ADAPTATIONS FOR FOLIVORY AND INSECTIVORY IN THE DIGESTIVE ENZYMES OF NON-HUMAN PRIMATES

by

MAREIKE CORA JANIAK

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Anthropology

Written under the direction of

Robert S. Scott

And approved by

_____

_____

_____

_____

New Brunswick, New Jersey

May, 2018

ABSTRACT OF THE DISSERTATION

Adaptations for folivory and insectivory in the digestive enzymes of non-human primates

By MAREIKE CORA JANIAK

Dissertation Director:

Robert S. Scott

Primates have a particularly diverse array of dietary ecologies, from exclusively insectivorous species to grass-eating monkeys, the primate digestive system has evolved in response to a multitude of pressures. Endogenous digestive enzymes are an important part of the dietary adaptations found in primates, but comparatively little research has sought to understand how primate digestive enzymes have adapted to the variety of challenges posed by primate foods. This dissertation examines the genes coding for the digestive enzymes lysozyme C (*LYZ*), pancreatic ribonuclease (*RNASE1*), and acidic mammalian chitinase (*CHIA*) in a comparative non-human primate sample ($n = 7$, $n = 25$, $n = 35$, respectively). Specifically, I investigate evidence for inter-specific variation and adaptive evolution in these digestive enzyme genes related to folivorous and insectivorous diets. Datasets were assembled through gene sequencing and genome mining, and evaluated with phylogenetic analyses, including PAML and RELAX.

My results show evidence for diet-related changes in pancreatic ribonuclease and acidic mammalian chitinase genes in primates. In the folivorous New World monkey *Alouatta palliata*, the *RNASE1* gene was duplicated twice and the daughter genes exhibit changes that are indicative of a reduced efficiency against double-stranded RNA,

suggesting a novel, and possibly digestive function. This had previously only been shown in colobine primates, but these findings suggest that in both foregut and caeco-colic fermenting primates pancreatic ribonuclease has convergently evolved a new role for digesting the products of microbial fermentation.

For acidic mammalian chitinase, results are consistent with the hypothesis that the enzyme is used for the digestion of insect exoskeletons. Early primates likely had three *CHIA* genes, in congruence with the theory that insects were an important component of the ancestral primate diet. Most extant primate species retain only one functional *CHIA* paralog. Exceptions include two colobine species (non-insectivorous), in which all *CHIA* genes have premature stop codons, and several New World monkey species that retain two functional genes. The more insectivorous species in the sample also have the largest number of functional *CHIA* genes, retaining three functional *CHIA* paralogs. *Tarsius syrichta*, the most insectivorous primate, has a total of five *CHIA* genes, two of which may be duplications specific to the tarsier lineage. Selection analyses indicate that *CHIA* genes are under more intense selection in species with higher insect consumption, as well as in smaller-bodied species (<500 g), providing molecular support for Kay's Threshold, a well-established component of primatological theory. *CHIA* genes are not subject to copy number variation in primates.

Overall, these results provide evidence that pancreatic ribonuclease and acidic mammalian chitinase are important digestive enzyme adaptations for folivorous howler monkeys and insectivorous primates, respectively. These proteins may provide crucial adaptive benefits by improving the digestion of foliage and insects, and thus increasing energy and amino acid returns to the animal.

Acknowledgements

I want to, first and foremost, thank Dr. Rob Scott for being my advisor, for adopting me into the Scott lab and convincing me to stay in graduate school, for his unwavering support in the face of many grant rejections, and for giving me the freedom to pursue a project completely unrelated to microwear. Rob's friendship and encouragement were central to the completion of this dissertation, but I also owe many thanks to the rest of my committee. During my first semester, Dr. Erin Vogel kick-started my interest in nutritional ecology and dietary adaptations, the seed from which this dissertation grew. She has opened countless doors for me, giving me opportunities to teach in Costa Rica and Indonesia, and has been an invaluable role model, both as a scientist and a woman in science. I am indebted to Dr. Ryne Palombit for accepting me into the program and for the many thought-provoking discussions (and pizza slices) during Simian Seminar. As director of CHES, he also played an integral part in giving me the financial means to conduct this research, for which I am eternally thankful. I am immensely grateful to Dr. Todd Disotell for providing me and my work with a home at the NYU Molecular Anthropology Lab and for his incredible generosity, in the lab and across the street from it.

Dr. PJ Perry's work on salivary amylase was my introduction to the world of primate digestive enzymes. His support and encouragement have meant a lot to me and I am very thankful for his advice and for the primate DNA samples he provided so generously. Morgan Chaney and Dr. Tony Tosi have been wonderful collaborators and I am very grateful for their participation in the *CHIA* project.

Finishing this PhD required persistence above all. I would not have been able to rebound from the many setbacks without the friendship and company of my fellow grad students at Rutgers. Shauhin Alavi, Tim Bransford, Rebecca Brittain, Padmini Iyer, Sarah Hlubik, Alysse Moldawer, Alex Pritchard, Darcy Shapiro, Rene Studer-Halbach, and Melanie Fenton (among others) provided comic relief and moral support over many a beer at BS&M (bones, stones, and monkeys) Fridays. Special thanks are due to my academic big sister, Dr. Susan Coiner-Collier, for always listening, and to Fred Foster, for being a true friend through it all, even when I may not have deserved it.

Many, many thanks also to everyone in the Disotell & Higham Lab at NYU for giving me a desk and an academic home in the city. I owe special thanks to Dr. Andrew Burrell for his patience, guidance, and assistance.

Shellie Harden, Jovani Reaves, Marilyn Reyes, Ginny Caputo, Maydelle Romero and, most recently, Bibi Salaman were indispensable when it came to traversing the endless minefield that is the Rutgers bureaucracy.

Last, but certainly not least, I need to thank my world outside of academia. Bailey Pollack has been my rock through most of my grad school journey and her friendship has meant more to me than I can adequately express here. Diana and Ralph Eyhorn welcomed me into their family when I was 16 and gave me a home in the US. I can never thank them enough for their kindness and warmth. I want to thank my mother for supporting my choices at every turn, even when the thought of her arachnophobic daughter wanting to go to the rainforest made her laugh. As I get older, I seem to become more and more like my father and I thank him for passing on his scientific mind to me, although I may still never match his *h*-index.

Finally, this work would not have been possible without the financial support provided by the Center for Human Evolutionary Studies at Rutgers University (Albert-Fellows Dissertation Research Award), the International Primatological Society, and the National Science Foundation (Award #1650864), and the DNA samples provided by the Coriell Biorepository.

Table of Contents

Abstract ..... **ii**

Acknowledgements ..... **iv**

Table of Contents ..... **vii**

List of Tables ..... **x**

List of Figures ..... **xi**

Chapter 1. Introduction **1**

Background 1

Dissertation Organization 3

References 6

Chapter 2. Duplication and convergent evolution of the pancreatic ribonuclease gene *RNASE1* in a non-colobine primate, the mantled howler monkey (*Alouatta palliata*) **10**

Abstract 10

Introduction 12

Methods 16

Results 19

Discussion 23

Conclusion 26

References 28

Chapter 3. Evolution of the acidic mammalian chitinase gene family (*CHIA*) is related to body mass and insectivory in primates **43**

Abstract 43
Introduction 44
Methods 48
Results 53
Discussion 59
Conclusion 63
References 65
Chapter 4. Investigating copy number variation in the acidic mammalian chitinase gene family (*CHIA*) in New World and Old World monkeys 87
Abstract 87
Introduction 88
Methods 93
Results 95
Discussion 96
Conclusion 100
References 101
Chapter 5. Conclusion 116
Summary and Conclusions 116
Future Directions 118
References 121
APPENDICES 125
Appendix A. Supplementary Table 1. Samples and genomes used in Chapter 3. 125
Appendix B. Supplementary Table 2. Primer sequences used in Chapter 3 127

Appendix C. Supplementary Table 3. Diet data used to calculate average insect consumption in Chapter 3. 130

Appendix D. Supplementary Table 4. Full QuantStudio 3D digital PCR results. 140

List of Tables

Table 2.1. Sequences used in this study with NCBI accession numbers. 33

Table 2.2. Pairwise identity (%) of lysozyme C (*LYZ*) amino acid sequences 35

Table 2.3. Results of CODEML analyses for primate *RNASE1* and duplicated sequences (n = 28) 36

Table 3.1. Species included in this study with annual average insect consumption and average body weight of adult females in grams 70

Table 3.2. CodeML results 72

Table 3.3. RELAX results 75

Table 4.1. Average annual insect consumption for species included in this study 108

Table 4.2. Primer and probe sequences used in TaqMan assays. 110

Table 4.3. QuantStudio 3D results for custom assays, *mCHIA* and *hCHIA*, and RNase P reference assay. 111

List of Figures

Fig. 2.1. Primate *LYZ* protein sequences aligned to cow (*Bos taurus*) reference sequence. 37

Fig 2.2. Phylogenies of primates based on coding sequences (474 bp) of *RNASE1* and duplications. 39

Fig. 2.3. Primate *RNASE1*, *RNASE1B*, and *RNASE1C* sequences aligned to human (*Homo sapiens*) reference sequence. 40

Fig. 2.4. Computed isoelectric points (*pI*) of RNase1 and the duplicated proteins RNase1B and RNase1C in colobines and *Alouatta palliata.* 41

Fig. 2.5. Evolutionary relationships and isoelectric point (*pI*) of the proteins encoded by *RNASE1*, *RNASE1B*, and *RNASE1C* in primates. 42

Figure 3.1. Correlation between primate diets and body size. 77

Figure 3.2. Evolutionary relationships of the *CHIA* genes in primates. 78

Figure 3.3. Evolutionary relationships as inferred from *CHIA* sequences, including timing of *CHIA* pseudogenization events. 80

Fig. 3.4. 81

Figure 3.5. Partial amino acid sequence alignment of functional *CHIA* genes. 83

Figure 3.6. Patterns of natural selection across *mCHIA* and *hCHIA*. 84

Figure 3.7. Branch specific relaxation parameters inferred for (a) *hCHIA* and (b) *mCHIA* genes under the General Descriptive model in RELAX. 86

Figure 4.1. Locations of primers and probes along the *mCHIA* and *hCHIA* gene..... 112

Figure 4.2. Copies per microliter measured for *mCHIA* and reference assays. 113

Figure 4.3. Copies per microliter measured for *hCHIA* and reference assays. 114

Figure 4.4. Relative copy number of *mCHIA* and *hCHIA* per diploid genome. 115

Chapter 1. Introduction

Background

All living organisms need to consume nutrients to grow, survive, and reproduce, making the successful acquisition of food resources a powerful selective pressure. However, acquiring food is only part of the challenge. While all animals spend much of their daily activity budget hunting, searching for, or otherwise procuring food, a large part of what is involved in overall nutrition occurs once the meal has been swallowed. Most nutritional components are too complex for immediate use and must be broken down into simpler compounds, which can then be absorbed by the body. This process is known as digestion and is catalyzed by enzymes that are either endogenous or produced by the host's microbial population (Stevens and Hume, 1995). Research shows that the nutritional value of food is partially constrained by the digestive abilities of the microbial community present in the host's gut, and that these microbes rapidly adapt to changes in diet and other environmental pressures (Graf et al., 2015). An accumulating body of evidence suggests that endogenously produced digestive enzymes also have been, and still are, common targets of natural selection, further cementing their crucial role in an organism's digestive system (Perry et al., 2007; Axelsson et al., 2013; Ranciaro et al., 2014).

Primates exhibit a particularly diverse array of dietary ecologies. From exclusively insectivorous species to grass-eating monkeys, the primate digestive system, including the enzymes contained therein, has evolved in response to a multitude of pressures. Recently, many research efforts have been focused on the gut microbiome, providing new insights into the interplay between diet and gut adaptation for a variety of

animals including human and non-human primates (Muegge et al., 2011; Yatsunenko et al., 2012; Amato et al., 2013; Macdonald et al., 2013; O'Sullivan et al., 2013; Albenberg and Wu, 2014). These are exciting new findings, but to get a full picture of an animal's digestive adaptations, the gut microbiome and endogenously produced digestive enzymes should be viewed as complementary parts of a system. While the genes coding for digestive enzymes will not change as quickly as those of the microbiome, the variety of endogenous digestive enzymes within primates nevertheless constitutes a major adaptive strategy, and warrants special attention. Changes in the expression of digestive enzymes are important dietary adaptations that may allow an organism to exploit food sources that were previously difficult or impossible to digest. There is evidence that these changes can occur quite rapidly in evolutionary time (Itan et al., 2009) and may thus be an important factor that allows animals to carve out separate dietary niches in environments where several species are competing for food resources.

Both South America and Madagascar were likely populated by a small number of primates rafting from the African mainland (Fleagle and Gilbert, 2006). Upon arrival, these primates rapidly diversified and filled the available dietary niches (Rosenberger, 1992), evolving a suite of physiological, morphological, and behavioral characteristics to process their respective diets (Nash, 1986; Janson and Boinski, 1992; Kinzey, 1992; Norconk et al., 2009). Changes in digestive enzymes were likely part of this adaptive suite, as digestive enzyme adaptations are not just important for the ability to tolerate new food resources, but also to maximize the energy obtained from them. Especially in human evolution maximizing the energy extracted from foods may have been a crucial factor in fueling the growth of our large brains (Aiello and Wheeler, 1995). In non-human

primates, many species depend on relatively low-quality foods (e.g., leaves), which can only be digested efficiently with specific gut adaptations, such as foregut fermentation and/or special digestive enzymes, as I describe below.

Recent work on primate nutritional ecology has highlighted the many challenges primates face to meet not just overall energy requirements, but also to balance micronutrients and protein intake (Felton et al., 2009; Rothman et al., 2011; Johnson et al., 2013; Lambert and Rothman, 2015), all while dealing with fiber, tannins and toxins contained in foods (Simmen et al., 2012; Garber et al., 2015). The ability to meet nutritional goals depends in part on foraging decisions and the nutritional composition of the food item (Mertl-Millhollen et al., 2003; Amato and Garber, 2014; Garber et al., 2015). However, it is also constrained by the gut's capability to extract these nutrients, which is where digestive enzymes and digestive enzyme variation undoubtedly play a key role (Stevens and Hume, 1995; Perry et al., 2007; Axelsson et al., 2013; Karasov and Douglas, 2013).

Dissertation Organization

My dissertation research focuses on the digestive enzymes lysozyme, pancreatic ribonuclease, and acidic mammalian chitinase. I use a comparative genetic approach in order to understand how these enzymes have evolved and how they differ in non-human primates with various diets.

In chapter 2, I investigate the enzymes lysozyme and pancreatic ribonuclease in the genome of the mantled howler monkey (*Alouatta palliata*). Lysozyme and pancreatic ribonuclease are two enzymes, originally involved in immune defense, that have evolved

new digestive functions in foregut fermenting animals, such as colobine monkeys and artiodactyl ruminants (Barnard, 1969; Jollès and Jollès, 1984; Stewart et al., 1987; Beintema, 1990; Swanson et al., 1991; Zhang, 2003; Cho et al., 2005; Zhang, 2006). These changes are adaptations for the digestion of an herbivorous or, in the case of colobines, folivorous diet (Stewart et al., 1987; Zhang, 2006). However, it has not been investigated whether non-foregut fermenting primates with a folivorous diet share any of these adaptations. I, therefore, analyze the genes coding for lysozyme and pancreatic ribonuclease in a newly assembled genome of *Alouatta palliata*, a folivorous New World primate with caeco-colic fermentation. The genome sequencing for this project was conducted in collaboration with Dr. Andrew Burrell and Dr. Todd Disotell of New York University.

In chapter 3, I present and analyze a comparative dataset of acidic mammalian chitinase gene sequences from 34 primate species. The enzyme acidic mammalian chitinase is of interest because insects are an important food resource for many primates, but the chitinous exoskeletons of arthropods have long been considered to be indigestible by the digestive enzymes of most mammals (Cork and Kenagy, 1989; Oftedal et al., 1991; Simunek and Bartonova, 2005; Strobel et al., 2013; Ohno et al., 2016). However, recently mice and insectivorous bats were found to produce the enzyme acidic mammalian chitinase and to use it to digest insect exoskeletons (Whitaker et al., 2004; Strobel et al., 2013; Ohno et al., 2016). To test whether non-human primates may share this adaptation for insectivory, I assemble a comparative genetic dataset of primate species with various levels of insect consumption here. Gene sequencing for this project

was done in collaboration with Morgan Chaney and Dr. Anthony Tosi of Kent State University.

In chapter 4, I test whether the acidic mammalian chitinase genes show evidence of copy number variation in non-human primates and whether any variation in copy number correlates with insect consumption. Numerous studies have provided evidence that gene copy number variation correlates with phenotypic variation (i.e. higher copy numbers correspond to increased expression of the protein and vice versa) (Hollox et al., 2003; Linzmeier and Ganz, 2006; Perry et al., 2006; 2007) and that phenotypic variation relating to digestive enzymes can strongly impact an organism's digestive abilities (Ingram et al., 2009; Mandel and Breslin, 2012). Therefore, I investigate here whether copy number variation is found in the digestive enzyme genes coding for acidic mammalian chitinase in a sample of primates with different levels of insect consumption.

A more complete understanding of inter- and intraspecific variation in primate digestive enzymes can provide insight into the evolution of dietary ecologies and dietary adaptations of primates past and present, as well as a better grasp of the digestive capabilities of different species.

References

Aiello LC, Wheeler P. 1995. The expensive-tissue hypothesis: the brain and the digestive system in human and primate evolution. Current Anthropology 36:199–221.

Albenberg LG, Wu GD. 2014. Diet and the Intestinal Microbiome: Associations, Functions, and Implications for Health and Disease. Gastroenterology 146:1564–1572.

Amato KR, Garber PA. 2014. Nutrition and foraging strategies of the black howler monkey (*Alouatta pigra*) in Palenque National Park, Mexico. Am J Primatol 76:774–787.

Amato KR, Yeoman CJ, Kent A, Righini N, Carbonero F, Estrada A, Gaskins HR, Stumpf RM, Yildirim S, Torralba M, Gillis M, Wilson BA, Nelson KE, White BA, Leigh SR. 2013. Habitat degradation impacts black howler monkey (*Alouatta pigra*) gastrointestinal microbiomes. The ISME Journal 7:1344–1353.

Axelsson E, Ratnakumar A, Arendt M-L, Maqbool K, Webster MT, Perloski M, Liberg O, Arnemo JM, Hedhammar Å, Lindblad-Toh K. 2013. The genomic signature of dog domestication reveals adaptation to a starch-rich diet. Nature 495:360–364.

Barnard EA. 1969. Biological function of pancreatic ribonuclease. Nature 221:340–344.

Beintema JJ. 1990. The primary structure of langur (*Presbytis entellus*) pancreatic ribonuclease: adaptive features in digestive enzymes in mammals. Mol Biol Evol 7:470–477.

Cho S, Beintema JJ, Zhang J. 2005. The ribonuclease A superfamily of mammals and birds: identifying new members and tracing evolutionary histories. Genomics 85:208–220.

Cork SJ, Kenagy GJ. 1989. Nutritional Value of Hypogeous Fungus for a Forest-Dwelling Ground Squirrel. Ecology 70:577–586.

Felton AM, Felton A, Raubenheimer D, Simpson SJ, Foley WJ, Wood JT, Wallis IR, Lindenmayer DB. 2009. Protein content of diets dictates the daily energy intake of a free-ranging primate. Behavioral Ecology 20:685–690.

Fleagle JG, Gilbert CC. 2006. The Biogeography of Primate Evolution: The Role of Plate Tectonics, Climate and Chance. In: Lehman SM, Fleagle JG, editors. Primate Biogeography. Springer US. p 375–418.

Garber PA, Righini N, Kowalewski MM. 2015. Evidence of Alternative Dietary Syndromes and Nutritional Goals in the Genus *Alouatta*. In: Garber PA, Cortés-Ortiz L, Urbani B, Youlatos D, editors. Howler Monkeys. New York, NY: Springer New York. p 85–109.

Graf D, Di Cagno R, Fåk F, Flint HJ, Nyman M, Saarela M, Watzl B. 2015. Contribution of diet to the composition of the human gut microbiota. Microb Ecol Health Dis 26:26164.

Hollox EJ, Armour JAL, Barber JCK. 2003. Extensive normal copy number variation of a beta-defensin antimicrobial-gene cluster. Am J Hum Genet 73:591–600.

Ingram CJE, Mulcare CA, Itan Y, Thomas MG, Swallow DM. 2009. Lactose digestion and the evolutionary genetics of lactase persistence. Hum Genet 124:579–591.

Itan Y, Powell A, Beaumont MA, Burger J, Thomas MG. 2009. The origins of lactase persistence in Europe. PLoS Comput Biol 5:e1000491.

Janson CH, Boinski S. 1992. Morphological and behavioral adaptations for foraging in generalist primates: the case of the cebines. Am J Phys Anthropol 88:483–498.

Johnson CA, Raubenheimer D, Rothman JM, Clarke D, Swedell L. 2013. 30 Days in the Life: Daily Nutrient Balancing in a Wild Chacma Baboon. PLoS ONE 8:e70383–7.

Jollès P, Jollès J. 1984. What's new in lysozyme research? Mol Cell Biochem 63:165–189.

Karasov WH, Douglas AE. 2013. Comparative digestive physiology. Compr Physiol 3:741–783.

Kinzey WG. 1992. Dietary and dental adaptations in the Pitheciinae. Am J Phys Anthropol 88:499–514.

Lambert JE, Rothman JM. 2015. Fallback Foods, Optimal Diets, and Nutritional Targets: Primate Responses to Varying Food Availability and Quality. Annu Rev Anthropol 44:493–512.

Linzmeier RM, Ganz T. 2006. Copy number polymorphisms are not a common feature of innate immune genes. Genomics 88:122–126.

Macdonald C, Barden S, Foley S. 2013. Isolation and characterization of chitin-degrading micro-organisms from the faeces of Goeldi's monkey, *Callimico goeldii*. J Appl Microbiol 116:52–59.

Mandel AL, Breslin PAS. 2012. High endogenous salivary amylase activity is associated with improved glycemic homeostasis following starch ingestion in adults. J Nutr 142:853–858.

Mertl-Millhollen AS, Moret ES, Felantsoa D, Rasamimanana H, Blumenfeld-Jones KC, Jolly A. 2003. Ring-tailed lemur home ranges correlate with food abundance and nutritional content at a time of environmental stress. Int J Primatol 24:969–985.

Muegge BD, Kuczynski J, Knights D, Clemente JC, González A, Fontana L, Henrissat B,

Knight R, Gordon JI. 2011. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science 332:970–974.

Nash LT. 1986. Dietary, behavioral, and morphological aspects of gummivory in primates. Am J Phys Anthropol 29:113–137.

Norconk MA, Wright BW, Conklin-Brittain NL, Vinyard CJ. 2009. Mechanical and nutritional properties of food as factors in platyrrhine dietary adaptations. In: Garber PA, Estrada A, Bicca-Marques JC, Heymann EW, Strier KB, editors. South American Primates. Springer. p 279–319.

Oftedal OT, Whiten A, Southgate DAT, Van Soest PJ. 1991. The nutritional consequences of foraging in primates: the relationship of nutrient intakes to nutrient requirements. Philos Trans R Soc Lond, B, Biol Sci 334:161–170.

Ohno M, Kimura M, Miyazaki H, Okawa K, Onuki R, Nemoto C, Tabata E, Wakita S, Kashimura A, Sakaguchi M, Sugahara Y, Nukina N, Bauer PO, Oyama F. 2016. Acidic mammalian chitinase is a proteases-resistant glycosidase in mouse digestive system. Sci Rep 6:37756.

O'Sullivan A, He X, McNiven EMS, Haggarty NW, Lönnerdal B, Slupsky CM. 2013. Early Diet Impacts Infant Rhesus Gut Microbiome, Immunity, and Metabolism. J Proteome Res 12:2833–2845.

Perry GH, Dominy NJ, Claw KG, Lee AS, Fiegler H, Redon R, Werner J, Villanea FA, Mountain JL, Misra R, Carter NP, Lee C, Stone AC. 2007. Diet and the evolution of human amylase gene copy number variation. Nat Genet 39:1256–1260.

Perry GH, Tchinda J, McGrath SD, Zhang J, Picker SR, Cáceres AM, Iafrate AJ, Tyler-Smith C, Scherer SW, Eichler EE, Stone AC, Lee C. 2006. Hotspots for copy number variation in chimpanzees and humans. Proc Natl Acad Sci USA 103:8006–8011.

Ranciaro A, Campbell MC, Hirbo JB, Ko W-Y, Froment A, Anagnostou P, Kotze MJ, Ibrahim M, Nyambo T, Omar SA, Tishkoff SA. 2014. Genetic origins of lactase persistence and the spread of pastoralism in Africa. Am J Hum Genet 94:496–510.

Rosenberger AL. 1992. Evolution of feeding niches in New World monkeys. Am J Phys Anthropol 88:525–562.

Rothman JM, Raubenheimer D, Chapman CA. 2011. Nutritional geometry: gorillas prioritize non-protein energy while consuming surplus protein. Biology Letters 7:847–849.

Simmen B, Tarnaud L, Hladik A. 2012. Leaf nutritional quality as a predictor of primate biomass: further evidence of an ecological anomaly within prosimian communities in Madagascar. J Trop Ecol 28:141–151.

Simunek J, Bartonova H. 2005. Effect of Dietary Chitin and Chitosan on Cholesterolemia

of Rats. Acta Veterinaria Brno 74:491–499.

Stevens CE, Hume ID. 1995. Comparative physiology of the vertebrate digestive system. 2nd ed. Cambridge: Cambridge University Press.

Stewart CB, Schilling JW, Wilson AC. 1987. Adaptive evolution in the stomach lysozymes of foregut fermenters. Nature 330:401–404.

Strobel S, Roswag A, Becker NI, Trenczek TE, Encarnação JA. 2013. Insectivorous Bats Digest Chitin in the Stomach Using Acidic Mammalian Chitinase. PLoS ONE 8:e72770.

Swanson KW, Irwin DM, Wilson AC. 1991. Stomach lysozyme gene of the langur monkey: tests for convergence and positive selection. J Mol Evol 33:418–425.

Whitaker JO, Dannelly HK, Prentice DA. 2004. Chitinase in Insectivorous Bats. J Mammal 85:15–18.

Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP, Heath AC, Warner B, Reeder J, Kuczynski J, Caporaso JG, Lozupone CA, Lauber C, Clemente JC, Knights D, Knight R, Gordon JI. 2012. Human gut microbiome viewed across age and geography. Nature 486:222–227.

Zhang J. 2003. Parallel Functional Changes in the Digestive RNases of Ruminants and Colobines by Divergent Amino Acid Substitutions. Mol Biol Evol 20:1310–1317.

Zhang J. 2006. Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys. Nat Genet 38:819–823.

Chapter 2. Duplication and convergent evolution of the pancreatic ribonuclease gene *RNASE1* in a non-colobine primate, the mantled howler monkey (*Alouatta palliata*)

Abstract

Pancreatic ribonuclease (encoded by *RNASE1*) and lysozyme C (*LYZ*) are two enzymes, originally involved in immune defense, that have evolved new digestive functions in foregut fermenting animals, such as colobine monkeys and artiodactyl ruminants. In all colobines, *RNASE1* was duplicated at least once, and in some colobine species twice, with the daughter genes (*RNASE1B*, *RNASE1C*) evolving new digestive roles. Howler monkeys (*Alouatta* spp.) are the most folivorous of the New World monkeys but, lacking the sacculated stomachs of colobines, digest foliage using caeco-colic fermentation. In this study, I report on the *RNASE1* and *LYZ* genes in the mantled howler monkey (*Alouatta palliata*). The results indicate that the *RNASE1* gene was duplicated twice in *A. palliata*, leading to two daughter genes, *RNASE1B* and *RNASE1C*. While the parent gene (*RNASE1*) is conserved, *RNASE1B* and *RNASE1C* have multiple amino acid substitutions that are convergent with those found in the duplicated *RNASE1* genes of colobines. As in colobines, the duplicated RNases in *A. palliata* have lower isoelectric points, a lower charge, and changes that are indicative of a reduced efficiency against double-stranded RNA, suggesting a novel, and possibly digestive function. Howler monkey *LYZ* is conserved and does not share the substitutions found in the colobine and bovine sequences. These findings suggest that in both foregut and caeco-colic fermenting primates pancreatic ribonuclease has convergently evolved a new role for digesting the

products of microbial fermentation. Energy gains from the digestion of these products can be substantial, therefore, these duplicated proteins may be crucial digestive enzyme adaptations allowing howler monkeys to survive on a folivorous diet during times of fruit scarcity.

Introduction

Howler monkeys (*Alouatta* spp.) are the only New World monkeys with a diet rich in leaves. While the amount of leaves consumed varies greatly between sites and populations (Garber et al., 2015), as well as across seasons (Milton, 1998), on average young and mature leaves make up half or more of the howler monkey diet (Glander, 1981; Milton, 1981; Chapman, 1987; Stoner, 1996; Williams-Guillén, 2003). Howler monkeys manage to persist on such a diet without the sacculated, foregut-fermenting stomachs found in folivorous Old World monkeys, the colobines. However, howler monkeys have a suite of other adaptations for a folivorous diet.

Adaptations for a leaf-rich diet include behavioral strategies, such as minimizing energy expenditure (Strier, 1992; Milton, 1998; Da Cunha and Byrne, 2006) and feeding preferentially on young leaves (Milton, 1979; Stoner, 1996; Amato and Garber, 2014) or leaves with a higher protein to fiber ratio (Glander, 1981), which are less tough and may be easier to digest (Matsuda et al., 2017). Howler monkeys are unique among platyrrhines in having routine trichromacy, a trait that has been proposed as an adaptation for detecting such young leaves (Dominy and Lucas, 2001; Lucas et al., 2003). Gut and digestive adaptations include very long gut transit times of approx. 20 hours (Milton et al., 1980; Espinosa-Gómez et al., 2013) and an enlarged caecum and colon (Chivers and Hladik, 1980).

These gut adaptations are important for folivorous mammals because the cell walls of plants are made of structural carbohydrates, cellulose and hemicellulose, that cannot be digested by the enzymes produced by vertebrates (Lambert, 1998). Instead, leaf-eating mammals rely on microbial fermentation to break down plant material in the

forestomach (foregut fermenters, like colobines) or in the large intestine ("hindgut" or caeco-colic fermenters) (Milton and McBee, 1983; Chivers and Langer, 1994; Lambert, 1998). The enlarged caecum and colon in howlers are important sites for microbial fermentation and are also enlarged in other folivorous or herbivorous caeco-colic fermenters, such as horses and elephants (Alexander, 1993). During microbial fermentation, the plant material is broken down by symbiotic bacteria found in the host's gut. This process releases volatile fatty acids, which are easily absorbed (Kay and Davies, 1994) but the bacteria themselves are also digested by the host and are an important source of nitrogen (Barnard, 1969; Kay and Davies, 1994). Both ruminant artiodactyls and colobine monkeys have convergent digestive enzyme adaptations for the digestion of fermenting gut bacteria, but we do not know whether howler monkeys share any of these adaptations. Endogenous digestive enzymes are a crucial component of an animal's digestive system and include important dietary adaptations (Janiak, 2016), such as the enzymes lysozyme and pancreatic ribonuclease.

RNase and lysozyme evolution in primates and ruminants

The enzyme lysozyme is found in many vertebrates and invertebrates where it has an immunological function (Jollès and Jollès, 1984). In cows and colobines this enzyme exhibits convergent amino acid changes that allow lysozyme to function at a much lower pH, an adaptation for bacteriolytic activity in the acidic stomach fluid (Stewart et al., 1987; Swanson et al., 1991). Similarly, the pancreatic ribonuclease enzyme (RNase1), which has an original function in pathogen defense (Cho et al., 2005), has acquired a new digestive function in both colobines and ruminant artiodactyls (Barnard, 1969; Beintema, 1990; Zhang, 2003; 2006). In this case, the *RNASE1* gene

underwent one or more duplications, and the duplicated gene(s) (*RNASE1B* and *RNASE1C*, or alternatively *RNASE1β* and *RNASE1γ*) evolved a new function in digesting the nucleic acids of fermenting microbes found in the digestive system. These duplications and subsequent functional changes evolved independently in artiodactyl ruminants and colobines(Zhang, 2003). They may have also evolved separately in African and Asian colobines (Zhang, 2006; Yu et al., 2010), although this has been questioned by some studies (Schienman et al., 2006; Xu et al., 2009) and the evolutionary history of *RNASE1* duplications in colobines has not yet been completely resolved.

Significantly, the duplicated RNases of both ruminants and colobines have amino acid changes leading to changes of the isoelectric point (*pI*), optimal pH, and charge of the resulting protein. The digestive RNases (RNase1B and RNase1C) exhibit a lower *pI*, a lower optimal pH, and a decreased charge at pH 7.0 (Zhang, 2003; 2006) compared to the ancestral RNase1. Previous research has shown that the charge of pancreatic ribonuclease affects the enzyme's ability to degrade double-stranded RNA (Libonati et al., 1976). A high *pI* and positive charge is indicative of a ribonuclease that is involved in defense against pathogens, while a decrease in *pI* and charge suggest a novel function for the enzyme (Goo and Cho, 2013; Liu et al., 2014), as shown in the case of cows and colobines (Zhang, 2003).

RNase and lysozyme evolution in other mammals

Duplications of the *RNASE1* gene have also been found in other mammal groups. In rodents, both rats (*Rattus* spp.) and guinea pigs (*Cavia* spp.) independently evolved two duplicated genes, *RNASE1B* and *RNASE1C*, from the ancestral *RNASE1* gene (Dubois et al., 2002; Goo and Cho, 2013; Lang et al., 2017). Whether these duplicated

genes remain involved in immune function or have acquired novel, possibly digestive, functions has not been determined. Lang and colleagues (2017) suggest that the high *pI* (9.66) of rat RNase1B and its expression in the spleen point to retention of the original immunological function, while rat RNase1C (*pI* = 7.71) may have acquired a novel function (Lang et al., 2017). Similar to cows and colobines, the duplicated ribonuclease in guinea pigs has a lower *pI* and decreased charge and is less effective at degrading double-stranded RNA than the ancestral enzyme, suggesting a novel and possibly digestive function (Libonati et al., 1976; Dubois et al., 2003). Like howler monkeys, guinea pigs are herbivorous animals with caeco-colic rather than foregut fermentation. This suggests that ribonuclease could have a function in the digestion of microbial fermentation products, regardless of the presence of a foregut.

RNASE1 also underwent multiple duplications independently in two families of bats, the Vespertilionidae and Molossidae, but the duplicate proteins have high isoelectric points suggesting an immunological rather than dietary function (Xu et al., 2013). Seven *RNASE1* genes were found in *Myotis lucifugus*, an insectivorous Vesper bat (Goo and Cho, 2013). The authors propose that this may be an immunological adaptation, as the communal roosting behavior of these bats potentially increases their exposure to pathogens and RNase1 may improve their resilience to them (Goo and Cho, 2013). In the superfamily Musteloidea, a group that includes red pandas, weasels, raccoons, and skunks, *RNASE1* was duplicated independently in four families but the functional significance of these duplicates is not yet clear (Liu et al., 2014). As summarized here, *RNASE1* genes have an interesting history of duplications and functional diversification

in mammals, including for novel dietary functions in herbivorous foregut and caeco-colic fermenters.

