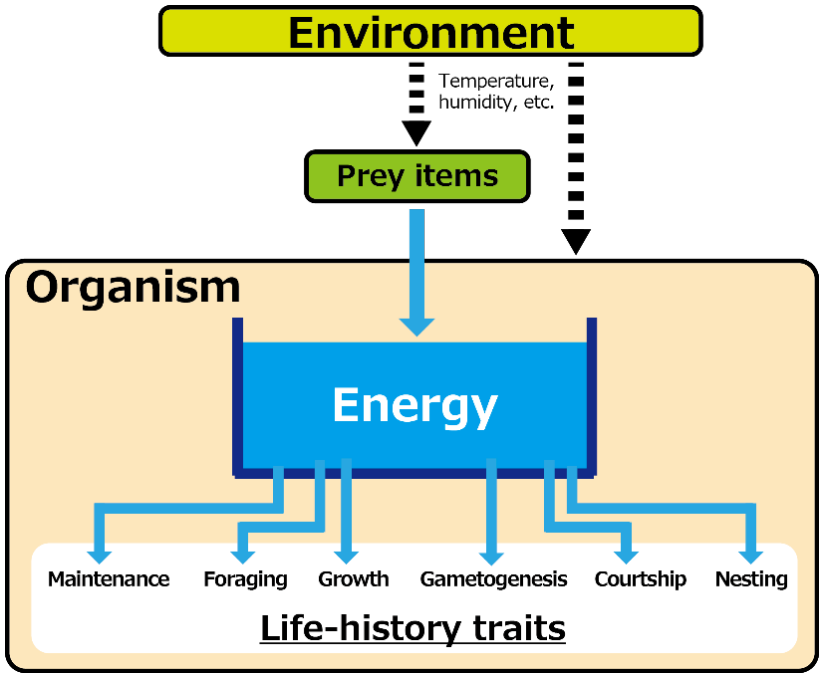


Figure I-1. The importance of energy metabolism for environmental adaptation

The amount of energy an organism obtains from prey items, together with its activity, is affected by environmental fluctuations. Therefore, optimization of energy partitioning among various life-history traits is the key to survival and successful reproduction in certain environment. This is especially true to ectotherms, whose body temperature and activity is largely dependent on ambient temperature. Among ectothermic amniotes, or reptiles, the order Squamata seems to have great flexibility in such energy partitioning.

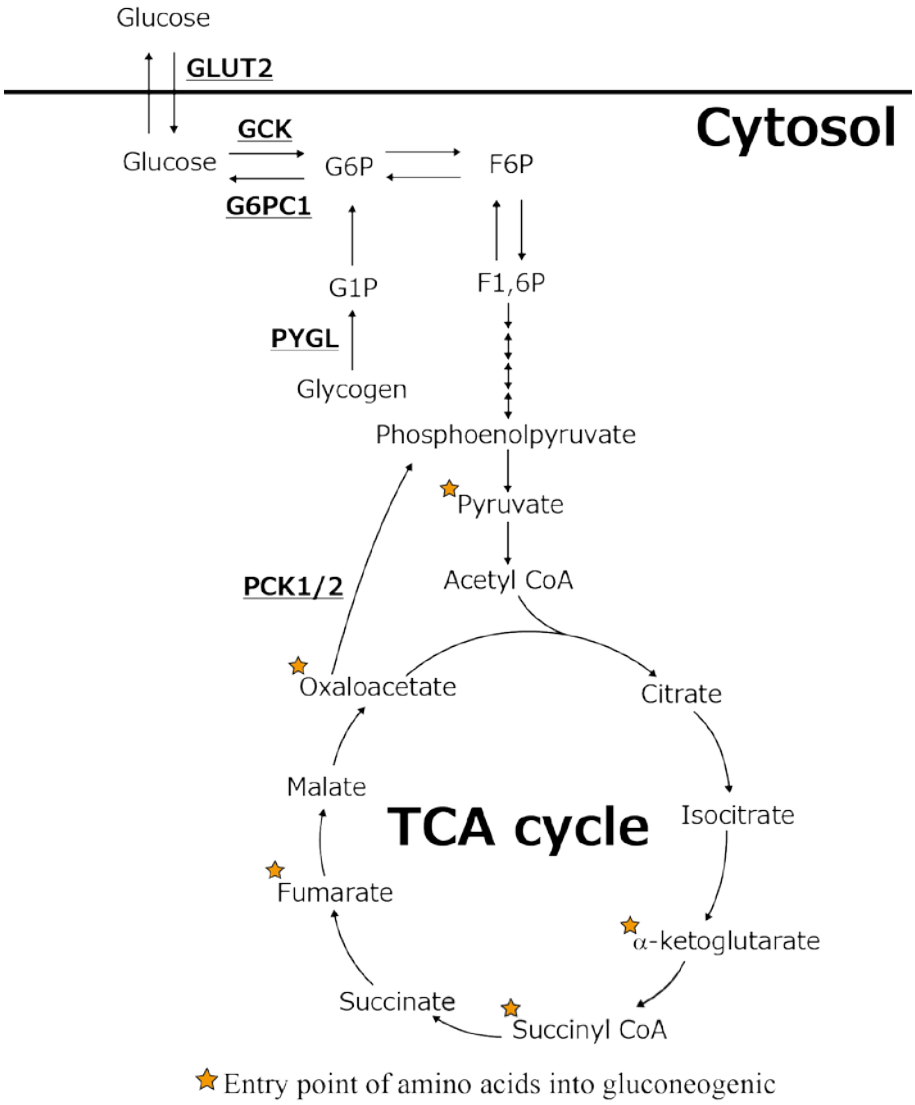


Figure I-2. Major pathways for glucose metabolism in the liver

Among multiple regulators involved, those investigated in this research are shown. Stars in the figure represent entry points for amino acids into the gluconeogenic pathway. F6P: fluctose-6-phosphate. F1,6P: fructose-1,6-bisphosphate. G1P: glucose-1-phosphate: G6P: glucose-6-phosphate. G6PC1: glucose-6-phosphatase catalytic subunit 1 (duplicates of the Japanese gecko are designated as G6PC1-1 and G6PC1-2). GCK: glucokinase. GLUT2: glucose transporter 2. PCK: phosphoenolpyruvate carboxykinase. PYGL: glycogen phosphorylase, liver form.

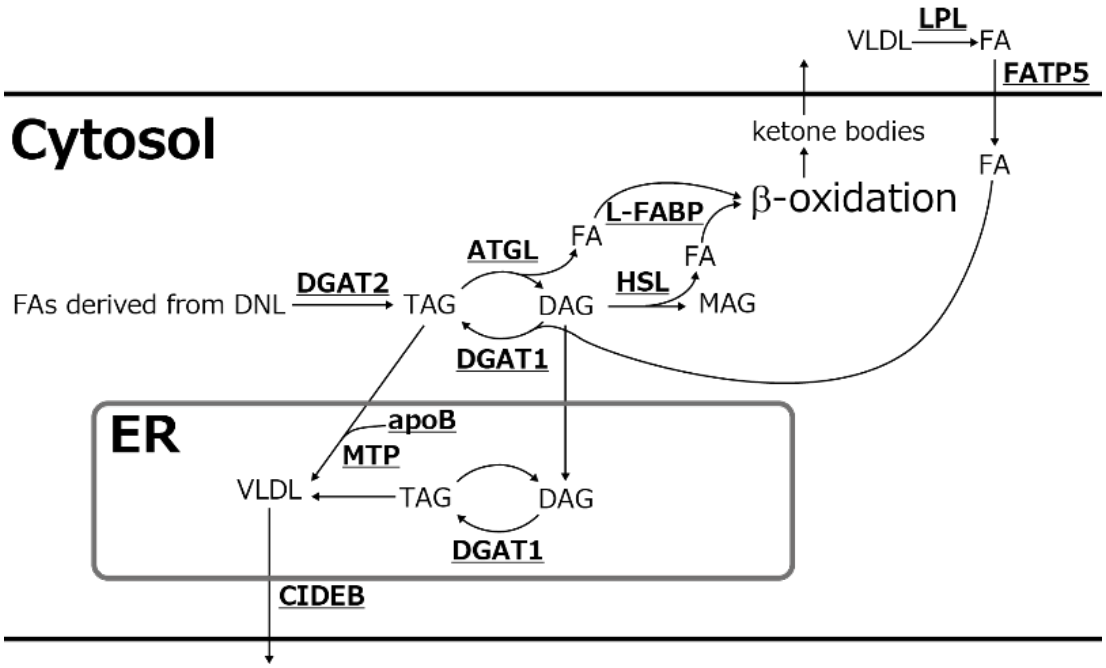


Figure I-3. Major pathways for TAG metabolism in the liver

Among multiple regulators involved, those investigated in this research are shown. apoB: apolipoprotein B. ATGL: adipose triglyceride lipase. CIDEB: Cell death-inducing DFF45-like effector protein family B. DAG: diacylglycerol. DGAT1: diacylglycerol acyltransferase 1. DGAT2: diacylglycerol acyltransferase 2. DNL: *de novo* lipogenesis. FA: fatty acid. FATP5: fatty acid transport protein 5. HSL: hormone sensitive lipase. L-FABP: Liver fatty acid binding protein. LPL: lipoprotein lipase. MAG: monoacylglycerol. MTP: microsomal triglyceride transfer protein. TAG: triacylglycerol. VLDL: very-low-density lipoprotein.

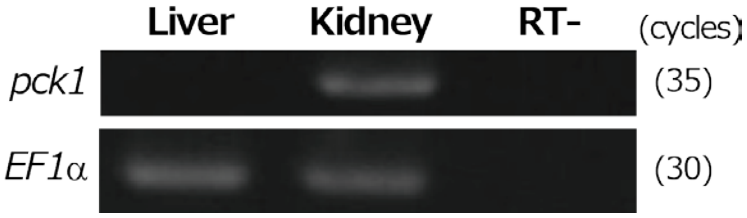


Figure I-4. Expression of *pck1* in the liver and kidney of the Japanese gecko

The expression of *pck1* in the liver and kidney was investigated by PCR. As shown in this figure, its expression was generally absent in the liver, whereas it was detected in the kidney of the Japanese gecko.

Chapter 1

Molecular characterization of insulin from squamate reptiles and analyses on molecular evolution

Introduction

Vertebrate insulin is generally released from pancreatic islets (β-cells) in response to food intake, and promotes the anabolism of substrates such as carbohydrates, lipids, and amino acids. Primary transcript of *ins* (insulin gene), preproinsulin, contains four regions: signal peptide, B-chain, C-peptide, and A-chain. Signal peptide is removed upon translocation of the product into endoplasmic reticulum (ER). The remaining polypeptide, comprising B-chain, C-peptide, and A-chain, is called proinsulin. In ER, two inter-chain disulfide bonds are formed between B7-A7 and B19-A20 to link B- and A-chains. An intra-chain bond is also formed between A6-A11. Proinsulin is transported into secretory granule, where cleavage of C-peptide takes place by prohormone convertase (PC) -1/3 and PC-2. Through these processes, mature insulin, comprising B-chain and A-chain connected by two disulfide bonds, is formed (Norris and Carr, 2013; Rhodes et al., 2005).

For the elucidation of molecular mechanisms regulating energy metabolism of squamates, insulin is an interesting topic for examination because what role it plays in the physiology of squamates currently remains unclear. Insulin has been generally regarded as the central anabolic hormone for vertebrates. The most significant role is its hypoglycemic effect, which is attributable only to this hormone (Norris and Carr, 2013; Sherwood et al., 2013). In contrast, several studies have suggested that the role of insulin in squamate species is unique among reptiles, with only a negligible effect on blood glucose regulation (Dessauer, 1970; Matty, 1966), though there is some controversy (Sidorkiewicz and Skoczylas, 1974).

Previous researches in our lab have isolated cDNAs coding for preproinsulin from two Gekkota species, the leopard gecko (*Eublepharis macularius*) and Japanese gecko (*Gekko japonicus*) in order to establish the molecular basis for further studies. When the deduced amino acid sequences were compared, the leopard gecko insulin sequence resembled those of other amniotes. However, the Japanese gecko insulin showed the accumulation of multiple amino acid substitutions within B-chain

and A-chain. Moreover, some of the substitutions were located at sites that were widely conserved among other vertebrates, from teleosts to mammals. Such variation of insulin sequences had never been reported from reptilian species.

It has been established that insulin shares receptors and intracellular signaling pathways with its paralogous genes, insulin-like growth factor (IGF) -1 and 2. This signaling network overall is called insulin and insulin-like signaling (IIS) network (Schwartz and Bronikowski, 2016). Recent studies on the evolution of this network have revealed possible involvement of this signaling network in the acquisition of life-history traits unique to squamates. Analyses on the molecular evolution of *igf1* among reptilian lineages (reptiles and aves) suggested rapid evolution of squamates' *igf1* driven by positive selection on sites comprising functional C-domain. Analyses by McGaugh et al. on molecular evolution of overall IIS network suggested multiple components of this network to have been under positive selection (McGaugh et al., 2015).

Unfortunately, little attention has been paid to the molecular evolution of squamate *ins* in the previous works. It seems that they have overlooked significance of insulin, because they investigated limited number of lineages among squamates. In this chapter, I isolated cDNAs coding for insulin from species that represent various groups of Squamata in order to fill the gap in the knowledge about squamates' insulin and to examine the origin of diversity in the insulin sequences. The deduced amino acid sequences of insulin isolated in this study (27 species in total), together with those of two species obtained from the GenBank Database (the green anole and king cobra) were compared with representative species of amniotes. The results suggested that amino acid sequences of squamate insulin were more divergent than those of the other amniotes. Further investigations into each substitution and evolutionary process suggested that squamate *ins* had undergone a unique evolutionary process among vertebrates, which may have been advantageous for their survival.

Materials and Methods

1. Animals

For the isolation of cDNAs coding for insulin, living samples of 27 squamate species (listed in Table 1-1) were obtained. Japanese geckos, Japanese five-lined skinks (*Plestiodon finitimus*), and Japanese grass lizards (*Takydromus tachydromoides*) were collected from a field in the Tokyo metropolitan area, Japan. Leopard geckos were bred over several generations in the laboratory. Other species of squamates were purchased from local Pet shops. All animals used in this research were treated according to the guidelines of the Bioscience Committee at the University of Tokyo (Approval no. P12-02, P16-04). Animals were anesthetized by an injection of sodium pentobarbital at a dose of 25 μg/g body weight, and were sacrificed by decapitation and exsanguination. Tissues were collected and frozen in nitrogen liquid. Frozen tissues were stored at -80 °C until used for RNA extraction.

2. RNA extraction and cDNA synthesis

Total RNA was extracted from the pancreas using ISOGEN (NIPPON GENE, Tokyo). For the construction of cDNA templates, 1~3μg of total RNA was used. cDNA templates for 3'-RACE were constructed as follows; denatured total RNA was reverse-transcribed using an oligo (dT)-adaptor primer and M-MLV Reverse Transcriptase (Promega, Madison, WI). The reaction volume contained oligo (dT)-adaptor primer at 5pM, each dNTP at 2mM, 200 units of M-MLV Reverse Transcriptase and 1×M-MLV buffer (Promega). The volume was brought up to 20 μl with Nuclease-Free Water (Quiagen, Tokyo, Japan). The reaction was performed under 42℃, 90 min. and then 70℃, 15min. In some cases, ThermoScript Reverse Transcriptase (Invitrogen, Tokyo) was used, instead. The reaction volume contained oligo (dT)-adaptor primer at 5pM, each dNTP at 2mM, 15units of ThermoScript Reverse Transcriptase and 1×cDNA Synthesis Buffer (Invitrogen). The volume was brought up to 20μl with Nuclease-Free Water (Quiagen). The reaction condition for this enzyme was 42℃, 30min., 55℃,

45min., and then 70℃, 15min. cDNA templates for 5'-RACE were constructed according to the following methods. Messenger RNA was purified from total RNA using Dynabeads Oligo $(dT)_{25}$ (Invitrogen) and then reverse transcribed using SuperScript III Reverse Transcriptase (Invitrogen). For this reverse transcription, oligo $(dT)_{25}$ on the surface of the beads worked as primers. The reaction volume contained DTT at 5mM, each dNTP at 500μM, 200units of SuperScript III Reverse Transcriptase and 1×First-Strand Buffer (Invitrogen). The reaction condition was room temperature,10~15min., 42℃, 30min., 55℃, 45min., and then 70℃, 15min. Synthesized cDNA was added with poly (C) at its 5'-terminus using Terminal Deoxynucleotidyl Transferase (TaKaRa, Shiga) and dCTP. The reaction volume contained dCTP at 200 μM, BSA (TaKaRa) at 0.01%, 14units of Terminal deoxynucleotidyl Transferase and 1×Terminal deoxynucleotidyl Transferase buffer (TaKaRa). The volume was brought up to 20 μl with RNase-Free Water (Quiagen). The reaction condition was 37℃, 10min. and then 70℃, 10 min. A PCR reaction was performed using TaKaRa EX Taq (TaKaRa) and an adaptor primer with poly (G) region at its tail, which was designed to bind to the poly (C) region of cDNA. The reaction volume contained the adaptor primer at 1pM, each dNTP at 250 μM, 0.25unit of TaKaRa EX Taq (TaKaRa) and 1×EX Taq buffer (TaKaRa). The volume was brought up to 20 μl with Nuclease-Free Water (Quiagen). The PCR conditions were as follows: 94°C for 2 min., 3 cycles of 94°C for 30sec., 73℃ for 30 sec., and 72°C for 10min., and then 72°C for 7 min.

3. Molecular cloning of insulin cDNA by RACE and RT-PCR

5'- or 3'- RACE, and RT-PCR was performed using the primers shown in Table 1-2. The PCR reaction was performed in a reaction volume that contained 1 μM of both sense and antisense primers, each dNTP at 250 μM, 0.25 unit of TaKaRa EX Taq (TaKaRa), and 1×EX Taq buffer (TaKaRa). The volume was brought up to 20 μl with Nuclease-Free Water (Quiagen, Tokyo, Japan). The PCR reaction

conditions were as follows: 94 °C for 5 min., 30 or 35 cycles of 94 °C for 40sec., 49-55 °C (according to which primer was used and from which species cDNA was isolated) for 30 sec., and 72 °C for 1min., and then 72 °C for 7 min. The PCR products were electrophoresed on a 1.5% agarose gel and visualized using ethidium bromide staining and a UV transilluminator. PCR products of the expected size were extracted from the gel and purified using the phenol and chloroform according to the methods described in Aitken, 2012. Some of the DNA fragments extracted were directly sequenced, while others were cloned into pTAC-2 vectors (BioDynamics Laboratory Inc., Tokyo) according to the manufacturer's protocols. DH5α competent cells prepared according to Inoue et al. (Inoue et al., 1990) were transformed with the recombinant vector and positive bacteria were selected by blue/white colony screening and colony PCR with vector specific M13 primers (BioDynamics Laboratory Inc.). The colonies selected were cultured in 2 ml of LB medium supplemented with ampicillin. Plasmids were extracted and purified from bacteria according to Borodina *et al*. (Borodina et al., 2003) and then sequenced.

