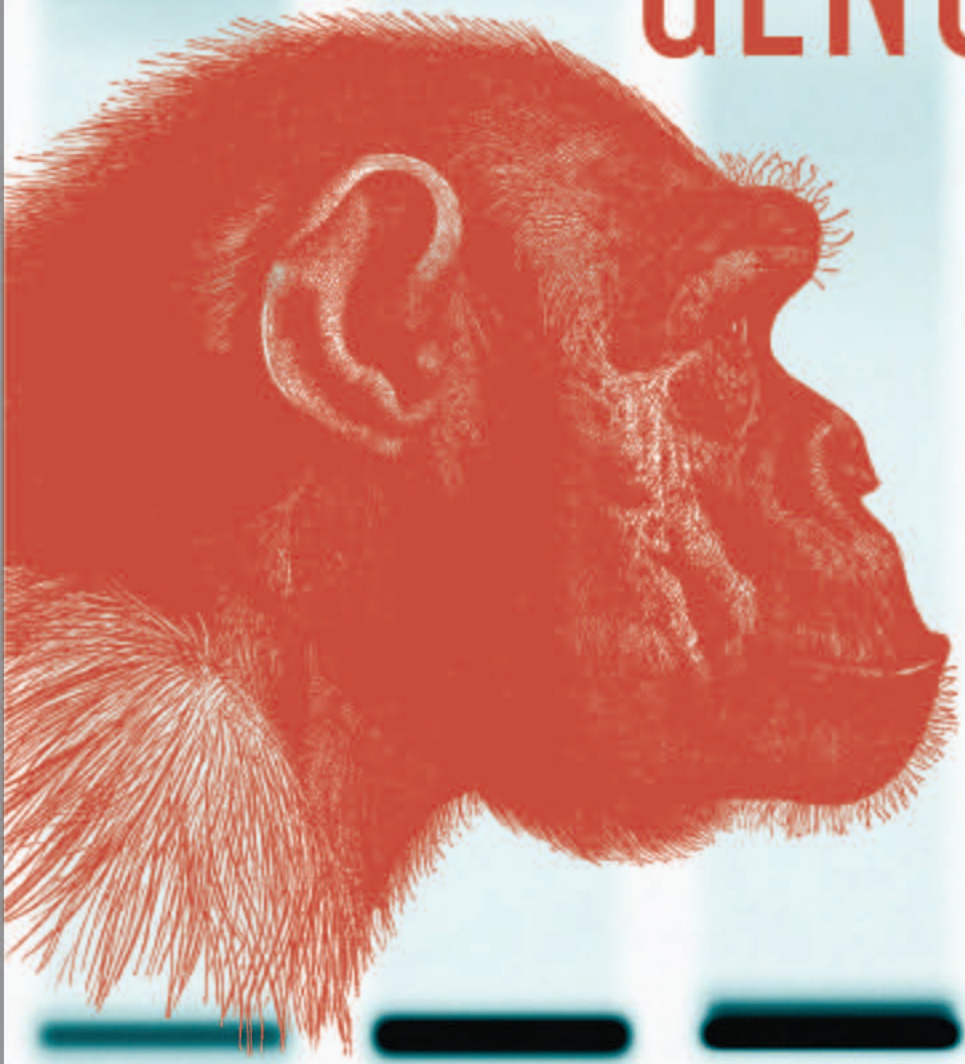
EUGENE E. HARRIS

ANCESTORS IN OUR GENOME



THE NEW SCIENCE OF HUMAN EVOLUTION

UPDATED AND EXPANDED CHAPTERS ON HUMAN ADAPTATION AND ANCIENT GENOMES

Ancestors in Our Genome

Ancestors in Our Genome

The New Science of Human Evolution

Eugene E. Harris

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide.

Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trademark of Oxford University Press
in the UK and certain other countries.