Present study

The *RNASE1* gene(s) of folivorous foregut-fermenting primates (colobines) and many non-folivorous primates are now well characterized (Schienman et al., 2006; Zhang, 2006; Yu et al., 2010), but it is not clear whether folivorous caeco-colic fermenting primates, such as howler monkeys, share the digestive enzyme adaptations of colobines. In this study, I therefore investigated both the *RNASE1* and *LYZ* genes in the mantled howler monkey (*Alouatta palliata*), a leaf-eating primate with caeco-colic fermentation.

Hypotheses and predictions

I hypothesize that howler monkey RNase1 and lysozyme exhibit similar changes in charge and isoelectric point as the proteins in colobines, as adaptations for a folivorous diet. To find the predicted shared amino acid changes, I assembled a comparative dataset of *RNASE1* and *LYZ* gene sequences across primates and translated and aligned the coding sequences. To better understand the evolutionary history and selective pressures acting on *RNASE1*, I tested for positive/purifying selection and reconstruct the ancestral gene sequences. Finally, the *pI* and charge at pH 7.0 were calculated for both extant and ancestral sequences, in order to identify the predicted shared biochemical properties of the proteins in different species.

Methods

Genome mining

To assemble a comparative dataset of pancreatic ribonuclease (*RNASE1*) and lysozyme (*LYZ*) gene, I mined primate genomes and gene sequences available on GenBank, as well as an unpublished draft genome assembly of the mantled howler monkey (*Alouatta palliata*). The *RNASE1* sequence from the common marmoset (*Callithrix jacchus*) reference genome and the *LYZ* sequence from the black-capped squirrel monkey (*Saimiri boliviensis*) reference genome were used as queries to run a BLASTN search (default search parameters) in the genome assembly of *Alouatta palliata*. Primate *RNASE1* sequences generated in previous studies (Zhang et al., 2002; Schienman et al., 2006; Zhang, 2006; Yu et al., 2010) were downloaded from the National Center for Biotechnology (NCBI). For a better understanding of the history of *RNASE1* in primates, especially in platyrrhines, the published reference genomes of *Aotus nancymaae* (Anan_1.0), *Cebus capucinus imitator* (Cebus_imitator-1.0), *Microcebus murinus* (Mmur_3.0), and *Tarsius syrichta* (Tarsius_syrichta-2.0.1) were searched for RNASE1 sequences using BLASTN with the same query and parameters as above. Primate sequences of *LYZ* generated in a previous study (Messier and Stewart, 1997) were retrieved from GenBank. The *LYZ* gene sequence found in the *Bos taurus* reference genome (Bos_taurus_UMD_3.1.1) differed from the sequence reported in (Stewart et al., 1987). The sequence for *Bos taurus LYZC2* was used, because it was most similar and almost identical to the bovine lysozyme sequence reported in Stewart et al. (1987). A full list of sequences used in this study and their accession numbers are presented in Table 1.

Sequence analyses

Coding regions of all sequences for *RNASE1* (471 bp) and *LYZ* (447 bp) were translated using Geneious 9.1.8. Coding regions and translated amino acid sequences were aligned using the MAFFT alignment server (http://mafft.cbrc.jp/alignment/server/) and Geneious 9.1.8.

The following analyses were only done with data for *RNASE1*, because no evidence of duplication or convergence between *Alouatta* and colobines was found for *LYZ*.

Phylogenetic trees were constructed from both nucleotide and protein alignments of *RNASE1* with MrBayes and PHYML programs using the Hasegawa-Kishino-Yano (HKY) substitution model (Hasegawa et al., 1985) with a discrete Gamma distribution (+G). The nucleotide substitution model was chosen based on the Akeike Information Criterion (AIC) statistics calculated by the jModelTest program (webserver: http://jmodeltest.org/) (Darriba et al., 2012). In PHYML, 1000 bootstrap replicates were completed. Because these relatively short sequences did not resolve the primate phylogeny accurately, trees based on the best available primate phylogeny (Perelman et al., 2011) were used in CODEML analyses and ancestral sequence reconstructions.

I used site-specific and branch-site models in the program CODEML in the PAML package (Yang, 2007) to test for evidence of positive selection acting on the *RNASE1* gene in primates in general and in *Alouatta palliata* in particular. Site-specific models (M0 – null, M1a – nearly neutral selection, M2a – positive selection, M3 – discrete, M7 – beta, and M8 – beta & $\omega>1$) are used to determine if there is variation in the values of ω across sites along the *RNASE1* alignment and to test for evidence of positively selected sites (Yang, 2007). Model fit is evaluated using likelihood ratio tests

(LRT). For branch-site models the duplicated *Alouatta RNASE1* genes (*RNASE1B* and *RNASE1C*) and the branch leading to them were designated as foreground branches, while all parent *RNASE1* genes were designated as background branches. In this model ω is allowed to vary both between sites and across branches to determine if sites are under positive selection in the foreground branches (Yang, 2007). This alternative model is compared to the null model in which ω is fixed at 1 and model fit is evaluated with a LRT.

Ancestral sequence reconstruction

To better understand the evolutionary history of *RNASE1* in primates, ancestral sequences were reconstructed using the FastML webserver (http://fastml.tau.ac.il/) (Ashkenazy et al., 2012). The following running parameters were used: sequence type = codons, model of substitution = yang, use gamma distribution = yes and probability cutoff to prefer ancestral indel over character = 0.5.

Protein properties

To compare the biochemical properties of proteins across species, the isoelectric points (*pI*) of lysozyme, extant RNAse1 and ancestral RNAse1 proteins were computed with ProteinCalculator v3.4 (http://protcalc.sourceforge.net/) and the Compute pI Tool on the ExPASy webserver (http://web.expasy.org/compute_pi/). Estimates differed slightly between the two tools, so an average of the two values is presented here.

Results

I was able to identify sequences in the *Alouatta palliata* draft genome putatively orthologous to the query *LYZ* and *RNASE1* gene sequences.

The *LYZ* nucleotide sequence of *A. palliata* was 96.42% and 96.20% identical to that of *Saimiri boliviensis* and *Callithrix jacchus*, respectively. The *LYZ* amino acid sequence was also very similar to those of other platyrrhines, being 91.22-91.89% identical. A protein alignment of the lysozyme C sequences from *A. palliata*, *S. boliviensis*, *C. jacchus*, *Saguinus oedipus*, *Papio anubis*, *Colobus guereza*, *Nasalis larvatus*, and the cow (*Bos taurus*) is shown in Figure 1. While colobines and cows have a number of convergent amino acid changes, these substitutions are not found in howler monkey lysozyme C (Fig. 1). Pairwise distances of lysozyme C amino acid sequences are shown in Table 2. As found in previous studies (Stewart et al., 1987), the two colobines (*C. guereza* and *N. larvatus*) have overall greater sequence similarity with the cow (111/148 amino acids, 75%) than another closely related catarrhine, *Papio anubis*, has with the cow (102/148 amino acids, 68.92%). The howler monkey, on the other hand, does not share the convergent changes found in the lysozyme sequences of colobines and cows and its sequence identity with these groups is comparable to those of other platyrrhines (Table 2). Like the other platyrrhines, the howler monkey shares 124-125/148 amino acids (83.78-84.46%) with the *LYZ* sequence of colobines and 102/148 amino acids (68.92%) with the *LYZ* sequence in cows.

The *RNASE1* BLAST search of the howler monkey draft genome returned three different *RNASE1*-like sequences that were 93.19-93.75% identical to the *Callithrix jacchus* query and 96.78-98.24% identical to each other. To exclude the possibility that these sequences were of another, closely related gene, I conducted BLAST searches with these sequences as queries against all published platyrrhine genomes. All searches only returned hits to *RNASE1* and no other platyrrhine genomes investigated here showed

evidence of a second *RNASE1*-like gene. Therefore, it is most likely that the additional *RNASE1*-like sequences found in the howler monkey genome are duplications of the ancestral *RNASE1* gene. I am going to refer to these two sequences as *RNASE1B* and *RNASE1C*.

Primate *RNASE1* genes generally appear to be conserved and lack premature stop codons in all species included here (n = 25). The overall amino acid sequence divergence across *RNASE1* in these species was low (mean difference of 15.02 amino acids out of 145 total sites = 10.36%, SE over 1000 bootstrap replicates = 2.23). When including the duplicated *RNASE1* genes found in colobines and *Alouatta*, overall divergence only increased slightly (17.46/145 = 12.04%, SE = 2.02).

Trees built from the coding region of *RNASE1* failed to accurately resolve the phylogenetic relationships of all primate species (Fig. 2). Different programs (MrBayes, PHYML) and approaches (neighbor-joining, maximum likelihood) gave different topologies and did not resolve the history of *RNASE1* duplications in colobines with confidence. This is likely due to the short size of (474 bp) and overall conservation of *RNASE1*. Sequence information for non-coding regions of *RNASE1* was not available for all species in my sample, so it was not possible to use a longer sequence to construct a phylogenetic tree.

Aligning all RNase1 protein sequences shows several convergent amino acid changes between the duplicated howler monkey and colobine genes, *RNASE1B* and *RNASE1C* (Fig. 3). These include a change from arginine to glutamine at site 32, a change from lysine to glutamic acid at site 35, and a change from arginine to tryptophan at site 68. Arginine and lysine are positively charged amino acids, while glutamic acid is

negatively charged, so these changes contribute to a change in the charge of the resulting protein. The isoelectric point (*pI*) of RNase1 and the duplicated RNase1 proteins are shown in Figure 4. Compared to the high *pI* (8.29-9.13) and higher charge of parent RNase1 proteins, the duplicated proteins in colobines have a lower *pI* (6.26-8.29) (Fig. 4) and a reduced charge at pH 7.0 (Fig. 3). Likewise, the parent howler monkey RNase1 has a higher *pI* (8.29) and charge (3.9) than the duplicated howler proteins (pI = 6.75-6.76, charge = -0.1) (Fig. 3, Fig. 4).

Reconstruction of the ancestral *RNASE1* sequences showed that isoelectric points of RNase1 proteins are consistently high across the primate phylogeny, with the exception of the duplicated proteins in colobines and *Alouatta* (Fig. 5). Another exception is the RNase1 protein of the owl monkey (*Aotus nancymaae*). While I found no evidence of a duplicated gene, the *Aotus RNASE1* sequence has two amino acid changes that are convergent with changes found in the *Alouatta RNASE1B* and *RNASE1C* sequences (K34E and R68W). Consistent with these changes, the *pI* of owl monkey RNAse1 is lower (7.16) than any other non-duplicated RNAse1 in primates (Fig. 3, Fig. 5).

The CODEML analyses of selective pressure did not provide evidence that any sites along *RNASE1* are positively selected (Site-specific M1 vs. M2: ΔLRT = 0, p =1; M7 vs. M8: ΔLRT = 0.124, p = 0.94) or that the duplicated howler monkey *RNASE1B* and *RNASE1C* genes are under positive selection (Table 3). As in most functional proteins, the model of variable ω values across sites was favored over the null model of a single ω (M0 vs. M3: ΔLRT = 73.274, $p < 0.00001$) and most sites in *RNASE1* appear to be under purifying selection (M3: ω_1 = 0.047(24.83%), ω_2 = 0.047(42.57%), ω_3 = 0.807(32.60%)).

Discussion

In many mammal groups *RNASE1* genes have a history of duplications and functional diversification. This study identified a previously unknown *RNASE1* duplication event in *Alouatta palliata* and found amino acid substitutions in the duplicated genes that are convergent with duplicated *RNASE1* genes in colobines and consistent with a new function in the digestive system. The howler monkey *LYZ* gene, however, was found to be conserved and not to have evolved changes consistent with a new role as a digestive enzyme.

While all other platyrrhine species studied so far only had one *RNASE1* gene (Zhang et al., 2002), three different *RNASE1*-like sequences were identified in the genome of *Alouatta palliata*. One sequence (*RNASE1*) retained a high *pI* and positive charge, while the two other sequences (*RNASE1B* and *RNASE1C*) had several amino acid changes (Fig. 3) that resulted in a reduction of the proteins' *pI* and charge (Fig. 3, Fig. 4). Such changes have also been found in *RNASE1* duplications in Asian and African colobine monkeys and artiodactyl ruminants (Zhang, 2003; Schienman et al., 2006; Zhang, 2006; Yu et al., 2010). Colobine RNase1B and RNase1C have up to nine amino acid changes that reduce the enzymes' effectiveness against double stranded RNA (Zhang et al., 2002) and howler monkey RNase1B and RNase1C share four (R32Q, K34E, R68W, D112E) and three (R32Q, K34E, R68W) of these substitutions, respectively (Fig. 3). It is therefore likely that the duplicated howler monkey proteins are not as effective against double-stranded RNA as the ancestral protein. Combined with the lowered *pI* and reduced charge, this supports the idea that these *RNASE1* duplications in

howler monkeys have diverged in function from the original, immunological role of the ancestral protein (Barnard, 1969; Libonati et al., 1976; Zhang et al., 2002; Goo and Cho, 2013; Liu et al., 2014). Zhang (2006) had previously calculated the probability of three or more parallel amino acid substitutions arising in two lineages by chance to be between 0.0001 and 0.0026 (Zhang, 2006). The more likely explanation is that the convergent substitutions in howler monkey and colobine *RNASE1B/C* arose due to shared selective pressures.

In colobines and ruminant artiodactyls the duplicated RNase1 proteins are thought to be adaptations for the digestion of bacteria, or short-chain fatty acids produced by these bacteria, that ferment leaves in the foregut of these (Beintema, 1990; Zhang, 2003). While howler monkeys do not have a sacculated forestomach like colobines and ruminants, they do rely on microbial fermentation in the caeco-colic region to break down the foliage they consume (Milton and McBee, 1983). The duplicated RNase1B and RNase1C may thus fill a role similar to the duplicated proteins in colobines and artiodactyl ruminants, by digesting the products of leaf-fermenting bacteria, such as short-chain fatty acids, in the caeco-colic region. In a study of fermentative digestion in *Alouatta palliata* the authors found that up to 31% of the monkeys' daily required energy may come from the digestion of fermentation end products (Milton and McBee, 1983). During times of fruit scarcity when the howler diet consists entirely of leaves, the ability to efficiently digest such products may therefore be crucial to their survival (Milton and McBee, 1983) and RNase1B and RNase1C enzymes may be a key factor ensuring digestive efficiency.

Unlike cows and colobines, howler monkeys did not have convergent amino acid changes in the *LYZ* gene (Fig. 1). Lysozyme is an enzyme found in many vertebrates and invertebrates and is thought to play a role in immune function (Jollès and Jollès, 1984). In cows and colobines, however, lysozyme exhibits convergent amino acid changes that allow the enzyme to function at a much lower pH, possibly as an adaptation for bacteriolytic activity in the acidic stomach fluid (Stewart et al., 1987; Swanson et al., 1991). In artiodactyl ruminants the *LYZ* gene underwent multiple duplications and some of the daughter genes acquired a novel digestive function (Irwin, 1995). Interestingly, in the colobines there is only a single *LYZ* gene that was adapted for a digestive function. In non-colobine primates, including howler monkeys, *LYZ* is conserved (Fig. 1, (Messier and Stewart, 1997). It may be that the utility of lysozyme as a digestive enzyme is tied to foregut-fermentation, while ribonuclease can be adaptive for microbial fermentation both in the foregut, as well as in the "hindgut." Some support for this is provided by a study of lysozyme in the only avian species with foregut-fermentation, the hoatzin (*Opisthocomus hoazin*), a leaf-eating bird from South America (Grajal et al., 1989). Despite arising from a different lysozyme gene family, a lysozyme expressed in the hoatzin stomach has biochemical properties and amino acid substitutions that are convergent with those found in artiodactyl ruminants and colobines (Kornegay et al., 1994). Another possible example of ribonuclease adaptation for a digestive purpose comes from the ancient DNA of the extinct subfossil lemur *Megaladapis* (Perry and colleagues, in prep). Based on studies of dental microwear and dental topography *Megaladapis* was likely folivorous (Scott et al., 2009; Godfrey et al., 2012) and its *RNASE1* gene shares several amino acid substitutions with the duplicated *RNASE1B* and *RNASE1C* genes of colobines and howler monkeys

(Perry and colleagues, in prep). Since no extant lemurs have a colobine-like digestive system (Lambert, 1998), it is most likely that *Megaladapis* relied on caeco-colic fermentation to digest leaves, like howler monkeys.

While there is a possibility that *RNASE1* was only duplicated once in howler monkeys and *RNASE1B* and *RNASE1C* are only different alleles of the same gene, this is unlikely due to the quality of the *Alouatta palliata* genome assembly. Additionally, even a single duplication of *RNASE1* with the amino acid substitutions observed here would provide important evidence of convergent evolution between howler monkeys and colobines. A bigger limitation of the current study is that data are lacking on where in the howler monkey body *RNASE1*, *RNASE1B*, and *RNASE1C* are expressed. Finding that *RNASE1B* and *RNASE1C* are expressed in the howler caecum and/or colon would provide strong evidence that these proteins have been repurposed as digestive enzymes. Future expression studies should therefore be a priority. The role of *RNASE1* in owl monkeys (*Aotus* spp.) likewise deserves additional study. While there was no evidence of a gene duplication, *A. nancymaae RNASE1* shared two amino acid substitutions with howler and colobine *RNASE1B/C* and consequently had a lower *pI* than RNase1 in other species (Fig. 3, Fig. 5).

Conclusion

The *RNASE1* gene family has a history of duplications and functional divergence in many mammals, including the colobine primates (Zhang, 2003; Schienman et al., 2006; Zhang, 2006; Yu et al., 2010; Goo and Cho, 2013; Xu et al., 2013; Liu et al., 2014; Lang et al., 2017). Here I present evidence that *RNASE1* has also been duplicated in a non-colobine primate, the mantled howler monkey (*Alouatta palliata*), and that the

duplicated genes (*RNASE1B*, *RNASE1C*) have biochemical properties and amino acid substitutions that are convergent with those found in foregut-fermenting primates. These proteins may therefore be used for an analogous function in howler monkeys, digesting the products of microbial fermentation in the caeco-colic region, a potentially substantial source of energy (Milton and McBee, 1983). Along with behavioral and morphological adaptations, these duplicated proteins may be crucial digestive enzyme adaptations allowing howler monkeys to survive on a folivorous diet during times of fruit scarcity.

References

Alexander R. 1993. The relative merits of foregut and hindgut fermentation. J Zoology 231:391–401.

Amato KR, Garber PA. 2014. Nutrition and foraging strategies of the black howler monkey (*Alouatta pigra*) in Palenque National Park, Mexico. Am J Primatol 76:774–787.

Ashkenazy H, Penn O, Doron-Faigenboim A, Cohen O, Cannarozzi G, Zomer O, Pupko T. 2012. FastML: a web server for probabilistic reconstruction of ancestral sequences. Nucleic Acids Research 40:W580–4.

Barnard EA. 1969. Biological function of pancreatic ribonuclease. Nature 221:340–344.

Beintema JJ. 1990. The primary structure of langur (*Presbytis entellus*) pancreatic ribonuclease: adaptive features in digestive enzymes in mammals. Mol Biol Evol 7:470–477.

Chapman C. 1987. Flexibility in diets of three species of Costa Rican primates. Folia Primatol 49:90–105.

Chivers DJ, Hladik CM. 1980. Morphology of the gastrointestinal tract in primates: comparisons with other mammals in relation to diet. Journal of Morphology 166:337–386.

Chivers DJ, Langer P. 1994. Gut form and function: variations and terminology. In: Chivers DJ, Langer P, editors. The Digestive System in Mammals: Food Form and Function. Food Form and Function. Cambridge: Cambridge University Press. p 3–8.

Cho S, Beintema JJ, Zhang J. 2005. The ribonuclease A superfamily of mammals and birds: identifying new members and tracing evolutionary histories. Genomics 85:208–220.

Da Cunha R, Byrne RW. 2006. Roars of black howler monkeys (*Alouatta caraya*): evidence for a function in inter-group spacing. Behaviour 143:1169–1199.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature Methods 9:772–772.

Dominy NJ, Lucas PW. 2001. Ecological importance of trichromatic vision to primates. Nature 410:363–366.

Dubois J-YF, Jekel PA, Mulder PPMFA, Bussink AP, Catzeflis FM, Carsana A, Beintema JJ. 2002. Pancreatic-type ribonuclease 1 gene duplications in rat species. J Mol Evol 55:522–533.

Dubois J-YF, Ursing BM, Kolkman JA, Beintema JJ. 2003. Molecular evolution of

mammalian ribonucleases 1. Molecular Phylogenetics and Evolution 27:453–463.

Espinosa-Gómez F, Gómez-Rosales S, Wallis IR, Canales-Espinosa D, Hernández-Salazar L. 2013. Digestive strategies and food choice in mantled howler monkeys *Alouatta palliata mexicana*: bases of their dietary flexibility. J Comp Physiol B.

Garber PA, Righini N, Kowalewski MM. 2015. Evidence of Alternative Dietary Syndromes and Nutritional Goals in the Genus *Alouatta*. In: Garber PA, Cortés-Ortiz L, Urbani B, Youlatos D, editors. Howler Monkeys. New York, NY: Springer New York. p 85–109.

Glander KE. 1981. Feeding patterns in mantled howling monkeys. In: Foraging behavior: Ecological, ethological, and psychological approaches. Garland Press. p 231–257.

Godfrey LR, Winchester JM, King SJ, Boyer DM, Jernvall J. 2012. Dental topography indicates ecological contraction of lemur communities. Am J Phys Anthropol 148:215–227.

Goo SM, Cho S. 2013. The expansion and functional diversification of the mammalian ribonuclease a superfamily epitomizes the efficiency of multigene families at generating biological novelty. Genome Biology and Evolution 5:2124–2140.

Grajal A, Strahl SD, Parra R, Gloria Dominguez M, Neher A. 1989. Foregut fermentation in the hoatzin, a neotropical leaf-eating bird. Science 245:1236–1238.

Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307–321.

Hasegawa M, Kishino H, Yano T-A. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22:160–174.

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

Irwin DM. 1995. Evolution of the bovine lysozyme gene family: changes in gene expression and reversion of function. J Mol Evol 41:299–312.

Janiak MC. 2016. Digestive enzymes of human and nonhuman primates. Evol Anthropol 25:253–266.

Jollès P, Jollès J. 1984. What's new in lysozyme research? Mol Cell Biochem 63:165–189.

Kay R, Davies AG. 1994. Digestive physiology. In: Davies AG, Oates JF, editors. Colobine Monkeys: Their Ecology, Behaviour and Evolution. Cambridge University Press. p 229–249.

Kornegay JR, Schilling JW, Wilson AC. 1994. Molecular adaptation of a leaf-eating bird: stomach lysozyme of the hoatzin. Mol Biol Evol 11:921–928.

Lambert JE. 1998. Primate digestion: interactions among anatomy, physiology, and feeding ecology. Evol Anthropol 7:8–20.

Lang D-T, Wang X-P, Wang L, Yu L. 2017. Molecular evolution of pancreatic ribonuclease gene (RNase1) in Rodentia. J Genet Genomics 44:219–222.

Libonati M, Furia A, Beintema JJ. 1976. Basic Charges on Mammalian Ribonuclease Molecules and the Ability to Attack Double-Stranded RNA. Eur J Biochem 69:445–451.

Liu J, Wang X-P, Cho S, Lim BK, Irwin DM, Ryder OA, Zhang Y-P, Yu L. 2014. Evolutionary and functional novelty of pancreatic ribonuclease: a study of Musteloidea (order Carnivora). Sci Rep 4:5070.

Lucas PW, Dominy NJ, Riba-Hernández P, Stoner KE, Yamashita N, Loría-Calderón E, Petersen-Pereira W, Rojas-Durán Y, Salas-Pena R, Solis-Madrigal S, Osorio D, Darvell BW. 2003. Evolution and function of routine trichromatic vision in primates. Evolution 57:2636–2643.

Matsuda I, Clauss M, Tuuga A, Sugau J, Hanya G, Yumoto T, Bernard H, Hummel J. 2017. Factors Affecting Leaf Selection by Foregut-fermenting Proboscis Monkeys: New Insight from in vitro Digestibility and Toughness of Leaves. Sci Rep 7:42774.

Messier W, Stewart CB. 1997. Episodic adaptive evolution of primate lysozymes. Nature 385:151–154.

Milton K, McBee RH. 1983. Rates of fermentative digestion in the howler monkey, *Alouatta palliata* (primates: ceboidea). Comp Biochem Physiol A Comp Physiol 74:29–31.

Milton K, Van Soest PJ, Robertson JB. 1980. Digestive efficiencies of wild howler monkeys. Physiological Zoology:402–409.

Milton K. 1979. Factors influencing leaf choice by howler monkeys: a test of some hypotheses of food selection by generalist herbivores. The American Naturalist:362–378.

Milton K. 1981. Food choice and digestive strategies of two sympatric primate species. The American Naturalist:496–505.

Milton K. 1998. Physiological ecology of howlers (*Alouatta*): energetic and digestive considerations and comparison with the Colobinae. Int J Primatol 19:513–548.

Perelman P, Johnson WE, Roos C, Seuánez HN, Horvath JE, Moreira MAM, Kessing B, Pontius J, Roelke M, Rumpler Y, Schneider MPC, Silva A, O'Brien SJ, Pecon-

Slattery J. 2011. A Molecular Phylogeny of Living Primates. PLoS Genet 7:e1001342–17.

Schienman JE, Holt RA, Auerbach MR, Stewart C-B. 2006. Duplication and divergence of 2 distinct pancreatic ribonuclease genes in leaf-eating African and Asian colobine monkeys. Mol Biol Evol 23:1465–1479.

Scott JR, Godfrey LR, Jungers WL, Scott RS, Simons EL, Teaford MF, Ungar PS, Walker A. 2009. Dental microwear texture analysis of two families of subfossil lemurs from Madagascar. Journal of Human Evolution 56:405–416.

Stewart CB, Schilling JW, Wilson AC. 1987. Adaptive evolution in the stomach lysozymes of foregut fermenters. Nature 330:401–404.

Stoner KE. 1996. Habitat selection and seasonal patterns of activity and foraging of mantled howling monkeys (*Alouatta palliata*) in northeastern Costa Rica. Int J Primatol 17:1–30.

Strier KB. 1992. Atelinae adaptations: behavioral strategies and ecological constraints. Am J Phys Anthropol 88:515–524.

Swanson KW, Irwin DM, Wilson AC. 1991. Stomach lysozyme gene of the langur monkey: tests for convergence and positive selection. J Mol Evol 33:418–425.

Williams-Guillén K. 2003. The behavioral ecology of mantled howling monkeys (*Alouatta palliata*) living in a Nicaraguan shade coffee plantation.

Xu H, Liu Y, Meng F, He B, Han N, Li G, Rossiter SJ, Zhang S. 2013. Multiple bursts of pancreatic ribonuclease gene duplication in insect-eating bats. Gene 526:112–117.

Xu L, Su Z, Gu Z, Gu X. 2009. Evolution of RNases in leaf monkeys: being parallel gene duplications or parallel gene conversions is a problem of molecular phylogeny. Molecular Phylogenetics and Evolution 50:397–400.

Yang Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. Mol Biol Evol 24:1586–1591.

Yu L, Wang X-Y, Jin W, Luan P-T, Ting N, Zhang Y-P. 2010. Adaptive evolution of digestive RNASE1 genes in leaf-eating monkeys revisited: new insights from ten additional colobines. Mol Biol Evol 27:121–131.

Zhang J, Zhang Y-P, Rosenberg HF. 2002. Adaptive evolution of a duplicated pancreatic ribonuclease gene in a leaf-eating monkey. Nat Genet 30:411–415.

Zhang J. 2003. Parallel Functional Changes in the Digestive RNases of Ruminants and Colobines by Divergent Amino Acid Substitutions. Mol Biol Evol 20:1310–1317.

Zhang J. 2006. Parallel adaptive origins of digestive RNases in Asian and African leaf

monkeys. Nat Genet 38:819–823.

Table 2.1. Sequences used in this study with NCBI accession numbers.

Gene	Species	NCBI Accession Number/Location
RNASE1	*Alouatta palliata*	TBD
	Aotus nancymaae	NW_012163535.1 [8186684-8187154]
	Ateles geoffroyi	AF449639.1
	Callithrix jacchus	NC_013905.1 [45071412-45071882]
	Cebus capucinus	NW_016107398.1 [302123-302593]
	Chlorocebus aethiops	AF449635.1
	Colobus guereza	DQ516063.1
	Gorilla gorilla	AF449629.1
	Homo sapiens	NC_000014.9 [20801598-20802069]
	Lagothrix lagotricha	AF449640.1
	Lemur catta	AF449641.1
	Macaca mulatta	AF449632.1
	Macaca nemestrina	AF449633.1
	Miopithecus talapoin	AF449636.1
	Nasalis larvatus	DQ494879.1
	Nomascus leucogenys	AF449631.1
	Pan troglodytes	AF449628.1
	Papio hamadryas	AF449634.1
	Pongo pygmaeus	DQ494868.1
	Procolobus badius	DQ494875.1
	Pygathrix nemaeus	AF449642.1
	Rhinopithecus bieti	GQ334693.1
	Saguinus oedipus	AF449638.1
	Saimiri sciureus	AF449637.1
	Tarsius syrichta	NW_007023513.1 [17837-18307]

RNASE1B	*Alouatta palliata*	TBD
	Colobus guereza	DQ516064.1
	Nasalis larvatus	DQ494863.1
	Pygathrix nemaeus	AF449643.1
	Procolobus badius	DQ494873.1
	Rhinopithecus bieti	GQ334696.1
RNASE1C	*Alouatta palliata*	TBD
	Colobus guereza	DQ516065.1
	Procolobus badius	DQ494874.1
LYZ	*Alouatta palliata*	TBD
	Callithrix jacchus	U76923.1
	Colobus guereza	U76916.1
	Nasalis larvatus	AH004928.2
	Papio anubis	U76919.1
	Saguinus oedipus	U76922.1
	Saimiri boliviensis	NW_003943629.1 [1119538-1124395]
LYZ2	*Bos taurus*	NM_180999.1

Table 2.2. Pairwise identity (%) of lysozyme C (*LYZ*) amino acid sequences

	Bos taurus LYZ2C	*Colobus guereza LYZ*	*Nasalis larvatus LYZ*	*Papio anubis LYZ*	*Alouatta palliata LYZ*	*Callithrix jacchus LYZ*	*Saguinus oedipus LYZ*
Colobus guereza LYZ	75.00						
Nasalis larvatus LYZ	75.00	95.95					
Papio anubis LYZ	68.92	91.22	91.22				
Alouatta palliata LYZ	68.92	84.46	83.78	86.49			
Callithrix jacchus LYZ	68.92	86.49	84.46	89.86	91.89		
Saguinus oedipus LYZ	68.24	85.81	83.78	89.19	91.22	97.97	
Saimiri boliviensis LYZ	68.92	86.49	84.46	89.86	91.22	97.97	97.3

Table 2.3. Results of CODEML analyses for primate *RNASE1* and duplicated sequences (n = 28).