Partial nucleotide sequences of cDNAs coding for insulin were obtained from the leopard gecko and Japanese gecko by 3'-RACE with SE01. Based on the sequences isolated, AS01 was constructed, and was used for 5'-RACE. These steps enabled isolation of full cDNA sequences of preproinsulin (completed with signal peptide, B-chain, C-peptide, and A-chain) from the leopard gecko and Japanese gecko. Based on the sequences of these two species, SE02, SE03 and SE04 were constructed in the region corresponding to signal peptide. 3'-RACE with SE02 and SE03 enabled isolation of proinsulin sequences (the region that consists of B-chain, C-peptide, and A-chain) from other squamates but the rat snake. Partial cDNA sequences coding for the rat snake insulin was isolated by 3'-RACE with SE01. AS02 was constructed based on this partial sequence. RT-PCR with SE04 and AS02 (this set of primers was designed to flank an intron) enabled isolation of the nucleotide sequence corresponding to proinsulin region from the rat snake.

The final sequence for each transcript was the result of the assembly of three clones that were amplified from independent cDNA synthesis, so as to avoid any errors that resulted from reverse-transcription and/or PCR amplification.

4. Sequence analysis and construction of molecular phylogenetic tree

The CLUSTAL W program embedded in MEGA (version 6.06) (Tamura et al., 2013) was used with default settings to align the nucleotide and deduced amino acid sequences of insulin from the squamates and species that represent other groups of vertebrates. Species names and GenBank accession numbers of representative species are summarized in Table 1-3. The alignment was checked by the naked eye and corrected manually based on the previous studies (Conlon, 2000, 2001). Amino acid identities were calculated using GeneDoc (version 2.7) (Nicholas and Nicholas, 1997). Signal peptide sequence was predicted using SignalP (version 4.1) (Petersen et al., 2011) in the CBS Prediction Servers (http://www.cbs.dtu.dk/services/SignalP/).

Molecular phylogenetic trees of *ins* were drawn with its paralogous genes, insulin-like growth factor (IGF)-1 and IGF-2. Codon alignment of the nucleotide sequences corresponding to proinsulin and mature IGF-1/2 (comprising following four domains; B-domain, C-domain, A-domain, and D-domain) were constructed using the CLUSTAL W program embedded in MEGA (version 6.06) (Tamura et al., 2013). The alignment was checked by the naked eye and corrected manually based on the previous studies (Chan et al., 1990; Conlon, 2000, 2001; Humbel, 1990; Upton et al., 1997). "A la Carte" mode of Phylogeny. fr (Dereeper et al., 2008) was used to build two Neighbor-Joining (N-J) trees and one Maximum Likelihood (ML) tree. Gaps in the alignment were deleted by selecting “Remove positions with gaps” for the alignment curation option. N-J trees were built by selecting “BioNJ” or “Neighbor” for phylogeny reconstruction. An ML tree was built by selecting “PhyML” for phylogeny reconstruction. Another ML tree was also drawn using MEGA. For this tree

construction, gaps in the alignment were deleted by the program. “Find Best DNA/Protein Models” was run first to find best settings for the “Substitution Model” and “Rates and Patterns”. The best setting was Kimura 2- parameter model for “Substitution Model” and Gamma distributed with Invariant sites (G+I) for “Rates and Patterns”. The number of bootstrap replicates for the estimation of robustness of the internal nodes was set to 1000 for N-J trees, and 100 for ML trees. Full species names and GenBank accession numbers of all the sequences used for tree construction are summarized in Table 1-3.

5. Analysis of molecular evolution by PAML CODEML

The evolutionary pattern of squamate *ins* was analyzed using the CODEML program implemented in PAML4 (version 4.7) (Yang, 2007, 2013a). This program estimates the values of the nonsynonymous-to-synonymous substitution (dN/dS) ratio (described hereafter as “ω”) of lineages among a phylogenetic tree or of sites (codons) comprising proteins. The value of ω is used as an indicator of selection pressure, with $\omega < 1$ suggesting negative (purifying) selection, $\omega = 1$ suggesting neutral evolution, and $\omega > 1$ suggesting (Darwinian) positive selection (Yang and Bielawski, 2000). In subsequent analyses, the alignment of 144bp nucleotide sequences that included the regions corresponding to B1-B27 of the insulin B-chain and A1-A21 of the insulin A-chain were used. B28-B30 and extensions at the C-terminus of the A-chain were excluded from the alignment in order to avoid problems caused by the possible ambiguity of the alignment in these regions. This dataset was first tested for saturation bias with DAMBE5 (Xia, 2013) using the index of saturation (Xia *et al.*, 2003). The tree topology shown in Figure 1-2 was used for the analyses. This topology is based on previous studies for phylogenetic relationship of reptilian and avian species (Iwabe et al., 2005; Livezey and Zusi, 2007; McCormack et al., 2013; Pyron et al., 2013; Vidal and Hedges, 2005). Therefore, it has nothing to do with the molecular phylogenetic trees drawn in this study. Sequences

from 2 amphibian species (the western clawed frog and leopard frog) were included in the tree (Figure 1-2) to obtain a trifurcated tree according to the recommendation by Yang (Yang, 2005).

5.1. Tests for heterogeneous values of dN/dS (ω) ratio among lineages by the Branch Model analysis

I used the Branch Model analysis, one of the analyses implemented in CODEML, to detect differences in the selection pressure imposed on the molecular evolution of *ins* among reptilian and avian lineages. In this analysis, several models were constructed according to assumptions as to the number of lineages with independent values of ω. In each model, branches or clades of interest among the phylogenetic tree were specified by branch labels according to the rules described in the User Guide (Yang, 2013a). Branches or clades specified with labels were called “foreground”, while those with no label were called “background” according to the User Guide. Nested models were compared with the log likelihood ratio test (LRT); twice the difference between the log likelihood values of the two models (2δ) was compared to a χ^2 distribution. The degree of freedom for the χ^2 distribution corresponded to the difference between the number of parameters in the two models tested. Each model was run three times with different starting values of ω (ω = 0.5, 1, and 2).

Description of the models constructed in this test is as follows. Free-ratio model assumed independent values of ω for every branch in the phylogeny. One-ratio model assumed constant value of ω across the phylogeny. Two-ratio model assumed different values of ω between Squamata clade and other background lineages. In this model, Squamata clade (ancestral branch of Squamata and all the branches descending from it) was specified as foreground. Three-ratio models 1-4 allowed one lineage of interest among Squamata clade to have independent ω value from those of other squamates and background lineages. Five-ratio model allowed three lineages among Squamata clade to have independent values of ω from each other, the rest of the squamates, and background lineages. Six-ratio models (1-3) were constructed by adding one more foreground branch (or clade, in Six-ratio 3) to the Five-ratio model.

5.2. Detection of positively selected sites (codons)

Sites (codons) that evolved under positive selection pressure are termed “positively selected site” (Yang, 2013a) or PSS in short. Detection of PSS is important for estimating functional changes in biological molecules. However, the Branch Model analysis cannot be applied to detect the PSS, because it cannot assume independent values of ω for each site comprising a molecule. In order to detect the PSS of a specific lineage among the phylogeny, I used two types of analyses, called Branch-Site Model and Clade model C, implemented in the CODEML program, both of which assumed independent values for ω at each site.

5.2.1. Test for specific lineage by the Branch-Site Model

The test by comparison between Branch-Site Model A and corresponding null model was used to estimate the PSS according to the recommendation by Yang (Yang, 2013a). Branch-Site Model A and the corresponding null model both accounted for the independent values of ω at each site. They both assumed two branch or clade types; “foreground” specified with a label and “background” with no label. The difference between these two models was that while Model A allowed for the existence of the PSS in the “foreground” branch or clade, the null model did not (Yang, 2013a). Both models did not allow PSS in the “background” branches/clades.

Model A was tested against the null model by LRT as follows; twice the difference between the log likelihood values of the two models (2δ) was compared with the null distribution (a 50:50 mixture of point mass 0 and χ_1^2) (Yang, 2013a).

The following five branches/clades were examined with this test: family Gekkonidae clade, yellow-headed gecko branch, savannah monitor branch, Squamata clade, Japanese gecko group clade (Japanese gecko and hokou gecko). Each model was run three times with different starting values of ω (ω = 0.5, 1, and 2). The posterior probability that each site had been under positive selection was calculated by the Bayes Empirical Bayes (BEB) procedure according to the recommendation by Yang

(Yang, 2013a).

5.2.2. Test for different evolutionary process between ancestors and descendants within the "foreground" clade by Clade model C

If a mutation in a gene is advantageous for the survival of a certain species, it is likely to be fixed in the descendant species. In this case, the mutation site is expected to be under negative selection pressure in the descendant species, and in a neutral state or under positive selection pressure in the ancestral species. Thus, the value of ω estimated for such site is likely to be smaller in the descendants than in the ancestor. However, differences between the ancestor and descendants within the "foreground" clade could not be examined with the Branch-Site Model analysis, because it cannot specify two or more branches or clades as "foreground" at the same time. Clade model C (CmC) is applicable for such assumptions. CmC allows for two or more types of "foreground" at the same time, and this analysis has been referred to as the "multi-clade approach" or "multi-clade analysis" (Weadick and Chang, 2012b; Yoshida et al., 2011). In the present study, I focused on the family Gekkonidae and Squamata clade, and constructed two models, CmC-2ratio and CmC-3ratio, for each clade. CmC-2ratio and CmC-3ratio both assumed independent values for ω at each site. They also assumed sites with different ω values between the foreground clade and background. The difference between these two models was that while CmC-2ratio assumed no difference in the evolutionary process between the ancestral branch and descendant branches within the "foreground" clade, CmC-3ratio assumed different processes between the ancestor and descendants. Thus, CmC-3ratio accounted for sites with different ω values not only between "foreground" and "background", but also between the ancestor and descendants within the "foreground" clade.

CmC-2ratio and CmC-3ratio models were tested against two types of null models, namely M1a and M2a_rel (Weadick and Chang, 2012a; Yang, 2013a), by LRT; twice the difference between the log likelihood values of the two models (2δ) was compared with the null distribution (χ^2 distribution).

The degree of freedom (d. f.) for the null distribution was as follows: CmC-2ratio vs M1a, d. f. = 3; CmC-2ratio vs M2a_rel, d. f. = 1; CmC-3ratio vs M1a, d. f. = 4; CmC-3ratio vs M2a_rel; d. f. = 2 (Weadick and Chang, 2012a, 2012b; Yang, 2013a). As the tests against M1a is old and prone to false-positive detection of positive selection, I mainly relied on the tests against M2a_rel as this test has been shown to be more accurate (Weadick and Chang, 2012a). The CmC-3ratio model was also tested against CmC-2ratio by LRT according to the method described by Weadick and Chang (Weadick and Chang, 2012b), with the null distribution of this test being χ^2 distribution with d. f. =1.

Each model was run six times with different starting values of ω (ω = 0.1, 0.2, 0.5, 1, 2, and 5). The posterior probability that each site had been under positive selection was calculated by the Bayes Empirical Bayes (BEB) procedure according to the recommendation by Yang (Yang, 2013a).

Table 1-1. Squamate species from which insulin was isolated.

Scientific name	Common name	in this article	NCBI accession
Gekkonidae, Gekkota			
Gekko japonicus	Japanese gecko	Japanese gecko	LC020029
Gekko hokouensis	hokou gecko	hokou gecko	LC020030
Gekko badenii	golden gecko	golden gecko	LC020032
Gekko vittatus	lined gecko	lined gecko	LC020033
Gekko gecko	tokay gecko	tokay gecko	LC020031
Gekko monarchus	spotted house gecko	spotted house gecko	LC020034
Hemidactylus frenatus	common house gecko	common house gecko	LC020041
Hemidactylus platyurus	flat-tailed house gecko	flat-tailed house gecko	LC020040
Cyrtodactylus quadrivirgatus	Taylor's bow-fingered gecko	bow-fingered gecko	LC020042
Stenodactylus sthenodactylus	elegant gecko	elegant gecko	LC020043
Lygodactylus kimhowelli	yellow headed dwarf gecko	dwarf gecko	LC020038
Phelsuma laticauda	gold dust day gecko	day gecko	LC020037
Homopholis fasciata	striped velvet gecko	velvet gecko	LC020039
Chondrodactylus turneri	Turner's thick-toed gecko	thick-toed gecko	LC020036
Paroedura picta	panther gecko	panther gecko	LC020035
Phyllodactylidae, Gekkota			
Ptyodactylus guttatus	Sinai fan-fingered gecko	fan-fingered gecko	LC020044
Tarentola mauritanica	common wall gecko	wall gecko	LC020045
Sphaerodactylidae, Gekkota			
Teratoscincus roborowskii	Roborowski's wonder gecko	wonder gecko	LC020046
Gonatodes albogularis	yellow-headed gecko	yellow-headed gecko	LC020047
Eublepharidae, Gekkota			
Eublepharis macularius	leopard gecko	leopard gecko	LC020048
Coleonyx mitratus	Central American banded gecko	banded gecko	LC020049
Scinciformata, Unidentata			
Plestiodon finitimus	Japanese five-lined skink	five-lined skink	LC020051
Scincus scincus	sandfish	sandfish	LC020050
Cordylus tropidosternum	dwarf sungazer	dwarf sungazer	LC020052

Lacertiformata, Unidentata			
Takydromus tachydromoides	Japanese grass lizard	grass lizard	LC020053
Anguimorpha, Unidentata			
Varanus exanthematicus	savannah monitor	savannah monitor	LC020054
Serpentes, Unidentata			
Elaphe climacophora	Japanese rat snake	rat snake	LC020055

Insulin was isolated from 27 squamate species listed above. For the analyses of amino acid sequences and molecular evolution, nucleotide sequences of green anole (NCBI accession: XM_003214757.1) and king cobra (AZIM01002201.1) were added. Thus, 29 squamate species in total were analyzed.

Table 1-2. Oligonucleotides (primers) used for the identification of squamate insulin.

Name	Nucleotide sequence	Used for:
SE01	GTRTGYGGRGANCGNGGNTTCTTCTA	degenerate PCR (3'-RACE)
SE02	ATGACATTCTGGATCAGATCTTTGCCT	3'-RACE, sequencing
SE03	ATGACATTCTGGATCAGATCTTTGCCTCT	3'-RACE (savannah monitor), sequencing
SE04	ATGACATTCTGGATCAGATCTTTGCCTCTCCT	PCR (rat snake), sequencing
AS01	AGGTTTCCTCTTCTAGTTGCAGTAGCT	5'-RACE
AS02	GAGCAGGTATTTTCACAGCATTGCT	PCR (rat snake)

Abbreviations: R = A or G; Y = C or T; N = A, T, G, or C.

Table 1-3. List of insulin, IGF-1 and IGF-2 sequences obtained from the GenBank Database

Molecule/ Taxonomy	Scientific name	Name in the figures	GenBank accession number	Purpose
Insulin/Mammalia				
	Cavia porcellus	Guinea pig	NM_001172891.1	Tree
	Homo sapiens	human	DQ778082.1	Alignment and Tree
	Mesocricetus auratus	Syrian hamster	M26328.1	Alignment and Tree
	Octodon degus	degu	M57671.1	Tree
Insulin/Squamata				
	Anolis carolinensis	Green anole	XM_003214757.1	Alignment, Tree and PAML
	Ophiophagus hannah	King cobra	AZIM01002201.1	Alignment, Tree and PAML
Insulin/Testudines				
	Chrysemis picta bellii	western painted turtle	XM_005312381.1	Alignment, Tree and PAML
	Pelodiscus sinensis	softshell turtle	XM_006134914.1	Alignment, Tree and PAML
Insulin/Crocodilia				
	Alligator mississippiensis	American alligator	XM_006272425.1	Alignment, Tree and PAML
	Alligator sinensis	Chinese alligator	XM_006033646.1	Alignment, Tree and PAML
Insulin/Aves				
	Anas platyrhynchos	wild duck	XM_005019748.1	Tree and PAML
	Columba livia	rock pigeon	XM_005499243.1	Tree and PAML
	Falco cherrug	saker falcon	XM_005439113.1	Tree and PAML
	Falco peregrinus	peregrine falcon	XM_005237317.1	Tree
	Ficedula albicollis	collared flycatcher	XM_005046804.1	Tree
	Gallus gallus	chicken	NM_205222.2	Alignment, Tree and PAML
	Geospiza fortis	medium ground finch	XM_005427205.1	Tree
	Meleagris gallopavo	turkey	XM_003206278.1	Tree and PAML
	Melopsittacus undulatus	budgerigar	XM_005149200.1	Tree and PAML
	Pseudopodoces humilis	ground tit	XM_005522339.1	Tree and PAML
	Taeniopygia guttata	zebra finch	XM_002198933.1	Alignment, Tree and PAML
	Zonotrichia albicollis	white-throated sparrow	XM_005486567.1	Tree
Insulin/Amphibia				
	Rana pipiens	leopard frog	AF227187.1	Tree and PAML
	Silurana tropicalis	western clawed frog	NM_001100236.1	Alignment, Tree and PAML
Insulin/Teleosts				
	Danio rerio	zebrafish	BC095013.1	Alignment and Tree

	Salmo salar	Atlantic salmon	BT049203.1	Tree
Insulin/Agnatha				
	Myxine glutinosa	Atlantic hagfish	V00649.1	Tree
IGF-1/Mammalia				
	Cavia porcellus	guinea pig	NM_001172966.1	Tree
	Homo sapiens	human	BC152321.1	Tree
	Mus musculus	mouse	AF440694.1	Tree
IGF-1/Squamata				
	Anolis carolinensis	green anole	XM_008110573.1	Tree
	Coelognathus helena	trinket snake	JN315417.1	Tree
	Python bivittatus	Burmese python	XM_007419940.1	Tree
	Pantherophis guttatus	corn snake	JN315416.1	Tree
	Sceloporus undulatus	eastern fence lizard	JN315420.1	Tree
	Scincella lateralis	little brown skink	JQ693411.1	Tree
IGF-1/Testudines				
	Chelydra serpentina	common snapping turtle	JN315421.1	Tree
	Pelusios castaneus	West African mud turtle	JQ693410.1	Tree
	Pelodiscus sinensis	softshell turtle	JN698984.1	Tree
IGF-1/Crocodilia				
	Alligator mississippiensis	American alligator	JQ693409.1	Tree
	Alligator sinensis	Chinese alligator	NM_001286846.1	Tree
IGF-1/Aves				
	Anas platyrhynchos	wild duck	EU031044.1	Tree
	Anser anser	greylag goose	DQ662932.1	Tree
	Gallus gallus	chicken	M32791.1	Tree
	Meleagris gallopavo	Turkey	XM_003202378.1	Tree
IGF-2/Mammalia				
	Cavia porcellus	guinea pig	XM_003468138.2	Tree
	Homo sapiens	human	BC000531.1	Tree
	Mus musculus	mouse	M14951.1	Tree
IGF-2/Squamata				
	Anolis carolinensis	green anole	XM_008107927.1	Tree
	Python bivittatus	Burmese python	XM_007424231.1	Tree
IGF-2/Testudines				

	Chrysemis picta bellii	western painted turtle	XM_005312382.1	Tree
	Pelodiscus sinensis	softshell turtle	XM_006134915.1	Tree
IGF-2/Crocodilia				
	Alligator mississippiensis	American alligator	XM_006272426.1	Tree
	Alligator sinensis	Chinese alligator	XM_006033647.1	Tree
IGF-2/Aves				
	Gallus gallus	chicken	NM_001030342.1	Tree
	Meleagris gallopavo	turkey	XM_003206279.1	Tree

Results

1. Preproinsulin sequence from the leopard gecko and Japanese gecko

I first isolated nucleotide sequences of full-length ORF of preproinsulin from the leopard gecko and Japanese gecko by 3'-RACE and 5'-RACE. The ORF of the leopard gecko preproinsulin comprised 303bp, while that of Japanese gecko preproinsulin comprised 294bp. The deduced amino acid sequences are shown in Figure 1-3 together with those of other vertebrates. As is the case in most vertebrate species, insulin of these two species both contained a 30aa B-chain and a 21aa A-chain. On the other hand, the C-peptide was shorter than those of other vertebrates. While most vertebrate species have a C-peptide of approximately 30aa, those of the leopard gecko and Japanese gecko comprised 26aa and 23aa, respectively.