Published in the United States of America by
Oxford University Press
198 Madison Avenue, New York, NY 10016

Library of Congress Cataloging-in-Publication Data

Harris, Eugene E. (Professor), author.
Ancestors in our genome : the new science of human evolution / Eugene E. Harris.
p. ; cm.
Includes bibliographical references and index.
ISBN 978–0–19–997803–8 (alk. paper)
I. Title. [DNLM: 1. Genome, Human. 2. Biological Evolution. 3. Hominidae—genetics.
QU 460] QP624
611'.018166—dc23
2014010296

9 8 7 6 5 4 3 2 1
Printed in the United States of America
on acid-free paper

To my parents, Joan and Whitney, and my son, Bryan

CONTENTS

LIST OF FIGURES AND TABLES

FIGURES

TABLE

ACKNOWLEDGMENTS

I thank the following colleagues for their generous time in commenting on chapters: Drs. Günter Bräuer, Henry Harpending, Jody Hey, Carles Lalueza-Fox, Diogo Meyer, Maryellen Ruvolo, Caro-Beth Stewart, as well as two anonymous reviewers. Rebecca Ham deserves very special thanks for her extremely thorough editorial advice, organization suggestions, and encouragement. All illustrations were rendered with the expert and very gracious help of Kei Hayashi. Drs. Sarah Tishkoff, Hopi Hoekstra, Günter Bräuer, Michael Day, and John Fleagle kindly allowed me to use research photos of theirs. Tim White and David Brill allowed me to use photos of fossil skulls. Stephen Nash generously allowed me to use his original illustrations of papionin monkeys. I thank Jeremy Lewis, Erik Hane and Michael Stein at Oxford University Press for their generous editorial advice. An extra special thanks goes to my agent Deirdre Mullane for her sound editorial advice and her steadfast and generous support in guiding this project along. I also owe a debt of gratitude to my previous mentors Terry Harrison, Clifford Jolly, Todd Disotell and Jody Hey. I also thank Dean Falk for her helpful advice and support, Lisa Schlotterhausen for her editorial suggestions, as well as my colleagues Avelin Malyango, Varsha Pilbrow, and Scott Sherman. Any inaccuracies or mistakes in interpretation of other people's work are my own.

PROLOGUE

We are now in an era of enormous potential for studies of our evolutionary past. With the determination of the full human genome sequence in 2001—the culmination of a scientific quest begun almost fifty years earlier when James Watson and Francis Crick first discovered the molecular structure of DNA—we are at the beginning of a genomic voyage back in time. The pace at which full genomes of our primate relatives are also being sequenced is exhilarating; genomes of the common chimpanzee, bonobo, gorilla, orangutan, and macaque monkey have already been fully determined and that of other primates is well under way. Armed with sophisticated new tools, researchers are starting to examine variations in our genome among diverse peoples of the world and compare it with those of our close primate relatives in order to answer age-old questions about where, when, and how we evolved. For the first time, we are seeing our ancestors in our genome, obtaining new and often astonishing views of our evolutionary past, and we are already beginning to identify genomic features in which humans are similar to other primates and features in which humans are unique. Numerous large-scale studies have also started to catalogue millions of differences in DNA among individuals from around the world, providing us with more finely detailed knowledge of genetic diversity within our own species. Among other remarkable insights, genomic analyses are enabling us to identify with great certainty evolutionary relationships among our ape cousins; to estimate more precisely the time and nature of the evolutionary process that produced the human lineage; to identify the genetic bases of our species' adaptations, such as increased brain size and language; and to determine when and by which genetic mechanisms human populations adapted to different environments around the world. In short, the unprecedented scale of genomic evidence now being collected is revolutionizing how and what we can learn about our origins.

The source of all this information is found in the nucleotides of the human genome, the strands of molecules that join together in varying but precise ways to form the ladderlike steps of hereditary information encoded in the helical structure of our DNA. Although each person's genome is unique, reflecting the individual biological blueprints we inherit from our parents, the basic structure of the genome is similar across our species. (This is the reason why some scientists refer to *the* human genome, even though this is not technically accurate.) The entire human genome is somewhat greater than three billion nucleotides in length, organized into segments that combine to form genes that code for the proteins in the human body and influence the development of our every physical feature, from eye color to blood type. Like pearls strung along a DNA necklace, genes can be of different lengths, ranging from a few hundred nucleotides for the smallest genes to a few million nucleotides for the longest.

A gene's A, C, G, and T nucleotides provide the instructions to our body about how to construct its different proteins. Proteins, of course, are macromolecules that play innumerable important functions and are the workhorses of our bodies. A protein like collagen plays a structural role in making our bones, ligaments, and tendons strong; the protein hemoglobin in our red blood cells helps transport oxygen and carbon dioxide in and out of our tissues; and the protein fibrin is essential for normal blood clotting. Proteins themselves are actually composed of chains of little building block molecules called amino acids—of which we have twenty different types—and each different protein has a unique chain of these amino acids. The type and order of DNA nucleotides in a gene provides the instructions on how to put these amino acids together.