Model	**InL**	**ΔLRT**	***p***
Site-specific model			
M0	-1898.31271	73.274	< 0.00001
M3	-1861.675721		
M1	-1862.385569	0	1
M2	-1862.385569		
M7	-1862.469421	0.124	0.940
M8	-1862.40745		
Branch-site model			
A. palliata RNASE1B & C			
M0	-1862.165721	0	1
MA	-1862.165721		

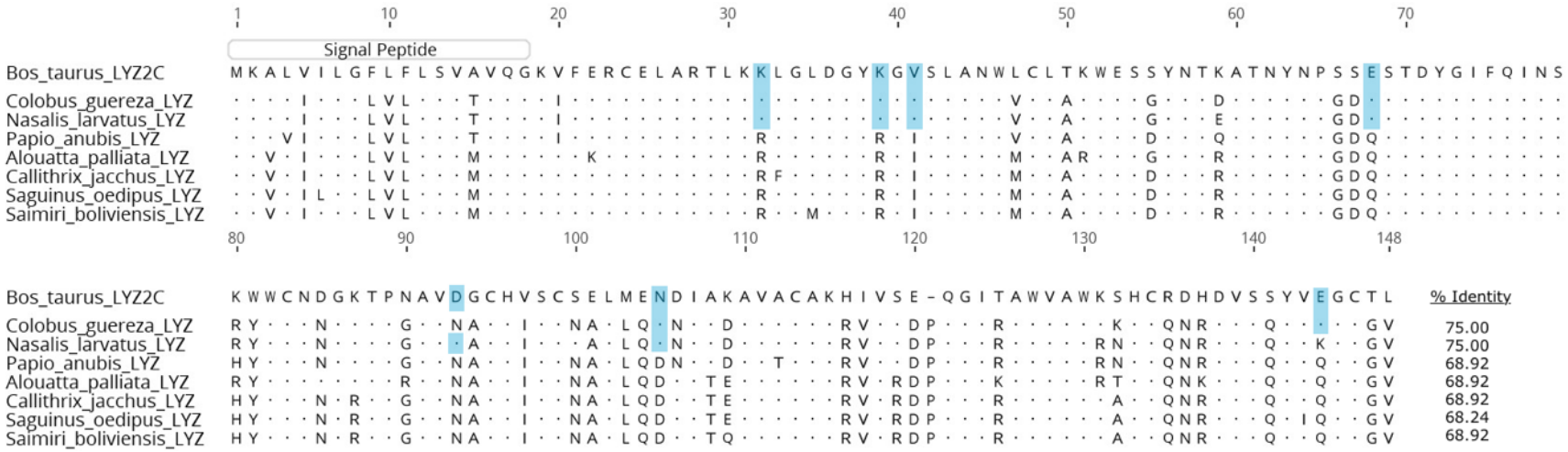
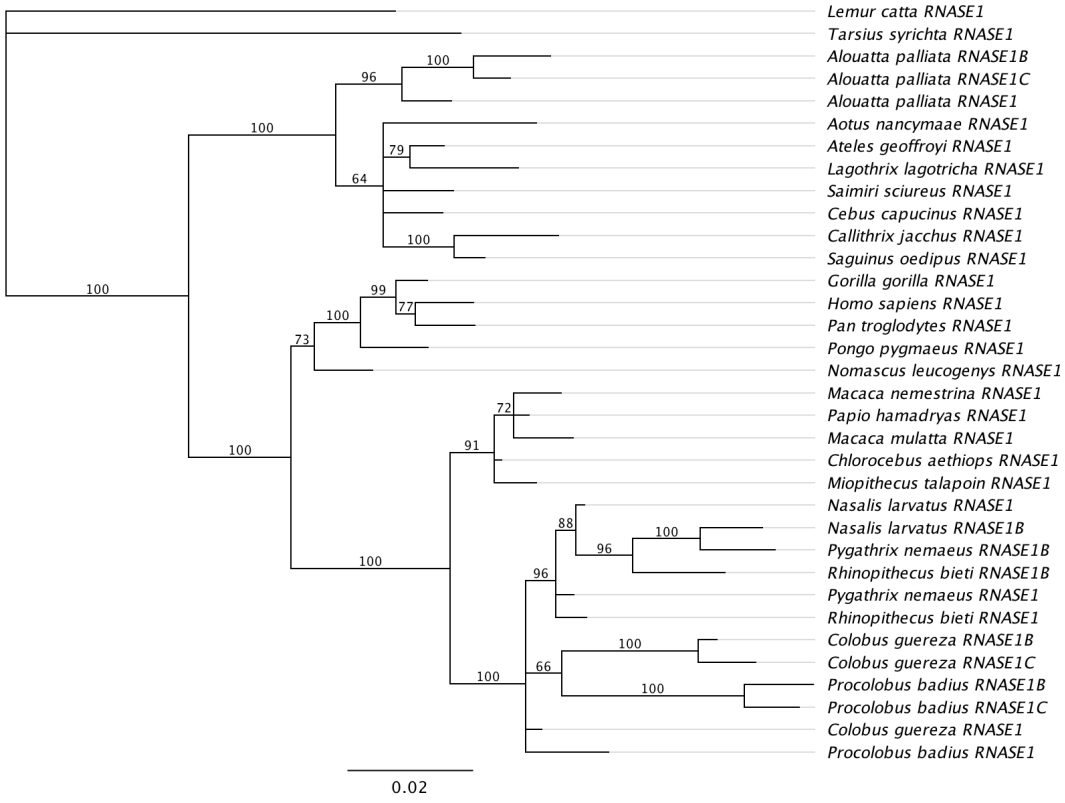


Fig. 2.1. Primate *LYZ* protein sequences aligned to cow (*Bos taurus*) reference sequence. Convergent amino acid changes between cow and colobines are highlighted. Percent identity to the cow reference sequence is indicated in the bottom right.

a)


Lemur catta RNASE1
Tarsius syrichta RNASE1
Alouatta palliata RNASE1B
Alouatta palliata RNASE1C
Alouatta palliata RNASE1
Aotus nancymaae RNASE1
Ateles geoffroyi RNASE1
Lagothrix lagotricha RNASE1
Saimiri sciureus RNASE1
Cebus capucinus RNASE1
Callithrix jacchus RNASE1
Saguinus oedipus RNASE1
Gorilla gorilla RNASE1
Homo sapiens RNASE1
Pan troglodytes RNASE1
Pongo pygmaeus RNASE1
Nomascus leucogenys RNASE1
Macaca nemestrina RNASE1
Papio hamadryas RNASE1
Macaca mulatta RNASE1
Chlorocebus aethiops RNASE1
Miopithecus talapoin RNASE1
Nasalis larvatus RNASE1
Nasalis larvatus RNASE1B
Pygathrix nemaeus RNASE1B
Rhinopithecus bieti RNASE1B
Pygathrix nemaeus RNASE1
Rhinopithecus bieti RNASE1
Colobus guereza RNASE1B
Colobus guereza RNASE1C
Procolobus badius RNASE1B
Procolobus badius RNASE1C
Colobus guereza RNASE1
Procolobus badius RNASE1
0.02

b)

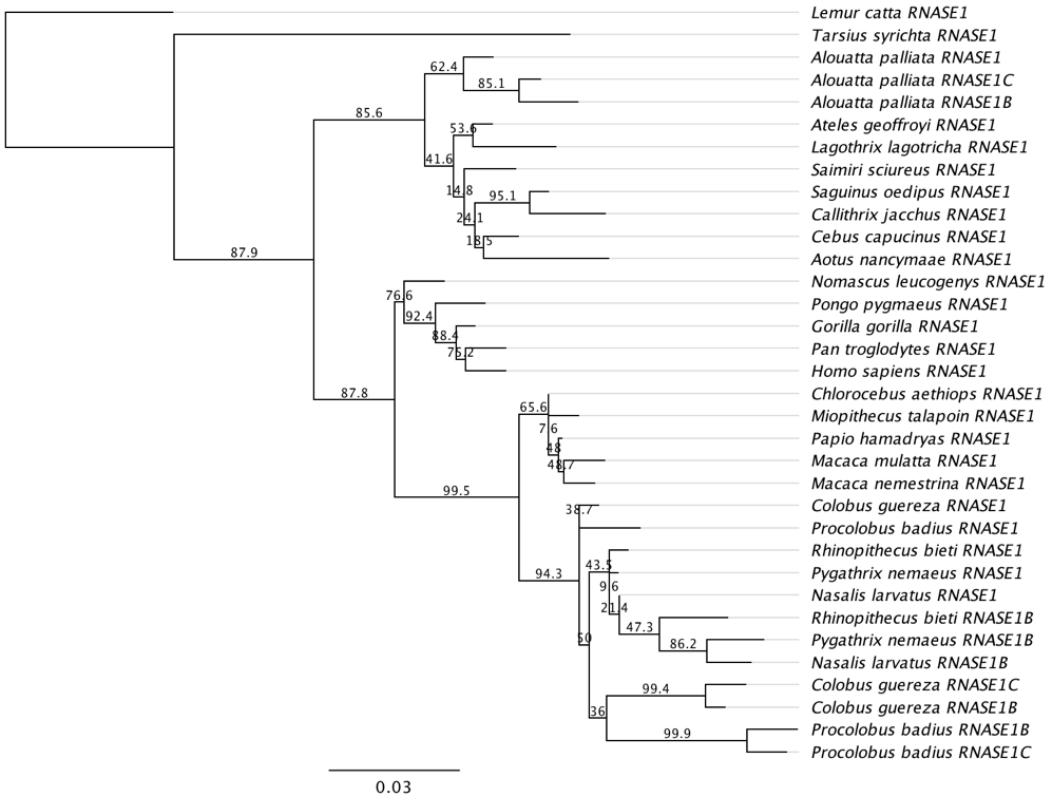


Fig 2.2. Phylogenies of primates based on coding sequences (474 bp) of *RNASE1* and duplications. Trees were built using a a) maximum-likelihood method with PHYML (Guindon et al., 2010) and a b) Bayesian approach with MrBayes (Huelsenbeck and Ronquist, 2001). Branch labels indicate a) bootstrap support in percent (1000 replicates) and b) the posterior probability in percent. Scale bars indicate rate of substitutions per site.

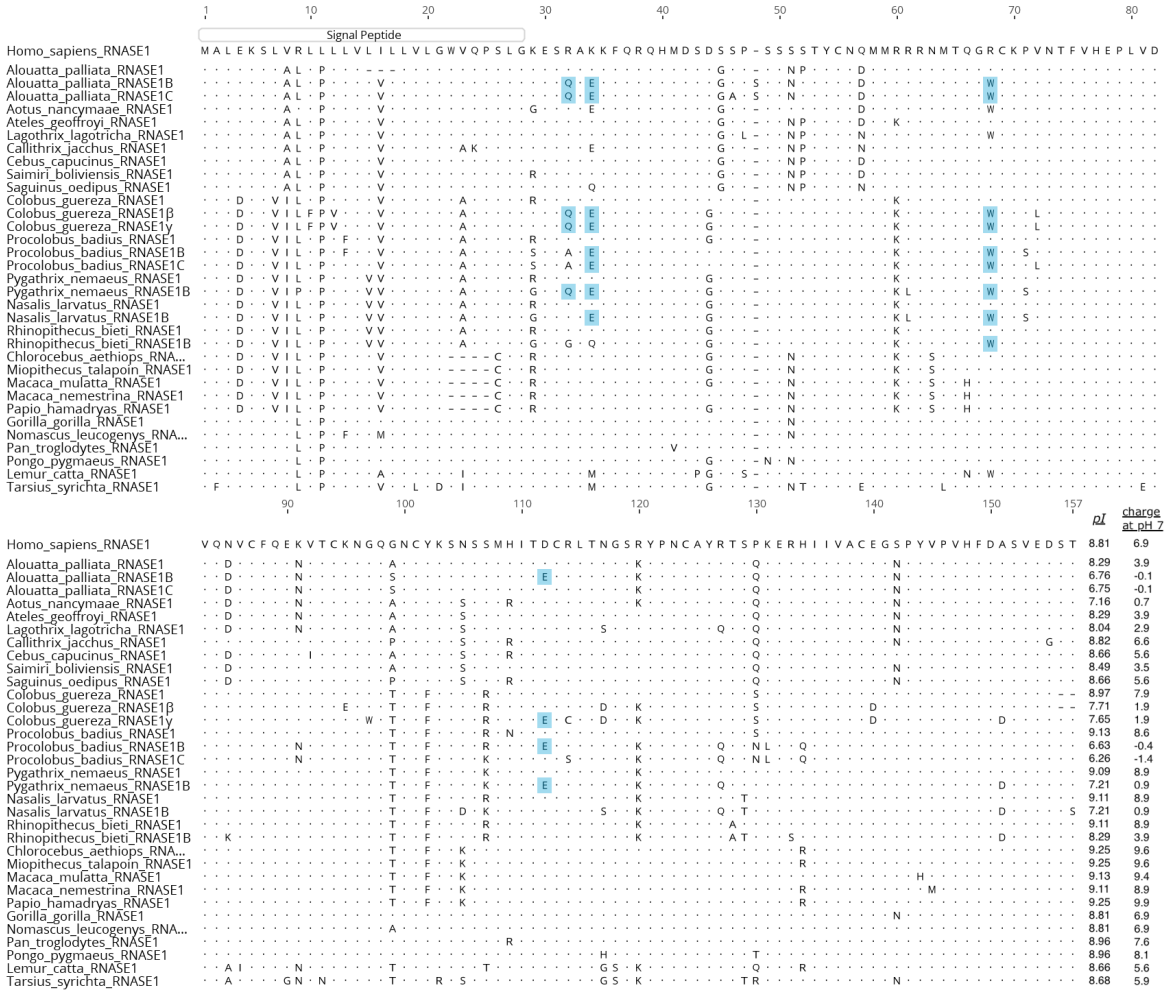


Fig. 2.3. Primate *RNASE1*, *RNASE1B*, and *RNASE1C* sequences aligned to human (*Homo sapiens*) reference sequence. Convergent amino acid changes between duplicated genes (*RNASE1B* and *RNASE1C*) in *Alouatta palliata* and colobines are highlighted. Average isoelectric point (*pI*) and charge at pH 7.0 are shown in the bottom right. *PI* and charge calculated with ProteinCalculator v3.4 (http://protcalc.sourceforge.net/) and ExPASy Compute pI tool (http://web.expasy.org/compute_pi/).

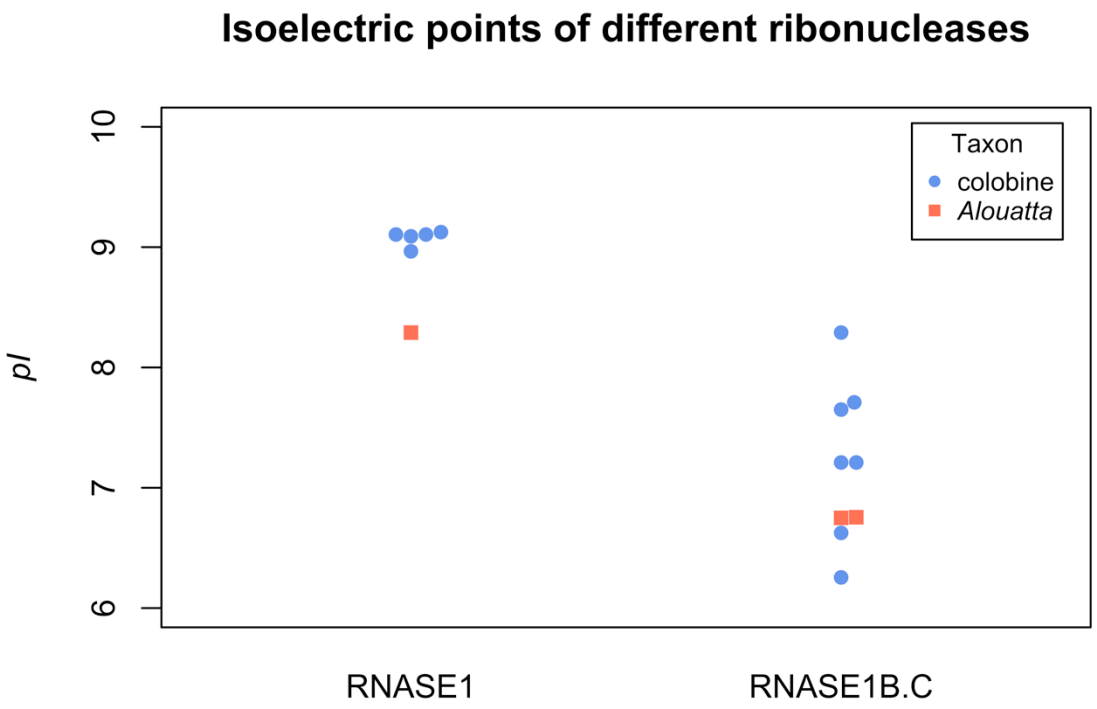


Fig. 2.4. Computed isoelectric points (*pI*) of RNase1 and the duplicated proteins RNase1B and RNase1C in colobines and *Alouatta palliata.*

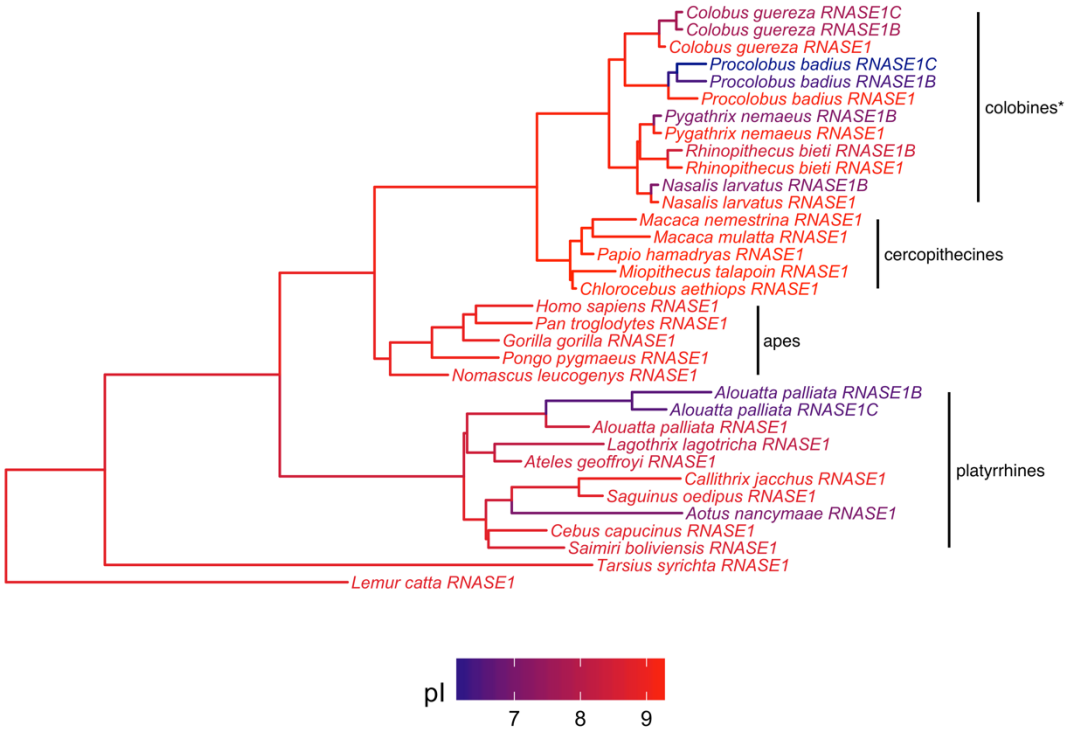


Fig. 2.5. Evolutionary relationships and isoelectric point (*pI*) of the proteins encoded by *RNASE1*, *RNASE1B*, and *RNASE1C* in primates. Branches are colored based on computed *pI* of extant primate protein sequences and reconstructed ancestral protein sequences.

*The colobine clade as shown here likely does not reflect the true evolutionary history of *RNASE1* and gene duplications. Because the evolutionary history of the *RNASE1* genes in colobines has not been fully resolved, the colobine genes are grouped by species here to illustrate the differences in *pI* between the parent and daughter proteins.

Chapter 3. Evolution of the acidic mammalian chitinase gene family (*CHIA*) is related to body mass and insectivory in primates

Abstract

Insects are an important food resource for many primates, but the chitinous exoskeletons of arthropods have long been considered to be indigestible by the digestive enzymes of most mammals. However, recently mice and insectivorous bats were found to produce the enzyme acidic mammalian chitinase (AMCase) to digest insect exoskeletons. Here, we report on the gene *CHIA* and its paralogs, which encode AMCase, in a comparative nonhuman primate sample. Our results show that early primates likely had three *CHIA* genes, suggesting that insects were an important component of the ancestral primate diet. With some exceptions, most extant primate species retain only one functional *CHIA* paralog. The exceptions include two colobine species, in which all *CHIA* genes have premature stop codons, and several New World monkey species that retain two functional genes. The most insectivorous species in our sample also have the largest number of functional *CHIA* genes. *Tupaia chinensis* and *Otolemur garnettii* retain three functional *CHIA* paralogs, while *Tarsius syrichta* has a total of five, two of which may be duplications specific to the tarsier lineage. Selection analyses indicate that *CHIA* genes are under more intense selection in species with higher insect consumption, as well as in smaller-bodied species (<500 g), providing molecular support for Kay's Threshold, a well-established component of primatological theory. These findings suggest that primates, like mice and insectivorous bats, may use the enzyme AMCase to digest the chitin in insect exoskeletons, providing potentially significant nutritional benefits.

Introduction

All primates include some insects in their diet, whether through accidental consumption or through active insectivory (Raubenheimer and Rothman, 2013). Degree of insectivory in living primates ranges from nearly exclusive (e. g., *Tarsius* spp.) to complementary (e. g., the Callitrichidae) to supplemental, such as in the great apes (McGrew, 2001). While insects are a significant source of energy and protein for many living primates, this is especially true for small-bodied primates (Kay, 1984). Due to their sparse distribution in most environments, insects are usually energetically costly to find and catch, making it difficult for a large-bodied primate to fill their nutritional demands solely with insects (Raubenheimer and Rothman, 2013). Small-bodied primates have relatively higher metabolic requirements per unit body mass, but are small enough that the insects they catch suffice to meet their nutritional needs (Fleagle, 2013). A classic concept in primatology, Kay's Threshold (Kay, 1984), suggests that only species below 500 g will be insectivores, while only species above this weight will be folivorous (Fig. 1). Among frugivorous primates, those that are smaller (≤ 1 kg) will typically rely on insects as their source of protein, while those that are larger will use leaves (Gingerich, 1980; Kay, 1984; Fleagle, 2013) (Fig. 1). Social insects, including ants and termites, represent an exception to this rule because they occur as clumped resources across time and space and can be efficiently preyed upon by larger primates (Isbell, 1998). In the case of termites, extractive foraging tools are often used (Goodall, 1986; McGrew, 1992; van Schaik, 2003; Souto et al., 2011). While the nutrient composition of insects varies widely, geometric analyses show that insects eaten by non-human primates tend to have high protein-to-fat ratios and are important sources of minerals that may not otherwise be

included in the diet (Raubenheimer and Rothman, 2013), making them a valuable resource for extant primates and a possible driving force in primate evolution.

The Visual Predation Hypothesis posits that insect predation was the adaptive pressure leading to the evolution of the primate visual system and other morphological features (Cartmill, 1972; 1992; 2012). A suite of adaptations is associated with insectivorous primates, including molars with crests that are used to masticate insect exoskeletons (Kay, 1975), simple guts with low stomach-to-small intestine ratios (Chivers and Hladik, 1980), relatively larger home ranges (Clutton Brock and Harvey, 1980) and small body size (Kay, 1984). The primate visual system and grasping hands have also been suggested as adaptations for preying on insects (Cartmill, 1972; 1992; 2012).

Despite the high nutritional value of insects, there are drawbacks to consuming them. One such drawback is that they are often protected by exoskeletons, which are made up of the structural carbohydrate chitin (Finke, 2007). Chitin makes up between 2-20% of an insect's drymatter and is considered to be indigestible by most primates, unless their digestive systems contain chitinolytic enzymes (Rothman et al., 2014). Given the digestive challenges posed by chitinous exoskeletons (Strait and Vincent, 1998; Rothman et al., 2014), paired with potentially significant energy and amino acid returns if they are digested (Finke, 2007; Rothman et al., 2014), endogenously producing such a chitinolytic enzyme could have important adaptive benefits for insectivorous primates, complementing the dental, behavioral, and morphological adaptations discussed above.

Indeed, mice and insectivorous bats have been shown to digest chitin using an enzyme produced in the stomach called acidic mammalian chitinase (AMCase) (Whitaker

et al., 2004; Strobel et al., 2013; Ohno et al., 2016). This chitinolytic digestive enzyme is produced in the gastric chief cells, where other digestive enzymes are also secreted (Strobel et al., 2013; Ohno et al., 2016). Studies in mice further showed that AMCase is resistant to degradation by the proteases found in the stomach, such as pepsin C, trypsin and chymotrypsin, and breaks down chitin in the presence of these enzymes, as well as at an acidic pH (Ohno et al., 2016). Even though chitinolytic activity has also been observed in the gastric juices of two primates (*Perodicticus potto* and *Cebus capucinus*) (Cornelius et al., 1976; Jeuniaux and Cornelius, 1978), it was long believed that primates (and most other mammals) did not produce such an enzyme and could not digest chitin (Cork and Kenagy, 1989; Oftedal et al., 1991; Simunek and Bartonova, 2005; Strobel et al., 2013; Ohno et al., 2016). Instead, it was thought that insect-eating primates had fast gut-transit times to quickly pass indigestible exoskeletons (Gaulin, 1979; Milton, 1984). More recently, one study found that some primates harbor chitin-digesting microbes (Macdonald et al., 2013) and another study identified an acidic mammalian chitinase in macaques (*Macaca fascicularis*) that is expressed in the stomach and effectively digests chitin at an acidic pH (Krykbaev et al., 2010). However, because macaques are not very insectivorous and in humans AMCase has been associated with type-2 immune response, such as asthma, allergies, eye diseases, and parasite defense (Zhu et al., 2004; Reese et al., 2007; Musumeci et al., 2009; Bucolo et al., 2011; Muzzarelli et al., 2012; Vannella et al., 2016), any potential benefit of AMCase for insectivorous primates remains unresolved.

AMCase is encoded by the gene *CHIA* or one of its paralogs. In primates, two functional *CHIA* genes have been identified (Krykbaev et al., 2010): *hCHIA* and *mCHIA*.

While *hCHIA* remains functional in humans, *mCHIA* has a premature stop codon. In macaques, this is reversed and *mCHIA* is functional, while *hCHIA* has a premature stop codon (Krykbaev et al., 2010).

Here, we present data on the acidic mammalian chitinase gene (*CHIA*) in a large comparative sample of nonhuman primates (n = 34) and one treeshrew (*Tupaia chinensis*). Primates are an order of mammals with over 230 species that have a range of different dietary ecologies (Groves, 2001). Across primates, insect consumption varies from virtually 0% (e.g., colobine monkeys) to almost 100% (*Tarsius* spp.), making them an ideal group for a comparative study of dietary adaptations associated with insectivory.

Hypotheses and Predictions

We hypothesize that production of AMCase is a digestive enzyme adaptation for insectivory and investigate this by comparing the paralogous *CHIA* gene sequences across primates with different levels of insectivory. We predict that primates that routinely consume at least some insects will retain a functional *CHIA* sequence, while *CHIA* pseudogenes will only be found in primates that do not feed on insects. We also test the strength of selection on *CHIA* as a function of insect consumption, predicting that the selection pressure acting on *CHIA* genes is stronger in primates with increased insect intake. Finally, we further explicate the relationship between insectivory and body size, as proposed by Kay's threshold (Kay, 1984), predicting that selective pressures acting on *CHIA* genes are stronger in smaller-bodied primates.

Methods

In this study, we collected and analyzed sequence data on the *CHIA* paralogs of 34 primate species from 27 genera with varying levels of insect consumption (Table 1). We mined the published genomes of 22 primates and one treeshrew (*Tupaia chinensis*) for *CHIA*-like sequences using BLAST and sequenced the *CHIA* genes in additional primate species that do not have publicly available genomes. DNA samples for *Callicebus moloch*, *Callithrix jacchus*, *Saguinus fuscicollis*, and *Saimiri sciureus* were obtained from Coriell Biorepositories (see suppl. Table 1). Dr. George Perry provided extracted DNA from *Sapajus apella*. Extracted DNA from *Erythrocebus patas* was provided by Dr. Todd Disotell. DNA for the following samples was provided by one of the collaborators (Anthony J. Tosi): *Allenopithecus nigroviridis, Allochrocebus lhoesti, Cercopithecus mitis, Chlorocebus aethiops, Colobus guereza kikuyensis,* and *Miopithecus ogouensis*.

Genome Mining

We conducted BLAST searches against the whole-genome sequences of the following taxa, using the *Macaca fascicularis CHIA* gene sequences (Krykbaev et al., 2010) as queries: *Callithrix jacchus*, *Saimiri boliviensis*, *Aotus nancymaae*, *Cebus capucinus imitator*, *Cercocebus atys*, *Mandrillus leucophaeus*, *Papio anubis*, *Macaca mulatta*, *Macaca nemestrina*, *Chlorocebus sabaeus*, *Nasalis larvatus*, *Rhinopithecus roxellana*, *Rhinopithecus bieti*, *Colobus angolensis*, *Gorilla gorilla gorilla*, *Pan paniscus*, *Pan troglodytes*, *Pongo abelii*, *Nomascus leucogenys*, *Tarsius syrichta*, *Microcebus murinus*, *Otolemur garnettii*, and *Tupaia chinensis* whole-genome sequences (GenBank accession numbers in Supplementary Table 1).

Amplification, Sequencing, and Assembly

We designed PCR primers for the amplification of each of the 11 exons of both *CHIA* paralogs *(mCHIA*, *hCHIA*) in Old World and New World monkeys. We first tried to find regions that were conserved across a wide range of primates, including both platyrrhines and catarrhines, however, it was difficult to find intron regions that were suitable as PCR primer sites. Separate sets of PCR primers therefore had to be designed for catarrhine species and for platyrrhine species for most exons (suppl. Table 2). After aligning the *CHIA* sequences of multiple catarrhines (*Macaca fascicularis*, *Mandrillus leucophaeus*, *Nasalis larvatus*, *Cercocebus atys*, *Papio anubis*, *Rhinopithecus roxellana*, *Chlorocebus sabaeus*) and platyrrhines (*Saimiri boliviensis*, *Callithrix jacchus*, *Aotus nancymaae*), ambiguities in the consensus sequence were masked and primers were designed using Primer3 (http://primer3.ut.ee/).

PCRs were carried out with one of two commercial kits. One kit was the Qiagen Fast Cycling PCR Kit MasterMix, which was used following manufacturer's protocols, but with a proportionally reduced volume of 15 µl. Reactions were run at 95˚C for 5 min followed by 35 cycles of 96˚C for 15 sec, 56-59˚C for 15 sec, and 68˚C for 35 sec, and a final step at 72˚C for 7 min. PCR products were purified with exonuclease I and shrimp alkaline phosphatase (ExoSAP) or, when gels showed the amplification of nontarget sequences (i.e. multiple bands), with the Qiagen QIAquick Gel Extraction Kit. Purified PCR products were sequenced directly on the Applied Biosystems 3500 Genetic Analyzer using the Applied Biosystems BigDye Terminator v3.1 Cycle Sequencing Kit and the same primers as used for PCR. Alternatively, PCRs were carried out in 25 µl with the Promega GoTaq G2 Master Mix and run under the following conditions: 95° for 2

min followed by 35-40 cycles of 95° for 20 sec, 55°-60° for 25 sec, and 72° for 30-60 sec. These cycles were followed by a final extension phase at 72° for 5 min. Some of the exons of the *hCHIA* gene were difficult to amplify in *Co*. *guereza*, *Ce. mitis*, *A. lhoesti*, *Ch. aethiops*, *M*. *ogouensis*, and *Al*. *nigroviridis*. For these taxa, we used a long range PCR approach with Promega GoTaq Long PCR Master Mix to amplify the entire *hCHIA* gene (8-9 kb) for these species under the following thermal cycling conditions: 95° for 2 min followed by 35 cycles of 95° for 20 sec, 57° for 25 sec, and 65° for 18 min. These long-range reactions were followed by a final extension phase of 72° for 10 min. These whole-gene products were then cleaned up using Millipore Microcon centrifugal filters by spinning the products at 500 rcf for 4 min, and then spinning the inverted filter to recover only dsDNA at 1000 rcf for 3 min. We then diluted the cleaned-up DNA concentrate with 10 µl of nuclease-free water. This step was taken in order to quantitate the DNA so that we could use an appropriate amount of these products for subsequent nested PCRs, which targeted the problematic exons mentioned above. The nested products were sequenced directly at the Molecular Cloning Laboratories (San Francisco, CA) on an ABI 3730XL sequencer.

Reads were assembled, mapped to the *Macaca fascicularis* reference sequence, concatenated into coding sequences, and translated using Geneious v. 9.1.8. Coding regions and translated amino acid sequences were aligned with MAFFT alignment server v. 7.

Sequence Analyses

Sequences were visually inspected for frameshift mutations causing premature stop codons; these were considered to be *CHIA* pseudogenes. To determine whether full-

length *CHIA* sequences were likely coding for a functional enzyme, we inspected the translated amino acid sequences for conservation of the catalytic site motif and chitin-binding domain, signatures of functional chitinases (Tjoelker et al., 2000; Krykbaev et al., 2010).

We looked for evidence of positive selection acting on sites along both *CHIA* genes in primates with the CODEML program in the PAML package (Yang, 2007). We used site-specific models (M0 – null, M1a – nearly neutral selection, M2a – positive selection, M3 – discrete, M7 – beta, and M8 – beta & ω>1) to determine if there is variation in the ratios of nonsynonymous to synonymous nucleotide substitutions (dN/dS or ω) across sites along the *mCHIA* and *hCHIA* alignments and to test for evidence of positively selected sites (Yang, 2007). Model fit was evaluated using likelihood ratio tests (LRT). The software program RELAX, which is part of the HyPhy package, tests for relaxed versus intensified selection in a codon-based phylogenetic framework (Wertheim et al., 2015). Given two sets of branches in a phylogeny (foreground and background, or test and reference branches), RELAX tests whether selection is intensified or relaxed in one set versus the other. To do this, RELAX uses a branch-site evolutionary (BS-REL) model to estimate the distribution of ω for each of the two branch sets and then compares this distribution using two models. In the null model the selection intensity parameter k is constrained to 1, causing the ω distribution to be the same on both test and reference branches, while under the alternative model k is allowed to vary. If the latter model is a significantly better fit (as determined by a LRT), this suggests that selection on the test branches is either relaxed ($k < 1$) or intensified ($k > 1$) compared to the reference branches (Wertheim et al., 2015). In addition to the null and alternative models, the

partitioned exploratory model can provide a quantitative measure of selection patterns in the test and reference branches. The partitioned exploratory model is a less constrained model that allows the proportion of sites in each category of ω to vary between test and reference branches. Here we used RELAX to test whether selection was relaxed or intensified along any of the branches in our phylogeny as a function of either insect consumption or body size. Specifically, we hypothesized that selection on *CHIA* genes would be relaxed in lineages with a) pseudogenizing mutations, b) below average insect consumption, and c) body size above Kay's threshold (500 g).

Phenotypic Data

Data on insect consumption and body size were taken from the literature. Most of the data on body size came from (Smith and Jungers, 1997) and only measures from wild specimens were used. To avoid skews introduced by large adult males in sexually dimorphic species we only used adult female body weight measures for all species. Average adult female body weight for each species is presented in Table 1. Species-specific information on average annual insect consumption was taken from studies of wild primates, the results of many of which are collected in (Campbell et al., 2011). Annual averages of insect consumption used in our analyses are presented in Table 1, while detailed information on all data used to calculate these averages is given in Supplementary Table 3. Species were classified as relatively more or less insectivorous depending on whether their average annual insect consumption was higher or lower than the average annual insect consumption across all species included in our sample (16.58%). Where possible we used data from (1) studies that covered at least 12 months to account for seasonal variation in insect consumption, and (2) multiple studies for each

species to account for inter-population differences. The diets of a few of the species included here have not been studied in the wild; in these cases, we used data from a closely related species where available (*Saimiri boliviensis*, *Aotus nancymaae*), or where unavailable (*Mandrillus leucophaeus*) left the branch "unclassified" in the RELAX analyses involving dietary data.

Results

We successfully sequenced two *CHIA* genes in 12 primate species for which whole-genome sequences are not available. The sequences we generated for *Callithrix jacchus* contained numerous differences to the reference genome sequence. Since our sequences were more parsimonious in the comparative context of our study, both compared to sequences generated by us and the whole-genome sequences of closely related species, we believe that our sequences most likely accurate. All sequences have been deposited in GenBank (accession numbers in Supplementary Table 1). We further found both homologous sequences (*mCHIA* and *hCHIA*) in the 23 genomes we surveyed, with the exception of *Microcebus murinus* in which we could only positively identify one complete *CHIA* sequence (*hCHIA*). In the tarsier (*Tarsius syrichta*), galago (*Otolemur garnettii*), and Chinese treeshrew (*Tupaia chinensis*) genomes we identified more than two *CHIA* sequences. Both the galago and the Chinese treeshrew genomes had a third *CHIA* gene sequence (*CHIA3*), while the tarsier genome included three additional *CHIA* sequences (*CHIA3, CHIA4, CHIA5*) for a total of five homologous genes.

Phylogenetic trees generated from an alignment of the coding sequences (Fig. 2) show a deep split between the *CHIA* genes, indicating that *mCHIA* and *hCHIA* (and likely *CHIA3*) arose in a duplication event that was ancestral to primates and treeshrews and are not independently duplicated genes. While we did not find complete *CHIA3* sequences in any primates other than the tarsier and galago, partial sequences were identified in some genomes.

CHIA Pseudogenizations

While all species except the mouse lemur (*Microcebus murinus*) had two complete *CHIA* sequences, one of these sequences often contained frameshift-causing indels or nonsense mutations leading to premature stop codons and likely rendering the gene nonfunctional. As a result, most primates only retain one full-length, and likely functional *CHIA* paralog (Fig. 3). With some exceptions, *hCHIA* remains functional only in apes, while only *mCHIA* is functional in most monkeys. Two species in our data set did not retain any functional *CHIA* genes. In *Rhinopithecus bieti* and *Nasalis larvatus* both the *mCHIA* and *hCHIA* sequences contained premature stop codons (Fig. 3, Fig. 4a). Several New World primates, on the other hand, retained full-length sequences of both *mCHIA* and *hCHIA*, including *Callithrix jacchus*, *Cebus capucinus*, *Saimiri boliviensis*, *S. sciureus*, and *Saguinus fuscicollis* (Fig. 3, fig. 4b). All three *CHIA* sequences in the galago and treeshrew, and all five sequences in the tarsier were free from any indels or premature stop codons (Fig. 3, Fig. 4c-d).

Interestingly, across our sample of primate species (n = 34), premature stop codons were independently introduced into the *hCHIA* or *mCHIA* sequences numerous times, through frameshift or nonsense mutations at various sites along the sequence. The

mCHIA gene lost function independently at least six times in primates: three times in the apes and three times in the colobine monkeys (Fig. 3). Premature stop codons in the *hCHIA* gene arose at least seven times: three times in New World monkeys, twice in colobine monkeys, and two (possibly three) times in cercopithecine monkeys (Fig. 3), a subfamily that includes the tribes Cercopithecini and Papionini (Table 1). The Papionini (*Macaca* spp, *Mandrillus leucophaeus*, *Cercocebus atys*, and *Papio anubis*) share a deletion in exon 8 that causes a frameshift and premature stop codon, while most of the Cercopithecini (*Cercopithecus mitis*, *Allenopithecus nigroviridis*, *Allochrocebus lhoesti*, *Erythrocebus patas*, and *Chlorocebus* spp.) share a frameshifting deletion and premature stop codon at the beginning of exon 11. Even though *Miopithecus ogouensis* is considered to be part of the tribe Cercopithecini (Tosi et al., 2002; 2005), this species shares the exon 8 deletion with the Papionini and lacks the exon 11 deletion characteristic of its tribe (Fig. 3). Exons 8 and 11 were sequenced again in another lab, using a different *Miopithecus* sample, confirming these results. At this time, it is unclear what accounts for this pattern. Possible explanations include polymorphism in the common ancestor of the Cercopithecini and Papionini, or ancient hybridization.