Focusing on the possible differences in molecular functions between species, the amino acid sequences of insulin (B-chain and A-chain) from these two species were aligned and compared with those of squamate species previously reported and representative species from non-squamate reptiles, aves and mammals. The leopard gecko insulin (B-chain and A-chain) showed 84% identity with that of the green anole (*Anolis carolinensis*), 82% with the king cobra (*Ophiophagus hannah*), 76% with the American alligator (*Alligator mississippiensis*), 78% with the western painted turtle (*Chrysemys picta bellii*), 78% with the chicken (*Gallus gallus*) and the zebra finch (*Taeniopygia guttata*), and 74% with human (*Homo sapiens*) and the Syrian hamster (*Mesocricetus auratus*). The sequence identity of the Japanese gecko was lower than these values: 60% with the green anole, 64% with the king cobra, 58% with the American alligator, 56% with the western painted turtle, 56% with the chicken and the zebra finch, 58% with human, and 54% with the Syrian hamster (Table 1-10). Such difference in the percentages of identity implies the high variability of amino acid sequences among squamate insulin.

2. Proinsulin sequences from species of Squamata

The nucleotide sequences containing the proinsulin region (B-chain, C-peptide and A-chain) were isolated by 3'-RACE from 25 squamate species other than the leopard gecko and Japanese gecko. The deduced amino acid sequences were aligned with the representative species of previously reported squamates (the green anole and the king cobra), non-squamate reptiles (the American alligator and the western painted turtle), aves (the chicken), and mammals (human) (Figure 1-4). This comparison enabled identification of recognition motifs for PC-1/3 and -2 (Seidah and Chretien, 1999) at the chain borders in all squamate species. Cysteine residues for the inter- and intra-chain disulfide bridges were also found to be fully conserved (Davidson, 2004) (Figure 1-4)

Amino acid sequences of insulin (B-chain and A-chain) were aligned and compared with 8 representative species, two species each from squamates, non-squamate reptiles, aves and mammals (Table 1-10). The percentages of identity were highly variable. Among squamate species, the savannah monitor showed the lowest percentages with every species compared: the values never exceeded 50%. In most cases, the five-lined skink showed the highest percentages among squamates, showing more than 82% identity with all the representative species compared. Only when compared with the king cobra, five-lined skink was second to the rat snake, which showed 98% identity. When compared with human insulin, the following four lineages showed particularly low values among squamates and the representatives of amniotes: family Gekkonidae, the yellow-headed gecko (*Gonatodes albogularis*), the dwarf sungazer (*Cordylus tropidosternum*), and the savannah monitor (*Varanus exanthematicus*). These lineages showed not more than 66% identity values, while other species of amniotes (including squamates) showed values over 74% (Figure 1-2, Table 1-10).

ML and N-J trees were drawn based upon the ORF nucleotide (codon) alignment of the proinsulin, IGF-1 and IGF-2 (Figure 1-1. Trees drawn by PhyML and BioNJ are not shown). Bootstrap values for the robustness of internal nodes were generally not good. In the ML trees, mammalian IGF-1 did not

cluster with that of other vertebrates. However, all insulin sequences, together with those identified from squamate species, separated from IGF sequences. Among insulin sequences, those of squamates formed a monophyletic cluster. The tree topology of this "squamate insulin clade" was not so consistent with phylogenetic relationship of squamate species. For example, species of Serpentes, one of the most descendant groups among Squamata (Pyron et al., 2013; Vidal and Hedges, 2005), sometimes branched off at the root of this clade (Figure 1-1). In addition, the yellow-headed gecko, dwarf sungazer, and savannah monitor were not in sister relationship with their respective closest relatives (the wonder gecko for yellow-headed gecko, the five-lined skink and sandfish for dwarf sungazer, and the green anole for savannah monitor) (Pyron et al., 2013). They also showed especially long branch length compared with the other squamate species. Considering low homology of these species with human insulin (Figure 1-2, Table 1-10), the branch length appeared to reflect divergent sequences.

Amino acid substitutions were observed in two binding surfaces, called the "classical binding surface" and the "novel binding surface", through which insulin interacts with the insulin receptor (IR) (De Meyts et al., 2000). Among 12 sites comprising the classical binding surface, substitutions were found in B16, B24, B25, B26, A5, and A21. On the other hand, substitutions were found in all sites comprising the "novel binding surface" (B10, B13, B17, A12, A13, and A17) (Tables 1-8, 1-9).

In summary, amino acid sequences of insulin were highly variable among squamate species, with some species showing substitutions at the sites involved in receptor binding.

3. Tests for heterogeneous ω among lineages by Branch Models

Before analyzing *ins* with PAML CODEML, saturation bias on the dataset was examined with DAMBE5. For our dataset, the index score (ISS) was significantly lower than the critical score (ISS.c) with $p < 0.0001$ when a symmetrical topology was assumed. Even when an extremely asymmetrical

topology was assumed, the ISS was still lower than the ISS.c with $p < 0.06$. Therefore, the effect of saturation on the dataset was likely to be subtle.

To examine changes in the selection pressure imposed on the *ins* in relation to the phylogenetic relationship, the heterogeneity of ω among lineages was examined with the Branch Model analysis.

First, the Free-ratio model was compared with the One-ratio model. The Free-ratio model assumed independent ω values for every branch in the phylogenetic tree, whereas the One-ratio model assumed a constant value across phylogeny. This test was used to confirm the heterogeneity of ω among lineages. The LRT ($2\delta = 208.41$; χ^2_{82}, see Table 1-3) supported the Free-ratio model with $p < 10^{-5}$ (Table 1-3). Then Two-ratio model (an independent value of ω was allowed for the Squamata clade) was compared with the One-ratio model. The Two-ratio model was a significantly better fit to our data than the One-ratio model with $p < 10^{-5}$ (Table 1-3), and a markedly higher ω was estimated for the Squamata clade ($\omega_{Squamata} = 0.34$) than the background branches/clades ($\omega_{background} = 0.04$) (Table 1-4).

I then tested the heterogeneity of ω among the lineages of the Squamata clade by comparing Three-ratio models with the Two-ratio model. Three-ratio 1 (the family Gekkonidae as an additional foreground clade) and Three-ratio 3 (the savannah monitor as an additional foreground branch) were a better fit than the Two-ratio model with $p = 0.02$ and $p < 10^{-2}$, respectively. Three-ratio 2 (the yellow-headed gecko as an additional foreground) was weakly supported with $p = 0.06$ (Table 1-3). In these three models, increased values of ω were estimated for the additional foreground clade or branches ($\omega_{Gekkonidae} = 0.44$ in Three-ratio 1, $\omega_{yellow\text{-}headed\ gecko} = 0.74$ in Three-ratio 2, and $\omega_{savanah\ monitor} = 1.19$ in Three-ratio 3) (Table 1-4). On the other hand, Three-ratio 4 (the dwarf sungazer as an additional foreground) was not a better fit than the Two-ratio model ($p = 0.34$) (Table 1-3).

I then constructed the Five-ratio model, assuming that the Gekkonidae clade, yellow-headed gecko, and savannah monitor to have independent values of ω among the Squamata clade. This model was compared with the One-ratio model and the Two-ratio model, and was a better fit than both with $p <$

10^{-5} (Table 1-3). The values of ω estimated for these three lineages were higher than that of the other squamates ($\omega_{Gekkonidae} = 0.44$, $\omega_{yellow\text{-}headed\ gecko} = 0.74$, $\omega_{savannah\ monitor} = 1.27$, and $\omega_{Squamata} = 0.16$) (Table 1-4). In addition, Six-ratio models were constructed to test the Japanese gecko and hokou gecko, by adding additional foreground branch/clade with the Five-ratio model. Six-ratio 1 (the Japanese gecko as an additional foreground branch) and Six-ratio 3 (both species together as an additional foreground clade) were a better fit than the Five-ratio model with $p < 10^{-2}$. Six-ratio 2 (the hokou gecko as an additional foreground) was weakly supported with $p = 0.07$ (Table 1-3). Six-ratio 1 and Six-ratio 2 estimated $\omega = 999.00$ for the additional foreground branch (Table 1-4). This appeared to be an erroneous value that resulted from $dS = 0$. In other words, no "synonymous substitution" was found in the foreground branch. As this was the case, the equation $\omega = dN/dS$ yielded an erroneous value. In contrast, Six-ratio 3, in which two species were grouped together as an additional foreground clade, estimated this foreground clade to have $\omega_{Japanese\ gecko\ group} = 2.44$ (Table 1-4).

4. Detection of positively selected sites

4.1. Tests for specific lineage

I tested following five lineages with the Branch-Site Model for the existence of the PSS: the Squamata clade, Gekkonidae clade, yellow-headed gecko branch, savannah monitor branch, and Japanese gecko group clade (a clade formed by the Japanese gecko and hokou gecko). These represent the lineages in which increase of ω value was supported by the Branch Model analyses. Models that specified the Squamata clade or Gekkonidae clade as foreground were not a better fit than the null models, as same log likelihood value was estimated for Model A and the null model. On the other hand, models for the yellow-headed gecko and that for the savannah monitor were a better fit than the null models with $p = 0.040$ and $p = 0.004$, respectively. The model for the Japanese gecko group clade was a relatively better fit with $p = 0.060$. The parameter estimates, LRTs and the PSS with more than

95% posterior probabilities (calculated with BEB) are summarized in Table 1-5.

4.2. Tests for heterogeneity between the ancestor and descendants

The Squamata clade and Gekkonidae clade were tested for different evolutionary processes and different PSS between the ancestral branch and descendant branches by the CmC multi-clade analysis. The parameter estimates of each model, together with the PSS with more than 95% posterior probabilities (calculated with BEB) are summarized in Table 1-6. The results of the LRTs are summarized in Table 1-7. In the Squamata clade, CmC-2ratio and CmC-3ratio were both a significantly better fit than the null models ($p < 10^{-7}$) (Table 1-7). CmC-3ratio was a relatively better fit than CmC-2ratio ($p = 0.07$) (Table 1-7), suggesting different evolutionary processes between the ancestral branch and descendant branches. However, the values of ω estimated for each site never exceeded 1, suggesting no PSS in this clade (Table 1-6). In the Gekkonidae clade, CmC-2ratio and CmC-3ratio were both a significantly better fit than the null models ($p < 10^{-7}$), whereas CmC-3ratio was not a better fit than CmC-2ratio ($p = 0.13$) (Table 1-7). However, it is worthy of note that CmC-3ratio suggested the existence of the PSS in the ancestral branch of Gekkonidae, as sites with $\omega > 1$ were detected ($\omega_3 = 2.09$) with high posterior probabilities (95% or more) (Table 1-6).

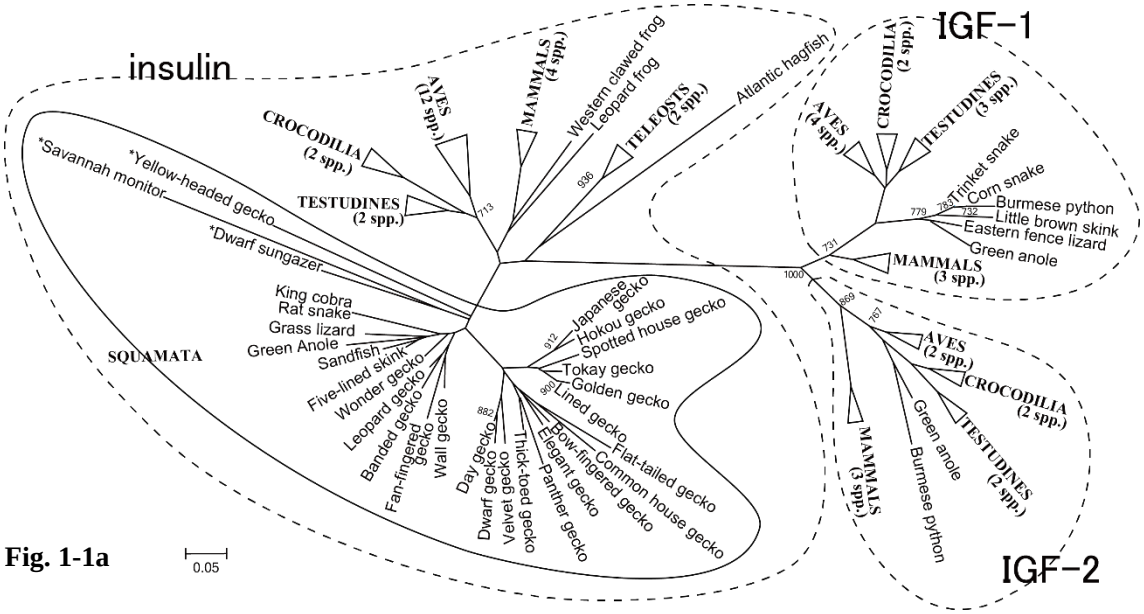


Fig. 1-1a

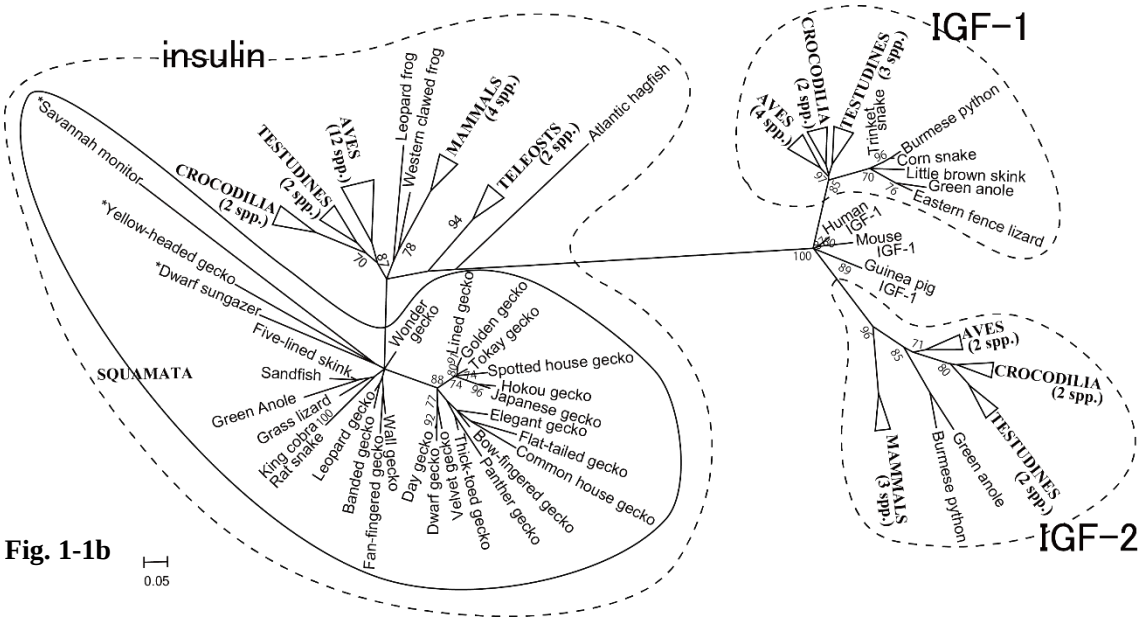


Fig. 1-1b

Figure 1-1. Molecular phylogenetic trees drawn based on nucleotide sequences of proinsulin, IGF-1 and IGF-2

In both trees, the scale bar in the figure indicates the estimated evolutionary distance units. (a) Bootstrap N-J tree drawn with Phylogeny. fr (+Neighbor). Bootstrap values (1000 replications) larger than 700 are indicated. Each hormone formed a monophyletic cluster. All squamate proinsulin clustered together and formed a distinct clade from other groups of vertebrates. Three squamate species (indicated with asterisk) branched off from the ancestral branch of this “Squamata clade”. They also showed especially long branch length among squamates. (b) Maximum likelihood tree drawn with MEGA. Bootstrap values (100 replications) larger than 70 are indicated. Mammalian IGF-

1 did not cluster with those of the other species in this tree. However, proinsulin formed a monophyletic cluster. Also, all squamate proinsulin clustered together and formed a distinct clade from other vertebrates. In addition, the same three squamate species (indicated with asterisk) as in the N-J tree showed especially long branch length among squamates in this tree, too.