Perhaps surprisingly, however, genes make up only about 1.5% percent of the entire human genome. This means that only a very small fraction of our vast genome is actually coding for the proteins that carry out important structural and functional roles that make our body work. The remainder of the genome was long considered to be almost entirely "junk DNA" (i.e. non-functional), but more recent studies have estimated that as much as 4% to 12% of this remaining DNA also has important functions and perhaps plays important roles in controlling the expression of genes.[1,2] For example, many short elements within the genome have been identified that can turn genes on and off, or can increase and decrease the amounts of protein genes produce. These gene switches are responsible for activating different sets of genes in different anatomical regions of our body, which helps to explain why our brains have very different characteristics from our livers.

The sequence of a genome is determined through a complex laboratory process using increasingly efficient technologies to resolve the precise order

of the nucleotide bases that pair together to form the steps of the ladder within our DNA. Since a human genome is so large, its sequence needs to be deciphered in small segments, which are then assembled together using sophisticated computer programs to produce the entire genomic sequence. One significant finding from initial analyses of our genome is that the number of genes it contains is far fewer than we had long believed. Instead of the 50,000 to 100,000 genes suggested in genetics textbooks as recently as fifteen years ago, the human genome is now considered to be made up of approximately 21,000 genes.

A FRESH LOOK AT OLD QUESTIONS ABOUT OUR PAST

But even as old mysteries are finally solved, the new DNA evidence leads us forward by causing us to revisit old and persistent questions about our evolutionary origins, as well as raising important new ones, many of which we would never have thought to ask. One of the greatest events in all of evolution—from a human-centric standpoint—is the one that led to the divergence of the human lineage from our great ape cousins—the chimpanzees, gorillas, and orangutans. Within these pages, I'll tell the story of the genetic quest, from small stretches of DNA to entire genomes, to trace our past to the origin of our lineage and find our closest ape relative. We now know, of course, that chimpanzees (and their close cousins, the bonobos) are our closest living relatives. Along the way, we'll discover some surprising aspects of our genome that show our deep evolutionary connectedness with all the apes, not just chimpanzees. These studies reveal to us that our genome is like a patchwork quilt, with new pieces added over the course of generations—a genome with segments that were picked up at different stages of our ancestry.

As recently as twenty-five years ago, most anthropologists believed our evolutionary separation from the apes occurred very deeply in the past, all the way back to fifteen million years ago, with the first human ancestor being an extinct robust-jawed and large-molared ape from northern India named *Ramapithecus*. Today, genetic analyses can make far more accurate estimates about the origin of our lineage, pointing to much more recent times for our separation from chimpanzees.

Another long-standing question concerns the evolutionary forces that led to our reproductive isolation from the ancestors we once shared with chimpanzees. Various theories have been put forward as answers, including the traditional and widely believed "savanna hypothesis" initially advanced in 1879 by Charles Darwin in his *Descent of Man* and later promulgated

by the paleoanthropologist Raymond Dart, who in 1924 made the first discovery of a human fossil in Africa—the Taung skull—formally named *Australopithecus africanus*. The savanna hypothesis and its variants suggest that a progressive warming and drying of the environment led to an expansion of the African savanna, prompting some early ape species to leave their forested habitats behind. On the savannas, these species developed bipedalism, eventually becoming a new species in the earliest beginnings of the human lineage, while their ancestors remained in the forests. Such geographic separation led to relatively quick reproductive isolation and speciation of early humans from their ancestors. Other hypotheses do not emphasize such clear geographic separation, suggesting that the earliest human ancestors still lived in very nearly the same forested habitats that their ancestors had always lived in and therefore geographic isolation was not the prime mover of the emergence of the human lineage. Researchers have now started comparing genomes of humans and chimpanzees to begin to evaluate the likelihood of these hypotheses and whether this ancient evolutionary split resembled a short and quickie divorce, or more like a long and drawn-out affair with mating between the two emerging species.

Throughout our quest in the last quarter of the twentieth century to determine the precise evolutionary relationship among the apes, a persistent debate—aptly captured by the phrase "molecules versus morphology"—was what evidence was best to use. Would we obtain the most accurate reconstruction of these relationships if we compared many anatomical features of skeletons and skulls among different species, or would comparisons of the biological molecules of DNA provide the most reliable evidence? Since the 1990s, this debate has subsided in favor of DNA, and I will explore the rationale that justifies using DNA and now full genomes to build very accurate reconstructions of species' relationships. Like opening a cloudy and jammed window to let in both light and fresh air, this new perspective has revealed that outward appearances and even the smallest of anatomical details of different species can deceive and even mislead us, and it continues to help us obtain more accurate understandings of the complex processes by which anatomical structures of organisms can evolve.