Signatures of Catalytically Active Chitinases

In all of the full-length *CHIA* amino acid sequences we found that the signatures of catalytically active chitinases were conserved (Fig. 5): these have a conserved glutamate and the consensus sequence DXXDXDXE at the active site (Synstad et al., 2004; Krykbaev et al., 2010). In addition, catalytically active chitinases further have a chitin-binding domain at the C-terminus, containing six cysteines, which are essential for attaching the enzyme to the chitin (Tjoelker et al., 2000). All of the full-length *CHIA*

sequences in our study had conserved chitin-binding domains that contained all six cysteines (Fig. 5).

CODEML Analyses

Results of the CODEML analysis for *mCHIA* suggested that all sites in this gene are under purifying or neutral selection (Table 2). Comparison of the models M0 and M3 falsified the null hypothesis that the same dN/dS ratio (ω) applies to all sites of the *mCHIA* gene ($\chi^2 = 79.54$; $df = 4$; $P < 0.00001$); however, this result is expected for most functional proteins. Interestingly, CODEML estimated all three discrete ω groups below 1.00, and it placed 75.5% of sites into categories that have a dN/dS ratio of 0.09. Comparison of models M1a and M2a ($\chi^2 = 0$, $df = 2$, $P = 1$) and models M7 and M8 ($\chi^2 = 2.85$, $df = 2$, $P = 0.24$) both failed to support any hypothesis of positive selection. Overall, these models indicate that most *mCHIA* codons appear to be under purifying selection with a smaller number of sites under neutral selection (16-19%, Table 2).

Results for the same analyses for the *hCHIA* gene indicated that a small number of sites (1.0-1.5%) may be subject to positive selection (Table 2). Both the comparison of models M1a and M2a ($\chi^2 = 8.70$, $df = 2$, $P = 0.013$) and of models M7 and M8 ($\chi^2 = 14.76$, $df = 2$, $P = 0.001$) favored the hypothesis that sites in *hCHIA* are under positive selection over the null hypothesis. Bayes Empirical Bayes analysis indicated two sites that had a significant probability of being under positive selection, 36P and 62Q (Table 2). As with the *mCHIA* gene, the majority of sites in *hCHIA* appear to be under purifying selection.

RELAX Analyses

While CODEML results suggested that most sites in *mCHIA* and *hCHIA* are under purifying selection, we also tested whether the strength of purifying selection acting on these sites varies across different branches of our phylogeny using the program RELAX. For *hCHIA*, RELAX results supported the hypothesis that selection was relaxed in species in which the gene has become pseudogenized ($k = 0.02$, $P = 0$, $LR = 74.26$, table 3). Compared to branches in which *hCHIA* remains functional, the ω values of pseudogene branches were shifted towards neutrality ($\omega = 1$) indicating relaxed selection in these species (Fig. 6a). Branch specific inferences of the selection intensity parameter (k) under the General Descriptive model in RELAX showed that the branches under more intense selection are ones with functional *hCHIA* genes, such as *Saimiri sciureus* ($k = 1.99$), *Otolemur garnettii* ($k = 2.07$), *Tupaia chinensis* (k = 2.47), and *Cebus capucinus* ($k = 4.74$) (Fig. 7a). These species also have some of the highest insect intakes in our sample (Table 1). Results of a RELAX test including only functional *hCHIA* sequences supported the hypothesis that *hCHIA* is under more intense selection in species with higher insect consumption than in species with lower insect consumption ($k = 0.20$, $P < 0.001$, $LR = 26.6$, table 3). The ω values of species with lower insect consumption were shifted toward neutrality compared to those of species with higher insect intake (Fig. 6b), but the majority of sites remained below $\omega = 1$ (0.165, 78%). We found similar results for our test of Kay's threshold (Fig. 6c). Selection on *hCHIA* was relaxed in species with body weights above this 500 g threshold ($k = 0.67$, $LR = 7.68$, $P = 0.006$, table 3).

The initial RELAX results for *mCHIA* suggested a more complex pattern. The RELAX test comparing pseudogene branches to branches with functional *mCHIA* genes was significant for selection intensification ($k = 1.88$, $P < 0.001$) acting on pseudogene

branches. However, the ω distributions of the best-fitting model, Partitioned Exploratory (LR = 53.24, P = <0.001, table 3), suggested that most sites (99.89%) of *mCHIA* pseudogenes are shifted toward neutrality (0.88-0.981) compared to most sites (97%) of functional genes (ω = 0.23-0.04). Only a very small number of *hCHIA* pseudogene sites (0.11%) were pushed far above neutrality (ω = 325) (Fig. 6d). Similarly, the RELAX test comparing *mCHIA* between species with high and low insect consumption indicated intensified selection on branches with lower insect intake (k = 2.06, LR = 17.50, P < 0.001, Table 3), but the best-fitting model, Partitioned Exploratory (Fig. 6e), suggested a pattern of relaxation. When comparing species above and below Kay's threshold, RELAX results reject the hypothesis that *mCHIA* is under relaxed selection in larger-bodied species, suggesting that they are instead under intensified selection (k = 3.81, LR = 20.8, P < 0.001, Table 3) with ω values shifted away from neutrality (Fig. 6e).

Interestingly, branch specific inferences of k under the General Descriptive model show that three branches in our *mCHIA* sample that were part of the reference branches in all three RELAX tests have very strong selective regimes, two extremely so (Fig. 7b). *Otolemur garnettii* (k = 3.89), *Tarsius syrichta* (k = 34.59), and *Tupaia chinensis* (k = 50) all appear to be under intense selection. It is likely that these extreme values and the long branch lengths observed in these species were biasing the RELAX test (Pond SLK, pers. comm.). We therefore ran these models again, leaving the *O. garnettii*, *T. syrichta*, and *Tu. chinensis* branches "unclassified" to exclude them from the hypothesis test and eliminate any bias contributed by these outliers. In support of our predictions, the RELAX test indicated that *mCHIA* pseudogene branches are under relaxed selection compared to branches with functional genes (k = 0.58, LR = 15.08, P < 0.001), and that

selection on *mCHIA* is relaxed in species with lower insect consumption ($k = 0.42$, $LR = 11.78$, $P < 0.001$). The model comparing selection between smaller-bodied and larger-bodied species could not be retested, because only two species below 500 g remained in our sample after removing the aforementioned species.

Discussion

He we present the first comparative study of *CHIA* genes in primates, including species with a variety of dietary ecologies and different levels of insectivory. Our findings are consistent with the hypothesis that primates produce acidic mammalian chitinase (AMCase) as a digestive enzyme for the breakdown of insect exoskeletons and that this enzyme is under more intense selection in more insectivorous and smaller-bodied primates.

We have identified a number of homologs of the *CHIA* gene in primates, *mCHIA*, *hCHIA*, *CHIA3*, *CHIA4*, and *CHIA5*. Three of these, *mCHIA*, *hCHIA*, and *CHIA3* were previously identified by (Krykbaev et al., 2010) and are most likely ancestral to the Euarchonta, as they are also found in the treeshrew (*Tupaia chinensis*) (Fig. 2, fig. 3). Because we could only identify a complete and putatively functional sequence for *CHIA3* in the treeshrew, the galago (*Otolemur garnettii*), and the tarsier (*Tarsius syrichta*) (Fig. 4), we conclude that *CHIA3* was likely lost early in the anthropoid lineage (Fig. 3).The additional two genes are only found in the tarsier (*T. syrichta*), the most insectivorous of all primates (Table 1, (Gursky, 2011)). These genes likely arose as duplications after the tarsier lineage split from the anthropoids between 60-70 MYA (Pozzi et al., 2014; Di Fiore et al., 2015; Kistler et al., 2015), but it is unclear which of the three ancestral *CHIA*

genes gave rise to *CHIA4* and *CHIA5*, as they are most similar to each other and equidistant from tarsier *mCHIA*, *hCHIA*, and *CHIA3* (Fig. 2).

With the exception of *Microcebus murinus* (see Future Directions), the genes *mCHIA* and *hCHIA* are found in all species in our sample, but show a pattern of independent pseudogenization events that is consistent with our hypothesis that acidic mammalian chitinase is a digestive enzyme evolved and retained for insectivory. Most primates include at least some insects in their diet (Raubenheimer and Rothman, 2013) and we find that most primates have one functional *CHIA* gene (Fig. 3). Only two species in our sample did not retain any functional *CHIA* genes, the black snub-nosed monkey (*Rhinopithecus bieti*) and the proboscis monkey (*Nasalis larvatus*) (Fig. 3, Fig. 4a). Both of these monkeys are colobines, a subfamily of primates that are highly folivorous and generally do not include insects in their diet (Table 1). Further, the only species in which more than one *CHIA* gene remained functional are species with above-average levels of insectivory. This includes some of the New World monkeys, which are generally both smaller-bodied and more reliant on insects than Old World monkeys (Gaulin, 1979; Terborgh, 1983). In the common marmoset (*Callithrix jacchus*), saddleback tamarin (*Saguinus fuscicollis*), the common and Bolivian squirrel monkeys (*Saimiri sciureus* and *S. boliviensis*, respectively), and the white-faced capuchin (*Cebus capucinus*) (Fig. 4b), both *mCHIA* and *hCHIA* were free of premature stop codons and were conserved in the catalytic site and chitin-binding domain (Fig. 5).

Finally, the most insectivorous primate species in our sample are also the species with the greatest number of (putatively) functional *CHIA* genes, the galago and tarsier having three and five functional *CHIA* paralogs, respectively (Fig. 4c-d). Tarsiers are the

only primates that are entirely faunivorous (Gursky, 2011), feeding mostly on arthropods and sometimes small vertebrates (Niemitz, 1984; Anon, 2015). While there are no gene expression data for the tarsier digestive system, it is noteworthy that, despite consuming arthropods almost exclusively, tarsier feces do not contain visible insect exoskeletons. Instead, their feces are described as "powder-like" (Gursky S, personal communication), suggesting that the exoskeletons ingested by the tarsier are broken down completely during gut transit. Therefore, these unique *CHIA* duplications are likely to be important digestive adaptations.

The primates with more than one functional *CHIA* gene are not just more insectivorous but are also among the smaller species in our sample (Table 1), in accordance with what is predicted by Kay's threshold (Fig. 1). Because it is costly for small-bodied primates to fill their relatively smaller guts with indigestible bulk, such as leaves (Gaulin, 1979; Fleagle, 2013), the ability to efficiently digest insect exoskeletons using AMCase might be an especially valuable adaptation for these species. The results of our selection analyses show that *CHIA* genes are under intensified purifying selection in more insectivorous species (Fig. 7a, b) and also partially support the hypothesis that selection on *CHIA* genes is relaxed in larger-bodied primates (Fig. 6c). While the initial *mCHIA* RELAX test did not find support for this hypothesis (Fig. 6e) and had inconsistent results for the other two tests (pseudogenes vs. functional genes [Fig. 6c], above-average vs. below-average insectivory [Fig. 6d]) it is worth noting that, in accordance with our predictions, *mCHIA* appears to be under extreme selection in the three most insectivorous species (*Tupaia chinensis, Tarsius syrichta*, and *Otolemur garnettii*). Further, it is likely that the extremely high *k* values and long branch lengths of

these lineages (Fig. 7b) masked other selective patterns across the alignment and biased the RELAX model (Pond, personal communication). This is supported by the results of additional analyses in which *Tu. chinensis, T. syrichta*, and *O. garnettii* were left "unclassified" and not included in the reference branches. Here, we found support for the hypotheses that *mCHIA* is under relaxed selection in (1) pseudogene branches and (2) less insectivorous species.

Our results may also shed some light on primate origins. Two commonly cited and competing hypotheses to explain the evolution of primate features, such as the visual system and grasping hands, are the Visual Predation Hypothesis (Cartmill, 1972; 1992; 2012) and the Terminal Branch Feeding Hypothesis (Sussman, 1991). According to the former, primate vision and hands evolved to facilitate the predation on insects, while the latter proposes that primate features evolved to allow for the exploitation of angiosperms at the ends of terminal branches. While these two hypotheses do not have to be mutually exclusive, our results show that early primates likely had three *CHIA* genes, suggesting that insects were an important component of the ancestral primate diet.

Limitations and Future Directions

We were unable to confirm complete *CHIA* sequences in most of the currently available lemur genomes. In the mouse lemur (*Microcebus murinus*) genome we identified the full *hCHIA* sequence and found partial sequences of one, or possibly two, similar genes. However, we could not identify all exons of these additional genes with confidence, an issue that was also encountered with other lemur genomes. We therefore chose to focus on other taxa, but look forward to revisiting the evolutionary history of chitinase genes in the strepsirrhines as additional high-quality genomes become available.

One limitation of the current study is the available data on primate diets. During our review of the literature, we encountered many inconsistencies in the ways that data on food intake are collected, making it difficult to compare levels of insect consumption across species and even populations. In addition, identifying and observing insect-feeding is notoriously difficult (Frazão, 1991), so it is very likely that insect consumption in many primates, and especially arboreal primates, is underestimated.

Finally, without gene expression data for primate digestive systems, we cannot be certain whether the various *CHIA* genes are actually expressed in the stomachs of all primates. Krykbaev and colleagues report that *mCHIA* is highly expressed in the stomach of *Macaca fascicularis*, but expression data for humans suggests that the role of *hCHIA* may be more complicated. While some studies show that *hCHIA* is expressed in the human stomach (Boot et al., 2005; Krykbaev et al., 2010), there is disagreement over whether the gene actually translates into a *functional* chitinase in the stomach. One study found that chitinolytic activity in the gastric juice was present in 80% of their participants, but absent in the other 20% (Paoletti et al., 2007), while another study failed to detect any evidence of chitinase in the human digestive system (Goto et al., 2003). It is plausible that this may be due to dietary difference across human populations and invites further study in groups with long histories of insect consumption.

Conclusion

Even though several studies from the 1970s suggested that primates have chitinolytic enzymes (Cornelius et al., 1976; Jeuniaux and Cornelius, 1978; Kay and Sheine, 1979), the notion that chitin was indigestible by the endogenous digestive enzymes of primates and other mammals has persisted (Cork and Kenagy, 1989; Oftedal

et al., 1991; Simunek and Bartonova, 2005; Strobel et al., 2013; Ohno et al., 2016). Here we present evidence that suggests insect-eating primates share an adaptation found in insectivorous bats (Vespertilionidae) and mice (*Mus musculus*) (Strobel et al., 2013; Ohno et al., 2016) and use the enzyme acidic mammalian chitinase to digest the chitin in insect exoskeletons. The efficient digestion of insect exoskeletons is likely to have important adaptive benefits for all insect-eating primates, through the potentially significant energy and amino acid returns from the digestion of the polysaccharide chitin (Finke, 2007; Rothman et al., 2014) and, for small-bodied primates, by reducing the amount of indigestible bulk in their guts.

References

Boot RG, Bussink AP, Verhoek M, de Boer PAJ, Moorman AFM, Aerts JMFG. 2005. Marked differences in tissue-specific expression of chitinases in mouse and man. J Histochem Cytochem 53:1283–1292.

Bucolo C, Musumeci M, Musumeci S, Drago F. 2011. Acidic Mammalian chitinase and the eye: implications for ocular inflammatory diseases. Front Pharmacol 2:43.

Campbell CJ, Fuentes A, MacKinnon KC, Bearder SK, Stumpf RM. 2011. Primates in Perspective. 2nd ed. (Campbell CJ, Fuentes A, MacKinnon KC, Bearder SK, Stumpf RM, editors.). Oxford University Press.

Cartmill M. 1972. Arboreal adaptations and the origin of the order Primates. In: Tuttle RH, editor. The functional and evolutionary biology of primates. Aldine-Atherton. p 97–122.

Cartmill M. 1992. New views on primate origins. Evol Anthropol 1:105–111.

Cartmill M. 2012. Primate origins, human origins, and the end of higher taxa. Evol Anthropol 21:208–220.

Chivers DJ, Hladik CM. 1980. Morphology of the gastrointestinal tract in primates: comparisons with other mammals in relation to diet. Journal of Morphology 166:337–386.

Clutton Brock TH, Harvey PH. 1980. Primates, brains and ecology. J Zoology 190:309–323.

Cork SJ, Kenagy GJ. 1989. Nutritional Value of Hypogeous Fungus for a Forest-Dwelling Ground Squirrel. Ecology 70:577–586.

Cornelius C, Dandrifosse G, Jeuniaux C. 1976. Chitinolytic enzymes of the gastric mucosa of *Perodicticus potto* (primate prosimian): Purification and enzyme specificity. International Journal of Biochemistry 7:445–448.

Frazão E. 1991. Insectivory in free-ranging bearded saki (*Chiropotes satanas chiropotes*). Primates 32:243–245.

Di Fiore A, Chaves PB, Cornejo FM, Schmitt CA, Shanee S, Cortés-Ortiz L, Fagundes V, Roos C, Pacheco V. 2015. The rise and fall of a genus: Complete mtDNA genomes shed light on the phylogenetic position of yellow-tailed woolly monkeys, *Lagothrix flavicauda*, and on the evolutionary history of the family Atelidae (Primates: Platyrrhini). Molecular Phylogenetics and Evolution 82 Pt B:495–510.

Finke MD. 2007. Estimate of chitin in raw whole insects. Zoo Biol 26:105–115.

Fleagle JG. 2013. Primate adaptation and evolution. 3rd ed. Academic Press.

Gaulin SJ. 1979. A Jarman/Bell model of primate feeding niches. Human Ecology 7:1–20.

Gingerich PD. 1980. Dental and cranial adaptations in Eocene Adapidae. Z Morphol Anthropol 71:135–142.

Goodall J. 1986. The chimpanzees of Gombe: Patterns of behavior. Cambridge, MA: Belknap Press of Harvard University Press.

Goto M, Fujimoto W, Nio J, Iwanaga T, Kawasaki T. 2003. Immunohistochemical demonstration of acidic mammalian chitinase in the mouse salivary gland and gastric mucosa. Arch Oral Biol 48:701–707.

Groves CP. 2001. Primate taxonomy. Washington, D.C.: Smithsonian.

Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307–321.

Gursky SL. 2007. The Spectral Tarsier. Abingdon, UK: Routledge.

Gursky SL. 2011. Tarsiiformes. In: Campbell CJ, Fuentes A, MacKinnon KC, Bearder SK, Stumpf RM, editors. Primates in Perspective. Oxford University Press. p 79–90.

Isbell LA. 1998. Diet for a small primate: Insectivory and gummivory in the (large) patas monkey (*Erythrocebus patas pyrrhonotus*). Am J Primatol 45:381–398.

Jeuniaux C, Cornelius C. 1978. Distribution and activity of chitinolytic enzymes in the digestive tract of birds and mammals. In: Muzzarelli RAA, Pariser ER, editors. p 542–549.

Jones KE, Bielby J, Cardillo M, Fritz SA, O'Dell J, Orme CDL, Safi K, Sechrest W, Boakes EH, Carbone C, Connolly C, Cutts MJ, Foster JK, Grenyer R, Habib M, Plaster CA, Price SA, Rigby EA, Rist J, Teacher A, Bininda-Emonds ORP, Gittleman JL, Mace GM, Purvis A. 2009. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. Ecology 90:2648–2648.

Kay RF, Sheine WS. 1979. On the relationship between chitin particle size and digestibility in the primate Galago senegalensis. Am J Phys Anthropol 50:301–308.

Kay RF. 1975. The functional adaptations of primate molar teeth. Am J Phys Anthropol 43:195–216.

Kay RF. 1984. On the use of anatomical features to infer foraging behavior in extinct primates. In: Cant J, Rodman PS, editors. Adaptations for Foraging in Nonhuman Primates. p 21–53.

Kistler L, Ratan A, Godfrey LR, Crowley BE, Hughes CE, Lei R, Cui Y, Wood ML, Muldoon KM, Andriamialison H, McGraw JJ, Tomsho LP, Schuster SC, Miller W, Louis EE, Yoder AD, Malhi RS, Perry GH. 2015. Comparative and population mitogenomic analyses of Madagascar's extinct, giant 'subfossil' lemurs. Journal of Human Evolution 79:45–54.

Krykbaev R, Fitz LJ, Reddy PS, Winkler A, Xuan D, Yang X, Fleming M, Wolf SF. 2010. Evolutionary and biochemical differences between human and monkey acidic mammalian chitinases. Gene 452:63–71.

Macdonald C, Barden S, Foley S. 2013. Isolation and characterization of chitin-degrading micro-organisms from the faeces of Goeldi's monkey, *Callimico goeldii*. J Appl Microbiol 116:52–59.

McGrew WC. 1992. Chimpanzee material culture: Implications for human evolution. Cambridge: Cambridge University Press.

McGrew WC. 2001. The other faunivory: primate insectivory and early human diet. In: Stanford CB, Bunn HT, editors. Meat-eating and human evolution. Oxford: Meat-eating and human evolution. p 160–178.

Milton K. 1984. The role of food-processing factors in primate food choice. Adaptations for Foraging in Nonhuman Primates, Columbia University Press, New York:249–279.

Musumeci M, Aragona P, Bellin M, Maugeri F, Rania L, Bucolo C, Musumeci S. 2009. Acidic mammalian chitinase in dry eye conditions. Cornea 28:667–672.

Muzzarelli RAA, Boudrant J, Meyer D, Manno N, DeMarchis M, Paoletti MG. 2012. Current views on fungal chitin/chitosan, human chitinases, food preservation, glucans, pectins and inulin: A tribute to Henri Braconnot, precursor of the carbohydrate polymers science, on the chitin bicentennial. Carbohydrate Polymers 87:995–1012.

Niemitz C. 1984. Biology of Tarsiers. Stuttgart: Gustav Fischer.

Oftedal OT, Whiten A, Southgate DAT, Van Soest PJ. 1991. The nutritional consequences of foraging in primates: the relationship of nutrient intakes to nutrient requirements. Philos Trans R Soc Lond, B, Biol Sci 334:161–170.

Ohno M, Kimura M, Miyazaki H, Okawa K, Onuki R, Nemoto C, Tabata E, Wakita S, Kashimura A, Sakaguchi M, Sugahara Y, Nukina N, Bauer PO, Oyama F. 2016. Acidic mammalian chitinase is a proteases-resistant glycosidase in mouse digestive system. Sci Rep 6:37756.

Paoletti MG, Norberto L, Damini R, Musumeci S. 2007. Human gastric juice contains chitinase that can degrade chitin. Ann Nutr Metab 51:244–251.

Pozzi L, Hodgson JA, Burrell AS, Sterner KN, Raaum RL, Disotell TR. 2014. Primate phylogenetic relationships and divergence dates inferred from complete mitochondrial genomes. Molecular Phylogenetics and Evolution 75:165–183.

Raubenheimer D, Rothman JM. 2013. Nutritional ecology of entomophagy in humans and other primates. Annu Rev Entomol 58:141–160.

Reese TA, Liang H-E, Tager AM, Luster AD, Van Rooijen N, Voehringer D, Locksley RM. 2007. Chitin induces accumulation in tissue of innate immune cells associated with allergy. Nature 447:92–96.

Rothman JM, Raubenheimer D, Bryer MAH, Takahashi M, Gilbert CC. 2014. Nutritional contributions of insects to primate diets: Implications for primate evolution. Journal of Human Evolution 71:59–69.

Simunek J, Bartonova H. 2005. Effect of Dietary Chitin and Chitosan on Cholesterolemia of Rats. Acta Veterinaria Brno 74:491–499.

Smith RJ, Jungers WL. 1997. Body mass in comparative primatology. Journal of Human Evolution 32:523–559.

Souto A, Bione CBC, Bastos M, Bezerra BM, Fragaszy D, Schiel N. 2011. Critically endangered blonde capuchins fish for termites and use new techniques to accomplish the task. Biology Letters 7:532–535.

Strait SG, Vincent J. 1998. Primate faunivores: physical properties of prey items. Int J Primatol 19:867–878.

Strobel S, Roswag A, Becker NI, Trenczek TE, Encarnação JA. 2013. Insectivorous Bats Digest Chitin in the Stomach Using Acidic Mammalian Chitinase. PLoS ONE 8:e72770.

Sussman RW. 1991. Primate origins and the evolution of angiosperms. Am J Primatol 23:209–223.

Synstad B, Gaseidnes S, van Aalten DMF, Vriend G, Nielsen JE, Eijsink VGH. 2004. Mutational and computational analysis of the role of conserved residues in the active site of a family 18 chitinase. European Journal of Biochemistry 271:253–262.

Terborgh J. 1983. Five New World Primates: A Study in Comparative Ecology. Princeton University Press.

Tjoelker LW, Gosting L, Frey S, Hunter CL, Trong HL, Steiner B, Brammer H, Gray PW. 2000. Structural and functional definition of the human chitinase chitin-binding domain. J Biol Chem 275:514–520.

Tosi AJ, Buzzard PJ, Morales JC, Melnick DJ. 2002. Y-chromosome Data and Tribal Affiliations of Allenopithecus and Miopithecus. Int J Primatol 23:1287–1299.

Tosi AJ, Detwiler KM, Disotell TR. 2005. X-chromosomal window into the evolutionary history of the guenons (Primates: Cercopithecini). Molecular Phylogenetics and Evolution 36:58–66.

van Schaik CP. 2003. Orangutan Cultures and the Evolution of Material Culture. Science 299:102–105.

Vannella KM, Ramalingam TR, Hart KM, de Queiroz Prado R, Sciurba J, Barron L, Borthwick LA, Smith AD, Mentink-Kane M, White S, Thompson RW, Cheever AW, Bock K, Moore I, Fitz LJ, Urban JF, Wynn TA. 2016. Acidic chitinase primes the protective immune response to gastrointestinal nematodes. Nat Immunol 21:73–9.

Wertheim JO, Murrell B, Smith MD, Kosakovsky Pond SL, Scheffler K. 2015. RELAX: detecting relaxed selection in a phylogenetic framework. Mol Biol Evol 32:820–832.

Whitaker JO, Dannelly HK, Prentice DA. 2004. Chitinase in Insectivorous Bats. J Mammal 85:15–18.

Yang Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. Mol Biol Evol 24:1586–1591.

Zhu Z, Zheng T, Homer RJ, Kim Y-K, Chen NY, Cohn L, Hamid Q, Elias JA. 2004. Acidic mammalian chitinase in asthmatic Th2 inflammation and IL-13 pathway activation. Science 304:1678–1682.

Table 3.1. Species included in this study with annual average insect consumption and average body weight of adult females in grams.[1]

Species	Common name	Average insect consumption (%)	Average body weight (adult female, in g)
Aotus nancymaae	Nancy Ma's night/owl monkey	<15	780
Callicebus moloch	Red-bellied titi monkey	12	956
Callithrix jacchus	Common marmoset	7.2	381
Saguinus fuscicollis	Saddleback tamarin	28.3	358
Cebus capucinus	White-faced capuchin	31.4	2540
Saimiri boliviensis	Black-capped squirrel monkey	no wild data	711
Saimiri sciureus	Common squirrel monkey	53.4	662
Sapajus apella	Tufted capuchin	32.6	2520
Allenopithecus nigroviridis	Allen's swamp monkey	9	3180
Allochrocebus lhoesti	L'Hoest's monkey	8.8	3450
Cercocebus atys	Sooty mangabey	26	6200
Cercopithecus mitis	Blue monkey	17.5	4250
Chlorocebus aethiops	Grivet	15.4	2980
Chlorocebus sabaeus	Green monkey	15.4	3300
Erythrocebus patas	Patas monkey	23.5	6500
Macaca fascicularis	Long-tailed macaque	4.1	3590
Macaca mulatta	Rhesus macaque	0	5370
Macaca nemestrina	Pig-tailed macaque	12.2	6500
Mandrillus leucophaeus	Drill	no wild data	12500

Miopithecus talapoin	Talapoin	35	1120
Papio anubis	Olive baboon	2	13300
Colobus angolensis	Tanzanian black&white colobus	0	6935
Colobus guereza	Guereza	0	8550
Nasalis larvatus	Proboscis monkey	0	9820
Rhinopithecus bieti	Black snub-nosed monkey	0	9960
Rhinopithecus roxellana	Golden snub-nosed monkey	0	11600
Nomascus leucogenys	Northern white-cheeked gibbon	4	7320
Gorilla gorilla gorilla	Western lowland gorilla	7.7	71500
Pan paniscus	Bonobo	2	33200
Pan troglodytes	Common chimpanzee	6.4	41600
Pongo abelii	Sumatran orangutan	11.1	35600
Tarsius syrichta	Philippine tarsier	90	117
Microcebus murinus	Gray mouse lemur	8	63
Otolemur garnettii	Northern greater galago	50	734
Tupaia chinensis	Northern treeshrew	<50	200

[1]Primate body weight data from (Smith and Jungers, 1997); only data from wild primates were used. *Tupaia chinensis* data from PanTHERIA (Jones et al., 2009). Detailed dietary data and references can be found in Supplementary Table 3

Table 3.2. CodeML results[1]

Gene	Model	ln(L)	Parameter estimates	Test	LR	p-value	positively selected sites
mCHIA	M0	-5377.772	$\omega = 0.251$, $k = 3.835$				
	M1a	-5338.362	$k = 3.907$; $\omega 0 = 0.114$ (81.12%); $\omega 1 = 1.00$ (18.87%)				
	M2a	-5338.362	$k = 3.907$; $\omega 0 = 0.114$ (81.12%); $\omega 1 = 1.00$ (13.69%); $\omega 2 = 1.00$ (5.19%)	M1-M2	0		
	M3	-5338.001	$k = 3.861$; $\omega 0 = 0.1$ (77.85%); $\omega 1 = 0.354$ (0.0002%); $\omega 2 = 0.866$ (22.14%)	M0-M3	79.542	< 0.001	
	M7	-5339.595	$k = 3.830$, $\alpha = 0.348$; $\beta = 0.965$				
	M8	-5338.170	$k = 3.872$, $\alpha = 1.341$; $\beta = 8.405$; $p0 = 0.841$; $\omega s = 1.00$ (15.89%)	M7-M8	2.851	0.240	

hCHIA	M0	-6138.723	ω = 0.320, k = 4.132				
	M1a	-6079.211	k = 4.177; $\omega 0$ = 0.140 (77.26%); $\omega 1$ = 1.00 (22.74%)				
	M2a	-6074.863	k = 4.283; $\omega 0$ = 0.146 (77.36%); $\omega 1$ = 1.00 (21.68%); $\omega 2$ = 4.327 (0.96%)	M1-M2	8.696	0.013	
	M3	-6074.675	k = 4.219; $\omega 0$ = 0.00 (31.07%); $\omega 1$ = 0.345 (60.07%) ; $\omega 2$ = 1.639 (8.86%)	M0-M3	128.097	< 0.001	
	M7	-6081.171	k = 4.160, α = 0.369; β = 0.758				
	M8	-6073.789	k = 4.235, α = 0.457; β = 1.025; p0 = 0.985; ω = 3.503 (1.52%)	M7-M8	14.763	0.001	36P(0.953*), 62Q(0.971*), 164R(0.915), 280H(0.941)

[1]According to our hypothesis-testing framework, no *mCHIA* sites were found to be under positive selection; the vast majority (>75%) appear to be under purifying selection. Regarding *hCHIA*, models assuming positive selection outperformed null models, with 1.0-1.5% of sites found to be under positive selection.

Table 3.3. RELAX results.[2]

Gene	Test Branches	Reference Branches	Model	log L	AIC_c	LR	p-value
mCHIA	Pseudogenes	Functional genes	Null	-5892.80	11956.49		
			Alternative	-5876.86	11926.63	31.88	<0.001
			Partitioned Exploratory	-5866.18	11913.36	53.24	<0.001
	Below average insect consumption	Above average insect consumption	Null	-4778.17	9691.10		
			Alternative	-4769.42	9675.61	17.50	<0.001
			Partitioned Exploratory	-4764.26	9673.38	27.82	<0.001
	Above Kay's threshold	Below Kay's threshold	Null	-4778.17	9691.10		
			Alternative	-4767.75	9672.29	20.84	<0.001
			Partitioned Exploratory	-4765.17	9675.22	26.00	<0.001

hCHIA	Pseudogenes	Functional genes	Null	-6824.57	13820.01		
			Alternative	-6787.44	13747.78	74.26	<0.001
			Partitioned Exploratory	-6787.46	13755.90	74.22	<0.001
	Below average insect consumption	Above average insect consumption	Null	-4795.67	9685.95		
			Alternative	-4782.37	9661.39	26.60	<0.001
			Partitioned Exploratory	-4781.84	9668.44	27.66	<0.001
	Above Kay's threshold	Below Kay's threshold	Null	-4795.68	9685.98		
			Alternative	-4791.84	9680.33	7.68	0.006
			Partitioned Exploratory	-4788.79	9682.33	13.78	0.017

[2]Model-fits of null, alternative, and partitioned exploratory models inferred by the program RELAX (Wertheim et al., 2015). For each *CHIA* paralog, selective patterns were compared between species with (1) pseudogenes and functional genes, (2) below and above average insect consumption, and (3) body weights above and below Kay's threshold (500 g).

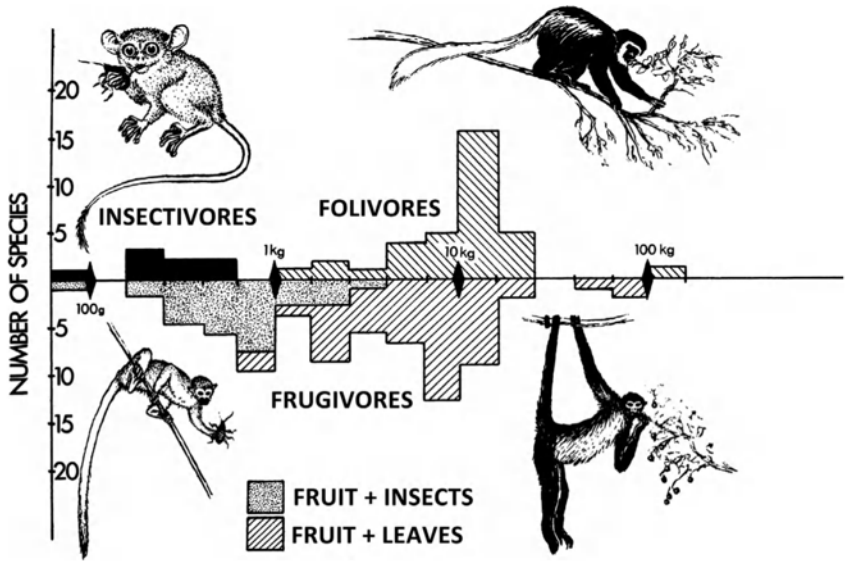


Figure 3.1. Correlation between primate diets and body size. Based on Kay (1984), modified in Fleagle (2013), reproduced with permission.