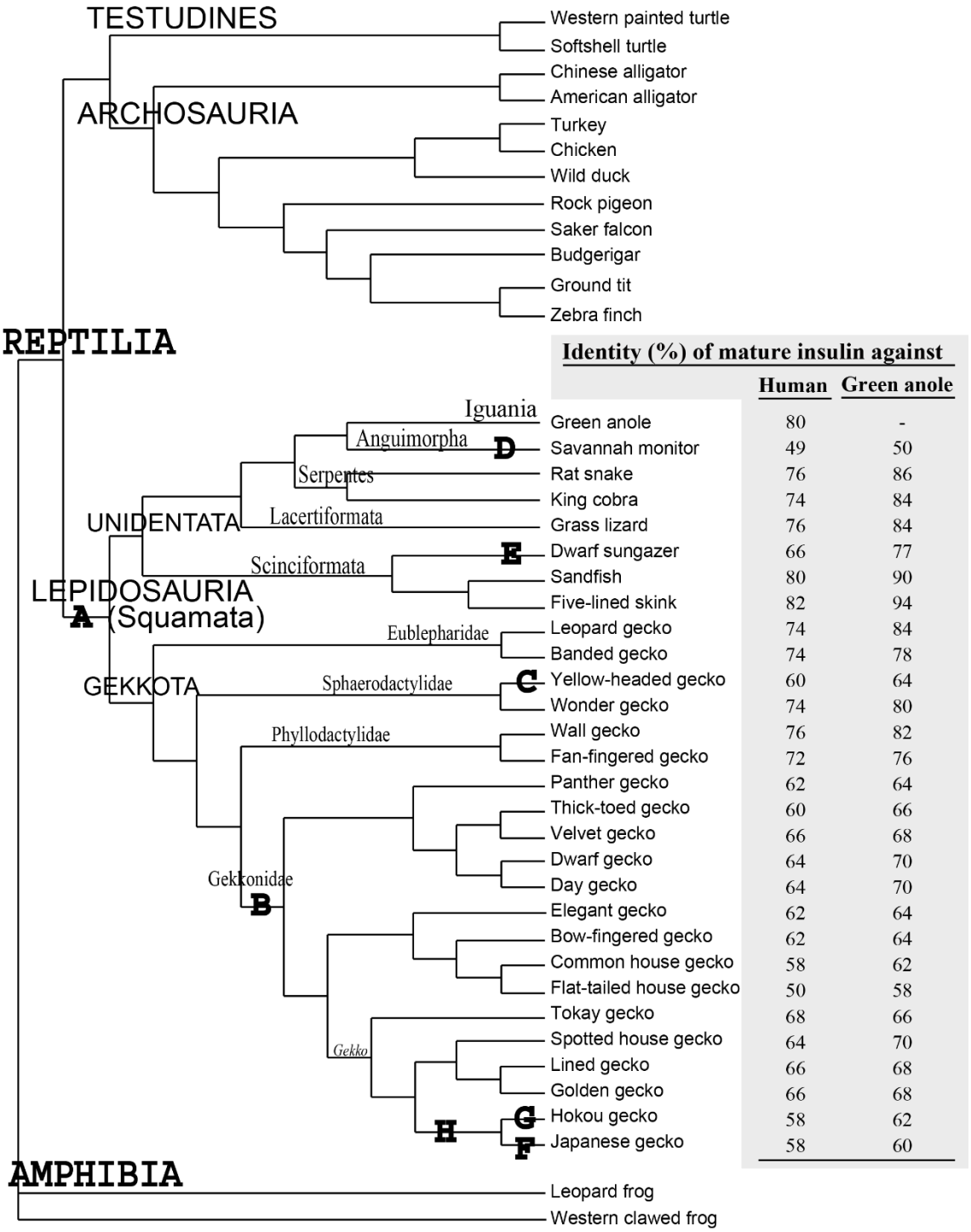
Species condensed to the taxonomic groups are as follows;

Mammalian insulin: degu, guinea pig, human, and Syrian hamster. Avian insulin: budgerigar, chicken, collared flycatcher, ground tit, medium ground finch, peregrine falcon, rock pigeon, saker falcon, turkey, white-throated sparrow, wild duck, and zebra finch. Testudine insulin: Chinese softshell turtle and western painted turtle. Crocodilian insulin: American alligator and Chinese alligator. Teleost insulin: Atlantic salmon and zebrafish.

Mammalian IGF-1: guinea pig, human, and mouse. Avian IGF-1: chicken, greylag goose, turkey, and wild duck. Testudine IGF-1: Chinese softshell turtle, common snapping turtle, and West African mud turtle. Crocodilian IGF-1: American alligator and Chinese alligator.

Mammalian IGF-2: guinea pig, human, and mouse. Avian IGF-2: chicken and turkey. Testudine IGF-2: Chinese softshell turtle and western panited turtle. Crocodilian IGF-2: American alligator and Chinese alligator.

The Genbank accession numbers of the insulin, IGF-1 and IGF-2 nucleotide sequences are listed in Table 1-1 and Table 1-3.



Identity (%) of mature insulin against	Human	Green anole
Green anole	80	-
Savannah monitor	49	50
Rat snake	76	86
King cobra	74	84
Grass lizard	76	84
Dwarf sungazer	66	77
Sandfish	80	90
Five-lined skink	82	94
Leopard gecko	74	84
Banded gecko	74	78
Yellow-headed gecko	60	64
Wonder gecko	74	80
Wall gecko	76	82
Fan-fingered gecko	72	76
Panther gecko	62	64
Thick-toed gecko	60	66
Velvet gecko	66	68
Dwarf gecko	64	70
Day gecko	64	70
Elegant gecko	62	64
Bow-fingered gecko	62	64
Common house gecko	58	62
Flat-tailed house gecko	50	58
Tokay gecko	68	66
Spotted house gecko	64	70
Lined gecko	66	68
Golden gecko	66	68
Hokou gecko	58	62
Japanese gecko	58	60

Figure 1-2. Phylogenetic relationship of the vertebrate species that were used to characterize the evolution of the *ins*

Tree topology is based upon following studies on phylogenetic relationship of reptilian and/or avian species: (Iwabe et al., 2005; Livezey and Zusi, 2007; McCormack et al., 2013; Pyron et al., 2013; Vidal and Hedges, 2005). The tree topology is trifurcated with two amphibian species (leopard frog

and western clawed frog) according to the recommendation by Yang (Yang, 2005). The taxonomic classification of each group is indicated above branches. Branches labelled in the analyses are indicated with characters A to H. The homology of squamate mature insulin (corresponding to concatenated insulin B-chain and A-chain) with those of human and the green anole is listed in the right column (shaded).

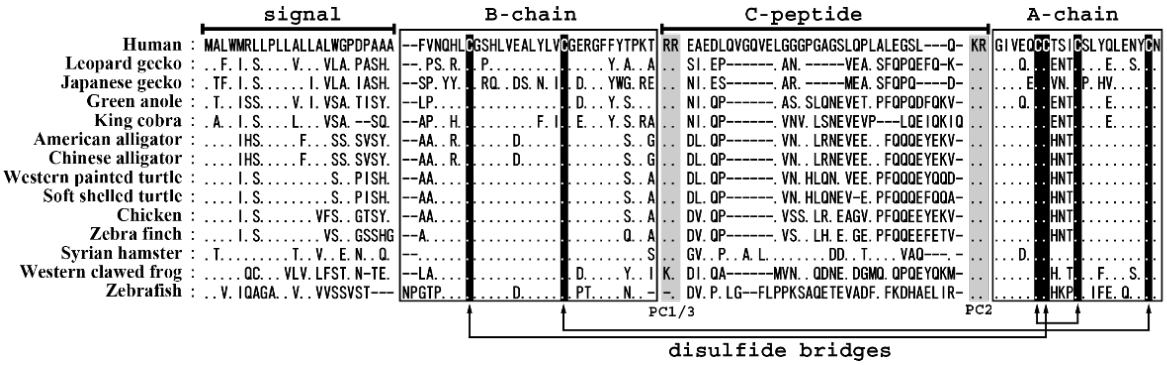


Figure 1-3. Amino acid sequence alignment of preproinsulin from the leopard gecko and Japanese gecko with other vertebrates.

The B-chain and A-chain, the regions comprising mature insulin, are boxed. Signal peptide and C-peptide are indicated by overlines. Sites that have the same amino acids as human preproinsulin are represented with dots (.). Recognition motifs for Prohormone Convertase (PC) -1/3 and PC-2 are indicated with gray shading. Cysteine residues responsible for inter- and intra-chain disulfide bridges are shown with outlined characters on blackened background, and are connected with arrows.

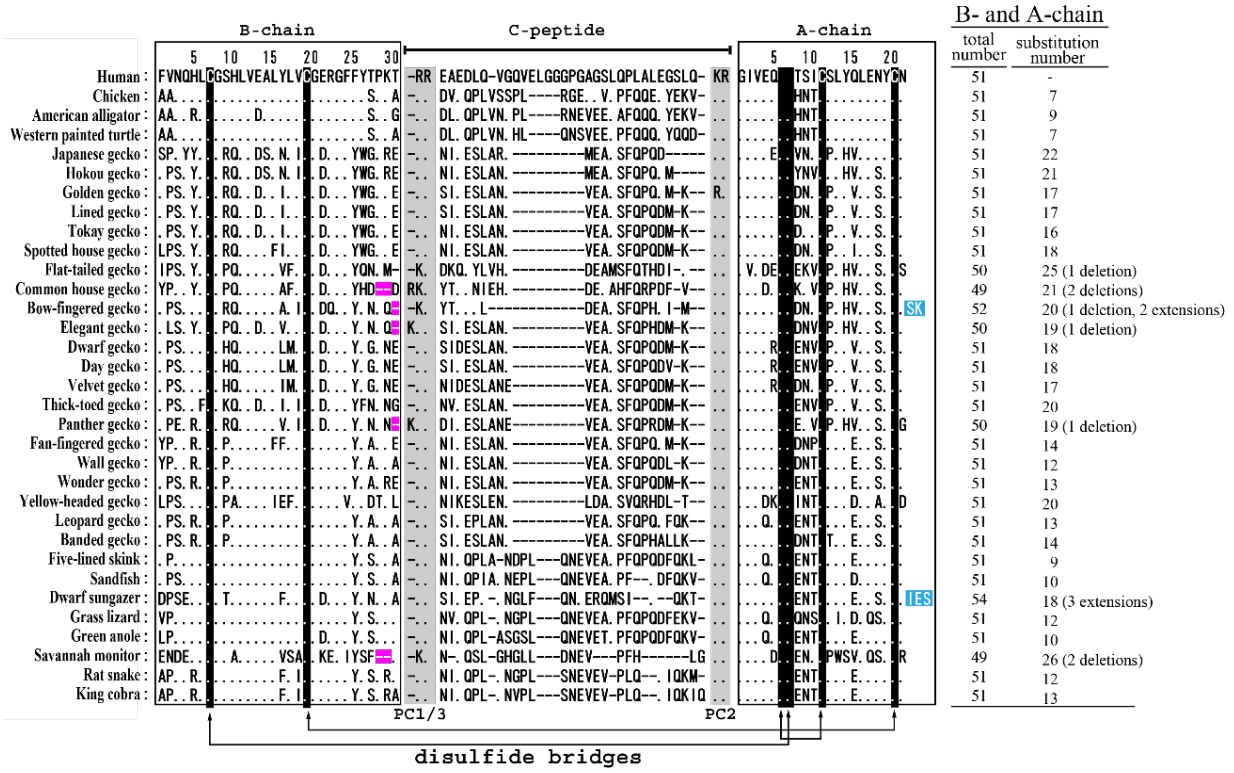


Figure 1-4. Amino acid alignment of proinsulin from squamate species with other groups of amniotes.

The B-chain and A-chain, the regions comprising mature insulin, are boxed. C-peptide is indicated by an overline. Recognition motifs for Prohormone Convertase (PC) -1/3 and PC-2 are indicated with shading. Cysteine residues responsible for inter- and intra-chain disulfide bridges are shown with outlined characters on blackened background, and are connected with arrows. The total number of amino acids comprising B- and A-chains and the number of amino acid substitutions within these chains compared with human insulin are annotated in the right column. Residue numbers of B-chain and A-chain are represented above each chain.

The sequences of human (DQ778082.1), chicken (NM_205222.2), American alligator (XM_006272425.1), western painted turtle (XM_005312381.1), green anole (XM_003214757.1) and king cobra (AZIM01002201.1) were obtained from the GenBank Database.

Following four species had deletions at the C-terminal region of the B-chain (represented with magenta background); the flat-tailed gecko, the common house gecko, the bow-fingered gecko, the elegant gecko, the panther gecko, and the savanah monitor. The bow-fingered gecko and the dwarf sungazer had extensions at the C-terminus of the A-chain (represented with blue background).

Table 1-3. Log likelihood Ratio Tests between Branch Models.

Model	Free-ratio		Two-ratio		Three-ratio 1 (Gekkonidae)		Three-ratio 2 (yellow-headed gecko)		Three-ratio 3 (savannah monitor)		Three-ratio 4 (dwarf sungazer)		Five-ratio		Six-ratio 1 (Japanese gecko)		Six-ratio 2 (hokou gecko)		Six-ratio 3 (Japanese gecko group)	
	2δ	*p*	2δ	*p*	2δ	*p*	2δ	*p*	2δ	*p*	2δ	*p*	2δ	*p*	2δ	*p*	2δ	*p*	2δ	*p*
One-ratio	208.41	$< 10^{-5}$	67.02	$< 10^{-5}$	72.20	$< 10^{-5}$	70.50	$< 10^{-5}$	75.58	$< 10^{-5}$	67.92	$< 10^{-5}$	96.27	$< 10^{-5}$	104.34	$< 10^{-5}$	99.65	$< 10^{-5}$	104.02	$< 10^{-5}$
Two-ratio	-	-	-	-	5.18	0.02	3.48	0.06	8.56	$< 10^{-2}$	0.90	0.34	29.25	$< 10^{-5}$	37.32	$< 10^{-5}$	32.63	$< 10^{-5}$	37.00	$< 10^{-5}$
Five-ratio	-	-	-	-	-	-	-	-	-	-	-	-	-	-	8.07	$< 10^{-2}$	3.39	0.07	7.75	$< 10^{-2}$

δ: difference between log likelihood values of models compared; *p*: critical values.

The degree of freedom of the null distribution for each test was as follows: Free-ratio vs One-ratio, d. f. = 82; Two-ratio vs One-ratio, d. f. = 1; Three-ratio vs One-ratio, d. f. = 2; Three-ratio vs Two-ratio, d. f. = 1; Five-ratio vs One-ratio, d. f. = 4; Five-ratio vs Two-ratio, d. f. = 3; Six-ratio vs One-ratio, d. f. = 5; Six-ratio vs Two-ratio, d. f. = 4; Six-ratio vs Five-ratio, d. f. = 1.

Table 1-4. Parameter estimates of the Branch Model analysis.

Model	lnL	np	Kappa	TL	Estimated values of nonsynonymous-to-synonymous substitution ratio (ω)								
					background	Squamata	Gekkonidae	yellow-headed gecko	savannah monitor	dwarf sungazer	Japanese gecko	hokou gecko	Japanese gecko group
Free-ratio	-2214.08	167	2.44	13.92									
One-ratio	-2318.29	85	2.44	13.32	0.21								
Two-ratio	-2284.78	86	2.47	14.03	0.04	0.34							
Three-ratio1 (Gekkonidae)	-2282.19	87	2.48	14.15	0.04	0.26	0.44						
Three-ratio2 (yellow-headed gecko)	-2283.04	87	2.47	14.01	0.04	0.32		0.74					
Three-ratio3 (savannah monitor)	-2280.50	87	2.48	13.91	0.04	0.31			1.19				
Three-ratio4 (dwarf sungazer)	-2284.33	87	2.47	14.05	0.04	0.35				0.21			
Five-ratio (lineages of Three-ratio 1-3)	-2270.15	89	2.48	13.86	0.04	0.16	0.44	0.74	1.27				
Six-ratio1	-2266.12	90	2.48	13.87	0.04	0.16	0.41	0.74	1.27		999.00		
Six-ratio2	-2268.46	90	2.48	13.86	0.04	0.16	0.43	0.74	1.27			999.00	
Six-ratio3	-2266.28	90	2.48	13.87	0.04	0.16	0.38	0.74	1.27				2.44

lnL: log likelihood value; np: number of parameters; Kappa: transition/transversion ratio; TL: tree length.

Labeling of tree branches in each model was as follows. Characters A-H correspond to specification of branches in Figure 1-2. While "#" specifies single branch, "$" specifies the branch labelled itself and all the branches descending from it at the same time (Yang, 2013a): One-ratio model, no label; Two-ratio model, "$1" to branch A; Two-ratio model, "$1" to branch A; Three-ratio 1, "$1" to branch A, "$2" to branch B; Three-ratio 2, "$1" to branch A, "$2" to branch C; Three-ratio 3, "$1" to branch A, "$2" to branch D; Three-ratio 4, "$1" to branch A, "$2" to branch E; Five-ratio model, "$1" to branch A, "$2" to branch B; "#3" to branch C, "#4" to branch D; Six-ratio 1, "$1" to branch A, "$2" to branch B; "#3" to branch C, "#4" to branch D, "#5" to branch F; Six-ratio 2, "$1" to branch A, "$2" to branch B; "#3" to branch C, "#4" to branch D, "#5" to branch G; Six-ratio 3, "$1" to branch A, "$2" to branch B; "#3" to branch C, "#4" to branch D, "$5" to branch H. Free-ratio model was designated by the setting files, thus without labeling in the tree.

The lnL value of Free-ratio model was obtained with initial omega = 1. In other models, the lnL value was constant across all initial values of omega.

Table 1-5. Parameter estimates for Branch-Site Model A and Null models.

Models	**lnL (Model A)**	**lnL (Null)**	**2δ**	***p***	**Kappa**	**TL**	**Parameter estimates for site classes**			**PSS (BEB, > 95%)**
							proportion	background ω	foreground ω	
Japanese gecko group	-2276.48	-2277.70	2.43	0.060	2.87	13.82	p_0 : 0.53249	ω_0 : 0.13	ω_0 : 0.13	B4
							p_1 : 0.13115	ω_1 : 1.00	ω_1 : 1.00	
							p_{2a}: 0.26989	ω_{2a}: 0.13	ω_{2a}: 3.49	
							p_{2b}: 0.06647	ω_{2b}: 1.00	ω_{2b}: 3.49	
Gekkonidae clade	-2261.23	-2261.23	0.00	0.500	2.82	15.13	p_0 : 0.56613	ω_0 : 0.10	ω_0 : 0.10	-
							p_1 : 0.06846	ω_1 : 1.00	ω_1 : 1.00	
							p_{2a}: 0.32599	ω_{2a}: 0.10	ω_{2a}: 1.00	
							p_{2b}: 0.03942	ω_{2b}: 1.00	ω_{2b}: 1.00	
yellow-headed gecko	-2274.63	-2276.16	3.08	0.040	2.97	14.15	p_0 : 0.64546	ω_0 : 0.13	ω_0 : 0.13	B10, A12
							p_1 : 0.17635	ω_1 : 1.00	ω_1 : 1.00	
							p_{2a}: 0.13995	ω_{2a}: 0.13	ω_{2a}: 4.43	
							p_{2b}: 0.03824	ω_{2b}: 1.00	ω_{2b}: 4.43	
savannah monitor	-2267.49	-2270.92	6.85	0.004	2.94	13.90	p_0 : 0.47012	ω_0 : 0.12	ω_0 : 0.12	B2, B10,
							p_1 : 0.12255	ω_1 : 1.00	ω_1 : 1.00	B22,A13,
							p_{2a}: 0.32311	ω_{2a}: 0.12	ω_{2a}: 4.28	A14, A21
							p_{2b}: 0.08423	ω_{2b}: 1.00	ω_{2b}: 4.28	
Squamata clade	-2245.04	-2245.04	0.00	0.500	2.90	14.20	p_0 : 0.47292	ω_0 : 0.05	ω_0 : 0.05	-
							p_1 : 0.00000	ω_1 : 1.00	ω_1 : 1.00	
							p_{2a}: 0.52708	ω_{2a}: 0.05	ω_{2a}: 1.00	
							p_{2b}: 0.00000	ω_{2b}: 1.00	ω_{2b}: 1.00	

lnL: log likelihood value; δ: difference between lnL (Model A) and lnL (Null model); *p*: critical values; Kappa: transition/transversion ratio; TL: tree length; PSS: positively selected sites. The value of lnL in each model was constant across initial omegas.

Table 1-6. Parameter estimates for the Clade model C analysis

Model	lnL	Kappa	TL	Site Class1		Site Class2		Site Class3		PSS (BEB, > 95%)
				prop	ω_0	prop	ω_1	prop	ω_2, ω_3, ω_4	
Gekkonidae										
CmC-2ratio	-2235.77	2.82	15.71	0.29874	0.00	0.06682	1.00	0.63443	ω_2: 0.15 ω_3: 0.57	-
CmC-3ratio	-2234.59	2.69	15.72	0.43553	0.02	0.02654	1.00	0.53793	ω_2: 0.19 ω_3: 2.09 ω_4: 0.73	B4, B9, B10, B16, B17, B18, B26, A5, A9, A10, A12, A14, A18, A21
Squamata										
CmC-2ratio	-2225.49	2.73	14.27	0.29794	0.00	0.05036	1.00	0.6517	ω_2: 0.05 ω_3: 0.45	-
CmC-3ratio	-2223.89	2.83	14.59	0.34677	0.01	0.07756	1.00	0.57567	ω_2: 0.06 ω_3: 0.03 ω_4: 0.48	-
Null models										
M1a	-2280.73	2.87	13.75	0.79659	0.14	0.20341	1.00	Class 3 Not assumed		Not allowed
M2a_rel	-2251.03	2.89	13.89	0.29126	0.00	0.15596	1.00	0.55278	0.24	Not allowed

lnL: log likelihood value; Kappa: transition/transversion ratio; TL: tree length; prop: proportion of each site class; PSS: positively selected sites.

Notes for site class 3: ω_2 represents the value estimated for site class 3 in background branches/clades. ω_3 represents the value in the entire foreground in CmC-2ratio models, while it represents the value in the ancestral branch within the foreground clade in CmC-3ratio models. ω_4 represents the value in the descendant branches within the foreground clade.

The values of initial omega with which the lnL values were obtained are as follows;

Gekkonidae-2ratio: 0.1, 0.2, 0.5, and 2; Gekkonidae-3ratio: 0.2, 0.5, 1, and 2; Squamata-2ratio: 0.1, 0.2, 0.5, 1, and 2; Squamata-3ratio: 0.2, 0.5, and 1; M1a: lnL was constant across all initial omegas; M2a_rel: 0.1, 0.2, 0.5, 1, and 2.

Table 1-7. Log likelihood ratio tests between CmC and null models

Alternative models	Gekkonidae				Squamata			
	CmC-2ratio		**CmC-3ratio**		**CmC-2ratio**		**CmC-3ratio**	
	2δ	*p*	2δ	*p*	2δ	*p*	2δ	*p*
Null models								
M1a	89.92	$< 10^{-7}$	92.28	$< 10^{-7}$	110.48	$< 10^{-7}$	113.68	$< 10^{-7}$
M2a_rel	30.52	$< 10^{-7}$	32.88	$< 10^{-7}$	51.08	$< 10^{-7}$	54.28	$< 10^{-7}$
CmC-2ratio (Gekkonidae)	-	-	2.35	0.13	-	-	-	-
CmC-2ratio (Squamata)	-	-	-	-	-	-	3.19	0.07

δ: difference between log likelihood values of models compared; *p*: critical values.

The degree of freedom of the null distribution for each test was as follows: CmC-2ratio vs M1a, d. f. = 3; CmC-2ratio vs M2a_rel, d. f. = 1; CmC-3ratio vs M1a, d. f. = 4; CmC-3ratio vs M2a_rel; d. f. = 2.

Table 1-8. Amino acids comprising “classical binding surface”

Species	Site											
	B12	B16	B23	B24	B25	B26	A1	A2	A3	A5	A19	A21
Human	Val	Tyr	Gly	Phe	Phe	Tyr	Gly	Ile	Val	Gln	Tyr	Asn
Japanese gecko	-	Asn	-	-	Tyr	Trp	-	-	-	Glu	-	-
Hokou gecko	-	Asn	-	-	Tyr	Trp	-	-	-	-	-	-
Golden gecko	-	Ile	-	-	Tyr	Trp	-	-	-	-	-	-
Lined gecko	-	Ile	--	-	Tyr	Trp	-	-	-	-	-	-
Tokay gecko	-	Ile	-	-	Tyr	Trp	-	-	-	-	-	-
Spotted-house gecko	-	Ile	-	-	Tyr	Trp	-	-	-	-	-	-
Flat-tailed gecko	-	Val	-	-	Tyr	Gln	-	Val		Glu	-	Ser
Common house gecko	-	Ala	-	-	Tyr	His	-	-	-	-	-	-
Bow-fingered gecko	-	Ala	-	-	Tyr	-	-	-	-	-	-	-
Elegant gecko	-	Val	-	-	Tyr	-	-	-	-	-	-	-
Dwarf gecko	-	Leu	-	-	Tyr	-	-	-	-	Arg	-	-
Day gecko	-	Leu	-	-	Tyr	-	-	-	-	Arg	-	-
Velvet gecko	-	Ile	-	-	Tyr	-	-	-	-	Arg	-	-
Thick-toed gecko	-	Ile	-	-	Tyr	-	-	-	-	-	-	-
Panther gecko	-	Val	-	-	Tyr	-	-	-	-	-	-	Gly
Fan-fingered gecko	-	Phe	-	-	Tyr	Phe	-	-	-	-	-	-
Wall gecko	-	-	-	-	Tyr	-	-	-	-	-	-	-
Wonder gecko	-	-	-	-	Tyr	-	-	-	-	-	-	-
Yellow-headed gecko	-	Gln	-	Val	-	-	-	-	-	Lys	-	Asp
Leopard gecko	-	-	-	-	Tyr	-	-	-	-	-	-	-
Banded gecko	-	-	-	-	Tyr	-	-	-	-	-	-	-
Five-lined skink	-	-	-	-	Tyr	-	-	-	-	-	-	-
Sandfish	-	-	-	-	Tyr	-	-	-	-	-	-	-
Dwarf sungazer	-	Phe	-	-	Tyr	-	-	-	-	-	-	-
Grass lizard	-	-	-	-	Tyr	-	-	-	-	-	-	-
Green anole	-	-	-	-	Tyr	-	-	-	-	-	-	-
Savannah monitor	-	Val	-	Ile	Tyr	Ser	-	-	-	Asp	-	Arg
Rat snake	-	Phe	-	-	Tyr	-	-	-	-	-	-	-
King cobra	-	Phe	-	-	Tyr	-	-	-	-	-	-	-

Table 1-9. Amino acids comprising “novel binding surface”

Species	Site					
	B10	B13	B17	A12	A13	A17
Human	His	Glu	Leu	Ser	Leu	Glu
Japanese gecko	Gln	Asp	-	Pro	-	-
Hokou gecko	Gln	Asp	-	-	-	-
Golden gecko	Gln	Asp	-	Pro	-	-
Lined gecko	Gln	Asp	-	Pro	-	-
Tokay gecko	Gln	Asp	-	Pro	-	-
Spotted-house gecko	Gln	Asp	-	Pro	-	-
Flat-tailed gecko	Gln	-	-	Pro	-	-
Common house gecko	Gln	-	Phe	Pro	-	-
Bow-fingered gecko	Gln	-	Phe	Pro	-	-
Elegant gecko	Gln	Asp	-	Pro	-	-
Dwarf gecko	Gln	-	Met	Pro	-	-
Day gecko	Gln	-	Met	Pro	-	-
Velvet gecko	Gln	-	Met	Pro	-	-
Thick-toed gecko	Gln	Asp	-	Pro	-	-
Panther gecko	Gln		-	Pro	-	-
Fan-fingered gecko	-	-	-	-	-	-
Wall gecko	-	-	-	-	-	-
Wonder gecko	-	-	-	-	-	-
Yellow-headed gecko	Ala	-	Phe	-	-	-
Leopard gecko	-	-	-	-		-
Banded gecko	-	-	-	Thr	-	-
Five-lined skink	-	-	-	-	-	-
Sandfish	-	-	-	-	-	-
Dwarf sungazer	-	-	-	-	-	-
Grass lizard	-	-	-	-	Ile	Gln
Green anole	-	-	-	-	-	-
Savannah monitor	Ala	-	Ser	Pro	Trp	Gln
Rat snake	-	-	-	-	-	-
King cobra	-	-	-	-	-	-

Table 1-10. Amino acid sequence identity and similarity (within parentheses) of insulin

Species compared	**Squamata**		**Other Reptiles**		**Birds**		**Mammals**	
	Green anole	King cobra	American alligator	Western painted turtle	Chicken	Zebra finch	Human	Syrian hamster
Squamata								
Japanese gecko	60 (80)	64 (80)	58 (78)	56 (80)	56 (80)	56 (80)	58 (86)	54 (80)
Hokou gecko	62 (86)	66 (82)	60 (82)	58 (84)	58 (84)	58 (84)	58 (86)	56 (86)
Golden gecko	68 (86)	64 (82)	66 (78)	64 (80)	64 (80)	64 (80)	66 (84)	64 (84)
Lined gecko	68 (86)	64 (82)	68 (78)	64 (80)	64 (80)	64 (80)	66 (84)	64 (84)
Tokay gecko	66 (86)	62 (82)	64 (78)	62 (80)	62 (80)	62 (80)	68 (84)	66 (84)
Spotted-house gecko	70 (86)	64 (82)	62 (78)	64 (80)	64 (80)	64 (80)	64 (84)	62 (84)
Flat-tailed gecko	58 (80)	54 (78)	49 (74)	50 (76)	50 (76)	50 (74)	50 (76)	52 (76)
Common house gecko	62 (78)	58 (78)	54 (74)	56 (76)	56 (76)	56 (76)	58 (80)	60 (80)
Bow-fingered gecko	64 (84)	64 (81)	56 (77)	60 (79)	60 (79)	60 (77)	62 (81)	60 (81)
Elegant gecko	64 (86)	62 (82)	64 (80)	62 (82)	62 (82)	62 (82)	62 (86)	60 (86)
Dwarf gecko	70 (84)	66 (82)	60 (76)	64 (78)	64 (78)	64 (78)	64 (82)	62 (82)
Day gecko	70 (84)	66 (82)	60 (76)	64 (78)	64 (78)	64 (78)	64 (82)	62 (82)
Velvet gecko	68 (84)	64 (80)	60 (76)	64 (78)	64 (78)	64 (78)	66 (82)	64 (82)
Thick-toed gecko	66 (86)	66 (82)	62 (80)	60 (80)	60 (80)	60 (78)	60 (82)	58 (82)
Panther gecko	64 (80)	68 (80)	58 (76)	58 (74)	58 (74)	58 (72)	62 (76)	60 (76)
Fan-fingered gecko	76 (90)	80 (92)	74 (88)	74 (86)	74 (86)	74 (84)	72 (86)	70 (86)
Wall gecko	82 (92)	84 (96)	80 (90)	82 (90)	82 (90)	82 (88)	76 (86)	74 (88)
Wonder gecko	80 (94)	84 (94)	76 (90)	76 (88)	76 (88)	76 (86)	74 (86)	72 (86)
Yellow-headed gecko	64 (80)	58 (78)	60 (74)	64 (76)	64 (76)	64 (76)	60 (76)	62 (76)
Leopard gecko	84 (94)	82 (96)	76 (90)	78 (90)	78 (90)	78 (88)	74 (86)	74 (86)
Banded gecko	78 (94)	80 (96)	76 (90)	78 (90)	78 (90)	78 (88)	74 (86)	72 (88)
Five-lined skink	94 (98)	88 (96)	82 (90)	88 (94)	88 (94)	86 (92)	82 (92)	82 (92)
Sandfish	90 (98)	84 (96)	80 (88)	86 (92)	86 (92)	84 (90)	80 (90)	80 (90)
Dwarf sungazer	77 (90)	75 (90)	66 (85)	72 (88)	72 (88)	72 (87)	66 (83)	64 (85)
Grass lizard	84 (100)	74 (94)	74 (90)	78 (92)	78 (92)	76 (90)	76 (90)	74 (88)
Green anole	-	84 (94)	80 (90)	84 (92)	84 (92)	82 (90)	80 (94)	78 (92)
Savannah monitor	50 (60)	49 (58)	43 (56)	47 (58)	47 (58)	47 (58)	49 (62)	45 (62)
Rat snake	86 (96)	98 (98)	82 (94)	82 (92)	82 (92)	80 (90)	76 (90)	72 (90)
King cobra	84 (94)	-	82 (94)	84 (94)	84 (94)	82 (92)	74 (88)	72 (90)
Other Reptiles								

American alligator	80 (90)	82 (94)	-	94 (96)	94 (96)	90 (92)	82 (88)	80 (88)
western painted turtle	84 (92)	84 (94)	94 (96)	-	100 (100)	96 (96)	86 (90)	84 (92)
Aves								
Chicken	84 (92)	84 (94)	94 (96)	100 (100)	-	96 (96)	86 (90)	84 (92)
Zebra finch	82 (90)	82 (92)	90 (92)	96 (96)	96 (96)	-	88 (90)	86 (92)
Mammalia								
Human	80 (94)	74 (88)	82 (88)	86 (90)	86 (88)	88 (90)	-	96 (100)
Syrian hamster	78 (92)	72 (90)	80 (88)	84 (92)	84 (92)	86 (92)	96 (100)	-

Discussion

I firstly focused on the deduced amino acid sequences of leopard gecko and Japanese gecko insulin, both of which were previously isolated in our lab. Compared with human insulin (comprising B-chain and A-chain), the deduced amino acid sequence of leopard gecko insulin showed 13 substitutions among 51 amino acids comprising the molecule. This was similar to previously reported squamate insulin (species of Iguania and Serpentes), which showed 9-13 substitutions compared with human insulin (Conlon, 2001). In contrast, Japanese gecko mature insulin showed a markedly higher number of substitution; it had 22 substitutions, which amounted to more than 40% of all amino acids comprising insulin. Number of substitutions equivalent to this had not been reported in any other reptiles nor aves. Even among tetrapods, only hystricomorph rodents are comparable (Conlon, 2001).

Among seven families comprising the infraorder Gekkota, the leopard gecko and Japanese gecko belong to Eublepharidae and Gekkonidae, respectively. Eublepharidae is phylogenetically basal to Gekkonidae and other two families, namely Phyllodactylidae and Sphaerodactylidae (Pyron et al., 2013). Considering such relationship between these four families, I hypothesized that amino acid substitution of insulin had accumulated after the common ancestor of Gekkonidae, Phyllodactylidae and Sphaerodactylidae separated from the lineage leading to Eublepharidae. I examined species of these four families to elucidate the evolutionary process that resulted in the high substitution number in Japanese gecko insulin. I especially focused on Gekkonidae and examined 15 species representatives of 9 genera, because the Japanese gecko belongs to this family. In addition to species of the infraorder Gekkota, 6 species of Unidentata, a sister group of Gekkota and Pygopodinae (legless lizards) (Vidal and Hedges, 2005), were also examined. Therefore, insulin was isolated from a total of 27 squamate species in the current study. In addition, preproinsulin sequences of the green anole and king cobra were obtained from the GenBank Database. Thus, in total 29 squamate species were investigated.

When amino acid sequences of insulin (B-chain and A-chain) were compared with that of human, particularly low identity values were observed in the following four lineages, not only of Gekkota, but also of Unidentata: the family Gekkonidae, yellow-headed gecko, savannah monitor, and dwarf sungazer. These lineages showed less than 66% identity values, while other squamates showed more than 74%. Some species of these lineages had deletions at the C-terminal region of the B-chain, and some had extensions of the A-chain at its C-terminus. This is the first report of deletions and extensions in reptilian insulin. Deletions and extensions are rarely observed in other tetrapods, with some amphibians and hystricomorph rodents being the only exception (Conlon, 2000, 2001). Although the deletions at C-terminal region of B-chain was reported to have little effect on insulin's functions (Nakagawa and Tager, 1986), the extension of A-chain at its C-terminus may affect by disrupting formation of ionic bond between $_{Arg}$B22 and $_{Asn}$A21 (Conlon et al., 1997), which is important for maintaining biological function (Markussen et al., 1988).

In order to surmise the evolutionary process that resulted in the diversity of squamate insulin sequences, I analyzed changes in the selection pressure imposed on the evolution of the *ins* by estimating the values of the nonsynonymous-to-synonymous substitution ratio, or ω in short.

I first tested heterogeneity between Squamata clade and other reptilian and avian lineages. Insulin was reported to be less effective for blood glucose regulation in squamates than in other reptiles (Dessauer, 1970; Matty, 1966). Thus I hypothesized that squamate *ins* had evolved under relaxed negative selection pressure, being liberated from evolutionary constraint (Tourasse and Li, 2000) that had maintained the hypoglycemic function. The Branch Model analysis suggested a higher ω value for the Squamata clade than the other reptilian and avian lineages. On the other hand, neither the Branch-Site Model nor Clade model C detected any positively selected sites. These results suggested the relaxation negative selection pressure rather than the imposition of positive selection in the Squamata clade.

I next tested following lineages among the Squamata clade: the family Gekkonidae, the yellow-headed gecko, and the savannah monitor. These lineages had substitution of $_{His}$B10 (Table 1-9). The substitution of this site was also reported from hystricomorph rodents, and was attributed to an increased number of amino acid substitution in insulin among vertebrates (Beintema and Campagne, 1987; Conlon, 2001). Thus, I hypothesized similar process in these squamates. In addition to these lineages, the dwarf sungazer was also tested. Although this species conserved His at B10, the identity value (66% identity with human insulin) was almost equivalent to those of Gekkonidae species. The Branch Model analyses suggested increased ω values for lineages with $_{His}$B10 substitution. On the other hand, the increased ω was not supported for the dwarf sungazer. This species was excluded from further analyses with Branch-Site Model and CmC, as I had no clear hypothesis that can explain the increased amino acid substitution in this species. However, it should be noted that deletions and extensions at the C-terminus of B- and A-chains were underestimated because these regions were excluded from the analyses by PAML.