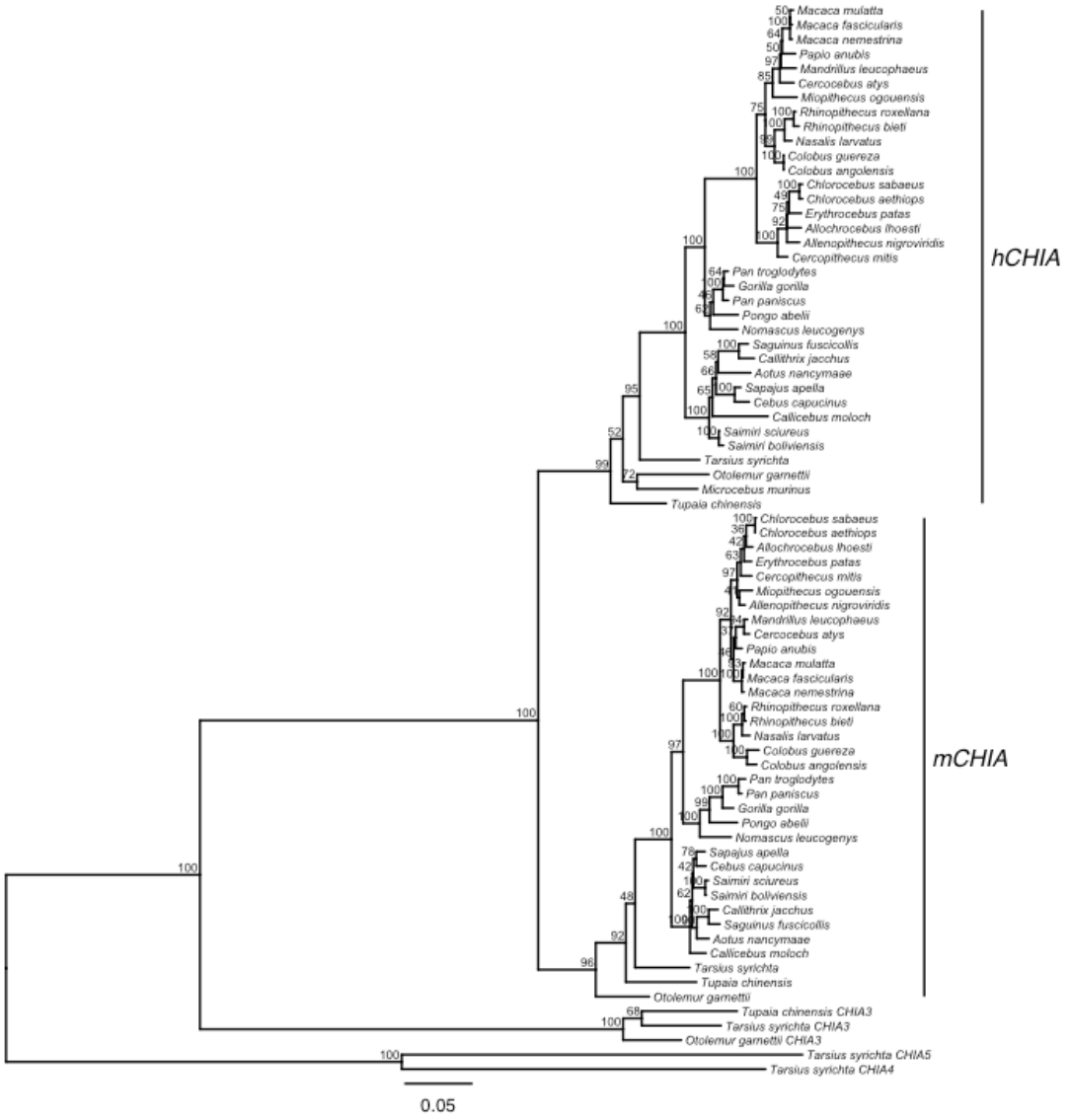


Figure 3.2. Evolutionary relationships of the *CHIA* genes in primates. The tree was inferred with PhyML (Guindon et al., 2010) using the HKY85 nucleotide substitution model. Node labels indicate percent bootstrap support (1000 replicates) and branches are scaled by number of substitutions per site. Tree is rooted at the midpoint. Because this tree is based only on the *CHIA* loci, not all relationships are resolved in a way that is consistent with primate phylogeny. Notably, the placements of *Tupaia chinensis mCHIA*

and *CHIA3*, and Old World monkey *hCHIA* do not reflect the likely *organismal* relationships among some taxa.

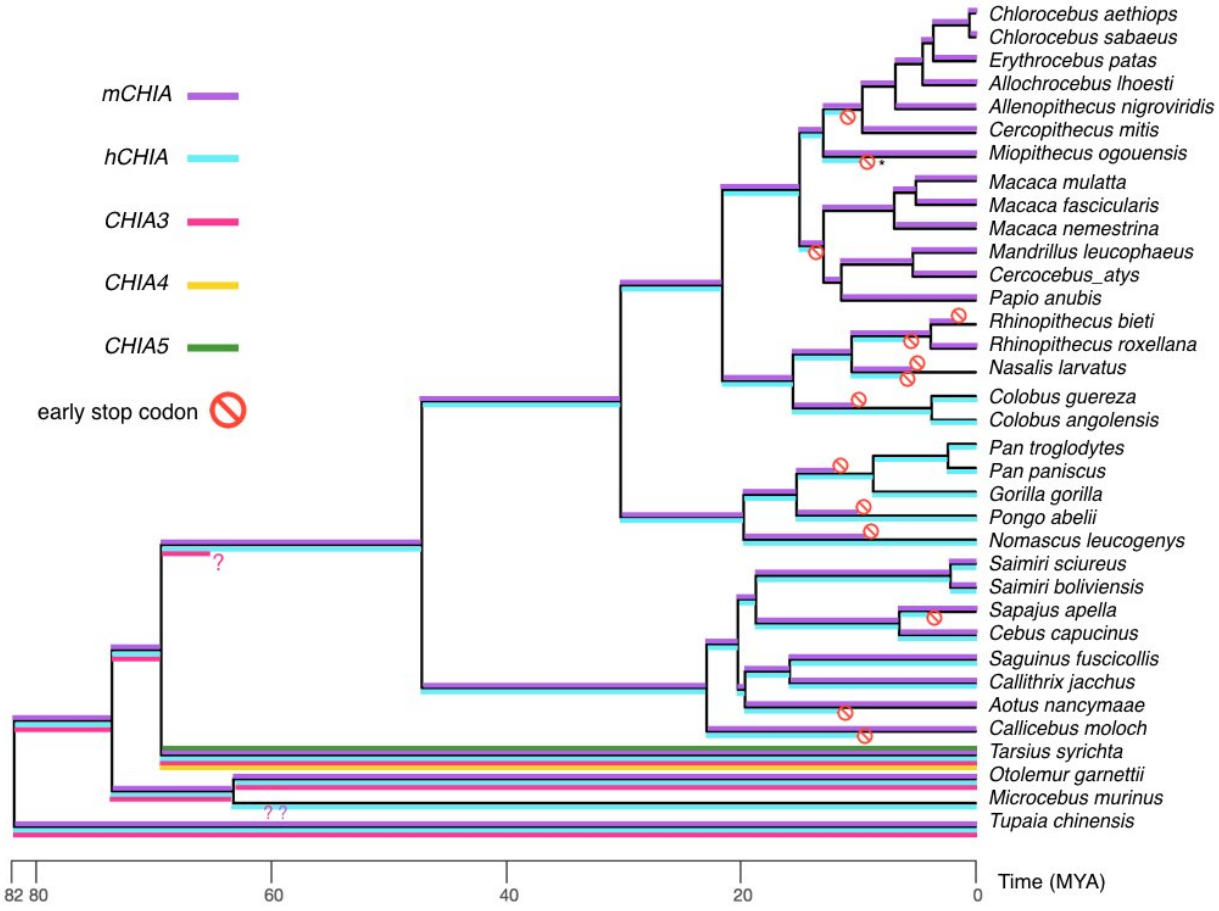


Figure 3.3. Evolutionary relationships as inferred from *CHIA* sequences, including timing of *CHIA* pseudogenization events. *The pseudogenizing mutation found in the *Miopithecus ogouensis hCHIA* sequence is the same as the one found in the Papionini (*Macaca* spp., *Papio* spp, etc.), but is not found in the other Cercopithecini. It is unclear what accounts for this unexpected pattern, but possible explanations include ancestral polymorphism or ancient hybridization.

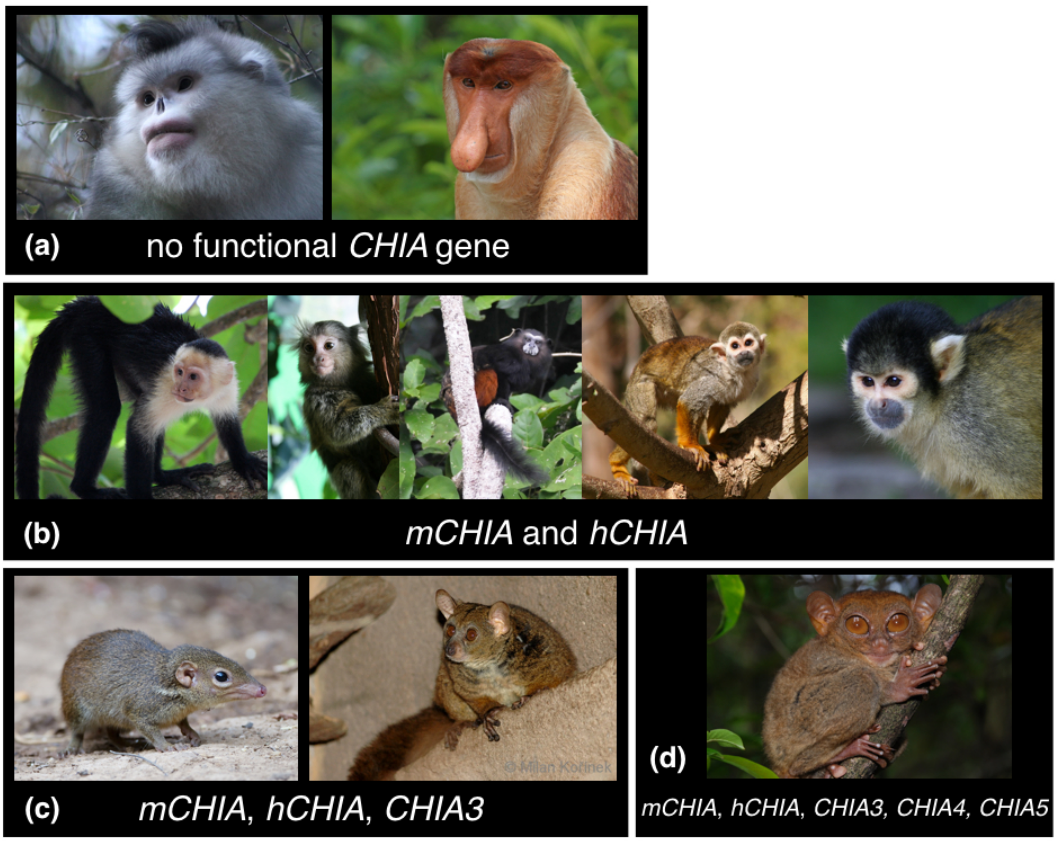


Fig. 3.4. Most primate species have one full-length *CHIA* sequence, with some exceptions. (a) In the colobine monkeys *Rhinopithecus bieti* and *Nasalis larvatus* both *mCHIA* and *hCHIA* sequences have a premature stop codon. (b) In some New World monkeys (in order: *Cebus capucinus*, *Callithrix jacchus*, *Saguinus fuscicollis*, *Saimiri sciureus*, *S. boliviensis*) both *mCHIA* and *hCHIA* are full-length. (c) The treeshrew (*Tupaia chinensis*) and the Northern greater galago (*Otolemur garnettii*), two insectivores, have three full-length *CHIA* sequences, while (d) the tarsier (*Tarsius syrichta*), the most insectivorous of all primates, has a total of five *CHIA* genes. Photos by (in order) Israel Didham (with pers. permission), Charles J. Sharp, Steven G. Johnson, Manfred Werner, Marie de Carne, Dave Pape, Julie Langford, JJ Harrison, Milan

Kořínek (with pers. permission), and Pierre Fidenci; reproduced with permission via Wikimedia Commons unless otherwise noted.

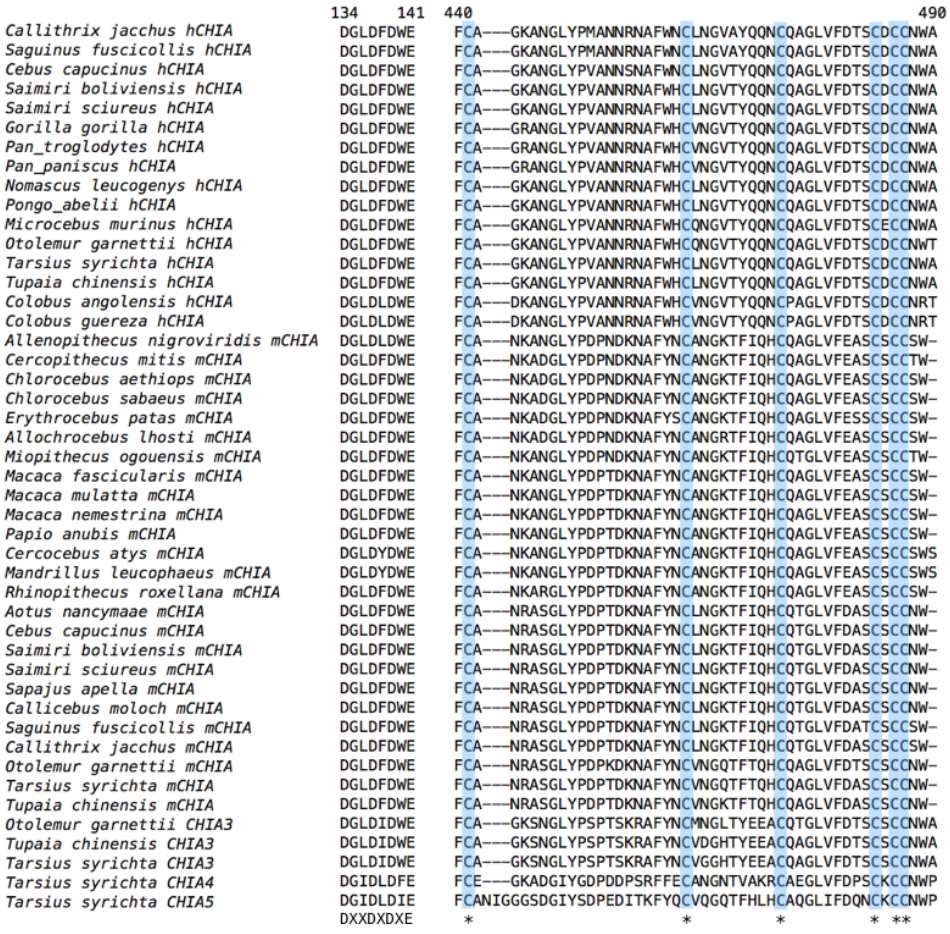


Figure 3.5. Partial amino acid sequence alignment of functional *CHIA* genes. The conserved motif for the chitinase catalytic site (DXXDXDXE) from residue 134-141 and the chitin-binding domain from residue 440-490 are shown. The cysteines in the chitin-binding domain are highlighted.

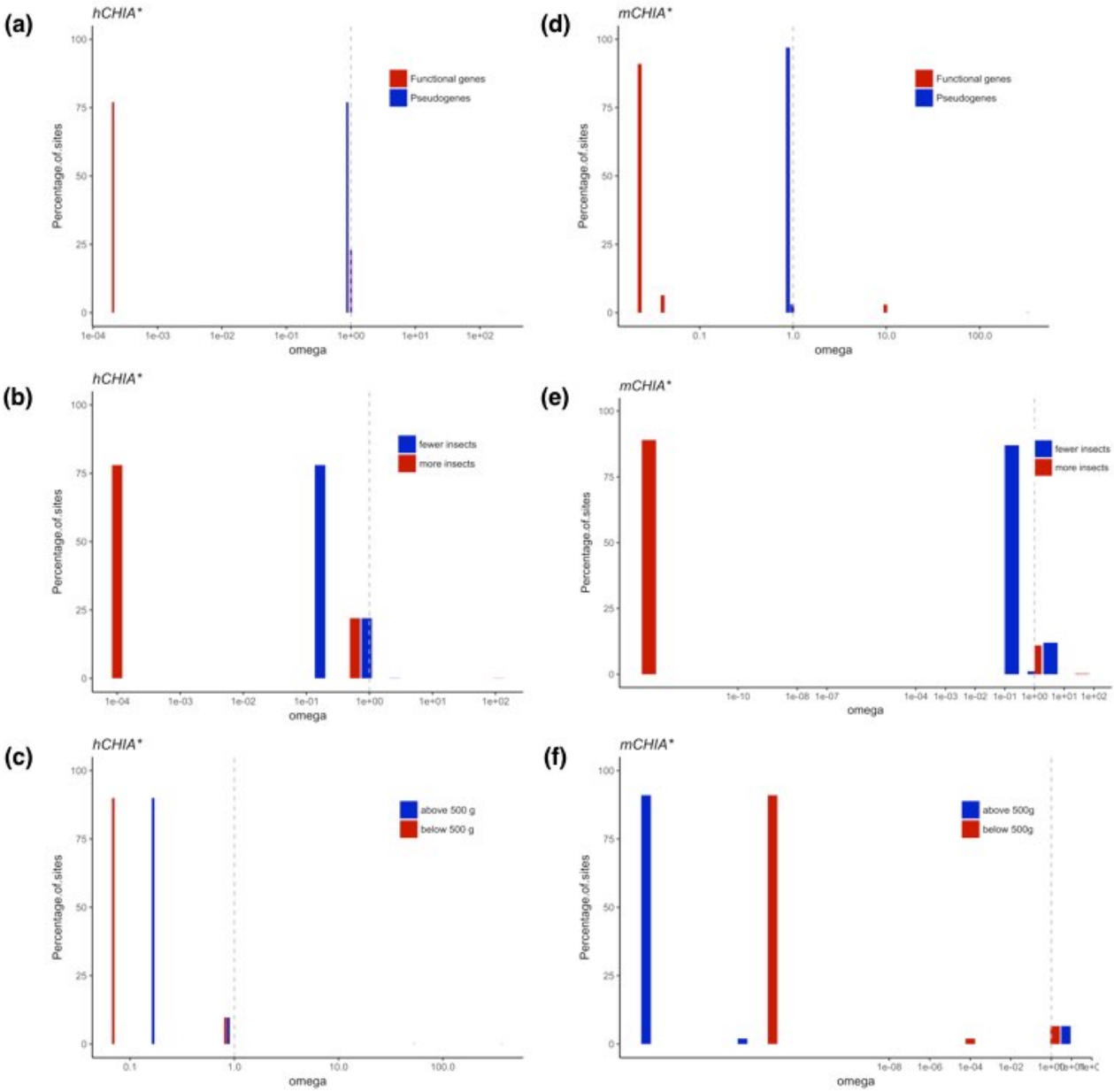
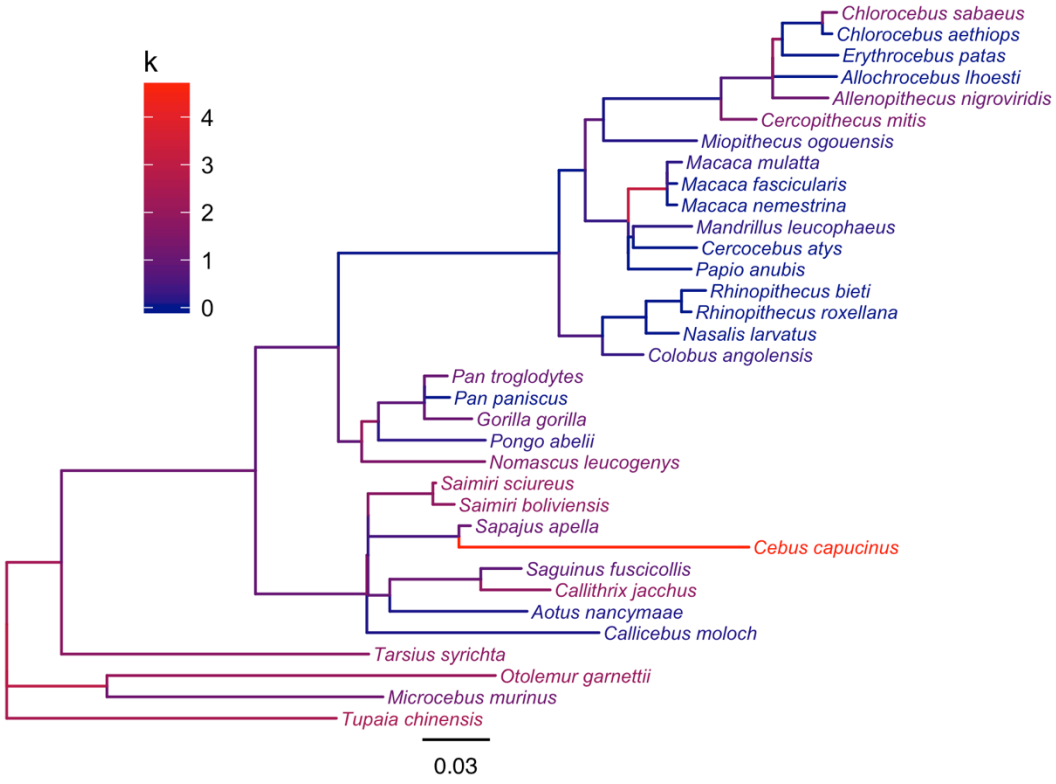


Figure 3.6. Patterns of natural selection across *mCHIA* and *hCHIA*. The best fitting model (as determined by AIC_c) for each RELAX analysis is shown. Three ω parameters and the percentage of sites they represent are plotted for test (blue) and reference (red) branches. The vertical gray and dashed line at $\omega = 1$ indicates neutral evolution. Asterisks indicate significant differences ($p < 0.05$) between test and reference branches.

(a)


k
4
3
2
1
0
Chlorocebus sabaeus
Chlorocebus aethiops
Erythrocebus patas
Allochrocebus lhoesti
Allenopithecus nigroviridis
Cercopithecus mitis
Miopithecus ogouensis
Macaca mulatta
Macaca fascicularis
Macaca nemestrina
Mandrillus leucophaeus
Cercocebus atys
Papio anubis
Rhinopithecus bieti
Rhinopithecus roxellana
Nasalis larvatus
Colobus angolensis
Pan troglodytes
Pan paniscus
Gorilla gorilla
Pongo abelii
Nomascus leucogenys
Saimiri sciureus
Saimiri boliviensis
Sapajus apella
Cebus capucinus
Saguinus fuscicollis
Callithrix jacchus
Aotus nancymaae
Callicebus moloch
Tarsius syrichta
Otolemur garnettii
Microcebus murinus
Tupaia chinensis
0.03

(b)

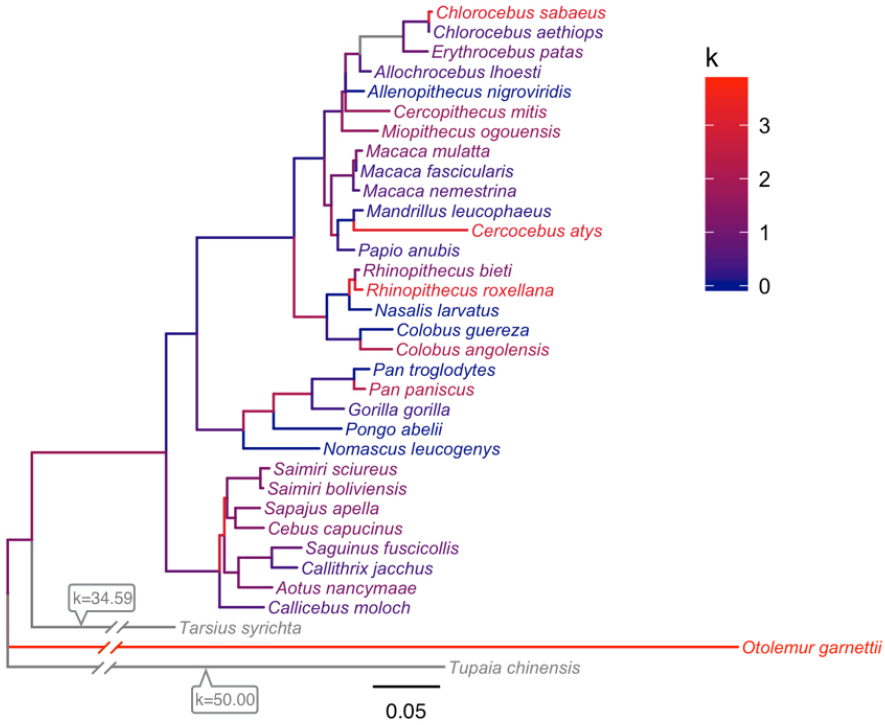


Figure 3.7. Branch specific relaxation parameters inferred for (a) *hCHIA* and (b) *mCHIA* genes under the General Descriptive model in RELAX (Wertheim et al., 2015). Branches are colored based on the selection intensity parameter *k*. A higher *k* value (red) indicates intensified selection, while a lower *k* value (blue) indicates relaxed selection. Scale bars indicate number of substitutions per site. Very long branches were truncated (indicated by breaks) to avoid obscuring the variation present in the remaining branches. Branches with extremely high *k* values are highlighted in gray.

Chapter 4. Investigating copy number variation in the acidic mammalian chitinase gene family (*CHIA*) in New World and Old World monkeys

Abstract

Copy number variation may be the most common form of structural genetic variation in the genome and numerous studies have shown that gene copy number variation correlates with phenotypic variation, where higher copy numbers correspond to increased expression of the protein and vice versa. Examples include some digestive enzyme genes, where variation in copy numbers and protein expression may be related to dietary differences. Increasing the expression of a digestive enzyme through higher gene copy numbers may thus be a useful mechanism for altering an organism's digestive capabilities. This study investigates copy number variation in genes coding for acidic mammalian chitinase, a chitinolytic digestive enzyme that may be used for the digestion of insect exoskeletons, in non-human primates with varying levels of insect consumption. Copy number variation was assessed with the QuantStudio 3D digital PCR platform, in a comparative sample of Old World and New World primate species. Contrary to predictions, no evidence of copy number variation was found and all species tested had two gene copies per diploid genome. These findings suggest that if acidic mammalian chitinase expression varies according to insect consumption in primates, it may be up- or downregulated through another mechanism.

Introduction

Copy number variation (CNV) is a form of structural variation in the genome, in which a genomic region is present in higher or lower numbers compared to a reference genome or another species' genome (Stankiewicz and Lupski, 2010; Clop et al., 2012). Research into CNVs has emerged over the course of the past 10 years, as cheaper next-generation sequencing methods and other techniques have allowed for the identification of these structural variants on a broader scale (Clop et al., 2012). While it is still unclear how common CNVs are, some research suggests that it might be the most common form of structural genetic variation, being found in up to 13% of the human genome (Stankiewicz and Lupski, 2010). Some work shows an accelerated rate of CNV in the great apes (Cheng et al., 2005; Marques-Bonet et al., 2009), possibly affecting 16% of hominid genome (Sudmant et al., 2013). Little is known about CNV in the genomes of primates outside of the great apes, and the phenotypic or functional effects of many CNVs are not yet understood (Zarrei et al., 2015).

Numerous studies have provided evidence that gene copy number variation correlates with phenotypic variation (i.e. higher copy numbers correspond to increased expression of the protein and vice versa) (Hollox et al., 2003; Linzmeier and Ganz, 2006; Perry et al., 2006; 2007) and some CNVs have been associated with diseases and developmental disorders in humans (Pinto et al., 2010; Jacquemont et al., 2011; Malhotra and Sebat, 2012; Cantsilieris and White, 2013). The β-defensin gene family is involved in immune function, host defense, and pigmentation, and is subject to CNV across mammals (Hardwick et al., 2011; Bickhart et al., 2012; Leonard et al., 2012; Wang et al., 2013). Interestingly, research on the β-defensin-2 gene in humans and macaques (*Macaca*

mulatta) suggests that this CNV evolved convergently, rather than from a common ancestor (Ottolini et al., 2014). What the precise adaptive benefits of CNV at this locus are remains unknown, but the authors speculate that it may confer benefits for group-living animals or help during the colonization of novel environments (Ottolini et al., 2014). Work on the genomes of domestic animals has described some of the phenotypic effects of CNV. For example, different CNVs cause various coat color phenotypes in sheep, goats, and pigs (Giuffra et al., 1999; Fontanesi et al., 2009; 2011) and are related to the timing of feather growth (Elferink et al., 2008) and comb size (Wright et al., 2009) in chickens.

Copy number variation has also been implicated in digestive enzyme expression, as discussed below. Previous research indicates that there is variation in the types and amounts of digestive enzymes that are secreted among primates (and other mammals), and that the driving selective pressure on this variation is diet (Zhang et al., 2002; Axelsson et al., 2013; Behringer et al., 2013). Phenotypic variation relating to digestive enzymes can strongly impact an organism's digestive abilities (Ingram et al., 2009; Mandel and Breslin, 2012), which in humans is illustrated most famously through the example of lactase persistence (Tishkoff et al., 2007). While in some human populations a large percentage of individuals (up to 95%) continue to produce the digestive enzyme lactase throughout adulthood, in many other populations lactase production ceases after childhood, causing the inability to digest lactose, the main carbohydrate in milk (Itan et al., 2009). The ability to digest lactose in adulthood has evolved independently several times (Enattah et al., 2007; Tishkoff et al., 2007; Enattah et al., 2008), and a shared trait of the populations that exhibit lactase persistence is that they have traditionally been

pastoralists with a long history of dairying, suggesting that this is a dietary adaptation to facilitate the consumption of fresh milk (Tishkoff et al., 2007; Itan et al., 2009; 2010; Schlebusch et al., 2013).

While variation in the lactase persistence phenotype is due to single nucleotide polymorphisms (SNPs) within the promoter region (Tishkoff et al., 2007; Ranciaro et al., 2014), CNV has been found in other digestive enzyme genes. The enzyme pepsin A hydrolyzes the peptide bonds of proteins, creating smaller chains of amino acids that can then be further digested by the enzymes trypsin, chymotrypsin, and elastase (Stevens and Hume, 1995). Pepsin A has been shown to be highly polymorphic at both the protein level and the genetic level in various primates. While humans have a cluster of three genes that are known to be present in variable copy numbers (Taggart et al., 1985), the other great apes may exhibit even greater variation than humans. Narita and colleagues (2000) purified numerous forms of the pepsin A precursor, pepsinogen A, from the gastric mucosa of a gibbon (*Hylobates lar*), an orangutan (*Pongo pygmaeus*), a gorilla (*Gorilla gorilla*), and a chimpanzee (*Pan troglodytes*). The number of pepsinogen A types found ranged from seven in the gorilla, to eight in the gibbon, thirteen in the chimpanzee, and fourteen in the orangutan (Narita et al., 2000). The genetic basis of this variation has not been resolved yet, but one study suggests that in orangutans pepsin A is encoded by two different genes, that may have three and five copies, respectively (Narita et al., 2010).

Human populations with high-starch diets produce higher levels of salivary amylase, an enzyme that digests starch, than populations that do not commonly include starch in their diets (Perry et al., 2007). This increased secretion correlates with a larger

number of copies of the salivary amylase gene, *AMY1* (Perry et al., 2007; Mandel et al., 2010). Notably, a relationship between amylase secretion and gene copy number has also been identified in dogs. The secretion of pancreatic amylase is variable and positively correlates with high copy numbers of the *AMY2* gene in dogs, but not in wolves, which only have the standard two copies (Axelsson et al., 2013). It is possible that this change allowed dogs to digest potentially starch-rich food scraps provided by humans (Axelsson et al., 2013). Further evidence that increased salivary amylase may be an adaptation for a high-starch diet comes from a more recent study linking *AMY1* gene copy numbers to a change in the sensory perception of starch – individuals with higher copy numbers and higher salivary amylase levels perceived oral starch viscosity to decrease more rapidly during mastication than individuals with low copy numbers (Mandel et al., 2010).

This work suggests that increasing the expression of a digestive enzyme through higher copy numbers may be a useful mechanism for altering an organism's digestive capabilities, and provides an incentive for exploring CNV in other digestive enzyme genes, such as the acidic mammalian chitinase genes (*CHIA*). The *CHIA* genes, *hCHIA* and *mCHIA*, encode the chitinolytic enzyme acidic mammalian chitinase (AMCase) in primates (Janiak et al., in review, Chapter 3). The enzyme AMCase is produced in the stomach (Krykbaev et al., 2010; Strobel et al., 2013; Ohno et al., 2016) and previous work on mice and insectivorous bats has shown that it is used for the digestion of chitinous insect exoskeletons (Strobel et al., 2013; Ohno et al., 2016). In primates, most species retain at least one functional *CHIA* gene, while some of the more insectivorous species have between two and five functional *CHIA* paralogs. The only species in which all *CHIA* paralogs have been pseudogenized are some of the folivorous colobine monkeys

(*Rhinopithecus bieti*, *Nasalis larvatus*), which are not known to feed on insects (Janiak et al., in review, Chapter 3). However, there is also significant variation in insect consumption across species that share the same *CHIA* genotype, such as the cercopithecines, in which only *mCHIA* remains functional. Likewise, platyrrhines, in some of which both *mCHIA* and *hCHIA* remain functional, vary in their average insect intake (Table 1). For example, insects reportedly make up only 7.20% of the diet of the common marmoset (*Callithrix jacchus*), while it is 53.38% for the common squirrel monkey (*Saimiri sciureus*) (Table 1). Another platyrrhine, the tufted capuchin monkey (*Sapajus apella*) only retains one functional *CHIA* gene, despite an average insect consumption of over 30% (Table 1). It is, therefore, possible that in primate species with higher insect consumption, the production of AMCase is increased via higher copy numbers of the *mCHIA* or *hCHIA* genes, similar to what has been found for salivary amylase and the *AMY1* gene (Perry et al., 2007).

Hypotheses and Predictions

This study investigates the potential of CNV at the *CHIA* locus in non-human primates, hypothesizing that *CHIA* copy number correlates with level of insectivory. I predict that primates with relatively more insectivorous diets (*Erythrocebus patas*, *Miopithecus talapoin*, *Macaca nigra*, *Saguinus fuscicollis*, *Saimiri sciureus*, and *Sapajus apella*) have higher copy numbers of the *mCHIA* gene and the *hCHIA* gene (only in species with a functional *hCHIA* sequence) than primates that are relatively less insectivorous (*Callicebus moloch*, *Callithrix jacchus*, *Macaca mulatta*, *Macaca nemestrina*).

Methods

Samples and sample preparation

High-quality DNA samples from seven non-human primate species were obtained from Coriell Biorepositories: *Callithrix jacchus*, *Callicebus moloch*, *Saguinus fuscicollis*, *Saimiri sciureus*, *Macaca mulatta*, *Macaca nemestrina*, and *Macaca nigra*. Dr. George Perry (Penn State University) provided extracted DNA from *Sapajus apella* and Dr. Todd Disotell (New York University) provided extracted DNA from *Erythrocebus patas* and *Miopithecus talapoin*. All DNA samples were quantified using the Qubit dsDNA BR assay kit (Invitrogen) and a Qubit 2.0 fluorometer (Invitrogen). Samples were diluted with water to a final concentration of 10 ng/µl.