Since the interaction between histidine at B10 and the Zn^{2+} ion is necessary for the aggregation of (pro)insulin monomers into a hexamer in β-cell granules, the substitution of this site results in the loss of the hexameric state (Beintema and Campagne, 1987). Hexamerization is important for insulin for several reasons: thermodynamic stabilization, improved processing efficiency by regulating the exposure of residues to processing enzymes (Dunn, 2005), and the efficient delivery of insulin molecules to peripheral tissues by protecting them from hepatic clearance (Tamaki *et al.*, 2013). Considering these stabilizing effect of hexamerization and its physiological importance, it has been suggested that the increased number of substitution in hystricomorph insulin was the result of adaptive evolution to make monomeric insulin more stable and tolerant to degradation. For example, residues such as $_{Gln}$B14, $_{Leu}$B17 and $_{Gly}$B20, which are buried inside the hexamer, but are exposed outside the monomer, are substituted with more hydrophilic amino acids in some hystricomorph species. The

substitution of trypsin-sensitive $_{Arg}$B22 with Asp or Ser, which is observed in several hystricomorphs, makes monomers more tolerant to enzymatic degradation (Blundell and Wood, 1975). Similar substitutions have also been observed in some of the squamates examined in the present study. However, since these substitutions were not common features for either of hystricomorph rodents nor squamate species (Conlon, 2001; Opazo et al., 2005), stabilization of the monomer alone cannot fully explain the increased number of substitution.

Another explanation for the increased number of the substitution is enhancement of mitogenicity (stimulatory effect of cellular growth and proliferation). Insulin binds to its receptor through two surfaces, one comprising B12, B16, B23, B24, B25, B26, A1, A2, A3, A5, A19 and A21 and the other comprising B10, B13, B17, A12, A13, and A17. The former one, which was presented to be the binding surface firstly in the history of studies on insulin, is called “classical binding surface”. The latter one, which was identified later, is called “novel binding surface” (De Meyts et al., 2000). The substitutions of sites comprising either of these surfaces may reduce binding affinity to the insulin receptor (Kristensen et al., 1997; Schaffer, 1994). However, the substitution of some sites comprising the “novel binding surface” has been shown to increase the mitogenic potency of insulin relative to binding affinity to the insulin receptor, while metabolic potency (such as the anabolic regulation of carbohydrates and lipids) relative to binding affinity is decreased (De Meyts, 1994). Such features have been reported from human analogues with a substitution at $_{Leu}$B17 or $_{Leu}$A13 (De Meyts, 1994). The reason for this remains currently unclear; however, conformational changes in the ligand-bound insulin receptor (De Meyts, 1994) or prolongation of the period in which the ligand-bound receptor is in an active state (Kiselyov *et al.*, 2009) have been suggested, resulting in selective activation of a mitogenic pathway rather than a metabolic pathway. The substitution of B17 and/or A13 has been reported in hystricomorph rodents and jawless fishes (Atlantic hagfish and river lamprey), and has been attributed to the enhanced mitogenicity of their insulin (De Meyts, 1994; King and Kahn, 1981;

Sajid *et al.*, 2009). In squamate insulin, the substitutions of B17 or A13 was detected in 5 Gekkonidae species, the yellow-headed gecko, savannah monitor, and Japanese grass lizard. It has also been reported that the substitution of $_{His}$B10 with Asp, Glu, and Gln increases mitogenicity in human analogues. Considering the cell types used in the previous study, this phenomenon appears to be caused by increased affinity for the IGF-1 receptor, which mainly activates a mitogenic pathway (Glendorf *et al.*, 2012). Among these substitutions, $_{His}$B10$_{Gln}$ was observed in all Gekkonidae species examined in this study.

Further analyses by PAML estimated that some of these substitutions occurred under positive selection pressure. CmC-3ratio suggested B10 and B17 as the PSS for the ancestor of Gekkonidae, while the Branch-Site Model suggested A13 as the PSS for the savannah monitor (Table1-5, 1-6). These results suggested that squamate *ins* had undergone adaptive evolution, possibly in an attempt to enhance mitogenicity. It should be noted that the CmC-3ratio of Gekkonidae was not a better fit than CmC-2ratio. Furthermore, some of the Branch-Site tests were performed without data from close-relatives of the yellow-headed gecko or savannah monitor. Thus, the evolutionary process of these lineages, together with the locations of PSS, should be validated precisely with more data from the genus *Gonatodes* (the genus to which the yellow-headed gecko belongs) and Anguimorpha (the group to which the savannah monitor belongs). However, these results are still important because the *ins* of hystricomorph rodents was suggested to have undergone a similar evolutionary process. Opazo and his co-workers analyzed the molecular evolution of hystricomorph *ins* by PAML, and concluded that B17 and A13 were under positive selection pressure (Opazo et al., 2005). Considering the unique features of growth regulation by IGF-1 and IGF-2 in hystricomorph rodents, they implied that hystricomoroh *ins* had evolved in relation to growth regulation by these hormones. Interestingly, squamate IGF-1 was also reported to have acquired a distinctive structure among vertebrates through a unique evolutionary process (Sparkman et al., 2012). These previous findings for insulin and IGFs,

together with our results for squamate insulin, strongly suggest that a relationship exists between the molecular evolution of *ins* and IGFs in squamates.

In addition to the relationship with IGFs, physiological regulation unique to squamates also suggests that alterations in the biological properties of insulin were adaptive characteristics. It is well known that the organs of squamates, especially those involved in digestion and nutrients assimilation, undergo marked morphological changes in response to food availability (Buono et al., 2006; Godet et al., 1984; Secor et al., 1994; Secor and Diamond, 1998). Secor suggested that such morphological changes were regulated by gastrointestinal hormones, including insulin, because they are secreted in response to food intake (Secor, 2008). As already explained, the importance of insulin for the regulation of blood glucose levels in squamates is suggested to have been reduced (Dessauer, 1970; Matty, 1966). If this is the case, enhancing mitogenic potency at the cost of metabolic potency (such as hypoglycemic effect) makes sense. Such changes in biological properties are exactly what has been reported from insulin with substitutions in the "novel binding surface".

For the validation of these hypotheses, however, further studies on ligand-receptor interactions are necessary. I currently have no data regarding the structure of squamate insulin receptor or IGF-1 receptor, both of which are determinant factors for the endogenous effects of insulin. In addition, the effects of substitutions in the "classical binding surface" need to be validated. Some of the sites comprising this surface have a large impact on the affinity of insulin for its receptor, and are well conserved among vertebrates, including hystricomorph rodents (Conlon, 2001). However, some squamate species had substitutions in this surface, which may have impaired binding affinity. The substitution of $_{Tyr}$B16 with the amino acids observed in Gekkonidae insulin (Ala, Asn, Ile, Leu, and Val) was previously shown to impair binding affinity of human analogues (Glendorf *et al.*, 2008). $_{Ile}$A2$_{Val}$, which was also observed in the flat-tailed house gecko (*Hemidactylus platyurus*), was attributed to the low affinity of owl monkey (*Aotus tivirgatus*) insulin for the human insulin receptor

(Seino *et al.*, 1987). Phe at B24 is of critical importance not only for receptor binding (Hua *et al.*, 2009; Mirmira and Tager, 1998), but also for the self-assembly and stability of insulin (Pandyarajan *et al.*, 2014). This site had been considered as an “invariant site” among vertebrates (Conlon, 2001), which was not the case for the yellow-headed gecko nor savannah monitor. They had Val and Ile at this site, respectively, both of which impair affinity for the insulin receptor in human analogues (Pandyarajan *et al.*, 2014). Since such substitutions in the “classical binding surface” are rarely, or not observed among other vertebrate species, the properties of squamate insulin may be different from those of the insulin from any other vertebrates. The combined effect of substitutions in the “classical binding surface” with those in the “novel binding surface” is also of interest, especially for further elucidation of relationship between the molecular structure and the biological activities of insulin.

Another interesting topic for further examination is insulin from the Japanese gecko and hokou gecko, both of which have distinct amino acid substitutions from other species of the genus *Gekko* (a genus of Gekkonidae). When B- and A-chains were compared with that of the tokay gecko (*Gekko gecko*), the Japanese gecko and hokou gecko had 12 and 9 amino acid substitutions, respectively. In contrast, other three species of *Gekko* examined in this study had only 5 substitutions or less. Phylogenetically, the Japanese gecko and hokou gecko, together with several other close-relatives, are referred to as the “*Gekko japonicus* group” and distinguished from other species of *Gekko*. In contrast to tropical habitats of most species of *Gekko*, this group inhabits subtropical and/or temperate zones (Rösler et al., 2011). Since seasonal climate changes and subsequent changes in food availability are harsh in these climate zones, this group is likely to have adapted to these harsh circumstances. It is possible that the roles of insulin, whether anabolic regulation or growth promotion, contributed to these adaptations. It has also been reported that molecular evolution of insulin can contribute to environmental adaptation of vertebrates. Conlon suggested that aberrant structure and function of the wood frog (*Rana sylvatica*) insulin is involved in acquisition of freeze tolerance, thereby enabling this

species to live in the north of the Arctic Circle (Conlon et al., 1998). Thus, I presumed that distinct amino acid substitutions of insulin in the two *Gekko* species had been involved in environmental adaptation. This was supported by the Branch Model analyses. Six-ratio 3, which was a better fit than the Five-ratio model, estimated the largest value of ω for these two species among all the species examined. This value, $\omega_{\text{Japanese gecko group}} = 2.44$, is a strong indicator of positive selection pressure, suggesting the possible adaptive evolution of *ins* in these species. Unfortunately, the characteristics acquired could not be determined based on the results of the present study. Although the Branch-Site Model suggested B4 as the PSS for these two species (Table 1-5), this result does not appear to be accurate. Among the *Gekko* species examined, only the Japanese gecko had Tyr at this site, while others, including the hokou gecko, had Gln (Figure 1-4). Previous studies reported that the estimation of an extremely high ω value for single species (the Japanese gecko, in this case) among the foreground clade resulted in false positive estimations (Yang, 2013b). In the future studies, what properties were acquired should be determined with biochemical procedures. This is an interesting topic to examine, especially as a representative case in which the evolution of *ins* may have contributed to the radiation and environmental adaptation of squamate species.

Conclusion

I herein demonstrated that the amino acid sequences of squamate insulin were more divergent than previously regarded (Conlon, 2000, 2001). The locations of the substitutions, combined with the evolutionary process estimated by PAML CODEML, suggests that this divergence may have resulted from adaptive evolution to optimize the biological functions of insulin to physiological regulation in squamates.

Of particular interest is the value of dN/dS ratio estimated for the species of *Gekko japonicus* group. The value, $\omega = 2.44$, strongly suggests adaptive evolution of insulin in this group after divergence from other groups of *Gekko*. Considering the geographical distribution of this group among *Gekko*, such high value of dN/dS ratio may be the result of environmental adaptation to subtropical and temperate zones, in which climate and food availability fluctuate seasonally.

Further analyses of evolutionary processes, interactions with receptors, and endogenous effects are likely to yield a meaningful explanation for the relationship between the molecular evolution of insulin and the physiology of squamates. It may also open the way for the elucidation of the adaptive mechanisms responsible for prosperity of squamates.

Chapter 2

Analyses on the expressional change of key regulators for hepatic glucose and lipid metabolism in the Japanese gecko

第２章については、５年以内に雑誌等で刊行予定のため、非公開。

.

.

General discussion

As explained in general introduction, energy metabolism is of critical importance for the fitness of organisms. This is because adaptation to a new environment is achieved by acquisition of novel life-history traits, which in turn requires remodeling of metabolic regulation. Exceptional diversity of the order Squamata among reptiles suggests that this group is the most successful in such remodeling. Therefore, the mechanisms that regulate energy metabolism of squamates is an interesting topic for the better understanding about adaptivity of squamates. Considering their flexibility in acquiring novel life-history traits, it may also provide meaningful insights into the evolution of life-history traits.

In this research, I aimed at establishing the basic viewpoints for investigation of metabolic regulation of squamates at the molecular level. Considering the importance of coding sequences and transcriptional regulation of genes for the acquisition of novel phenotypes (Little and Seebacher, 2016), I focused on variation of insulin sequences and transcriptional regulation of key regulators of glucose and TAG metabolism.

Chapter 1

In chapter 1, I explored evolutionary background of insulin among squamate lineages. Molecular characterization of insulin from multiple lineages revealed that insulin sequences of squamates are much more divergent than previously regarded. The values of dN/dS ratio (ω) calculated based on the nucleotide sequences suggest that squamate *ins* (insulin gene) have evolved under relaxed negative (purifying) selection. Among squamates, some with lineage-specific increase of ω (e.g. Japanese gecko) have possibly evolved under positive selection. This in turn suggests that sequence divergence is the result of processes that have increased fitness of each lineage (Figure D-1).

For the validation of this hypothesis, investigation into biochemical character of insulin, together with its *in vivo* effect is necessary so as to determine the exact contribution of sequence divergence to fitness. This is especially true to the Japanese gecko. The value of ω calculated for this species well

exceeded 1, implying that *ins* of this species had undergone adaptive evolution. Unfortunately, what functional change had occurred to Japanese gecko insulin could not be inferred in this study. This is because this species did not have substitution at sites with critical importance for hormonal function (although this species substitutes His at B10 with Gln, a site with critical importance for hexamerization, this seems to be an ancestral character because all Gekkonidae species have Gln at this site). It seems that additive effect of substitutions unique to this species has brought functional alteration, which is possibly advantageous for this species. Since this species is likely to have evolved from a tropical ancestor and subsequently adapted to temperate climate zone (Ikeuchi, 2004; Rösler et al., 2011), molecular evolution of insulin seems to be related to this process. To date, involvement of vertebrate insulin in environmental adaptation has not been reported, except a report of possible involvement in freeze tolerance of an anuran species (Conlon et al., 1998). Therefore, understanding about physiological function of Japanese gecko insulin may enhance current understanding about not only physiological, but also evolutionary aspects of vertebrate insulin.

Chapter 2

第２章の内容を含むため、非公開。

第２章の内容を含むため、
非公開。

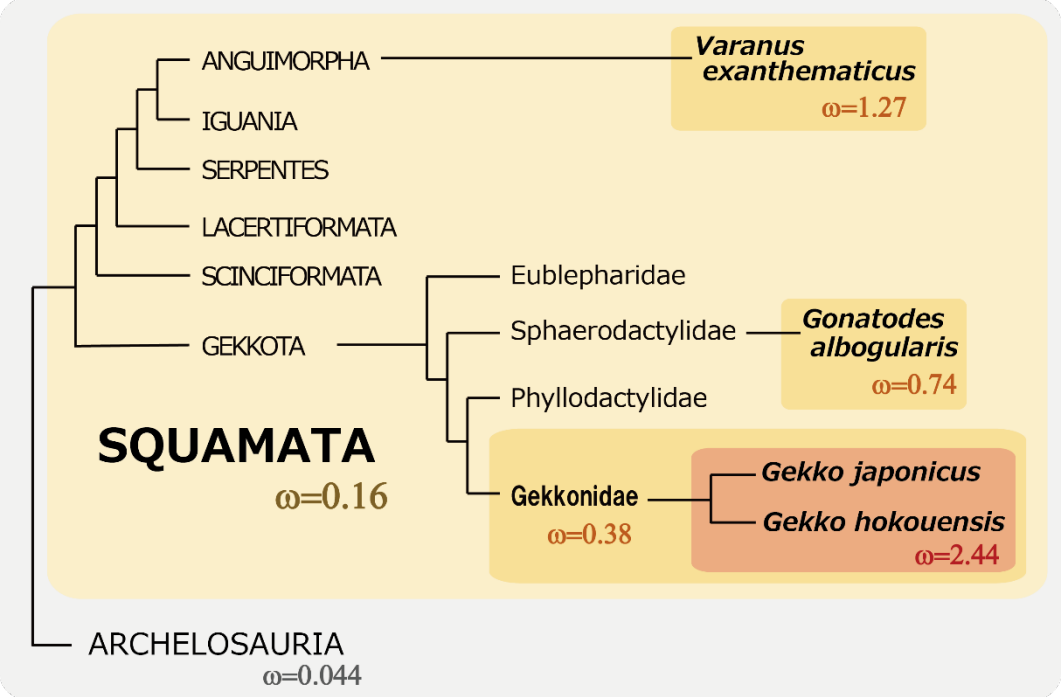


Figure D-1. Changes of selection pressure imposed on insulin gene among reptilian lineages

The values of dN/dS ratio estimated for each lineage by PAML Branch Models are denoted below taxon names. Larger value of Squamata clade (ω=0.16) than that of archelosaurian (a group comprising chelonians, crocodilians and birds) clade (ω=0.044) suggests relaxation of negative (purifying) selection. Further increase among squamate linaeges suggests lineage-specific relaxation of negative selection or imposition of positive selection, both of which may have contributed to lineage-specific evolution of insulin. Notably, the value estimated for the Japanese gecko (*Gekko japonicus*) and hokou gecko (*G. hokouensis*) well exceeded 1 (ω=2.44), strongly implying adaptive evolution under positive selection.

第２章の内容を含むため、非公開。

Figure D-2

Conclusion

In this thesis, I aimed at establishing viewpoints for investigation of molecular mechanisms that regulate energy metabolism of squamates as a means to understand background of their prosperity. Since coding sequences and transcriptional regulation of genes have been reported to determine phenotypes of organisms (Little and Seebacher, 2016), I focused on variation of insulin sequences among squamates and transcriptional regulation of key metabolic factors in one of squamate species, the Japanese gecko.

In chapter 1, evolutionary background of squamate insulin was investigated based on the nucleotide sequences isolated from 27 squamate species. The deduced amino acid sequences revealed sequence divergence among squamate lineages, which had been overlooked by previous researches. Estimation of selection pressure based on the values of nonsynonymous-to-synonymous substitution ratio suggested relaxed negative selection on overall squamates. Furthermore, insulin genes of several lineages were suggested to have undergone adaptive evolution. These results imply that lineage-specific molecular evolution of insulin has contributed to adaptive radiation of squamates.

One of the lineages whose insulin was suggested to have undergone adaptive evolution was the Japanese gecko, whose geographical distribution and seasonal cycle of reproduction suggest adaptation to environment in temperate climate (Ikeuchi, 2004; Rösler et al., 2011).

第 2 章の内容を含むため、非公開。

第２章の内容を含むため、
非公開。

第２章の内容を含むため、非公開。

Figure C-1. Overview of this thesis

In this thesis, I aimed at establishing viewpoints from which to investigate molecular mechanisms regulating energy metabolism of squamates, whose flexibility presumably contributes to environmental adaptation of this group. In chapter 1, molecular evolution of insulin was suggested to have contributed to adaptive radiation of squamates through the acquisition of unique amino acid sequences and physiological functions. One of squamate species with possible adaptive evolution of insulin was the Japanese gecko, whose geographical distribution and annual gametogenic cycle imply tropical origin and subsequent adaptation to temperate environment. In chapter 2, hepatic gene expression of this species was investigated. The results obtained suggest that seasonal reproduction of this species is fueled by trade-off relationship between oogenesis and maternal activity, possibly through lineage- or species-specific transcriptional regulation of key metabolic factors. All these together, arrangement of both coding sequences and transcriptional regulation of genes are likely to contribute to environmental adaptation of squamates.