Assay design

In order to design primers for use across multiple species, I identified a conserved region around exons 9 and 10 of the *mCHIA* and *hCHIA* genes. To design specific primers, I aligned the *mCHIA* exon 9 and 10 sequences, including the intronic region, of 13 catarrhine primates (*Chlorocebus aethiops*, *C. sabaeus*, *Allenopithecus nigroviridis*, *Cercocebus atys*, *Colobus guereza*, *Allochrocebus lhoesti*, *Macaca fascicularis*, *M. nemestrina*, *M. mulatta*, *Miopithecus ogouensis*, *Cercopithecus mitis*, *Papio anubis*, *Rhinopithecus roxellana*) and, separately, of four platyrrhine primates (*Aotus nancymaae*, *Callithrix jacchus*, *Cebus capucinus*, *Saimiri boliviensis*). The corresponding region of the *hCHIA* gene was likewise aligned for *Callithrix jacchus*, *Saimiri sciureus*, and *Saguinus fuscicollis*. All sequences were either available via reference genomes on GenBank or had been generated de novo for another study (Janiak, Chaney & Tosi, in review). Any bases that were not identical across all sequences in the alignment were

masked in the consensus sequence by replacing the corresponding nucleotide with 'N.' The masked consensus sequences were entered in the Custom TaqMan Assay Design Tool (ThermoFisher Scientific) and submitted to the TaqMan design pipeline. The design algorithms are proprietary and assay sequences are not shared by ThermoFisher until the assay is purchased. I, therefore, had to rely on the ThermoFisher technical assistance staff to conduct primer BLASTs of the assays proposed by the Custom TaqMan Assay Design Tool against primate reference genomes to ensure that the assays were specific to the region of interest and would not amplify other, unrelated, loci. Primer sequences and assay IDs are listed in Table 2. Primer and probe locations along the gene sequence are shown in Figure 1.

The RNase P TaqMan copy number reference assay (ThermoFisher) was used as the reference assay in all experiments. This assay detects the ribonuclease P (RNase P) RNA component H1 gene, *RPPH1*; a gene with known copy number of two copies per diploid genome. Including a reference assay for a gene with known copy number is necessary to determine the number of copies of the gene of interest. The measures for the gene of interest are normalized to those of obtained for the reference gene to determine the number of gene copies present per diploid genome.

3D Digital PCR

Copy number variation was assessed using the QuantStudio 3D Digital PCR system, with QuantStudio 3D Digital PCR 20K chips and QuantStudio 3D Digital PCR Master Mix. Reactions were prepared in a total volume of 34.8 µl containing 17.4 µl Master Mix, 1.74 µl custom assay, 1.74 µl RNase P assay, 1.92 µl water, and 12 µl template (10 ng/µl). This super mix was incubated at room temperature for 15 minutes

prior to being loaded onto the chips. All reactions were run in duplicate with 14.5 µl of the super mix loaded onto each chip. Chips were loaded and sealed according to the manufacturer's protocol and run under the following conditions: 96°C for 10 minutes, followed by 39 cycles of 60°C for 2 minutes, 98°C for 30 seconds, and 60°C for 2 minutes. Chips remained at 10°C until reading. The chips were read on the QuantStudio3D Digital PCR Instrument and analyzed with QuantStudio 3D Analysis Suite Cloud Software. The number of gene copies per diploid genome (CN) was calculated as follows:

$$CN = 2\left(\frac{Copies\ per\ \mu l\ CHIA}{Copies\ per\ \mu l\ RNASE\ P}\right)$$

The measure of copies/microliter for the custom assay, *mCHIA* or *hCHIA* in this case, was divided by the copies/microliter measured by the reference assay, RNase P, and multiplied by 2 to arrive at the total number of copies per diploid genome.

Results

The *mCHIA* and *hCHIA* TaqMan assays successfully amplified the loci of interest in all species included in the study. Absolute numbers amplified by the custom and reference assays ranged from 224.21-1961.60 copies/microliter (Table 3). Measurements for the numbers of copies per microliter and per reaction were very variable between species, but remained stable within species across the different assays (Fig. 2, Fig. 3). There was no difference in the ratio of copies/microliter detected by the custom assay versus copies/microliters detected by the RNase P reference assay for any of the samples (Fig. 2, Fig. 3). For all species in the study, the total number of copies calculated for *mCHIA* and *hCHIA* was closest to two copies per diploid genome (Fig. 4).

Discussion

This study investigated copy number variation in the *CHIA* gene family in a comparative sample of non-human primate species with varying levels of insect consumption. I predicted that primates with relatively more insectivorous diets would have higher copy numbers of the *CHIA* genes, *mCHIA* and *hCHIA*. Contrary to this prediction, the results do not show evidence of CNV in the primate *mCHIA* and *hCHIA* genes and suggest that all primates, regardless of insect consumption, have only two gene copies per diploid genome (Fig. 4).

Given the variation in insect consumption observed across primate species (Raubenheimer and Rothman, 2013), it would be expected to see this variation reflected in the expression of primate digestive enzymes. The evolution of the AMCase gene family, *CHIA*, does broadly correlate with patterns of insectivory across primates (Janiak et al., in review, Chapter 3). While some folivorous primates with no active insect consumption do not retain any functional *CHIA* paralogs and most primates only retain one, some of the more insectivorous primates, such as some platyrrhines, the bushbaby (*Otolemur garnettii*), and the tarsier (*Tarsius syrichta*) have two, three, or five functional paralogs, respectively. However, across the large number of primates that share the genotype of one functional *CHIA* paralog, such as the Old World monkeys, there is still notable variation in insect consumption (Janiak et al., in review, Table 1). Even though previous work on other enzymes suggests that higher copy numbers may be a useful mechanism for increasing the expression of a digestive enzyme (Perry et al., 2007;

Mandel et al., 2010; Mandel and Breslin, 2012; Axelsson et al., 2013), my results show that this is not the case for AMCase.

One possible explanation for this negative result is that an increase in AMCase expression may not be necessary for some of the species in my sample. According to the Jarman-Bell principle, an animal's nutritional requirements negatively correlate with body size (Bell, 1971; Jarman, 1974). Larger-bodied animals, while needing an absolutely larger amount of energy, have a relatively lower nutrient and energy requirement compared to smaller-bodied animals, meaning that they can afford to subsist on poorer quality, but abundant, foods. Smaller-bodied animals, on the other hand, require less energy in absolute terms, but need to focus on high-quality foods to meet their relatively higher nutrient demands (Bell, 1971; Jarman, 1974). Hence, the ability to collect and consume a large amount of food is the main challenge for large-bodied animals, whereas for smaller-bodied animals it is the efficient and quick extraction of nutrients from their foods (Gaulin, 1979).

Insects are a high-quality food (Raubenheimer et al., 2014; Rothman et al., 2014), but with the exception of social insects, they are not abundantly distributed across the environment, and their chitinous exoskeletons may pose a digestive challenge (Strait and Vincent, 1998; Rothman et al., 2014). Hence, smaller primates are expected to rely heavily on insects, as long as they can effectively cope with their exoskeletons, while insectivory in larger primates should focus on social insects, like ants and termites (Gaulin, 1979; Kay, 1984; Isbell, 1998; Fleagle, 2013). The finding that all cercopithecines, regardless of insect consumption, have only one functional *CHIA* paralog, *mCHIA*, and no CNV may, therefore, be consistent with primatological theory

(Gaulin, 1979). Patas monkeys (*Erythrocebus patas*) and crested black macaques (*Macaca nigra*) are among the Old World monkeys with the highest average insect consumption, but adult female body weight averages 6500 g and 5470 g, respectively (Smith and Jungers, 1997). This is well above the 500 g threshold associated with insectivorous primates (Gingerich, 1980; Kay, 1984; Fleagle, 2013) and maximally efficient digestion of their insect prey is likely to be less important for these cercopithecines than for a smaller primate (Gaulin, 1979). Even the talapoin (*Miopithecus talapoin*), which is the smallest Old World monkey (Fleagle, 2013) at 1120 g (Smith and Jungers, 1997), is above this threshold and larger than many New World monkeys. The amount of AMCase expressed from a single *CHIA* paralog may thus be sufficient, even for more insectivorous Old World monkeys.

New World monkeys tend to be both more insectivorous and smaller-bodied than Old World monkeys (Gaulin, 1979; Fleagle, 2013). In line with what is predicted by the Jarman-Bell principle and Kay's threshold (Bell, 1971; Jarman, 1974; Gingerich, 1980; Kay, 1984), previous work showed that some species retain two functional *CHIA* paralogs, theoretically doubling their ability to digest insect exoskeletons with AMCase (Janiak et al., in review, Chapter 3). Generally, those platyrrhine species that do not retain the second *CHIA* paralog also tend to be less insectivorous (Janiak et al, in review, Chapter 3). One exception to this is the tufted capuchin monkey (*Sapajus apella*), which retains only one functional *CHIA* gene, *mCHIA*, despite a fairly insectivorous diet (Table 1). Results of the CNV analyses show that this is not made up for by additional copies of the *mCHIA* gene (Fig. 2, Fig. 4). Possibly, this is due to the tufted capuchins comparatively large body size. Adult females weigh on average 2520 g, much more than

similarly insectivorous platyrrhine species, such as the common squirrel monkey (*Saimiri sciureus*, 662 g) or the saddle-back tamarin (*Saguinus fuscicollis*, 358 g) (Smith and Jungers, 1997), which retain two functional *CHIA* paralogs (Janiak et al., in review, Chapter 3). It should be noted, however, that the closely-related white-faced capuchin (*Cebus capucinus*) retains two functional *CHIA* paralogs (Janiak et al., in review, Chapter 3), despite having similar levels of insect consumption (Chapman and Fedigan, 1990; Rose, 1994; McCabe and Fedigan, 2007; Mallott, 2016) and being almost identical in size to *Sapajus apella* (Smith and Jungers, 1997). What causes this difference in genotype is currently unclear.

Measurements of copies per microliter varied widely across the different samples included in the study, despite efforts to standardize the amount of DNA used in each reaction (Table 3, Fig. 2, Fig. 3). This is likely due to errors during the quantification or dilution process. Stock DNA samples were mixed carefully, by slowly pipetting up and down, rather than vortexing, to avoid shearing the DNA. However, insufficient mixing of DNA samples prior to pipetting may have led to erroneous Qubit quantification readings, affecting downstream dilutions, or inconsistent amounts of DNA may have been added to the reaction dilutions. That said, the variation in DNA quantity observed across samples does not call into question the overall results, since measures of copies/µl were consistent across duplicate reactions using the same sample (supplementary Table 4).

While the results of this study suggest that the *CHIA* genes in primates are not subject to CNV, it is plausible that AMCase enzyme expression is up- or downregulated through another mechanism. Future studies should investigate differences in *CHIA* promoter regions, mRNA expression in primate stomach tissues, and functional effects of

polymorphisms in *CHIA* paralogs. Previous work showed that there are small inter-species differences in the *mCHIA* and *hCHIA* sequences, from single nucleotide polymorphisms to multi-basepair insertions and deletions. What effects, if any, these changes have on the resulting protein and its functionality is unclear at this point, but remains an avenue for future research.

Conclusion

I investigated copy number variation in the *CHIA* paralogs, *mCHIA* and *hCHIA*, in Old World and New World primate species with varying levels of insect consumption. Contrary to the hypothesis, there was no evidence that the *CHIA* genes had higher copy numbers in more insectivorous primate species. All primate species included in this study had two gene copies per diploid genome, suggesting that there is no CNV in primate *CHIA* genes.

References

Alonso C, Langguth A. 1989. Ecologia e comportamento de *Callithrix jacchus* (Primates: Callitrichidae) numa ilha de floresta Atlântica. Revista Nordestina de Biologia 6:105–137.

Axelsson E, Ratnakumar A, Arendt M-L, Maqbool K, Webster MT, Perloski M, Liberg O, Arnemo JM, Hedhammar Å, Lindblad-Toh K. 2013. The genomic signature of dog domestication reveals adaptation to a starch-rich diet. Nature 495:360–364.

Behringer V, Borchers C, Deschner T, Möstl E, Selzer D, Hohmann G. 2013. Measurements of salivary alpha amylase and salivary cortisol in hominoid primates reveal within-species consistency and between-species differences. PLoS ONE 8:e60773.

Bell RHV. 1971. A Grazing Ecosystem in the Serengeti. Scientific American 225:86–93.

Bickhart DM, Hou Y, Schroeder SG, Alkan C, Cardone MF, Matukumalli LK, Song J, Schnabel RD, Ventura M, Taylor JF, Garcia JF, Van Tassell CP, Sonstegard TS, Eichler EE, Liu GE. 2012. Copy number variation of individual cattle genomes using next-generation sequencing. Genome Research 22:778–790.

Brown AD, Zunino GE. 1990. Dietary variability in *Cebus apella* in extreme habitats: Evidence for adaptability. Folia Primatol 54:187–195.

Caldecott JO. 1986. An ecological and behavioural study of the pig-tailed macaque. Contrib Primatol 21:1–259.

Cantsilieris S, White SJ. 2013. Correlating multiallelic copy number polymorphisms with disease susceptibility. Hum Mutat 34:1–13.

Chapman CA, Fedigan LM. 1990. Dietary differences between neighboring Cebus capucinus groups: local traditions, food availability or responses to food profitability. Folia Primatol 54:177–186.

Cheng Z, Ventura M, She X, Khaitovich P, Graves T, Osoegawa K, Church D, DeJong P, Wilson RK, Pääbo S, Rocchi M, Eichler EE. 2005. A genome-wide comparison of recent chimpanzee and human segmental duplications. Nature 437:88–93.

Clop A, Vidal O, Amills M. 2012. Copy number variation in the genomes of domestic animals. Anim Genet 43:503–517.

Crandlemire-Sacco J. 1988. An ecological comparison of two sympatric primates: *Saguinus fuscicollis* and *Callicebus moloch* of Amazonian Peru. Primates 29:465–475.

Digby LJ, Ferrari SF, Saltzman W. 2011. Callitrichines. In: Campbell CJ, Fuentes A, MacKinnon KC, Bearder SK, Stumpf RM, editors. Primates in Perspective. Oxford

University Press. p 91–107.

Elferink MG, Vallée AAA, Jungerius AP, Crooijmans RPMA, Groenen MAM. 2008. Partial duplication of the *PRLR* and *SPEF2* genes at the late feathering locus in chicken. BMC Genomics 9:391.

Enattah NS, Jensen TGK, Nielsen M, Lewinski R, Kuokkanen M, Rasinpera H, El-Shanti H, Seo JK, Alifrangis M, Khalil IF, Natah A, Ali A, Natah S, Comas D, Mehdi SQ, Groop L, Vestergaard EM, Imtiaz F, Rashed MS, Meyer B, Troelsen J, Peltonen L. 2008. Independent introduction of two lactase-persistence alleles into human populations reflects different history of adaptation to milk culture. Am J Hum Genet 82:57–72.

Enattah NS, Trudeau A, Pimenoff V, Maiuri L, Auricchio S, Greco L, Rossi M, Lentze M, Seo JK, Rahgozar S, Khalil I, Alifrangis M, Natah S, Groop L, Shaat N, Kozlov A, Verschubskaya G, Comas D, Bulayeva K, Mehdi SQ, Terwilliger JD, Sahi T, Savilahti E, Perola M, Sajantila A, Järvelä I, Peltonen L. 2007. Evidence of still-ongoing convergence evolution of the lactase persistence T-13910 alleles in humans. Am J Hum Genet 81:615–625.

Fleagle JG. 2013. Primate adaptation and evolution. 3rd ed. Academic Press.

Fontanesi L, Beretti F, Riggio V, González EG, Dall'Olio S, Davoli R, Russo V, Portolano B. 2009. Copy Number Variation and Missense Mutations of the Agouti Signaling Protein (*ASIP*) Gene in Goat Breeds with Different Coat Colors. Cytogenet Genome Res 126:333–347.

Fontanesi L, Dall'Olio S, Beretti F, Portolano B, Russo V. 2011. Coat colours in the Massese sheep breed are associated with mutations in the agouti signalling protein (*ASIP*) and melanocortin 1 receptor (*MC1R*) genes. Animal 5:8–17.

Garber PA. 1988. Diet, Foraging Patterns, and Resource Defense in a Mixed Species Troop of *Saguinus mystax* and *Saguinus fuscicollis* in Amazonian Peru. Behaviour 105:18–34.

Gaulin SJ. 1979. A Jarman/Bell model of primate feeding niches. Human Ecology 7:1–20.

Gautier-Hion A. 1988. Diet and Dietary Habits of Forest Guenons. In: A primate radiation: Evolutionary biology of the African guenons. Cambridge University Press. p 257–283.

Gingerich PD. 1980. Dental and cranial adaptations in Eocene Adapidae. Z Morphol Anthropol 71:135–142.

Giuffra E, Evans G, Törnsten A, Wales R, Day A, Looft H, Plastow G, Andersson L. 1999. The Belt mutation in pigs is an allele at the Dominant white (I/KIT) locus. Mamm Genome 10:1132–1136.

Goldstein SJ, Richard AF. 1989. Ecology of rhesus macaques (*Macaca mulatta*) in northwest Pakistan. Int J Primatol 10:531–567.

Hardwick RJ, Machado LR, Zuccherato LW, Antolinos S, Xue Y, Shawa N, Gilman RH, Cabrera L, Berg DE, Tyler-Smith C, Kelly P, Tarazona-Santos E, Hollox EJ. 2011. A worldwide analysis of beta-defensin copy number variation suggests recent selection of a high-expressing DEFB103 gene copy in East Asia. Hum Mutat 32:743–750.

Hollox EJ, Armour JAL, Barber JCK. 2003. Extensive normal copy number variation of a beta-defensin antimicrobial-gene cluster. Am J Hum Genet 73:591–600.

Ingram CJE, Mulcare CA, Itan Y, Thomas MG, Swallow DM. 2009. Lactose digestion and the evolutionary genetics of lactase persistence. Hum Genet 124:579–591.

Isbell LA. 1998. Diet for a small primate: Insectivory and gummivory in the (large) patas monkey (*Erythrocebus patas pyrrhonotus*). Am J Primatol 45:381–398.

Itan Y, Jones BL, Ingram CJ, Swallow DM, Thomas MG. 2010. A worldwide correlation of lactase persistence phenotype and genotypes. BMC Evol Biol 10:36.

Itan Y, Powell A, Beaumont MA, Burger J, Thomas MG. 2009. The origins of lactase persistence in Europe. PLoS Comput Biol 5:e1000491.

Jacquemont S, Reymond A, Zufferey F, Harewood L, Walters RG, Kutalik Z, Martinet D, Shen Y, Valsesia A, Beckmann ND, Thorleifsson G, Belfiore M, Bouquillon S, Campion D, de Leeuw N, de Vries BBA, Esko T, Fernandez BA, Fernández-Aranda F, Fernández-Real JM, Gratacòs M, Guilmatre A, Hoyer J, Jarvelin M-R, Kooy RF, Kurg A, Le Caignec C, Männik K, Platt OS, Sanlaville D, Van Haelst MM, Villatoro Gomez S, Walha F, Wu B-L, Yu Y, Aboura A, Addor M-C, Alembik Y, Antonarakis SE, Arveiler B, Barth M, Bednarek N, Béna F, Bergmann S, Beri M, Bernardini L, Blaumeiser B, Bonneau D, Bottani A, Boute O, Brunner HG, Cailley D, Callier P, Chiesa J, Chrast J, Coin L, Coutton C, Cuisset J-M, Cuvellier J-C, David A, de Freminville B, Delobel B, Delrue M-A, Demeer B, Descamps D, Didelot G, Dieterich K, Disciglio V, Doco-Fenzy M, Drunat S, Duban-Bedu B, Dubourg C, El-Sayed Moustafa JS, Elliott P, Faas BHW, Faivre L, Faudet A, Fellmann F, Ferrarini A, Fisher R, Flori E, Forer L, Gaillard D, Gerard M, Gieger C, Gimelli S, Gimelli G, Grabe HJ, Guichet A, Guillin O, Hartikainen A-L, Heron D, Hippolyte L, Holder M, Homuth G, Isidor B, Jaillard S, Jaros Z, et al. 2011. Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. Nature 478:97–102.

Janiak, MC, Chaney, ME, Tosi, AJ. In review. Evolution of acidic mammalian chitinase genes (*CHIA*) is related to body mass and insectivory in primates.

Jarman PJ. 1974. The Social Organisation of Antelope in Relation To Their Ecology. Behaviour 48:215–267.

Kay RF. 1984. On the use of anatomical features to infer foraging behavior in extinct primates. In: Cant J, Rodman PS, editors. Adaptations for Foraging in Nonhuman

Primates. p 21–53.

Knogge C, Heymann EW. 2003. Seed dispersal by sympatric tamarins, *Saguinus mystax* and *Saguinus fuscicollis*: Diversity and characteristics of plant species. Folia Primatol 74:33–47.

Krykbaev R, Fitz LJ, Reddy PS, Winkler A, Xuan D, Yang X, Fleming M, Wolf SF. 2010. Evolutionary and biochemical differences between human and monkey acidic mammalian chitinases. Gene 452:63–71.

Lawrence JM. 2007. Understanding the pair bond in brown titi monkeys (*Callicebus brunneus*): male and female reproductive interests.

Leonard BC, Marks SL, Outerbridge CA, Affolter VK, Kananurak A, Young A, Moore PF, Bannasch DL, Bevins CL. 2012. Activity, Expression and Genetic Variation of Canine ß-Defensin 103: A Multifunctional Antimicrobial Peptide in the Skin of Domestic Dogs. Journal of Innate Immunity 4:248–259.

Lima EM, Ferrari SF. 2003. Diet of a Free-Ranging Group of Squirrel Monkeys (*Saimiri sciureus*) in Eastern Brazilian Amazonia. Folia Primatol 74:150–158.

Linzmeier RM, Ganz T. 2006. Copy number polymorphisms are not a common feature of innate immune genes. Genomics 88:122–126.

Malhotra D, Sebat J. 2012. CNVs: harbingers of a rare variant revolution in psychiatric genetics. Cell 148:1223–1241.

Mallott EK. 2016. Social, ecological, and developmental influences on fruit and invertebrate foraging strategies and gut microbial communities in white-faced capuchins (*Cebus capucinus*). :1–168.

Mandel AL, Breslin PAS. 2012. High endogenous salivary amylase activity is associated with improved glycemic homeostasis following starch ingestion in adults. J Nutr 142:853–858.

Mandel AL, Peyrot des Gachons C, Plank KL, Alarcon S, Breslin PAS. 2010. Individual differences in *AMY1* gene copy number, salivary α-amylase levels, and the perception of oral starch. PLoS ONE 5:e13352.

Marques-Bonet T, Kidd JM, Ventura M, Graves TA, Cheng Z, Hillier LW, Jiang Z, Baker C, Malfavon-Borja R, Fulton LA, Alkan C, Aksay G, Girirajan S, Siswara P, Chen L, Cardone MF, Navarro A, Mardis ER, Wilson RK, Eichler EE. 2009. A burst of segmental duplications in the genome of the African great ape ancestor. Nature 457:877–881.

McCabe GM, Fedigan LM. 2007. Effects of Reproductive Status on Energy Intake, Ingestion Rates, and Dietary Composition of Female *Cebus capucinus* at Santa Rosa, Costa Rica. Int J Primatol 28:837–851.

Nakagawa N. 1989. Activity budget and diet of patas monkeys in Kala Maloue National Park, Cameroon: A preliminary report. Primates 30:27–34.

Narita Y, Oda S, Takenaka O, Kageyama T. 2000. Multiplicities and some enzymatic characteristics of ape pepsinogens and pepsins. J Med Primatol 29:402–410.

Narita Y, Oda S-I, Takenaka O, Kageyama T. 2010. Lineage-specific duplication and loss of pepsinogen genes in hominoid evolution. J Mol Evol 70:313–324.

O'Brien TG, Kinnaird MF. 1997. Behavior, diet, and movements of the Sulawesi crested black macaque (*Macaca nigra*). Int J Primatol 18:321–351.

Ohno M, Kimura M, Miyazaki H, Okawa K, Onuki R, Nemoto C, Tabata E, Wakita S, Kashimura A, Sakaguchi M, Sugahara Y, Nukina N, Bauer PO, Oyama F. 2016. Acidic mammalian chitinase is a proteases-resistant glycosidase in mouse digestive system. Sci Rep 6:37756.

Ottolini B, Hornsby MJ, Abujaber R, MacArthur JAL, Badge RM, Schwarzacher T, Albertson DG, Bevins CL, Solnick JV, Hollox EJ. 2014. Evidence of convergent evolution in humans and macaques supports an adaptive role for copy number variation of the β-defensin-2 gene. Genome Biology and Evolution 6:3025–3038.

Perry GH, Dominy NJ, Claw KG, Lee AS, Fiegler H, Redon R, Werner J, Villanea FA, Mountain JL, Misra R, Carter NP, Lee C, Stone AC. 2007. Diet and the evolution of human amylase gene copy number variation. Nat Genet 39:1256–1260.

Perry GH, Tchinda J, McGrath SD, Zhang J, Picker SR, Cáceres AM, Iafrate AJ, Tyler-Smith C, Scherer SW, Eichler EE, Stone AC, Lee C. 2006. Hotspots for copy number variation in chimpanzees and humans. Proc Natl Acad Sci USA 103:8006–8011.

Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, Regan R, Conroy J, Magalhaes TR, Correia C, Abrahams BS, Almeida J, Bacchelli E, Bader GD, Bailey AJ, Baird G, Battaglia A, Berney T, Bolshakova N, Bölte S, Bolton PF, Bourgeron T, Brennan S, Brian J, Bryson SE, Carson AR, Casallo G, Casey J, Chung BHY, Cochrane L, Corsello C, Crawford EL, Crossett A, Cytrynbaum C, Dawson G, de Jonge M, Delorme R, Drmic I, Duketis E, Duque F, Estes A, Farrar P, Fernandez BA, Folstein SE, Fombonne E, Freitag CM, Gilbert J, Gillberg C, Glessner JT, Goldberg J, Green A, Green J, Guter SJ, Hakonarson H, Heron EA, Hill M, Holt R, Howe JL, Hughes G, Hus V, Igliozzi R, Kim C, Klauck SM, Kolevzon A, Korvatska O, Kustanovich V, Lajonchere CM, Lamb JA, Laskawiec M, Leboyer M, Le Couteur A, Leventhal BL, Lionel AC, Liu X-Q, Lord C, Lotspeich L, Lund SC, Maestrini E, Mahoney W, Mantoulan C, Marshall CR, McConachie H, McDougle CJ, McGrath J, McMahon WM, Merikangas A, Migita O, Minshew NJ, Mirza GK, Munson J, Nelson SF, Noakes C, Noor A, Nygren G, Oliveira G, Papanikolaou K, Parr JR, Parrini B, Paton T, et al. 2010. Functional impact of global rare copy number variation in autism spectrum disorders. Nature 466:368–372.

Porter LM. 2001. Dietary differences among sympatric Callitrichinae in northern Bolivia:

Callimico goeldii, *Saguinus fuscicollis* and *S. labiatus*. Int J Primatol 22:961–992.

Ranciaro A, Campbell MC, Hirbo JB, Ko W-Y, Froment A, Anagnostou P, Kotze MJ, Ibrahim M, Nyambo T, Omar SA, Tishkoff SA. 2014. Genetic origins of lactase persistence and the spread of pastoralism in Africa. Am J Hum Genet 94:496–510.

Raubenheimer D, Rothman JM, Pontzer H, Simpson SJ. 2014. Macronutrient contributions of insects to the diets of hunter-gatherers: A geometric analysis. Journal of Human Evolution 71:70–76.

Raubenheimer D, Rothman JM. 2013. Nutritional ecology of entomophagy in humans and other primates. Annu Rev Entomol 58:141–160.

Rose LM. 1994. Sex differences in diet and foraging behavior in white-faced capuchins (*Cebus capucinus*). Int J Primatol 15:95–114.

Rothman JM, Raubenheimer D, Bryer MAH, Takahashi M, Gilbert CC. 2014. Nutritional contributions of insects to primate diets: Implications for primate evolution. Journal of Human Evolution 71:59–69.

Schlebusch CM, Sjödin P, Skoglund P, Jakobsson M. 2013. Stronger signal of recent selection for lactase persistence in Maasai than in Europeans. Eur J Hum Genet 21:550–553.

Smith RJ, Jungers WL. 1997. Body mass in comparative primatology. Journal of Human Evolution 32:523–559.

Stankiewicz P, Lupski JR. 2010. Structural variation in the human genome and its role in disease. Annu Rev Med 61:437–455.

Stevens CE, Hume ID. 1995. Comparative physiology of the vertebrate digestive system. 2nd ed. Cambridge: Cambridge University Press.

Stone AI. 2007. Responses of squirrel monkeys to seasonal changes in food availability in an eastern Amazonian forest. Am J Primatol 69:142–157.

Strait SG, Vincent J. 1998. Primate faunivores: physical properties of prey items. Int J Primatol 19:867–878.

Strobel S, Roswag A, Becker NI, Trenczek TE, Encarnação JA. 2013. Insectivorous Bats Digest Chitin in the Stomach Using Acidic Mammalian Chitinase. PLoS ONE 8:e72770.

Sudmant PH, Huddleston J, Catacchio CR, Malig M, Hillier LW, Baker C, Mohajeri K, Kondova I, Bontrop RE, Persengiev S, Antonacci F, Ventura M, Prado-Martinez J, Great Ape Genome Project, Marques-Bonet T, Eichler EE. 2013. Evolution and diversity of copy number variation in the great ape lineage. Genome Research 23:1373–1382.

Taggart RT, Mohandas TK, Shows TB, Bell GI. 1985. Variable numbers of pepsinogen genes are located in the centromeric region of human chromosome 11 and determine the high-frequency electrophoretic polymorphism. Proc Natl Acad Sci USA 82:6240–6244.

Tishkoff SA, Reed FA, Ranciaro A, Voight BF, Babbitt CC, Silverman JS, Powell K, Mortensen HM, Hirbo JB, Osman M, Ibrahim M, Omar SA, Lema G, Nyambo TB, Ghori J, Bumpstead S, Pritchard JK, Wray GA, Deloukas P. 2007. Convergent adaptation of human lactase persistence in Africa and Europe. Nat Genet 39:31–40.

Wang J, Wang H, Jiang J, Kang H, Feng X, Zhang Q, Liu J-F. 2013. Identification of genome-wide copy number variations among diverse pig breeds using SNP genotyping arrays. PLoS ONE 8:e68683.

Wright D, Boije H, Meadows JRS, Bed'hom B, Gourichon D, Vieaud A, Tixier-Boichard M, Rubin C-J, Imsland F, Hallböök F, Andersson L. 2009. Copy number variation in intron 1 of *SOX5* causes the pea-comb phenotype in chickens. PLoS Genet 5:e1000512.

Wright PC. 1985. The cost and benefits of nocturnality for *Aotus trivirgatus* (the night monkey). PhD thesis, City University of New York.

Zarrei M, MacDonald JR, Merico D, Scherer SW. 2015. A copy number variation map of the human genome. Nat Rev Genet 16:172–183.

Zhang J, Zhang Y-P, Rosenberg HF. 2002. Adaptive evolution of a duplicated pancreatic ribonuclease gene in a leaf-eating monkey. Nat Genet 30:411–415.

Table 4.1. Average annual insect consumption for species included in this study.

Species	Percent insects	Site	Reference	Average insect consumption
Callicebus moloch (brunneus)	17	Manu, Peru	(Wright, 1985)	12
	11	Tambopata, Peru	(Crandlemire-Sacco, 1988)	
	8	Los Amigos, Peru	(Lawrence, 2007)	
Callithrix jacchus	5.4	Joao Pessoa, Brazil	(Alonso and Langguth, 1989)	7.2
	9	Nisia Floresta, Brazil	(Digby et al., 2011)	
Erythrocebus patas	35	Segera, Kenya	(Isbell, 1998)	23.5
	12	Kala Maloue, Cameroon	(Nakagawa, 1989)	
Macaca mulatta	0	Murree Hills, Pakistan	(Goldstein and Richard, 1989)	0
Macaca nemestrina	12.2	Pasoh, Malaysia	(Caldecott, 1986)	12.2
Macaca nigra	32.1	Sulawesi, Indonesia	(O'Brien and Kinnaird, 1997)	32.1
Miopithecus ogouensis/talapoin	35	Mankokou, Gabon	(Gautier-Hion, 1988)	35

Saguinus fuscicollis	5.8	Quebrada Bl., Peru	(Knogge and Heymann, 2003)	28.3
	53.1	Rio Blanco, Peru	(Garber, 1988)	
	26	Pando, Bolivia	(Porter, 2001)	
Saimiri sciureus	45	Gunma, Brazil	(Lima and Ferrari, 2003)	53.38
	61.75	Ananim, Brazil	(Stone, 2007)	
Sapajus apella	40.3	El Rey, Brazil	(Brown and Zunino, 1990)	32.6
	24.9	Iguazu, Brazil	(Brown and Zunino, 1990)	

Table 4.2. Primer and probe sequences used in TaqMan assays.

Gene	**Taxon**	**Assay**	**Sequence**
hCHIA	Platyrrhines	hCHIA_NW_CD47VWA_F	ATGGTCTGGGCCATTGATCTG
		hCHIA_NW_CD47VWA_R	CCTTCTTCAGGGTGGAGATTAGG
		hCHIA_NW_CD47VWA_M	ATGACTTCACTGGCACTTTCT
mCHIA	Catarrhines	CHIA_OWM_CDGZE6W_F	GAGTGGCTTGGATATGATAACACCAA
		CHIA_OWM_CDGZE6W_R	GGGAACATGCTCACAGGCA
		CHIA_OWM_CDGZE6W_M	ACACAGTCTACCTTGATTTGGAAACT
	Platyrrhines	CHIA_NWM_CDPRJ2G_F	AATTATTACAGGCTGATTGGTTAAAGA
		CHIA_NWM_CDPRJ2G_R	AGGGAATTTTCCTTGGTTGCAGAA
		CHIA_NWM_CDPRJ2G_M	CCATTGACTTGGATGATTTCAC

Table 4.3. QuantStudio 3D results for custom assays, *mCHIA* and *hCHIA*, and RNase P reference assay.

Gene	Species	custom assay		reference assay	
		Copies/µl	CI Copies/µl	Copies/µl	CI Copies/µl
mCHIA	*Saimiri sciureus*	1042.00	1027.2 -- 1057	1061.80	1046.8 -- 1077
	Sapajus apella	1580.40	1559.7 -- 1601.4	1603.50	1582.5 -- 1624.8
	Saguinus fuscicollis	294.63	287.82 -- 301.61	285.11	278.42 -- 291.96
	Callicebus moloch	567.50	557.47 -- 577.71	579.61	569.45 -- 589.95
	Callithrix jacchus	1637.90	1616.5 -- 1659.5	1606.60	1585.7 -- 1627.9
	Erythrocebus patas	1961.60	1940.5 -- 1982.9	1864.50	1844.0 -- 1885.3
	Macaca nigra	580.39	570.18 -- 590.77	591.48	581.16 -- 601.99
	Macaca mulatta	231.92	225.95 -- 238.05	224.21	218.35 -- 230.23
	Macaca nemestrina	1376.60	1358.1 -- 1395.4	1402.30	1383.5 -- 1421.3
	Miopithecus talapoin	1547.50	1527.4 -- 1567.9	1567.00	1546.6 -- 1587.6
hCHIA	*Saimiri sciureus*	1116.40	1100.8 -- 1132.2	1119.00	1103.3 -- 1134.8
	Saguinus fuscicollis	313.86	306.88 -- 321.01	308.76	301.84 -- 315.84
	Callithrix jacchus	1345.00	1326.8 -- 1363.5	1348.20	1329.9 -- 1366.7

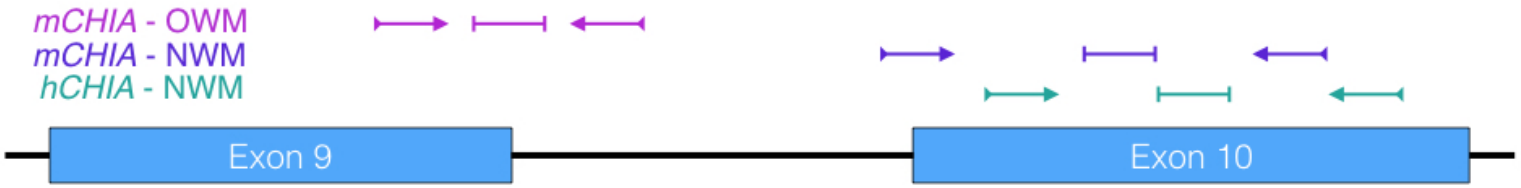


Figure 4.1. Locations of primers and probes along the *mCHIA* and *hCHIA* gene.