*1 Detailed in Figure D-1. *2 Detailed in Figure D-2.

Acknowledgements

I express my sincere gratitude to Associate Professor Min Kyun Park, Department of Biological Sciences, Graduate School of Science, The University of Tokyo, for constant guidance, encouragement, and valuable discussion during the course of my study.

I am grateful to Prof. Yoshitaka Oka and Associate Prof. Shinji Kanda for helpful advice and encouragement. I also thank Dr. Yasuhisa Akazome, Dr. Misaki Miyoshi, Dr. Yuko Mochizuki, Dr. Chie Umatani, Dr. Shigeo Otake, Dr. Masaharu Hasebe and Ms. Miho Kyokuwa, Dept. of Biological Sciences, Grad. Sch. of Science, The University of Tokyo, for valuable discussion throughout this study.

I am sincerely grateful to Prof. Yoshitaka Oka, Prof. Takeo Kubo and Associate Prof. Shinji Kanda, Department of Biological Sciences, Graduate School of Science, The University of Tokyo, and Prof. Susumu Hyodo, Atmosphere and Ocean Research Institute, The University of Tokyo, for critical reading of the present thesis and valuable discussion.

Warm thanks to my grandparents, Chizuko and Shirou Yamagishi for their encouragement.

Part of this work was supported by the Sasakawa Scientific Research Grant from The Japan Science Society (28-435) and a grant-in-aid for scientific research from the Ministry of Education, Science, Sports and Culture of Japan (26440162 and 23570069 to M.K.P.).

Part of chapter 1 has been published as “Yamagishi, G., et al. Molecular characterization of insulin from squamate reptiles reveals sequence diversity and possible adaptive evolution. Gen. Comp. Endocrinol. (2016)” (http://dx.doi.org/10.1016/j.ygcen.2015.08.021).

Lastly, I’d like to express my special thanks to my mother, Yumiko Yamagishi, Ph.D. I sincerely appreciate her patience, assistance and encouragement throughout the entire period of my study.

References

Aitken, A., 2012. DNA Purification from Agarose Gel. Distributed by the author <http://www.nhm.ac.uk/resources-rx/files/phenolchloroformgelextraction-118485.pdf>.

Andersen, C.L., Jensen, J.L., Ørntoft, T.F., 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res. 64, 5245-5250.

Andrade, D.V., 2016. Temperature effects on the metabolism of amphibians and reptiles: Caveats and recommendations. In: Andrade, D.V., Bevier, C.R., Carvalho, J.E. (Eds.), Amphibian and Reptile Adaptations to the Environment: Interplay Between Physiology and Behavior. CRC Press, Boca Raton, pp. 129–154.

Atshaves, B.P., Martin, G.G., Hostetler, H.A., McIntosh, A.L., Kier, A.B., Schroeder, F., 2010. Liver fatty acid-binding protein and obesity. J. Nutr. Biochem. 21, 1015–1032.

Baillie, J.E.M., Griffiths, J., Turvey, S.T., Loh, J., Collen, B., 2010. Evolution Lost: Status and Trends of the World's Vertebrates. Zoological Society of London, London.

Bali, D.S., Chen, Y.T., Austin, S., 2006 [Updated 2016 Aug 25]. Glycogen Storage Disease Type I. In: Adam, M.P., Ardinger, H.H., Pagon, R.A. (Eds.), GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2018. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1312/

Beintema, J.J., Campagne, R.N., 1987. Molecular evolution of rodent insulins. Mol. Biol. Evol. 4, 10–18.

Benton, M., 2014. Vertebrate Palaeontology, 4th Edition. John Wiley & Sons, Hoboken.

Berg J.M., Tymoczko, J.L., Stryer, L., 2002. Biochemistry. 5th edition. W. H. Freeman, New York. Available from: https://www.ncbi.nlm.nih.gov/books/NBK21154/

Bertile, F., Raclot, T., 2011. ATGL and HSL are not coordinately regulated in response to fuel partitioning in fasted rats. J. Nutr. Biochem. 22, 372–379.

Blundell, T.L., Wood, S.P., 1975. Is the evolution of insulin Darwinian or due to selectively neutral mutation? Nature 257, 197–203.

Borodina, T.A., Lehrach, H., Soldatov, A.V, 2003. DNA purification on homemade silica spin-columns. Anal. Biochem. 321, 135–137.

Brudno, M., Do, C.B., Cooper, G.M., Kim, M.F., Davydov, E., Program, N.C.S., Green, E.D., Sidow, A., Batzoglou, S., 2003. LAGAN and Multi-LAGAN: Efficient Tools for Large-Scale Multiple Alignment of Genomic DNA. Genome Res. 13, 721–731.

Buono, S., Odierna, G. 2nd, Putti, R., 2006. Morphology of the pancreas of some species belonging to the genera *Phelsuma* and *Gecko* (family Gekkonidae): evidence of apoptotic process during

the seasonal cycle. Anat. Embryol. (Berl). 211, 413–421.

Cahill, G.F.J., 2006. Fuel metabolism in starvation. Annu. Rev. Nutr. 26, 1–22.

Castoe, T.A., de Koning, A.P.J., Hall, K.T., Card, D.C., Schield, D.R., Fujita, M.K., Ruggiero, R.P., Degner, J.F., Daza, J.M., Gu, W., Reyes-Velasco, J., Shaney, K.J., Castoe, J.M., Fox, S.E., Poole, A.W., Polanco, D., Dobry, J., Vandewege, M.W., Li, Q., Schott, R.K., Kapusta, A., Minx, P., Feschotte, C., Uetz, P., Ray, D.A., Hoffmann, F.G., Bogden, R., Smith, E.N., Chang, B.S.W., Vonk, F.J., Casewell, N.R., Henkel, C. V, Richardson, M.K., Mackessy, S.P., Bronikowski, A.M., Yandell, M., Warren, W.C., Secor, S.M., Pollock, D.D., 2013. The Burmese python genome reveals the molecular basis for extreme adaptation in snakes. Proc. Natl. Acad. Sci. U.S.A. 110, 20645–20650.

Chan, S.J., Cao, Q.P., Steiner, D.F., 1990. Evolution of the insulin superfamily: cloning of a hybrid insulin/insulin-like growth factor cDNA from amphioxus. Proc. Natl. Acad. Sci. U.S.A. 87, 9319–9323.

Chou, J.Y., Jun, H.S., Mansfield, B.C., 2010. Glycogen storage disease type I and G6Pase-beta deficiency: etiology and therapy. Nat. Rev. Endocrinol. 6, 676–688.

Chou, J.Y., Mansfield, B.C., 2008. Mutations in the glucose-6-phosphatase-alpha (G6PC) gene that cause type Ia glycogen storage disease. Hum. Mutat. 29, 921–930.

Conlon, J.M., 2000. Molecular Evolution of Insulin in Non-Mammalian Vertebrates. Am. Zool. 40, 200–212.

Conlon, J.M., 2001. Evolution of the insulin molecule: insights into structure-activity and phylogenetic relationships. Peptides 22, 1183–1193.

Conlon, J.M., Platz, J.E., Nielsen, P.F., Vaudry, H., Vallarino, M., 1997. Primary structure of insulin from the African lungfish, *Protopterus annectens*. Gen. Comp. Endocrinol. 107, 421–427.

Conlon, J.M., Yano, K., Chartrel, N., Vaudry, H., Storey, K.B., 1998. Freeze tolerance in the wood frog *Rana sylvatica* is associated with unusual structural features in insulin but not in glucagon. J. Mol. Endocrinol. 21, 153–159.

Crawford, N.G., Parham, J.F., Sellas, A.B., Faircloth, B.C., Glenn, T.C., Papenfuss, T.J., Henderson, J.B., Hansen, M.H., Simison, W.B., 2015. A phylogenomic analysis of turtles. Mol. Phylogenet. Evol. 83, 250–257.

Croniger, C.M., Olswang, Y., Reshef, L., Kalhan, S.C., Tilghman, S.M., Hanson, R.W., 2002. Phosphoenolpyruvate carboxykinase revisited: Insights into its metabolic role. Biochem. Mol. Biol. Educ. 30, 14–20.

Davidson, H.W., 2004. (Pro)Insulin processing: a historical perspective. Cell Biochem. Biophys. 40, 143–158.

De Meyts, P., 1994. The structural basis of insulin and insulin-like growth factor-I receptor binding and negative co-operativity, and its relevance to mitogenic versus metabolic signalling.

Diabetologia 37 Suppl 2, S135-48.

De Meyts, P., Sajid, W., Palsgaard, J., Theede, A.M, Gauguin, L., Aladdin, H., Whittaker, J., 2000. Insulin and IGF-I Receptor Structure and Binding Mechanism. In: Madame Curie Bioscience Database [Internet]: Landes Bioscience, Austin, Texas. <http://www.ncbi.nlm.nih.gov/books/NBK6192/>

Deeley, R.G., Mullinix, D.P., Wetekam, W., Kronenberg, H.M., Meyers, M., Eldridge, J.D., Goldberger, R.F., 1975. Vitellogenin synthesis in the avian liver. Vitellogenin is the precursor of the egg yolk phosphoproteins. J. Biol. Chem. 250, 9060–9066.

Denzer, C., Thiere, D., Muche, R., Koenig, W., Mayer, H., Kratzer, W., Wabitsch, M., 2009. Gender-specific prevalences of fatty liver in obese children and adolescents: roles of body fat distribution, sex steroids, and insulin resistance. J. Clin. Endocrinol. Metab. 94, 3872–3881.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Guindon, S., Lefort, V., Lescot, M., Claverie, J.-M., Gascuel, O., 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 36, W465-9.

Derickson, W.K., 1976. Lipid Storage and Utilization in Reptiles. Am. Zool. 16, 711–723.

Dessauer, H.C., 1953. Hibernation of the lizard *Anolis carolinensis*. Proc. Soc. Exp. Biol. Med. 82, 351–353.

Dessauer, H.C., 1970. Blood chemistry of reptiles: physiological and evolutionary aspects. Biol. Reptil. 3, 1–72.

Di Maggio, A. 3rd, Dessauer, H.C., 1963. Seasonal changes in glucose tolerance and glycogen disposition in a lizard. Am. J. Physiol. 204, 677–680.

Dunn, M.F., 2005. Zinc-ligand interactions modulate assembly and stability of the insulin hexamer - a review. Biometals 18, 295–303.

Frazer, K.A., Pachter, L., Poliakov, A., Rubin, E.M., Dubchak, I., 2004. VISTA: computational tools for comparative genomics. Nucleic Acids Res. 32, W273-9.

Freeman, M.W., Walford, G.A., 2016. Lipoprotein Metabolism and the Treatment of Lipid Disorders. In: Jameson, J.L., De Groot, L.J., de Kretser, D.M., Giudice, L.C., Grossman, A.B., Melmed, S., Potts, J.T., Weir, G.C. (Eds.), Endocrinology: Adult and Pediatric (Seventh Edition). Saunders, Philadelphia, pp. 715-736.e7.

Gillooly, J.F., McCoy, M.W., 2014. Brain size varies with temperature in vertebrates. PeerJ. 2, e301.

Glendorf, T., Knudsen, L., Stidsen, C.E., Hansen, B.F., Hegelund, A.C., Sørensen, A.R., Nishimura, E., Kjeldsen, T., 2012. Systematic evaluation of the metabolic to mitogenic potency ratio for B10-substituted insulin analogues. PLoS One 7, e29198.

Glendorf, T., Sørensen, A.R., Nishimura, E., Pettersson, I., Kjeldsen, T., 2008. Importance of the solvent-exposed residues of the insulin B chain alpha-helix for receptor binding. Biochemistry 47, 4743–4751.

Godet, R., Mattei, X., Dupe-Godet, M., 1984. Alterations of endocrine pancreas B cells in a sahelian reptile (*Varanus exanthematicus*) during starvation. J. Morphol. 180, 173–180.

Goldberg, S., 2009. General notes: Reproduction in Smith's green-eyed gecko, *Gekko smithii* (Squamata: Gekkonidae). Texas J. Sci. 61, 225–228.

Hanson, R.W., Garber, A.J., 1972. Phosphoenolpyruvate carboxykinase. I. Its role in gluconeogenesis. Am. J. Clin. Nutr. 25, 1010–1021.

Hara, A., Hiramatsu, N., Fujita, T., 2016. Vitellogenesis and choriogenesis in fishes. Fish. Sci. 82, 187–202.

Hayward, A., Gillooly, J.F., 2011. The Cost of Sex: Quantifying Energetic Investment in Gamete Production by Males and Females. PLoS One 6, e16557.

Herczeg, G., Kovács, T., Hettyey, A., Merilä, J., 2003. To thermoconform or thermoregulate? An assessment of thermoregulation opportunities for the lizard *Zootoca vivipara* in the subarctic. Polar Biol. 26, 486–490.

Hisai, N., 1997. Ecological Observation on Hibernation of Japanese Common Gecko, *Gekko japonicus* (Dumeril et Bibron). Misc. reports Natl. Park Nat. Study 28, 1–5 (in Japanese).

Hua, Q., Xu, B., Huang, K., Hu, S.Q., Nakagawa, S., Jia, W., Wang, S., Whittaker, J., Katsoyannis, P.G., Weiss, M.A., 2009. Enhancing the activity of a protein by stereospecific unfolding: conformational life cycle of insulin and its evolutionary origins. J. Biol. Chem. 284, 14586–14596.

Humbel, R.E., 1990. Insulin-like growth factors I and II. Eur. J. Biochem. 190, 445–462.

Hutton, J.C., O'Brien, R.M., 2009. Glucose-6-phosphatase catalytic subunit gene family. J. Biol. Chem. 284, 29241–29245.

Ikeuchi, I., 2004. Male and Female Reproductive Cycles of the Japanese Gecko, *Gekko japonicus*, in Kyoto, Japan. J. Herpetol. 38, 269–274.

Inoue, H., Nojima, H., Okayama, H., 1990. High efficiency transformation of *Escherichia coli* with plasmids. Gene 96, 23–28.

Iwabe, N., Hara, Y., Kumazawa, Y., Shibamoto, K., Saito, Y., Miyata, T., Katoh, K., 2005. Sister group relationship of turtles to the bird-crocodilian clade revealed by nuclear DNA-coded proteins. Mol Biol Evol. 22, 810-813.

Ji, X., Wang, P., 1990. Annual cycles of lipid contents and caloric values of carcass and some organs of the gecko, *Gekko japonicus*. Comp. Biochem. Physiol. Part A Physiol. 96, 267–271.

Ji, X., Wang, P., Hong, W.X., 1991. The reproductive Ecology of the Gecko, *Gekko japonicus*. Acta Zool. Sin. 37, 192–192.

Jono, T., Inui, Y., 2012. Secret Calls from under the Eaves: Acoustic Behavior of the Japanese House Gecko, *Gekko japonicus*. Copeia 2012, 145–149.

King, G.L., Kahn, C.R., 1981. Non-parallel evolution of metabolic and growth-promoting functions

of insulin. Nature 292, 644–646.

Kiselyov, V.V, Versteyhe, S., Gauguin, L., De Meyts, P., 2009. Harmonic oscillator model of the insulin and IGF1 receptors' allosteric binding and activation. Mol. Syst. Biol. 5, 243.

Kristensen, C., Kjeldsen, T., Wiberg, F.C., Schaffer, L., Hach, M., Havelund, S., Bass, J., Steiner, D.F., Andersen, A.S., 1997. Alanine scanning mutagenesis of insulin. J. Biol. Chem. 272, 12978–12983.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. 33, 1870–1874.

Lacy, E.L., Sheridan, M.A., Moore, M.C., 2002. Sex differences in lipid metabolism during reproduction in free-living tree lizards (*Urosaurus ornatus*). Gen. Comp. Endocrinol. 128, 180–192.

Lenzen, S., 2014. A fresh view of glycolysis and glucokinase regulation: history and current status. J. Biol. Chem. 289, 12189–12194.

Lillywhite, H.B., 2016. Behavior and physiology: An ecological and evolutionary viewpoint on the energy and water relations of ectothermic amphibians and reptiles. In: Andrade, D.V., Bevier, C.R., Carvalho, J.E. (Eds.), Amphibian and Reptile Adaptations to the Environment: Interplay Between Physiology and Behavior. CRC Press, Boca Raton, pp. 1–39.

Little, A., Seebacher, F., 2016. Acclimation, Acclimatization and Seasonal Variation in Amphibians and Reptiles. In: Andrade, D.V., Bevier, C.R., Carvalho, J.E. (Eds.), Amphibian and Reptile Adaptations to the Environment: Interplay Between Physiology and Behavior. CRC Press, Boca Raton, pp. 41-62.

Liu, Y., Zhou, Q., Wang, Y., Luo, L., Yang, J., Yang, L., Liu, M., Li, Y., Qian, T., Zheng, Y., Li, M., Li, J., Gu, Y., Han, Z., Xu, M., Wang, Y., Zhu, C., Yu, B., Yang, Y., Ding, F., Jiang, J., Yang, H., Gu, X., 2015. *Gekko japonicus* genome reveals evolution of adhesive toe pads and tail regeneration. Nat. Commun. 6, 10033.

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

Luo, F., Huang, W., Guo, Y., Ruan, G., Peng, R., Li, X., 2017. 17β-estradiol lowers triglycerides in adipocytes via estrogen receptor α and it may be attenuated by inflammation. Lipids Health Dis. 16, 182.

Marandel, L., Panserat, S., Plagnes-Juan, E., Arbenoits, E., Soengas, J.L., Bobe, J., 2017. Evolutionary history of glucose-6-phosphatase encoding genes in vertebrate lineages: towards a better understanding of the functions of multiple duplicates. BMC Genomics 18, 342.

Markussen, J., Diers, I., Hougaard, P., Langkjaer, L., Norris, K., Snel, L., Sørensen, A.R., Sørensen, E., Voigt, H.O., 1988. Soluble, prolonged-acting insulin derivatives. III. Degree of protraction,

crystallizability and chemical stability of insulins substituted in positions A21, B13, B23, B27 and B30. Protein Eng. 2, 157–166.

Marler, C.A., Walsberg, G., White, M.L., Moore, M., Marler, C.A., 1995. Increased energy expenditure due to increased territorial defense in male lizards after phenotypic manipulation. Behav. Ecol. Sociobiol. 37, 225–231.

Marnocha, E., Pollinger, J., Smith, T.B., 2011. Human-induced morphological shifts in an island lizard. Evol. Appl. 4, 388–396.

Mashek, D.G., 2013. Hepatic fatty acid trafficking: multiple forks in the road. Adv. Nutr. 4, 697–710.

Mashek, D.G., Khan, S.A., Sathyanarayan, A., Ploeger, J.M., Franklin, M.P., 2015. Hepatic lipid droplet biology: Getting to the root of fatty liver. Hepatology 62, 964–967.

Matty, A.J., 1966. Endocrine Glands in Lower Vertebrates. Int. Rev. Gen. Exp. Zool. 2, 43–138.

McCormack, J.E., Harvey, M.G., Faircloth, B.C., Crawford, N.G., Glenn, T.C., Brumfield, R.T., 2013. A phylogeny of birds based on over 1,500 loci collected by target enrichment and high-throughput sequencing. PLoS One 8, e54848.

McGaugh, S.E., Bronikowski, A.M., Kuo, C.H., Reding, D.M., Addis, E.A., Flagel, L.E., Janzen, F.J., Schwartz, T.S., 2015. Rapid molecular evolution across amniotes of the IIS/TOR network. Proc. Natl. Acad. Sci. U.S.A. 112, 7055–7060.

Mirmira, R.G., Tager, H.S., 1989. Role of the phenylalanine B24 side chain in directing insulin interaction with its receptor. Importance of main chain conformation. J. Biol. Chem. 264, 6349–6354.

Moritz, C., Case, T.J., Bolger, D.T., Donnellan, S., 1993. Genetic diversity and the history of pacific island house geckos (*Hemidactylus* and *Lepidodactylus*). Biol. J. Linn. Soc. 48, 113–133.

Mullinix, K.P., Wetekam, W., Deeley, R.G., Gordon, J.I., Meyers, M., Kent, K.A., Goldberger, R.F., 1976. Induction of vitellogenin synthesis by estrogen in avian liver: relationship between level of vitellogenin mRNA and vitellogenin synthesis. Proc. Natl. Acad. Sci. U. S. A. 73, 1442–1446.

Mutel, E., Abdul-Wahed, A., Ramamonjisoa, N., Stefanutti, A., Houberdon, I., Cavassila, S., Pilleul, F., Beuf, O., Gautier-Stein, A., Penhoat, A., Mithieux, G., Rajas, F., 2011. Targeted deletion of liver glucose-6 phosphatase mimics glycogen storage disease type 1a including development of multiple adenomas. J. Hepatol. 54, 529–537.

Nakagawa, S.H., Tager, H.S., 1986. Role of the phenylalanine B25 side chain in directing insulin interaction with its receptor. Steric and conformational effects. J. Biol. Chem. 261, 7332–7341.

Nicholas, K.B., Nicholas, H.B.J., 1997. GeneDoc: a tool for editing and annotating multiple sequence alignments. Distributed by the author. (http://www.nrbsc.org/gfx/genedoc/ebinet.htm)

Nishi, M., Nanjo, K., 2011. Insulin gene mutations and diabetes. J. Diabetes Investig. 2, 92–100.

Norris, D.O., Carr, J.A., 2013. Vertebrate Endocrinology. Academic Press, Waltham, Massachusetts.

Ong, K.T., Mashek, M.T., Bu, S.Y., Greenberg, A.S., Mashek, D.G., 2011. Adipose triglyceride lipase is a major hepatic lipase that regulates triacylglycerol turnover and fatty acid signaling and partitioning. Hepatology 53, 116–126.

Ong, K.T., Mashek, M.T., Davidson, N.O., Mashek, D.G., 2014. Hepatic ATGL mediates PPAR-alpha signaling and fatty acid channeling through an L-FABP independent mechanism. J. Lipid Res. 55, 808–815.

Opazo, J.C., Palma, R.E., Melo, F., Lessa, E.P., 2005. Adaptive evolution of the insulin gene in caviomorph rodents. Mol. Biol. Evol. 22, 1290–1298.

Orrell, K.S., Congdon, J.D., Jenssen, T.A., Michener, R.H., Kunz, T.H., 2004. Intersexual differences in energy expenditure of *Anolis carolinensis* lizards during breeding and postbreeding seasons. Physiol. Biochem. Zool. 77, 50–64.

Ota, H., 1994. Female reproductive cycles in the northernmost populations of the two gekkonid lizards, *Hemidactylus frenatus* and *Lepidodactylus lugubris*. Ecol. Res. 9, 121–130.

Pandyarajan, V., Smith, B.J., Phillips, N.B., Whittaker, L., Cox, G.P., Wickramasinghe, N., Menting, J.G., Wan, Z., Whittaker, J., Ismail-Beigi, F., Lawrence, M.C., Weiss, M.A., 2014. Aromatic anchor at an invariant hormone-receptor interface: function of insulin residue B24 with application to protein design. J. Biol. Chem. 289, 34709–34727.

Patterson, J.W., Davies, P.M.C., Veasey, D.A., Griffiths, J.R., 1978. The influence of season on glycogen levels in the lizard *Lacerta vivipara* jacquin. Comp. Biochem. Physiol. B, Biochem. Mol. Biol. 60, 491-493.

Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods. 8, 785-786.

Postic, C., Burcelin, R., Rencurel, F., Pegorier, J.P., Loizeau, M., Girard, J., Leturque, A., 1993. Evidence for a transient inhibitory effect of insulin on GLUT2 expression in the liver: studies in vivo and in vitro. Biochem. J. 293, 119–124.

Pough, F., 2016. Herpetology. Sinauer Associates, Inc., Publishers, Sunderland.

Price, E.R., 2017. The physiology of lipid storage and use in reptiles. Biol. Rev. Camb. Philos. Soc. 92, 1406–1426.

Pyron, R.A., Burbrink, F.T., Wiens, J.J., 2013. A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. BMC Evol. Biol. 13, 93.

Qian, B., Xue, L., Huang, H., 2016. Liver Transcriptome Analysis of the Large Yellow Croaker (*Larimichthys crocea*) during Fasting by Using RNA-Seq. PLoS One 11, e0150240.

Qiu, S., Vazquez, J.T., Boulger, E., Liu, H., Xue, P., Hussain, M.A., Wolfe, A., 2017. Hepatic estrogen receptor alpha is critical for regulation of gluconeogenesis and lipid metabolism in males. Sci. Rep. 7, 1661.

Rastogi, S.C., 1971. Essentials of Animal Physiology. New Age International (P) Ltd., Publishers, New Delhi.

Rencurel, F., Waeber, G., Antoine, B., Rocchiccioli, F., Maulard, P., Girard, J., Leturque, A., 1996. Requirement of glucose metabolism for regulation of glucose transporter type 2 (GLUT2) gene expression in liver. Biochem. J. 314, 903–909.

Rhodes, C.J., Shoelson, S., Halban, P.A., 2005. Insulin biosynthesis, processing, and chemistry, in: Kahn, C.R., Weie, G.C., King, G.L., Jacobson, A.M., Smith, R.J. (Eds.), Joslin's Diabetes Mellitus. Lippincott Williams & Wilkins, Philadelphia, pp. 65–82.

Rogue, A., Spire, C., Brun, M., Claude, N., Guillouzo, A., 2010. Gene Expression Changes Induced by PPAR Gamma Agonists in Animal and Human Liver. PPAR Res. 2010, 325183.

Rösler, H., Bauer, A.M., Heinicke, M.P., Greenbaum, E., Jackman, T., Nguyen, T.Q., Ziegler, T., 2011. Phylogeny, taxonomy, and zoogeography of the genus *Gekko* Laurenti, 1768 with the revalidation of *G. reevesii* Gray, 1831 (Sauria: Gekkonidae) Zootaxa. 2989, 1–50.

Rui, L., 2014. Energy metabolism in the liver. Compr. Physiol. 4, 177–197.

Sajid, W., Holst, P.A., Kiselyov, V.V, Andersen, A.S., Conlon, J.M., Kristensen, C., Kjeldsen, T., Whittaker, J., Chan, S.J., De Meyts, P., 2009. Structural basis of the aberrant receptor binding properties of hagfish and lamprey insulins. Biochemistry 48, 11283–11295.

Schaffer, L., 1994. A model for insulin binding to the insulin receptor. Eur. J. Biochem. 221, 1127–1132.

Schmoll, D., Allan, B.B., Burchell, A., 1996. Cloning and sequencing of the 5' region of the human glucose-6-phosphatase gene: transcriptional regulation by cAMP, insulin and glucocorticoids in H4IIE hepatoma cells. FEBS Lett. 383, 63–66.

Schwartz, T.S., Bronikowski, A.M., 2016. Evolution and Function of the Insulin and Insulin-like Signaling Network in Ectothermic Reptiles: Some Answers and More Questions. Integr. Comp. Biol. 56, 171–184.

Secor, S.M., 2008. Digestive physiology of the Burmese python: broad regulation of integrated performance. J. Exp. Biol. 211, 3767–3774.

Secor, S.M., Diamond, J., 1998. A vertebrate model of extreme physiological regulation. Nature 395, 659–662.

Secor, S.M., Stein, E.D., Diamond, J., 1994. Rapid upregulation of snake intestine in response to feeding: a new model of intestinal adaptation. Am. J. Physiol. 266, 695-705.

Seidah, N.G., Chretien, M., 1999. Proprotein and prohormone convertases: a family of subtilases generating diverse bioactive polypeptides. Brain Res. 848, 45–62.

Seino, S., Steiner, D.F., Bell, G.I., 1987. Sequence of a New World primate insulin having low biological potency and immunoreactivity. Proc. Natl. Acad. Sci. U.S.A. 84, 7423–7427.

She, P., Shiota, M., Shelton, K.D., Chalkley, R., Postic, C., Magnuson, M.A., 2000.

Phosphoenolpyruvate carboxykinase is necessary for the integration of hepatic energy metabolism. Mol. Cell. Biol. 20, 6508–6517.

Shelness, G.S., Sellers, J.A., 2001. Very-low-density lipoprotein assembly and secretion. Curr. Opin. Lipidol. 12, 151–157.

Shen, M., Shi, H., 2015. Sex Hormones and Their Receptors Regulate Liver Energy Homeostasis. Int. J. Endocrinol. 2015, 294278.

Sherwood, L., Klandorf, H., Yancey, P.H., 2013. Animal Physiology: From Genes to Organisms. Brooks/Cole, Boston.

Sidorkiewicz, E., Skoczylas, R., 1974. Effect of insulin on the blood sugar level in the grass snake (*Natrix natrix* L.). Comp. Biochem. Physiol. Part A Physiol. 48, 457–464.

Sørensen, P.G., Petersen, I.M., Sand, O., 1995. Activities of carbohydrate and amino acid metabolizing enzymes from liver of mink (*Mustela vison*) and preliminary observations on steady state kinetics of the enzymes. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 112, 59–64.

Sparkman, A.M., Schwartz, T.S., Madden, J.A., Boyken, S.E., Ford, N.B., Serb, J.M., Bronikowski, A.M., 2012. Rates of molecular evolution vary in vertebrates for insulin-like growth factor-1 (IGF-1), a pleiotropic locus that regulates life history traits. Gen. Comp. Endocrinol. 178, 164–173.

Stark, R., Kibbey, R.G., 2014. The mitochondrial isoform of phosphoenolpyruvate carboxykinase (PEPCK-M) and glucose homeostasis: has it been overlooked? Biochim. Biophys. Acta 1840, 1313–1330.

Streeper, R.S., Eaton, E.M., Ebert, D.H., Chapman, S.C., Svitek, C.A., O'Brien, R.M., 1998. Hepatocyte nuclear factor-1 acts as an accessory factor to enhance the inhibitory action of insulin on mouse glucose-6-phosphatase gene transcription. Proc. Natl. Acad. Sci. U. S. A. 95, 9208–9213.

Streeper, R.S., Svitek, C.A., Chapman, S., Greenbaum, L.E., Taub, R., O'Brien, R.M., 1997. A multicomponent insulin response sequence mediates a strong repression of mouse glucose-6-phosphatase gene transcription by insulin. J. Biol. Chem. 272, 11698–11701.

Stukey, J., Carman, G.M., 1997. Identification of a novel phosphatase sequence motif. Protein Sci. 6, 469–472.

Tamaki, M., Fujitani, Y., Hara, A., Uchida, T., Tamura, Y., Takeno, K., Kawaguchi, M., Watanabe, T., Ogihara, T., Fukunaka, A., Shimizu, T., Mita, T., Kanazawa, A., Imaizumi, M.O., Abe, T., Kiyonari, H., Hojyo, S., Fukada, T., Kawauchi, T., Nagamatsu, S., Hirano, T., Kawamori, R., Watada, H., 2013. The diabetes-susceptible gene SLC30A8/ZnT8 regulates hepatic insulin clearance. J. Clin. Invest. 123, 4513–4524.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular

Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.

Tanaka, A., Inoue, A., Takeguchi, A., Washizu, T., Bonkobara, M., Arai, T., 2005. Comparison of expression of glucokinase gene and activities of enzymes related to glucose metabolism in livers between dog and cat. Vet. Res. Commun. 29, 477–485.

Tiwari, S., Siddiqi, S., Siddiqi, S.A., 2013. CideB protein is required for the biogenesis of very low density lipoprotein (VLDL) transport vesicle. J. Biol. Chem. 288, 5157–5165.

Tourasse, N.J., Li, W.H., 2000. Selective constraints, amino acid composition, and the rate of protein evolution. Mol. Biol. Evol. 17, 656–664.

Upton, Z., Francis, G.L., Chan, S.J., Steiner, D.F., Wallace, J.C., Ballard, F.J., 1997. Evolution of insulin-like growth factor (IGF) function: production and characterization of recombinant hagfish IGF. Gen. Comp. Endocrinol. 105, 79–90.

Utez, P. The Reptile Database. http://www.reptile-database.org/ Accessed on 7th February, 2015.

Vidal, N., Hedges, S.B., 2005. The phylogeny of squamate reptiles (lizards, snakes, and amphisbaenians) inferred from nine nuclear protein-coding genes. C. R. Biol. 328, 1000–1008.

Vitt, L.J., Caldwell, J.P., 2013. Herpetology, Fourth Edition: An Introductory Biology of Amphibians and Reptiles. Academic Press, Waltham.

Wallace, R.A., Bergink, E.W., 1974. Amphibian Vitellogenin: Properties, Hormonal Regulation of Hepatic Synthesis and Ovarian Uptake, and Conversion to Yolk Proteins. Am. Zool. 14, 1159–1175.

Walzem, R.L., Hansen, R.J., Williams, D.L., Hamilton, R.L., 1999. Estrogen induction of VLDLy assembly in egg-laying hens. J. Nutr. 129, 467S–472S.

Wang, H., Eckel, R.H., 2009. Lipoprotein lipase: from gene to obesity. Am. J. Physiol. Endocrinol. Metab. 297, E271-288.

Wang, T., Hung, C.C.Y., Randall, D.J., 2006. The comparative physiology of food deprivation: from feast to famine. Annu. Rev. Physiol. 68, 223–251.

Weadick, C.J., Chang, B.S., 2012a. An improved likelihood ratio test for detecting site-specific functional divergence among clades of protein-coding genes. Mol Biol Evol. 29, 1297-1300.

Weadick, C.J., Chang, B.S., 2012b. Complex patterns of divergence among green-sensitive (RH2a) African cichlid opsins revealed by Clade model analyses. BMC Evol Biol. 12, 206.

Weiser, H., Starostova, Z., Kubicka, L., Kratochvil, L., 2012. Overlap of female reproductive cycles explains shortened interclutch interval in a lizard with invariant clutch size (Squamata: Gekkonidae: *Paroedura picta*). Physiol. Biochem. Zool. 85, 491–498.

Wieser, W., 1985. A new look at energy conversion in ectothermic and endothermic animals. Oecologia 66, 506–510.

Xia, X., 2013. DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. Mol Biol Evol. 30, 1720-1728.

Xia, X., Xie, Z., Salemi, M., Chen, L., Wang, Y., 2003. An index of substitution saturation and its application. Mol Phylogenet Evol. 26, 1-7.

Xiong, Y., Lei, Q.Y., Zhao, S., Guan, K.L., 2011. Regulation of glycolysis and gluconeogenesis by acetylation of PKM and PEPCK. Cold Spring Harb. Symp. Quant. Biol. 76, 285–289.

Xiong, Z., Li, F., Li, Q., Zhou, L., Gamble, T., Zheng, J., Kui, L., Li, C., Li, S., Yang, H., Zhang, G., 2016. Draft genome of the leopard gecko, *Eublepharis macularius*. Gigascience 5, 47.

Yang, Z., 2005. PAML FAQ. Distributed by the author <http://abacus.gene.ucl.ac.uk/software/pamlFAQs.pdf>.

Yang, Z., 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24, 1586–1591.

Yang, Z., 2013a. User Guide, "PAML: Phylogenetic Analysis by Maximum Likelihood" Version 4.7 (January 2013). (Downloaded from: http://abacus.gene.ucl.ac.uk/software/paml.html in March 3, 2014.)

Yang, Z., 2013b. "Result with branch-site model" posted to PAML discussion group: <https://groups.google.com/forum/#!searchin/pamlsoftware/label$20/pamlsoftware/i8Qyl_zEK2U/RvOb8BrvMQsJ>

Yang, Bielawski, 2000. Statistical methods for detecting molecular adaptation. Trends Ecol. Evol. 15, 496–503.

Yoshida, I., Sugiura, W., Shibata, J., Ren, F., Yang, Z., Tanaka, H., 2011. Change of positive selection pressure on HIV-1 envelope gene inferred by early and recent samples. PLoS One 6, e18630.

Zammit, V.A., 2013. Hepatic triacylglycerol synthesis and secretion: DGAT2 as the link between glycaemia and triglyceridaemia. Biochem. J. 451, 1–12.

Zani, P.A., Irwin, J.T., Rollyson, M.E., Counihan, J.L., Healas, S.D., Lloyd, E.K., Kojanis, L.C., Fried, B., Sherma, J., 2012. Glycogen, not dehydration or lipids, limits winter survival of side-blotched lizards (*Uta stansburiana*). J. Exp. Biol. 215, 3126–3134.

Zechner, R., Kienesberger, P.C., Haemmerle, G., Zimmermann, R., Lass, A., 2009. Adipose triglyceride lipase and the lipolytic catabolism of cellular fat stores. J. Lipid Res. 50, 3–21.

Zhang, F., Xu, X., Zhou, B., He, Z., Zhai, Q., 2011. Gene expression profile change and associated physiological and pathological effects in mouse liver induced by fasting and refeeding. PLoS One 6, e27553.

Zhu, L., Brown, W.C., Cai, Q., Krust, A., Chambon, P., McGuinness, O.P., Stafford, J.M., 2013. Estrogen Treatment After Ovariectomy Protects Against Fatty Liver and May Improve Pathway-Selective Insulin Resistance. Diabetes 62, 424–434.