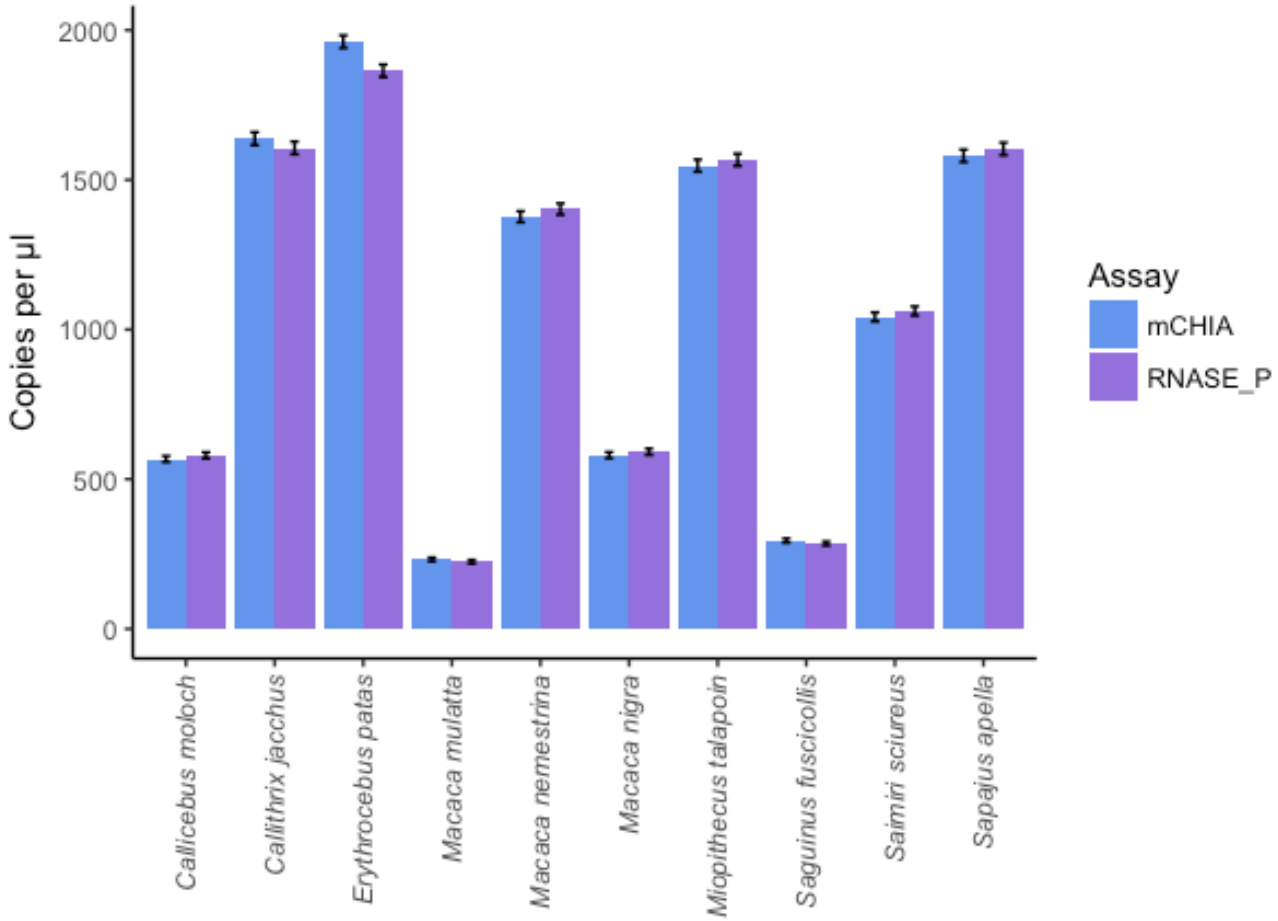


Figure 4.2. Copies per microliter measured for *mCHIA* and reference assays.

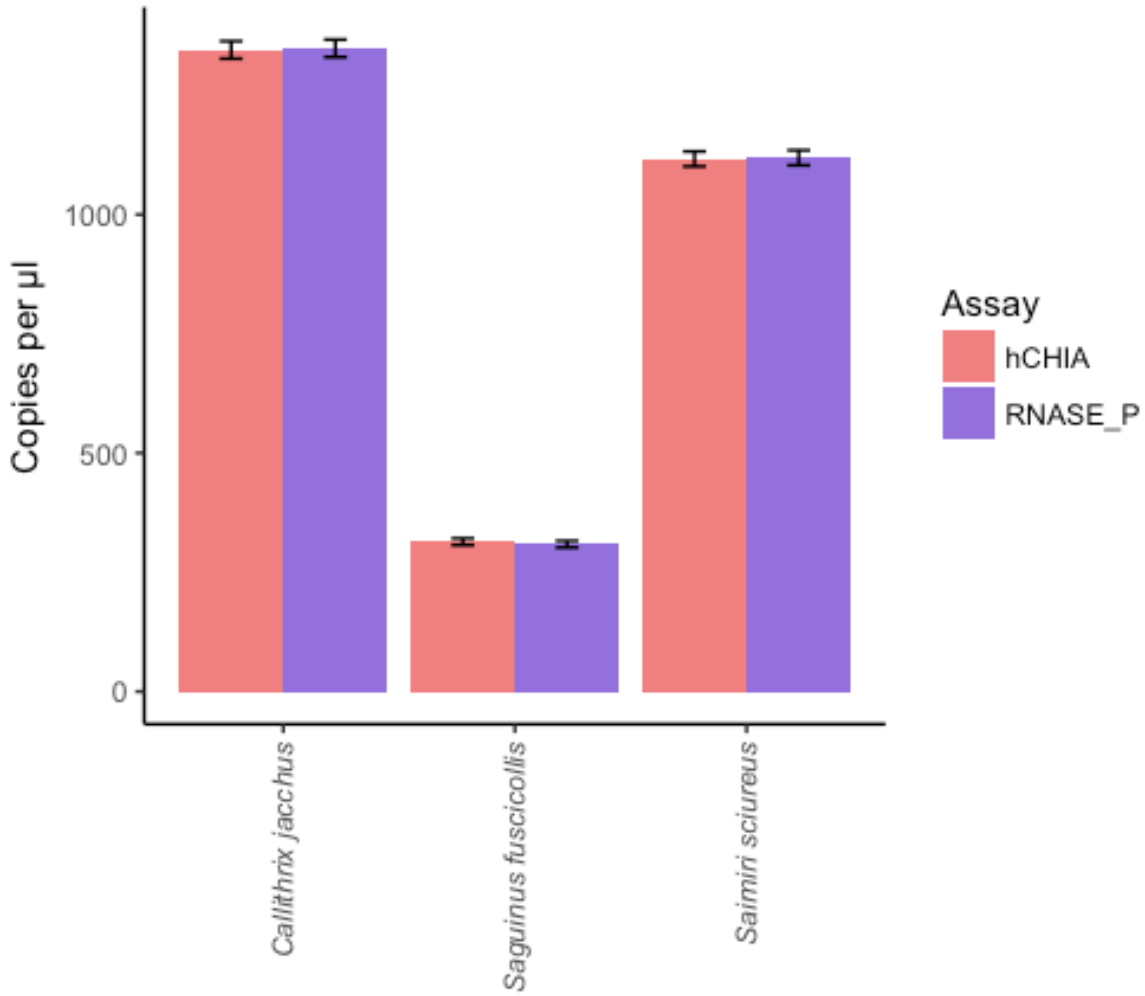


Figure 4.3. Copies per microliter measured for *hCHIA* and reference assays.

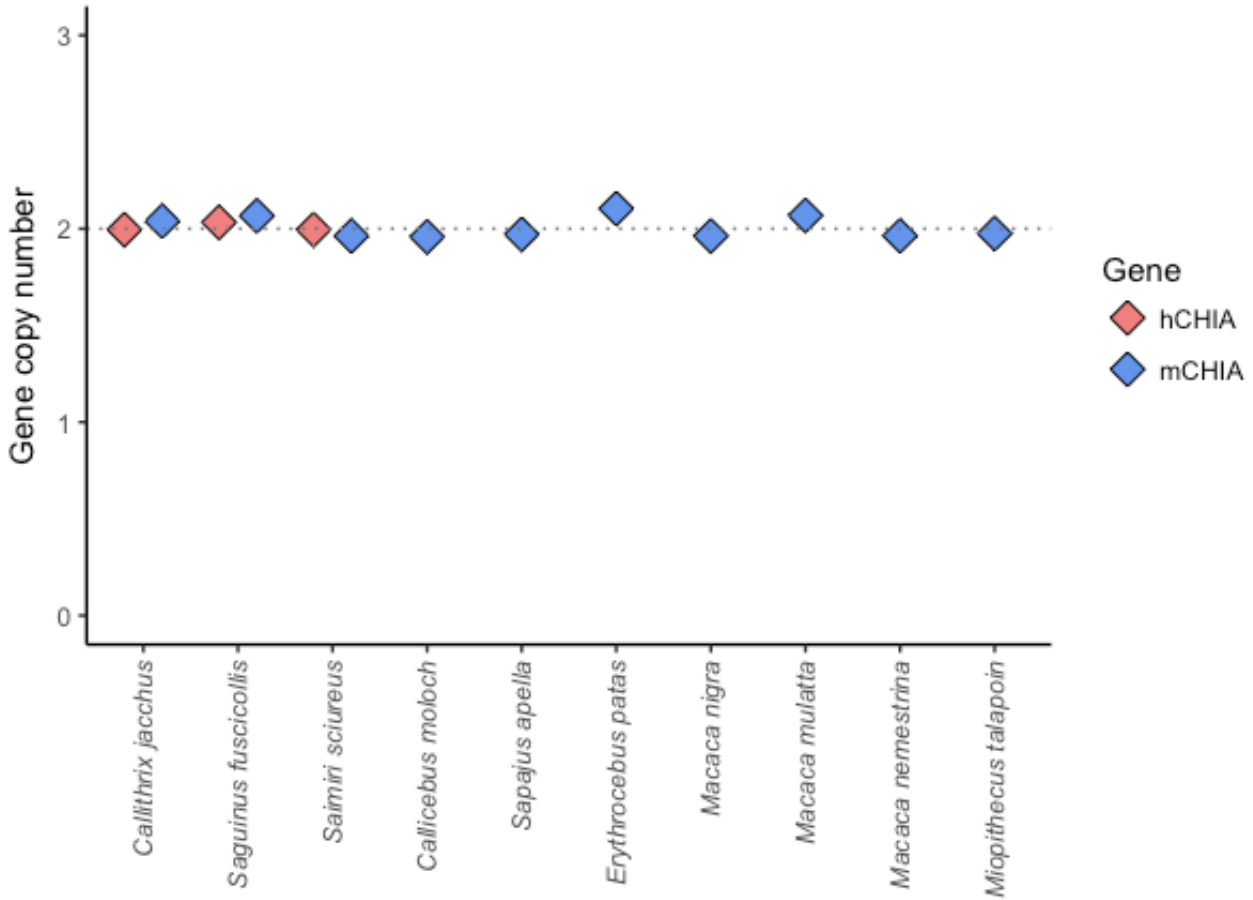


Figure 4.4. Relative copy number of *mCHIA* and *hCHIA* per diploid genome. Based on comparison with the RNase P reference assay, which detects *RPPH1*, a gene with a known number of copies per diploid genome.

Chapter 5. Conclusion

Summary and Conclusions

Digestive enzymes play a crucial role in an organism's digestive system and cannot be overlooked when discussing an animal's suite of dietary adaptations (Stevens and Hume, 1995; Zhang, 2006; Perry et al., 2007; Tishkoff et al., 2007; Axelsson et al., 2013; Ranciaro et al., 2014; Janiak, 2016). Some primate foods present digestive challenges, such as leaves, which may contain fiber, tannins and toxins (Simmen et al., 2012; Garber et al., 2015), or insects, whose exoskeletons are made up of chitin (Finke, 2007; Raubenheimer and Rothman, 2013). As primates diversified and adapted to available dietary niches (Rosenberger, 1992), changes in digestive enzymes allowed them to both tolerate new food resources and maximize the energy obtained from them (Stevens and Hume, 1995; Karasov and Douglas, 2013).

In chapter 2, I show that the pancreatic ribonuclease gene, *RNASE1*, has been duplicated in a non-colobine primate, the mantled howler monkey (*Alouatta palliata*), and that the duplicated genes (*RNASE1B*, *RNASE1C*) have biochemical properties and amino acid substitutions that are convergent with those found in foregut-fermenting primates and ruminants (Zhang et al., 2002; Zhang, 2003a; b; Schienman et al., 2006; Zhang, 2006; Yu et al., 2010). These proteins may therefore be used for an analogous function in howler monkeys, digesting the products of microbial fermentation in the caeco-colic region, a potentially substantial source of energy (Milton and McBee, 1983). Along with behavioral and morphological adaptations, these duplicated proteins may be crucial digestive enzyme adaptations allowing howler monkeys to survive on a folivorous diet during times of fruit scarcity.

Chapter 3 presents the first large, comparative dataset on the acidic mammalian chitinase gene family (*CHIA*) in primates. The notion that chitin was indigestible by the endogenous digestive enzymes of primates and other mammals has persisted (Cork and Kenagy, 1989; Oftedal et al., 1991; Simunek and Bartonova, 2005; Strobel et al., 2013; Ohno et al., 2016), despite several studies from the 1970s that suggested that primates have chitinolytic enzymes (Cornelius et al., 1976; Jeuniaux and Cornelius, 1978; Kay and Sheine, 1979). The work in chapter 3 provides evidence to suggest that insect-eating primates share an adaptation found in insectivorous bats (Vespertilionidae) and mice (*Mus musculus*) (Strobel et al., 2013; Ohno et al., 2016) and use the enzyme acidic mammalian chitinase to digest the chitin in insect exoskeletons. While most primates retain one functional *CHIA* paralog, more insectivorous and smaller-bodied species have between two and five functional paralogs. The efficient digestion of insect exoskeletons is likely to have important adaptive benefits for all insect-eating primates, through the potentially significant energy and amino acid returns from the digestion of the polysaccharide chitin (Finke, 2007; Rothman et al., 2014) and, for small-bodied primates, by reducing the amount of indigestible bulk in their guts.

To further investigate dietary adaptations for insectivory in the acidic mammalian chitinase genes, chapter 4 focuses on copy number variation in the *CHIA* paralogs. Copy number variation is a common form of structural genetic variation (Stankiewicz and Lupski, 2010; Clop et al., 2012; Sudmant et al., 2013). Higher gene copy numbers have been found to correlate with increased expression of the protein (Hollox et al., 2003; Linzmeier and Ganz, 2006; Perry et al., 2007) and have phenotypic effects (Giuffra et al.,

1999; Elferink et al., 2008; Fontanesi et al., 2009; Wright et al., 2009; Fontanesi et al., 2011). This includes digestive enzymes, such as salivary and pancreatic amylase (Perry et al., 2007; Mandel et al., 2010; Mandel and Breslin, 2012; Axelsson et al., 2013) and pepsin A (Taggart et al., 1985; Narita et al., 2010). However, contrary to the hypothesis, there is no evidence that the *CHIA* genes had higher copy numbers in more insectivorous primate species. All primate species included in this study had two gene copies per diploid genome, suggesting that there is no CNV in primate *CHIA* genes (Chapter 4). However, the expression of acidic mammalian chitinase may be regulated via other mechanisms to correlate with insect consumption in primates.

Future Directions

Future research is needed to fully understand digestive adaptations in primates and other mammals. Without gene expression data for primate digestive tracts, we cannot be certain where the genes investigated here (*RNASE1, RNASE1B, RNASE1C, CHIA*) are expressed and whether they are expressed in the stomach or intestines of primates. One study has found that *mCHIA* is expressed in the stomach of *Macaca fascicularis* (Krykbaev et al., 2010), but data from other primates are not available. It would be especially important to know how the various *CHIA* paralogs are expressed in species that retain more than one functional gene. The tarsier (*Tarsius syrichta*) is most intriguing in this regard, because it is the only primate that is exclusively faunivorous, consuming mostly insects (Gursky, 2007; 2011), and it has five (presumably) functional *CHIA* paralogs.

My work has identified inter-specific variation in *CHIA* sequences, ranging from single nucleotide polymorphisms to larger indels, but it is not clear what effects, if any, this variation has on the functionality of the acidic mammalian chitinase protein. Likewise, changes in the biochemical properties of the howler monkey *RNASE1B* and *RNASE1C* genes are inferred based on the results of previous experiments (Zhang, 2003b; 2006), but need to be confirmed. Cloning the different *CHIA* and *RNASE1* genes into expression vectors to study their effects on appropriate substrates *in vitro* is a clear avenue for future research (Zhang, 2006; Carrigan et al., 2015).

A larger comparative context of acidic mammalian chitinase evolution could further elucidate its role as a digestive enzyme adaptation for insectivory in mammals. The order Chiroptera exhibits exceptionally great dietary diversity, with species variably feeding on blood, insects, small vertebrates, nectar, fruit, or pollen (Schondube et al., 2001). Bats and primates share a gene (or genes) for acidic mammalian chitinase and appear to have a shared adaptation for the digestion of insect exoskeletons (Strobel et al., 2013). However, non-insectivorous species of both orders may have likewise lost the enzyme convergently. Vampire bats, fruit bats, and folivorous monkeys would provide a useful study for convergent evolution in this case.

Finally, I have started a project to investigate the expression of acidic mammalian chitinase in the saliva of insectivorous primates. Salivary enzymes are the first step in enzymatic digestion (Stevens and Hume, 1995), may correlate with diet (Lambert, 1998; Perry et al., 2007; Janiak, 2016), and affect our perception of food (Mandel et al., 2010). Mice have been shown to express chitinase in their saliva (Goto et al., 2003), but it is

unknown whether insectivorous primates, such as bushbabies or tarsiers, share this adaptation.

References

Axelsson E, Ratnakumar A, Arendt M-L, Maqbool K, Webster MT, Perloski M, Liberg O, Arnemo JM, Hedhammar Å, Lindblad-Toh K. 2013. The genomic signature of dog domestication reveals adaptation to a starch-rich diet. Nature 495:360–364.

Carrigan MA, Uryasev O, Frye CB, Eckman BL, Myers CR, Hurley TD, Benner SA. 2015. Hominids adapted to metabolize ethanol long before human-directed fermentation. Proc Natl Acad Sci USA 112:458–463.

Clop A, Vidal O, Amills M. 2012. Copy number variation in the genomes of domestic animals. Anim Genet 43:503–517.

Cork SJ, Kenagy GJ. 1989. Nutritional Value of Hypogeous Fungus for a Forest-Dwelling Ground Squirrel. Ecology 70:577–586.

Cornelius C, Dandrifosse G, Jeuniaux C. 1976. Chitinolytic enzymes of the gastric mucosa of *Perodicticus potto* (primate prosimian): Purification and enzyme specificity. International Journal of Biochemistry 7:445–448.

Elferink MG, Vallée AAA, Jungerius AP, Crooijmans RPMA, Groenen MAM. 2008. Partial duplication of the *PRLR* and *SPEF2* genes at the late feathering locus in chicken. BMC Genomics 9:391.

Finke MD. 2007. Estimate of chitin in raw whole insects. Zoo Biol 26:105–115.

Fontanesi L, Beretti F, Riggio V, González EG, Dall'Olio S, Davoli R, Russo V, Portolano B. 2009. Copy Number Variation and Missense Mutations of the Agouti Signaling Protein (*ASIP*) Gene in Goat Breeds with Different Coat Colors. Cytogenet Genome Res 126:333–347.

Fontanesi L, Dall'Olio S, Beretti F, Portolano B, Russo V. 2011. Coat colours in the Massese sheep breed are associated with mutations in the agouti signalling protein (*ASIP*) and melanocortin 1 receptor (*MC1R*) genes. Animal 5:8–17.

Garber PA, Righini N, Kowalewski MM. 2015. Evidence of Alternative Dietary Syndromes and Nutritional Goals in the Genus *Alouatta*. In: Garber PA, Cortés-Ortiz L, Urbani B, Youlatos D, editors. Howler Monkeys. New York, NY: Springer New York. p 85–109.

Giuffra E, Evans G, Törnsten A, Wales R, Day A, Looft H, Plastow G, Andersson L. 1999. The Belt mutation in pigs is an allele at the Dominant white (I/KIT) locus. Mamm Genome 10:1132–1136.

Goto M, Fujimoto W, Nio J, Iwanaga T, Kawasaki T. 2003. Immunohistochemical demonstration of acidic mammalian chitinase in the mouse salivary gland and gastric mucosa. Arch Oral Biol 48:701–707.

Gursky SL. 2007. The Spectral Tarsier. Abingdon, UK: Routledge.

Gursky SL. 2011. Tarsiiformes. In: Campbell CJ, Fuentes A, MacKinnon KC, Bearder SK, Stumpf RM, editors. Primates in Perspective. Oxford University Press. p 79–90.

Hollox EJ, Armour JAL, Barber JCK. 2003. Extensive normal copy number variation of a beta-defensin antimicrobial-gene cluster. Am J Hum Genet 73:591–600.

Janiak MC. 2016. Digestive enzymes of human and nonhuman primates. Evol Anthropol 25:253–266.

Jeuniaux C, Cornelius C. 1978. Distribution and activity of chitinolytic enzymes in the digestive tract of birds and mammals. In: Muzzarelli RAA, Pariser ER, editors. p 542–549.

Karasov WH, Douglas AE. 2013. Comparative digestive physiology. Compr Physiol 3:741–783.

Kay RF, Sheine WS. 1979. On the relationship between chitin particle size and digestibility in the primate Galago senegalensis. Am J Phys Anthropol 50:301–308.

Krykbaev R, Fitz LJ, Reddy PS, Winkler A, Xuan D, Yang X, Fleming M, Wolf SF. 2010. Evolutionary and biochemical differences between human and monkey acidic mammalian chitinases. Gene 452:63–71.

Lambert JE. 1998. Primate digestion: interactions among anatomy, physiology, and feeding ecology. Evol Anthropol 7:8–20.

Linzmeier RM, Ganz T. 2006. Copy number polymorphisms are not a common feature of innate immune genes. Genomics 88:122–126.

Mandel AL, Breslin PAS. 2012. High endogenous salivary amylase activity is associated with improved glycemic homeostasis following starch ingestion in adults. J Nutr 142:853–858.

Mandel AL, Peyrot des Gachons C, Plank KL, Alarcon S, Breslin PAS. 2010. Individual differences in *AMY1* gene copy number, salivary α-amylase levels, and the perception of oral starch. PLoS ONE 5:e13352.

Milton K, McBee RH. 1983. Rates of fermentative digestion in the howler monkey, *Alouatta palliata* (primates: ceboidea). Comp Biochem Physiol A Comp Physiol 74:29–31.

Narita Y, Oda S-I, Takenaka O, Kageyama T. 2010. Lineage-specific duplication and loss of pepsinogen genes in hominoid evolution. J Mol Evol 70:313–324.

Oftedal OT, Whiten A, Southgate DAT, Van Soest PJ. 1991. The nutritional consequences of foraging in primates: the relationship of nutrient intakes to nutrient

requirements. Philos Trans R Soc Lond, B, Biol Sci 334:161–170.

Ohno M, Kimura M, Miyazaki H, Okawa K, Onuki R, Nemoto C, Tabata E, Wakita S, Kashimura A, Sakaguchi M, Sugahara Y, Nukina N, Bauer PO, Oyama F. 2016. Acidic mammalian chitinase is a proteases-resistant glycosidase in mouse digestive system. Sci Rep 6:37756.

Perry GH, Dominy NJ, Claw KG, Lee AS, Fiegler H, Redon R, Werner J, Villanea FA, Mountain JL, Misra R, Carter NP, Lee C, Stone AC. 2007. Diet and the evolution of human amylase gene copy number variation. Nat Genet 39:1256–1260.

Ranciaro A, Campbell MC, Hirbo JB, Ko W-Y, Froment A, Anagnostou P, Kotze MJ, Ibrahim M, Nyambo T, Omar SA, Tishkoff SA. 2014. Genetic origins of lactase persistence and the spread of pastoralism in Africa. Am J Hum Genet 94:496–510.

Raubenheimer D, Rothman JM. 2013. Nutritional ecology of entomophagy in humans and other primates. Annu Rev Entomol 58:141–160.

Rosenberger AL. 1992. Evolution of feeding niches in New World monkeys. Am J Phys Anthropol 88:525–562.

Rothman JM, Raubenheimer D, Bryer MAH, Takahashi M, Gilbert CC. 2014. Nutritional contributions of insects to primate diets: Implications for primate evolution. Journal of Human Evolution 71:59–69.

Schienman JE, Holt RA, Auerbach MR, Stewart C-B. 2006. Duplication and divergence of 2 distinct pancreatic ribonuclease genes in leaf-eating African and Asian colobine monkeys. Mol Biol Evol 23:1465–1479.

Schondube JE, Herrera-M LG, Martinez del Rio C. 2001. Diet and the evolution of digestion and renal function in phyllostomid bats. Zoology (Jena) 104:59–73.

Simmen B, Tarnaud L, Hladik A. 2012. Leaf nutritional quality as a predictor of primate biomass: further evidence of an ecological anomaly within prosimian communities in Madagascar. J Trop Ecol 28:141–151.

Simunek J, Bartonova H. 2005. Effect of Dietary Chitin and Chitosan on Cholesterolemia of Rats. Acta Veterinaria Brno 74:491–499.

Stankiewicz P, Lupski JR. 2010. Structural variation in the human genome and its role in disease. Annu Rev Med 61:437–455.

Stevens CE, Hume ID. 1995. Comparative physiology of the vertebrate digestive system. 2nd ed. Cambridge: Cambridge University Press.

Strobel S, Roswag A, Becker NI, Trenczek TE, Encarnação JA. 2013. Insectivorous Bats Digest Chitin in the Stomach Using Acidic Mammalian Chitinase. PLoS ONE 8:e72770.

Sudmant PH, Huddleston J, Catacchio CR, Malig M, Hillier LW, Baker C, Mohajeri K, Kondova I, Bontrop RE, Persengiev S, Antonacci F, Ventura M, Prado-Martinez J, Great Ape Genome Project, Marques-Bonet T, Eichler EE. 2013. Evolution and diversity of copy number variation in the great ape lineage. Genome Research 23:1373–1382.

Taggart RT, Mohandas TK, Shows TB, Bell GI. 1985. Variable numbers of pepsinogen genes are located in the centromeric region of human chromosome 11 and determine the high-frequency electrophoretic polymorphism. Proc Natl Acad Sci USA 82:6240–6244.

Tishkoff SA, Reed FA, Ranciaro A, Voight BF, Babbitt CC, Silverman JS, Powell K, Mortensen HM, Hirbo JB, Osman M, Ibrahim M, Omar SA, Lema G, Nyambo TB, Ghori J, Bumpstead S, Pritchard JK, Wray GA, Deloukas P. 2007. Convergent adaptation of human lactase persistence in Africa and Europe. Nat Genet 39:31–40.

Wright D, Boije H, Meadows JRS, Bed'hom B, Gourichon D, Vieaud A, Tixier-Boichard M, Rubin C-J, Imsland F, Hallböök F, Andersson L. 2009. Copy number variation in intron 1 of *SOX5* causes the pea-comb phenotype in chickens. PLoS Genet 5:e1000512.

Yu L, Wang X-Y, Jin W, Luan P-T, Ting N, Zhang Y-P. 2010. Adaptive evolution of digestive RNASE1 genes in leaf-eating monkeys revisited: new insights from ten additional colobines. Mol Biol Evol 27:121–131.

Zhang J, Zhang Y-P, Rosenberg HF. 2002. Adaptive evolution of a duplicated pancreatic ribonuclease gene in a leaf-eating monkey. Nat Genet 30:411–415.

Zhang J. 2003a. Evolution by gene duplication: an update. Trends in Ecology & Evolution 18:292–298.

Zhang J. 2003b. Parallel Functional Changes in the Digestive RNases of Ruminants and Colobines by Divergent Amino Acid Substitutions. Mol Biol Evol 20:1310–1317.

Zhang J. 2006. Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys. Nat Genet 38:819–823.

APPENDICES

Appendix A. Supplementary Table 1. Samples and genomes used in Chapter 3.

Species	Sample ID/Source	WGS accession number	GenBank accession number
Allenopithecus nigroviridis	Dr. Anthony Tosi		TBD
Allochrocebus lhoesti	Dr. Anthony Tosi		TBD
Aotus nancymaae		JYKP00000000.2	
Callicebus moloch	NG06115, Coriell Biorepository		TBD
Callithrix jacchus			TBD
Callithrix jacchus		ACFV00000000.1	
Cebus capucinus imitator		LVWQ00000000.1	
Cercocebus atys		JZLG00000000.1	
Cercopithecus mitis	Dr. Anthony Tosi		TBD
Chlorocebus aethiops	Dr. Anthony Tosi		TBD
Chlorocebus sabaeus		AQIB00000000.1	
Colobus angolensis		JYKR00000000.1	
Colobus guereza	Dr. Anthony Tosi		TBD
Erythrocebus patas	Dr. Todd Disotell		TBD
Gorilla gorilla gorilla		CABD000000000.3	
Macaca fascicularis	Krykbaev et al., 2010	AQIA00000000.1	
Macaca mulatta		JSUE00000000.3	
Macaca nemestrina		JZLF00000000.1	
Mandrillus leucophaeus		JYKQ00000000.1	

Microcebus murinus		ABDC00000000.3	
Miopithecus talapoin	Dr. Anthony Tosi, Dr. Todd Disotell		TBD
Nasalis larvatus		JMHX00000000.1	
Nomascus leucogenys		ADFV00000000.1	
Otolemur garnettii		AAQR00000000.3	
Pan paniscus		AJFE00000000.2	
Pan troglodytes		AACZ00000000.4	
Papio anubis		AHZZ00000000.2	
Pongo abelii		ABGA00000000.1	
Rhinopithecus bieti		MCGX00000000.1	
Rhinopithecus roxellana		JABR00000000.1	
Saguinus fuscicollis	NG05313, Coriell Biorepository		TBD
Saimiri boliviensis		AGCE00000000.1	
Saimiri sciureus	NG05311, Coriell Biorepository		TBD
Sapajus apella	Dr. George Perry		TBD
Tarsius syrichta		ABRT00000000.2	
Tupaia chinensis		ALAR00000000.1	

Appendix B. Supplementary Table 2. Primer sequences used in Chapter 3.

mCHIA	**Forward**		**Reverse**	
Platyrrhines				
Exon 1	CHIA_Ex1_F	AGGCAATCTCAATAAAGGAGAATCT	CHIA_Ex1_R	GAAGGTAAAAGAGTTTTAAAGGGGC
Exon 2	CHIA_Ex2_F	AGCAAAAGCTTCCTACTGCTAAA	CHIA_Ex2_R	AACAATATAGTGGTTCATTCCAA
Exon 3	CHIA_Ex3_F	TCTATGACTCACTAGGGAGCTTGTC	CHIA_Ex3_R	TCAGCCTGAAGTCTGCTGTG
	or CHIA_3_F	GGAAATCGTTTGGAAATCAGTGGT	CHIA_3_R	TTGTTTAAACTGTCCTTGAAGGTT
Exon 4	CHIA_Ex4_F	CACTTTTCCTGGCTTCTTGC	CHIA_Ex4_R	GGAAACCATGGCAGTGAAAC
Exon 5	CHIA_Ex5_F	GATTTTACTCCAAGGTGTTGGG	CHIA_Ex5_R	GTGTATGCTTTCCAGTCAAAATGA
Exon 6	CHIA_Ex6_F	GAGTTCCAGATCATTTTGACTGG	CHIA_Ex6_R	CGTAGGTCATGACATGGATGTAG
Exon 6 & 7	CHIA_6_7_F	TTCACAGACTTTCTCCTGCTCCTG	CHIA_6_7_R	TCTTGCATCTGCACAGGGACT
Exon 7	CHIA_Ex7_F	ATCCAGTCTGGCTATGAGATCC	CHIA_Ex7_R	CTTGCATCTGCACAGGGAC
Exon 8	CHIA_Ex8_F	AATAATTAATCACTAGGCCATCAGC	CHIA_Ex8_R	GGTTTACTAAATGGACCAAGAACTG
Exon 9	CHIA_Ex9_F	CGTGTCCTTTTACCTCTGCCA	CHIA_Ex9_R	CAGGGCTATAATGGGAACT
Exon 10	CHIA_Ex10_F	CCTGTGAGCAAGTTCCCATTA	CHIA_Ex10_R	GGCATTTCTCTCAAGTCTGAGG
Exon 11	CHIA_Ex11_F	AGCCAAGAACAAACTCCAGG	CHIA_Ex11_R	AGAAGCAGAAGCCATTGTACAGTAA
	or CHIA_11_F	AAGCAAGGAGGTTCAGGATTTAT	CHIA_11_R	TTATTGCACAAGAACTACTGAGG

Catarrhines				
Exon 1	CHIA_OWM_Ex1_F	AATGGCAGGTTGGATGAGGG	CHIA_OWM_Ex1_R	GGGGCAGAAAAACAAGAGGG
Exon 2	CHIA_OWM_Ex2_F	GGCTGCTCTACTCACATTCCT	CHIA_OWM_Ex2_R	GGAAGGACACAGGGCCATAC
Exon 3	CHIA_OWM_Ex3_F	AGGGAGCTTGTCTCTCACCT	CHIA_OWM_Ex3_R	GCCTGAAGTCAGCTGTGGAA
Exon 4 & 5	CHIA_OWM_Ex4+5_F	TTCCTGGCTTCCTCCTCACT	CHIA_OWM_Ex4+5_R	TGGGTGTATGCTTTCCAGTCA
Exon 6 & 7	CHIA_OWM_Ex6+7_F	TGAATGGAAGGATGGAAGAGG	CHIA_OWM_Ex6+7_R	TCCCCATTTGAGAAAACCACT
Exon 8	CHIA_OWM_Ex8_F	TAGAGTGGGCACAACCAAACC	CHIA_OWM_Ex8_R	ACATGTACCAAGAGGAAGCCA
Exon 9 & 10	CHIA_OWM_Ex9+10_F	CCTCTACTTCCCCTTCCCCA	CHIA_OWM_Ex9+10_R	TCCCTGACTTGCAAGGAACC
Exon 11	CHIA_OWM_Ex11_F	TCCAGGGTATTTGCTCTCACAC	CHIA_OWM_Ex11_R	CACAACAACTTCTGCGGCTG
hCHIA				
Platyrrhines				
Exon 1	CHIA2_NWM_1_F	CCCTTGGAACTGCCAAAAA	CHIA2_NWM_1_R	AACAAGAGGCAGACTGATTTG
Exon 2	CHIA2_2_F	CTTGTCTGTTGGAAGGCTGT	CHIA2_2_R	GACAATGGCCTGGTTTAGGA
Exon 3	CHIA2_NWM_3_F	CATTGTCCCTCCTGCTGATT	CHIA2_NWM_3_R	TCCCGCCTAGGTGACAGA
Exon 4	CHIA2_NWM_4_F	GCAATGTATTTTAAACAGAGGGACT	CHIA2_NWM_4_R	TCCTCAGGGTTTGTACTCTTGAA
Exon 5	CHIA2_5_F	GATACCATGGCTGGGAAGAG	CHIA2_5_R	AGCCACCCAATTCTCCTTG
Exon 6 & 7	CHIA2_NWM_6+7_F	TTTAAGGAGCTAAAATCAGCATCA	CHIA2_NWM_6+7_R	CCCCGTTGTCCTTCCAGTAG

Exon 8	CHIA2_NWM_8_F	TCAATGTGGTGAGTCCCTGT	CHIA2_NWM_8_R	ACAGAGCCCTTTTCAGTCCA
Exon 9 & 10	CHIA2_NWM_9+10_F	GGACTGAAAAGGGCTCTGTG	CHIA2_NWM_9+10_R	GGCAAGAGCAGTGAGGTGA
Exon 11	CHIA2_NWM_11_F	TCACCACTTATGCCCAGATTC	CHIA2_NWM_11_R	CCAGTCAATCAGCTCTAGGTCTT
Catarrhines				
Exon 1	CHIA2_OWM_1_F	CATTTGGAAGCCTTTGTGAT	CHIA2_OWM_1_R	CCACTTCATGATTGTCAGAGC
Exon 2	CHIA2_2_F	CTTGTCTGTTGGAAGGCTGT	CHIA2_2_R	GACAATGGCCTGGTTTAGGA
Exon 3	CHIA2_OWM_3_F	ACCTCCCTCCACACAGAGAT	CHIA2_OWM_3_R	AAATGAAATGCAGAGTCTTGGTC
Exon 4	CHIA2_OWM_4_F	TGCATTCATGACAAGAACAACA	CHIA2_OWM_4_R	CAGTGATCACATGGGGCTAA
Exon 5	CHIA2_5_F	GATACCATGGCTGGGAAGAG	CHIA2_5_R	AGCCACCCAATTCTCCTTG
Exon 6 & 7	CHIA2_OWM_6+7_F	CTGAGACCTGGGTTTTGGTC	CHIA2_OWM_6+7_R	TCATCATCTCTGGTCTGCACT
Exon 8	CHIA2_OWM_8_F	TGCAATTAACAGCATACAGTTTCA	CHIA2_OWM_8_R	AAATGCTCTTCATTTGGATTATG
Exon 9 & 10	CHIA2_OWM_9+10_F	AACATGTTTTTCTTTAATGGGAGT	CHIA2_OWM_9+10_R	GTGGGAAGACATCAGGGTTG
Exon 11	CHIA2_OWM_11_F	TGGCTTGACACAATAGCTTTACTT	CHIA2_OWM_11_R	CCTAAGCAAAAGGGACTGGA

Appendix C. Supplementary Table 3. Diet data used to calculate average insect consumption in Chapter 3.

Species	Fruit	Leaves	Flowers	Seeds (when not incl. in fruit)	Insects	Other	Notes	Site	References
Allenopithecus nigroviridis	81	5			18			Zaire	Gautier-Hion, 1988
	40		20		0	20	6.7% gum, 13.3% other plant material	Lomako, DRC	Zeeve, 1991
Aotus nancymaae	45	41	14				*Aotus azarai azarai*	Guaycolec, Formosa, Argentina	Arditi & Placci, 1990
	60	"not extensive"	Nectar consumed in July and August				*Aotus nigriceps*	Manu, Peru	Wright, 1985
	75	10			15		*Aotus trivirgatus*	Peru	Wright, 1985
	16	40			11	33	*Aotus trivirgatus*	Paraguay	Wright, 1985
Callicebus moloch	54	28	2		17	1	*Callicebus brunneus*		Wright, 1985
	50	39			11		*Callicebus brunneus*		Crandlemire-Sacco, 1988
	70	14	1		8	7	*Callicebus brunneus*		Lawrence, 2007
Callithrix jacchus	18.1				5.4	89.3	76.4% exudates, 12.9% fungus	Joao Pessoa, Brazil	Alonso & Langguth, 1989
	23				9	68	68% exudates	Nisia Floresta, Brazil	Digby et al., 2011
Cebus capucinus					23		Paper focused on animal foods	Santa Rosa, Costa Rica	Rose, 1994
	62.14	1.13	0	10.57	21.36		adults only	Santa Rosa, Costa Rica	MacKinnon, 2006

	47.8	0.5	1.2		49.8	0.7		La Suerte, Costa Rica	Mallott, 2016
Cercocebus atys	68	2	2		26	1	insects = "animal food"	Taï Forest, Cote d'Ivoire	Bergmüller, cited in McGraw 1998
Cercopithecus lhoesti	24.5	35.2	4	17.7	8.8	9.8		Nyungwe, Rwanda	Kaplin & Moermond, 2000; Kaplin, 2001
Cercopithecus mitis	54.6	18.9	3.7	2.5	16.8	3.5		Kakamega, Kenya	Cords, 1986
	47.4	6.2	6.2	9.3	24.9	6.2		Nyungwe, Rwanda	Kaplin & Moermond, 2000; Kaplin, 2001
	57	26	13		6			Cape Vidal, South Africa	Lawes 1991
	53.5	32.6	10.2		1	2.9		Zomba Plateau, Malawi	Beeson et al., 1996
	91	3	2		0			Ngoye Forest, South Africa	Lawes et al., 1990
	27.7	33	6.9		37.7	0.6		Kibale, Uganda	Butynski, 1990
	30.1	22.8	9.8		35.9	1.3		Kibale, Uganda	Butynski, 1990
Chlorocebus aethiops	11.1	26.6	14.3	2.6	7.7	0.2		Amboseli, Kenya	Wrangham & Waterman, 1981
	5.8	0	44.7	19.6		9.6		Samburo-Isiolo, Kenya	Whitten, 1983
	10	2	8	8	23	10		Segera, Kenya	Isbell et al., 1998
Colobus angolensis	28	50	7		0	15		Ituri, DRC	Bocian, 1997
	23	38	1		0	37	32% lichen	Nyungwe, Rwanda	Vedder & Fashing, 2002
	17	72	5		0	6		Nyungwe, Rwanda	Fimbel et al., 2001
	15	81	2		0	2		Kibale, Uganda	Oates, 1977; 1994

Colobus guereza	5	86	1		0	8		Kibale, Uganda	Wasserman & Chapman, 2003
	12	80	6		0	2		Kibale, Uganda (fragment)	Wasserman & Chapman, 2003
	10	87	3		0	0		Kibale, Uganda (logged)	Wasserman & Chapman, 2003
	25	58	3		0	15		Ituri, DRC	Bocian, 1997
	39	54	1		0	8		Kakamega, Kenya	Fashing, 2001
	26	62	9		0	4		Budongo, Uganda	Plumptre, 2006
	36	53	8		0	3		Budongo, Uganda (logged)	Plumptre, 2006
Erythrocebus patas	10	2	7		35	38	37% exudates	Segera, Kenya	Isbell, 1998
		6	65		12	7	7% gum	Kala Maloue, Cameroon	Nakagawa, 1989
Gorilla gorilla	36.2	30.6			7.7	26.8	other = herbs	Mondika, CAR/Congo	Doran-Sheehy et al., 2009
Macaca fascicularis	66.7	17.2	8.9		4.1	3.1		Tanjung Puting, Kalimantan	Yeager, 1996
	87	1.6	5.4		4	2		Kutai, Kalimantan	Wheatley, 1980
Macaca mulatta	8.5	84.4	5.9		0	1.1		Murree Hills, Pakistan	Goldstein & Richard, 1989
Macaca nemestrina	74.2	11.1	1.1		12.2	1.4		Pasoh, Malaysia	Caldecott, 1986
Mandrillus leucophaeus							not well known in the wild		

Microcebus murinus	63		22		11	4		Mandena, Madagascar	Lahann, 2006
	25				5	70	60% gum (numbers taken from figure)	Ankarafantsika, Madagascar	Thoren et al., 2011
Miopithecus	43	5			35				Gautier-Hion, 1988
Nasalis larvatus	11	74	8		0	8		Kinabatangan-Sukai	Boonratana, 1993
	40	52	3		0	5		Tanjung Puting	Yeager, 1989, 1990
	58	41			0			Samunsam	Bennett & Sebastian, 1988
Nomascus leucogenys	39	53	5		4			Meng La, Yunnan, China	Hu et al., 1989, 1990
Otolemur garnettii	50				50		based on stomach contents		Charles-Dominique, 1977
Pan paniscus	55	14	2		2	27	25% terrestrial herbaceous vegetation		Conklin-Brittain et al., 2006
Pan troglodytes	60.8	4.3	5.5		23.9	1.6		Fongoli, Senegal	Bogart & Pruetz, 2011
	49.2	36.1	6.3	3.4		5.9	other = invertebrates, vertebrates, bark	Bossou, New Guinea	Sugiyama & Koman, 1987
	56	16	8	7	2		insects = "animal food"	Goualougo, Republic of Congo	Morgan & Sanz, 2006
	88	3	2		3		insects = "animal food"	Ndoki, Republic of Congo	Kuroda et al., 1996
	69.2	9.9	6.4	7.7		1		Lope, Gabon	Tutin et al., 1997
	36	22	7.8	4.8	6.6	8		Gombe, Tanzania	Wrangham, 1977
	54.3	23			4.1	2.9		Mahale, Tanzania	Matsumoto-Oda, 2002

	64.5	19.7	8.8		4		Insects = "animal food"	Budongo, Uganda	Newton-Fisher, 1999
	79	2.6			0.9		Insects = "animal food"	Kibale, Uganda	Wrangham et al., 1996
Papio anubis	55	33	7		3	3		Bole Valley, Ethiopia	Dunbar & Dunbar, 1974
	24.3	7.2	29.5			41	other = USOs, grasses, stems	Laikipia, Kenya	Barton et al., 1993
	47	17				27	17% garbage	Budongo Forest, Uganda	Okecha & Newton-Fisher, 2006
	47	5	6		1	41	29% grasses	Comoe, Cote d'Ivoire	Kunz & Linsenmair, 2008
Pongo abelii	67.5	16.5			8.8	4.8		Ketambe, Sumatra	Wich et al., 2006
	66.2	15.5			13.4	3.8		Suaq, Sumatra	Wich et al., 2006
Rhinopithecus bieti	11	20			0	69	67% lichen		Grueter et al. 2009
Rhinopithecus roxellana	29.4	24			0	46.6	29% lichen		Guo, Li, Watanabe 2007
Saguinus fuscicollis	64.2				5.8	30.3	30.3% exudates	Quebrada Bl., Peru	Knogge & Heymann, 2003
	39.2				53.1	7.6	7.6% exudates	Rio Blanco, Peru	Garber, 1988
	63				26	18	12% exudates	Pando, Bolivia	Porter, 2001
Saimiri boliviensis							no wild studies		
Saimiri sciureus	55		(incl in fruit)		45			Gunma, Brazil	Lima & Ferrari, 2003
					61.75		contribution of other food items not easily available	Ananim, Brazil	Stone, 2007
Sapajus apella	37.2	18.6	3.9		40.3			El Rey, Brazil	Brown & Zunino, 1990

	2.9	72.3	0		24.9		Iguazu, Brazil	Brown & Zunino, 1990
Tarsius syrichta					90	10	no observations of anything but insects and small vertebrates	Gursky, 2011
Tupaia belangeri/chinensis					80?		few wild studies, mostly based on stomach content	Langham, 1982

References

Alonso, C. & Langguth, A. (1989). Ecologia e comportamento de *Callithrix jacchus* (Primates: Callitrichidae) numa ilha de foresta Atlântica. Rev. Nordestina Biol. 6:107-137.

Arditi, S.I. & Placci, L.G. (1990). Hábitat y densidad de Aotus azarae y Alouatta caraya en Riacho Pilagá, Formosa. Boletín Primatológico Latinoamericano 2:29-47.

Barton, R.A. & Whiten, A. (1993). Feeding competition among female olive baboons, *Papio anubis*. Animal Behaviour 46:777–789.

Beeson, M. et al. (1996). Food habits of guenons (Cercopithecus spp.) in Afro-montane forest. African Journal of Ecology 34:202–210.

Bennett, E.L. & Sebastian, A.C. (1988). Social organization and ecology of proboscis monkeys (*Nasalis larvatus*) in mixed coastal forest in Sarawak. Int. J. Primatol. 9:233–255.

Bergmüller, R. (1998). Nahrungsökologie der Rauchgrauen Mangabe (*Cercocebus torquatus atys*), Ein Schlüssel zur sozialen Organisation? Diplomarbeit Universität Erlangen.

Bocian, C.M. (1997) Niche separation of black-and-white colobus monkeys (*Colobus angolensis* and *C. guereza*) in the Ituri Forest. PhD thesis, City University of New York.

Bogart, S.L. & Pruetz, J.D. (2011). Insectivory of savanna chimpanzees (*Pan troglodytes verus*) at Fongoli, Senegal. Am. J. Phys. Anthropol. 145:11–20.

Boonratana, R. (1993). The ecology and behaviour of the proboscis monkey (*Nasalis larvatus*) in the Lower Kinabatangan, Sabah. PhD Thesis. Mahidol University.

Brown, A.D. & Zunino, G.E. (1990). Dietary Variability in *Cebus apella* in Extreme Habitats: Evidence for Adaptability. Folia Primatol. 54:187–195.

Butynski, T.M. (1990). Comparative Ecology of Blue Monkeys (*Cercopithecus mitis*) in High- and Low-Density Subpopulations. Ecological Monographs 60:1–26.

Caldecott, J.O. (1986). An ecological and behavioural study of the pig-tailed macaque. *Contrib. Primatol.* 21: 1-259.

Charles-Dominique, P. (1977). Ecology and Behaviour of Nocturnal Primates. London: Duckworth Press.

Conklin-Brittain, N.L. et al. (2006). Energy intake by wild chimpanzees and orangutans: methodological considerations and a preliminary comparison. In: Feeding Ecology in Apes and Other Primates: Ecological, Physiological, and Behavioural Aspects. Cambridge: Cambridge University Press. pp. 445-472.

Cords, M. (1986). Interspecific and intraspecific variation in diet of two forest guenons, *Cercopithecus ascanius* and *C. mitis*. The Journal of Animal Ecology 55:811.

Crandlemire-Sacco J. 1988. An ecological comparison of two sympatric primates: *Saguinus fuscicollis* and *Callicebus moloch* of Amazonian Peru. Primates 29:465–475.

Digby, L.J. et al. (2011). Callitrichines. In: *Primates in Perspective*. New York: Oxford University Press. pp. 91-107.

Doran-Sheehy, D. et al. (2009). Male and female western gorilla diet: Preferred foods, use of fallback resources, and implications for ape versus old world monkey foraging strategies. Am. J. Phys. Anthropol. 140:727–738.

Dunbar, R.I.M & Dunbar, E.P. (1974). Ecological Relations and Niche Separation between Sympatric Terrestrial Primates in Ethiopia. Folia Primatol. 21:36-60.

Fashing, P. (2001). Male and female strategies during intergroup encounters in guerezas (*Colobus guereza*): evidence for resource defense mediated through males and a comparison with other primates. Behav. Ecol. Sociobiol. 50:219–230.

Fimbel, C. et al. (2001). An ecological basis for large group size in *Colobus angolensis* in the Nyungwe Forest, Rwanda. African Journal of Ecology 39:83–92.

Garber, P.A. (1988). Diet, Foraging Patterns, and Resource Defense in a Mixed Species Troop of *Saguinus mystax* and *Saguinus fuscicollis* in Amazonian Peru. Behaviour 105:18–34.

Gautier-Hion, A. (1988). Diet and dietary habits of forest guenons. In: *A primate radiation: Evolutionary Biology of the African Guenons*. New York: Cambridge University Press. pp. 257-283.

Goldstein, S.J. & Richard, A.F. (1989). Ecology of rhesus macaques (*Macaca mulatta*) in northwest Pakistan. International Journal of Primatology 10:531-567.

Grueter, C.C. et al. (2009). Dietary profile of *Rhinopithecus bieti* and its socioecological implications. Int. J. Primatol. 30:601–624.

Guo, S. et al. (2007). Diet and activity budget of *Rhinopithecus roxellana* in the Qinling Mountains, China. Primates 48:268–276.

Gursky, S. (2011). Tarsiiformes. In: *Primates in Perspective*. New York: Oxford University Press. pp. 79-90.

Hu, Y. et al. (1989). The studies on ecology of the *Hylobates leucogenys*. Zool. Res. 10(S):61-66.

Hu, Y. et al. (1990). Feeding ecology of the white-cheek gibbon (*Hylobates concolor leucogenys*). Acta. Ecol. Sin. 10:155-159.

Isbell, L.A. et al. (1998). Movements of vervets (*Cercopithecus aethiops*) and patas monkeys (*Erythrocebus patas*) as estimators of food resource size, density, and distribution. Behav. Ecol. Sociobiol. 42:123-133.

Isbell, L.A. (1998). Diet for a small primate: insectivory and gummivory in the (large) patas monkey (*Erythrocebus patas pyrrhonotus*). Am. J. Primatol. 45:381-389.

Kaplin, B.A. & Moermond, T.C. (2000). Foraging ecology of the mountain monkey (*Cercopithecus lhoesti*): implications for its evolutionary history and use of disturbed forest. Am. J. Primatol. 50:227-246.

Kaplin, B.A. (2001). Ranging behavior of two species of guenons (*Cercopithecus lhoesti* and *C. mitis doggetti*) in the Nyungwe Forest Reserve, Rwanda. Int. J. Primatol. 22:521-548.

Knogge, C. & Heymann, E.W. (2003). Seed dispersal by sympatric tamarins, *Saguinus mystax* and *Saguinus fuscicollis*: diversity and characteristics of plant species. Folia Primatol. 74:33-47.

Kunz, B.K. & Linsenmair, K.E. (2008). The disregarded west: diet and behavioral ecology of olive baboons in the Ivory Coast. Folia Primatol. 79:31-51.

Kuroda, S. et al. (1996). Sympatric chimpanzees and gorillas in the Ndoki Forest, Congo. In: *Great Ape Societies*. Cambridge: Cambridge University Press. pp. 71-81.

Lahann, P. (2007). Feeding ecology and seed dispersal of sympatric cheirogaleid lemurs (*Microcebus murinus*, *Cheirogaleus medius*, *Cheirogaleus major*) in the littoral rainforest of south-east Madagascar. J. Zoology 271:88–98.

Langham, N.P.E. (1982). The ecology of the Common tree shrew, *Tupaia glis* in peninsular Malaysia. J. Zoology 197:323–344.

Lawes, M.J. (1991). Diet of samango monkeys (*Cercopithecus mitis erthrarchus*) in the Cape Vidal dune forest, South Africa. J. Zool. Lond. 224:149-173.

Lawes, M. J. et al. (1990). Diet and feeding behavior of samango monkeys (*Cercopithecus mitis erthrarchus*) in Ngoye Forest, South Africa. Folia Primatol. 54:57-69.

Lawrence, J. M. (2007). Understanding the pair bond in brown titi monkeys (*Callicebus brunneus*): Male and female reproductive interests. PhD Thesis. Columbia University.

Lima, E.M. & Ferrari, S.F. (2003). Diet of a Free-Ranging Group of Squirrel Monkeys (*Saimiri sciureus*) in Eastern Brazilian Amazonia. Folia Primatol 74:150–158.

MacKinnon, K. (2006). Food choice by juvenile capuchin monkeys (*Cebus capucinus*) in a tropical dry forest. In: New Perspective in the Study of Mesoamerican Primates: Distribution, Ecology, Behavior, and Conservation. New York: Springer Science+Business Media. pp. 349-365.

Mallott, E. (2016) PhD dissertation. University of Illinois, Urbana-Champaign

Matsumoto-Oda, A. (2002). Behavioral seasonality in Mahale chimpanzees. Primates 43:103–117.
Morgan, D. & Sanz, C. (2006). Chimpanzee feeding ecology and comparisons with sympatric gorillas in the Goualougo Triange, Republic of Congo. In: Feeding ecology in apes and other primates. Cambridge: Cambridge University Press. p 97–122.
Nakagawa, N. (1989). Activity budget and diet of patas monkeys in Kala Maloue National Park, Cameroon: A preliminary report. Primates 30:27–34.
Newton-Fisher, N.E. (1999). Termite eating and food sharing by male chimpanzees in the Budongo Forest, Uganda. Afr J Ecol 37:369–371.
Oates, J.F. (1977). The guereza and its food. In: Primate Ecology: Studies of Feeding and Ranging Behavior in Lemurs, Monkeys, and Apes. New York: Academic Press. pp. 275-321.
Oates, J.F. (1994). The natural history of African colobines. In: Colobine Monkeys: Their Ecology, Behavior and Evolution. Cambridge: Cambridge University Press. pp. 75-128.
Okecha, A.A. & Newton-Fisher, N.E. (2006). The Diet of Olive Baboons (*Papio anubis*) in the Budongo Forest Reserve, Uganda. In: Primates of Western Uganda. Developments in Primatology: Progress and Prospects. New York, NY: Springer. pp 61–73.
Plumptre, A.J. (2006). The Diets, Preferences, and Overlap of the Primate Community in the Budongo Forest Reserve, Uganda. In: Primates of Western Uganda. Developments in Primatology: Progress and Prospects. New York, NY: Springer. pp 345–371.
Porter, L.M. (2001). Dietary differences among sympatric Callitrichinae in northern Bolivia: *Callimico goeldii*, *Saguinus fuscicollis* and *S. labiatus*. Int. J. Primatol. 22:961–992.
Rose, L.M. (1994). Sex differences in diet and foraging behavior in white-faced capuchins (*Cebus capucinus*). Int. J. Primatol. 15:95–114.
Stone, A.I. (2007). Responses of squirrel monkeys to seasonal changes in food availability in an eastern Amazonian forest. Am. J. Primatol. 69:142–157.
Sugiyama Y. & Koman J. (1987). A preliminary list of chimpanzees' alimentation at Bossou Guinea. Primates 28:133–147.
Thorén, S. et al. (2011). Seasonal changes in feeding ecology and activity patterns of two sympatric mouse lemur species, the gray mouse lemur (*Microcebus murinus*) and the golden-brown mouse lemur (*M. ravelobensis*), in northwestern Madagascar. Int. J. Primatol. 32:566–586.
Tutin, C.E.G. et al. (1997). The primate community of the Lope Reserve, Gabon: diets, responses to fruit scarcity, and effects on biomass. Am. J. Primatol. 42:1– 24.
Vedder, A. & Fashing, P.J. (2002). Diet of a 300-member Angolan colobus monkey (*Colobus angolensis*) supergroup in the Nyungwe Forest, Rwanda. Am. J. Phys. Anthropol. 117(S34):159-160.

Wasserman MD, Chapman CA. 2003. Determinants of colobine monkey abundance: the importance of food energy, protein and fibre content. Journal of Animal Ecology 72:650–659.
Wheatley, B. P. (1980). Feeding and ranging of East Bornean *Macaca fascicularis*. In: The Macaques: Studies in Ecology, Behavior and Evolution, New York: Van Nostrand Reinhold. pp. 215-246.
Whitten, P.L. (1982). Diet and dominance among female vervet monkeys (*Cercopithecus aethiops*). Am. J. Primatol. 5:139-159.
Wich, S.A. et al. (2006). Dietary and Energetic Responses of *Pongo abelii* to Fruit Availability Fluctuations. Int. J. Primatol. 27:1535–1550.
Wrangham, R.W. & Waterman, P.G. (1981). Feeding behaviour of vervet monkeys on *Acacia tortilis* and *Acacia xanthophloea*: with special reference to reproductive strategies and tannin production. The Journal of Animal Ecology 50:715.
Wrangham, R.W. et al. (1996). Social ecology of Kanyawara chimpanzees: implications for understanding the costs of great ape groups. In: Great ape societies. Cambridge: Cambridge University Press. pp. 45–57.
Wrangham, R.W. (1977). Feeding behaviour of chimpanzees in Gombe National Park, Tanzania. In: Primate ecology: studies of feeding and ranging behaviour in lemurs, monkeys and apes. New York: Academic Press. pp. 503–538.
Wright, P.C. (1985). The costs and benefits of nocturnality for *Aotus trivirgatus* (the night monkey). Phd Thesis. City University of New York.
Yeager, C.P. (1989). Feeding ecology of the proboscis monkey (*Nasalis larvatus*). Int. J. Primatol. 10:497–530.
Yeager, C.P. (1996). Feeding ecology of the long-tailed macaque (*Macaca fascicularis*) in Kalimantan Tengah, Indonesia. Int. J. Primatol. 17:51–62.
Zeeve, S.R. (1991). Behavior and ecology of primates in the Lomako Forest, Zaire. PhD Thesis. State University of New York, Stony Brook.

Appendix D. Supplementary Table 4. Full QuantStudio 3D digital PCR results.

Assay	Target dye	Sample	Dilution	Chip file	Copies/Rxn	CI Copies/Rxn	Copies/micro liter	CI Copies/microliter
mCHIA	FAM	Saimiri sciureus	1	170718_172755_C08 AA7.eds	0.793	0.777 -- 0.81	1050.8	1029.6 -- 1072.5
RNase P	VIC	Saimiri sciureus	1	170718_172755_C08 AA7.eds	0.807	0.791 -- 0.823	1068.6	1047.1 -- 1090.6
mCHIA	FAM	Saimiri sciureus	1	170718_172702_C07 TIX.eds	0.78	0.765 -- 0.796	1033.5	1013 -- 1054.5
RNase P	VIC	Saimiri sciureus	1	170718_172702_C07 TIX.eds	0.797	0.781 -- 0.813	1055.3	1034.4 -- 1076.5
mCHIA	FAM	Sapajus apella	1	170718_173059_C08 LWC.eds	1.171	1.15 -- 1.193	1551.4	1522.7 -- 1580.8
RNase P	VIC	Sapajus apella	1	170718_173059_C08 LWC.eds	1.187	1.165 -- 1.209	1571.8	1542.7 -- 1601.5
mCHIA	FAM	Sapajus apella	1	170718_173155_C07 ER5.eds	1.215	1.193 -- 1.238	1609.7	1580 -- 1640
RNase P	VIC	Sapajus apella	1	170718_173155_C07 ER5.eds	1.235	1.212 -- 1.258	1635.7	1605.5 -- 1666.4
mCHIA	FAM	Saguinus fuscicollis	1	170718_173412_C07 Y97.eds	0.217	0.21 -- 0.224	287.37	277.89 -- 297.18
RNase P	VIC	Saguinus fuscicollis	1	170718_173412_C07 Y97.eds	0.218	0.21 -- 0.225	288.31	278.81 -- 298.13
mCHIA	FAM	Saguinus fuscicollis	1	170718_173250_C08 B1E.eds	0.228	0.22 -- 0.235	301.67	291.98 -- 311.67
RNase P	VIC	Saguinus fuscicollis	1	170718_173250_C08 B1E.eds	0.213	0.206 -- 0.22	281.92	272.6 -- 291.56
mCHIA	FAM	Callicebus moloch	1	170718_172855_C07 Y7D.eds	0.403	0.393 -- 0.414	533.78	520.1 -- 547.83

RNase P	VIC	Callicebus moloch	1	170718_172855_C07Y7D.eds	0.413	0.403 -- 0.424	547.47	533.57 -- 561.73
mCHIA	FAM	Callicebus moloch	1	170718_172949_C088E9.eds	0.453	0.442 -- 0.464	599.35	584.84 -- 614.23
RNase P	VIC	Callicebus moloch	1	170718_172949_C088E9.eds	0.461	0.449 -- 0.472	610.04	595.36 -- 625.09
mCHIA	FAM	Callithrix jacchus	1	170718_173515_C07ES4.eds	1.247	1.224 -- 1.27	1651	1621 -- 1681.6
RNase P	VIC	Callithrix jacchus	1	170718_173515_C07ES4.eds	1.192	1.17 -- 1.214	1578.4	1549.6 -- 1607.8
mCHIA	FAM	Callithrix jacchus	1	170718_173618_C082E3.eds	1.226	1.204 -- 1.249	1624.2	1594.1 -- 1654.9
RNase P	VIC	Callithrix jacchus	1	170718_173618_C082E3.eds	1.235	1.212 -- 1.259	1636.2	1605.9 -- 1667.2
mCHIA	FAM	Macaca nigra	1	170717_164333_C07XZY.eds	0.427	0.417 -- 0.438	565.95	551.81 -- 580.45
RNase P	VIC	Macaca nigra	1	170717_164333_C07XZY.eds	0.438	0.427 -- 0.449	580.14	565.78 -- 594.87
mCHIA	FAM	Macaca nigra	1	170717_164232_C07TQ3.eds	0.449	0.438 -- 0.46	594.67	580.07 -- 609.65
RNase P	VIC	Macaca nigra	1	170717_164232_C07TQ3.eds	0.455	0.444 -- 0.466	602.74	588.01 -- 617.84
mCHIA	FAM	Erythrocebus patas	1	170717_163743_C07FG3.eds	0.775	0.759 -- 0.791	1026.6	1005.3 -- 1048.3
RNase P	VIC	Erythrocebus patas	1	170717_163743_C07FG3.eds	0.65	0.636 -- 0.665	861.26	842.5 -- 880.44
mCHIA	FAM	Erythrocebus patas	1	170717_163857_C07SCG.eds	0.745	0.729 -- 0.761	986.53	966.01 -- 1007.5
RNase P	VIC	Erythrocebus patas	1	170717_163857_C07SCG.eds	0.625	0.612 -- 0.639	828.29	810.13 -- 846.85

mCHIA	FAM	Erythrocebus patas 2	1	170926_150540_C07FU3.eds	3.318	3.243 -- 3.395	4394.8	4295.1 -- 4496.9
RNase P	VIC	Erythrocebus patas 2	1	170926_150540_C07FU3.eds	3.298	3.223 -- 3.374	4367.7	4269 -- 4468.6
mCHIA	FAM	Erythrocebus patas 2	1	170926_150637_C082MP.eds	3.012	2.947 -- 3.078	3989.1	3903.5 -- 4076.7
RNase P	VIC	Erythrocebus patas 2	1	170926_150637_C082MP.eds	2.999	2.935 -- 3.065	3972.7	3887.6 -- 4059.7
mCHIA	FAM	Macaca mulatta	1	170717_164127_C07I8N.eds	0.183	0.177 -- 0.19	242.9	234.32 -- 251.8
RNase P	VIC	Macaca mulatta	1	170717_164127_C07I8N.eds	0.173	0.166 -- 0.179	228.83	220.52 -- 237.45
mCHIA	FAM	Macaca mulatta	1	170717_164022_C08KM6.eds	0.166	0.16 -- 0.173	220.35	212.16 -- 228.85
RNase P	VIC	Macaca mulatta	1	170717_164022_C08KM6.eds	0.166	0.16 -- 0.172	219.46	211.29 -- 227.94
mCHIA	FAM	Macaca nemestrina	1	170717_164508_C085SH.eds	1.043	1.023 -- 1.063	1381.3	1354.8 -- 1408.2
RNase P	VIC	Macaca nemestrina	1	170717_164508_C085SH.eds	1.066	1.045 -- 1.086	1411.3	1384.3 -- 1438.7
mCHIA	FAM	Macaca nemestrina	1	170717_164616_C08HTL.eds	1.036	1.017 -- 1.056	1372.2	1346.5 -- 1398.5
RNase P	VIC	Macaca nemestrina	1	170717_164616_C08HTL.eds	1.052	1.033 -- 1.072	1393.7	1367.6 -- 1420.2
mCHIA	FAM	Miopithecus talapoin	1	170926_150733_C07ELV.eds	1.172	1.151 -- 1.194	1552.8	1524.1 -- 1582.1
RNase P	VIC	Miopithecus talapoin	1	170926_150733_C07ELV.eds	1.184	1.162 -- 1.206	1567.8	1538.9 -- 1597.4
mCHIA	FAM	Miopithecus talapoin	1	170926_150844_C082J4.eds	1.165	1.143 -- 1.186	1542.4	1514.3 -- 1571

RNase P	VIC	Miopithecus talapoin	1	170926_150844_C082J4.eds	1.182	1.161 -- 1.204	1566.1	1537.7 -- 1595.2
hCHIA	FAM	Saimiri sciureus	1	170831_161353_C08AHG.eds	0.887	0.869 -- 0.904	1174.2	1151 -- 1197.9
RNase P	VIC	Saimiri sciureus	1	170831_161353_C08AHG.eds	0.884	0.866 -- 0.902	1170.6	1147.4 -- 1194.2
hCHIA	FAM	Saimiri sciureus	1	170831_161301_C07WL9.eds	0.802	0.786 -- 0.818	1062.2	1041.3 -- 1083.5
RNase P	VIC	Saimiri sciureus	1	170831_161301_C07WL9.eds	0.808	0.792 -- 0.824	1070.6	1049.6 -- 1092
hCHIA	FAM	Saguinus fuscicollis	1	170831_161456_C07WNV.eds	0.227	0.22 -- 0.235	301.29	291.67 -- 311.22
RNase P	VIC	Saguinus fuscicollis	1	170831_161456_C07WNV.eds	0.225	0.218 -- 0.232	297.79	288.24 -- 307.66
hCHIA	FAM	Saguinus fuscicollis	1	170831_161555_C07YAV.eds	0.246	0.239 -- 0.254	326.03	315.99 -- 336.38
RNase P	VIC	Saguinus fuscicollis	1	170831_161555_C07YAV.eds	0.241	0.234 -- 0.249	319.41	309.49 -- 329.64
hCHIA	FAM	Callithrix jacchus	1	170831_161106_C07H2K.eds	0.829	0.813 -- 0.846	1098.1	1076.2 -- 1120.4
RNase P	VIC	Callithrix jacchus	1	170831_161106_C07H2K.eds	0.829	0.813 -- 0.846	1098.6	1076.7 -- 1120.9
hCHIA	FAM	Callithrix jacchus	1	170831_161204_C07O3Q.eds	1.207	1.185 -- 1.23	1599.3	1569.8 -- 1629.2
RNase P	VIC	Callithrix jacchus	1	170831_161204_C07O3Q.eds	1.212	1.19 -- 1.235	1605.4	1575.9 -- 1635.5
hCHIA	FAM	control	1	170831_161659_C07EPI.eds	1.10E-04	2.76E-5 -- 4.41E-4	0.146	3.65E-2 -- 0.584
RNase P	VIC	control	1	170831_161659_C07EPI.eds	1.66E-04	5.34E-5 -- 5.13E-4	0.219	7.07E-2 -- 0.68

hCHIA	FAM	control	1	170831_161756_C07ER7.eds	1.18E-04	2.96E-5 -- 4.73E-4	0.157	3.92E-2 -- 0.626
RNase P	VIC	control	1	170831_161756_C07ER7.eds	5.91E-05	8.33E-6 -- 4.20E-4	7.83E-02	1.10E-2 -- 0.556