Ever since Darwin, we have wanted to understand how and why our unique features, which provided distinct advantages on our evolutionary journey since the separation from our common ancestor with chimpanzees, have evolved. Shaped by the mechanism Darwin described as natural selection, these human adaptations increased our ability to survive and enhanced our capability to reproduce and successfully raise our offspring. But what are the adaptations that separate us from chimpanzees, how many of these adaptations do we have, and where in the genome are they found? For example,

increased cognition and language almost certainly result from changes that occurred in numerous parts of our genome, but we don't know in how many regions, where they are located, and how these changes led to our ability for complex thought and communication. Beyond our species-wide adaptations, which all humans possess, there are also adaptations unique to peoples living in different parts of the world. These adaptations evolved as modern humans spread out from the site of their evolutionary origin and settled in different regions, encountering different environmental conditions, eating different foods, and facing different pathogens.

Having entire genomes to work with allows researchers to voyage across their vast uncharted nucleotide bases searching for locations that bear signs of having been shaped by the forces of natural selection, and that therefore might represent regions underlying our unique features. Once discovered, such regions become a point of departure for further research into determining exactly how these regions function and if they indeed contributed to our adaptations. After all, for most of the genome, and even for most of our genes, we still have only a relatively simple appreciation for how they influenced our biological features.

DNA analysis is also shedding light on another vexing evolutionary riddle. Ever since the first discoveries of Neandertals in the mid-1800s, there has been a preoccupation with their evolutionary fate. This mystery has grown deeper with time, especially with the introduction of "out of Africa" models of human evolution, from the mid-1980s into the 1990s, which in their strictest form suggested that anatomically modern humans originated in Africa and subsequently spread to Asia and Europe, replacing all archaic forms, including the peculiar Neandertals, and relegating them to the evolutionary dustbin.

In the early 21st century analyses of full nuclear genomes from diverse peoples from around the world overwhelmingly indicate that the evolutionary origin of modern humans was in Africa and that we subsequently dispersed into Europe, Asia, and eventually the Americas. But the theory is now decidedly shorn of the notion that these archaic hominins living in Europe and Asia were completely replaced by newly dispersing modern humans. One of the greatest triumphs in recent anthropological studies has been the recovery of ancient DNA from fossils of extinct relatives. The determination of the full nuclear genomes of Neandertals and of an ancient Denisovan of Asia, a previously unknown contemporary of Neandertals, confirms that modern humans interbred with both of these archaic human relatives. Accumulating genomic evidence presents a fundamental challenge to our previous views that the origin of our modern species in Africa was recent, quick, and regionally restricted.

Many questions still remain about the evolution of modern humans in Africa, however, and the vast size of the genome is providing researchers with ample evidence for rigorous testing of previous hypotheses about human origins. With the help of genomic analysis, we are beginning to have a fairly detailed understanding of our place in the primate evolutionary tree and how other primate species are related to one another. This knowledge provides the essential evolutionary framework by which we can trace our adaptations back in time to learn when on the primate tree they first started to emerge. While certain human features will likely prove to have uniquely evolved along the human lineage, many others presumed to be unique will have deeper origins within our shared past with other primates. We already know that a number of genes that show signatures of adaptation within the time frame of human evolutionary history (for example, several genes associated with our increased brain size) also show signatures of adaptation on the common ancestral stem we share with apes. Charles Darwin himself appreciated this early on, suggesting that "the difference in mind between man and the higher animals, great as it is, certainly is one of degree and not of kind.[3] Results from the analyses of the genomes of many other nonhuman primates, and even distantly related animals, will likely be a lesson in humility, revealing our deep connection to the rest of the animal world.

My aim in writing this book is not to provide definitive answers to the persistent questions we have about our evolutionary origins—we can only hope to approximate answers to these questions more accurately—but to introduce the reader to the ways in which genomes can be used to begin to reexamine old questions with new evidence. Our evolutionary past is a puzzle with disparate and fragmentary pieces remaining—including fossils, cultural artifacts, and now genomes—that ultimately all need to fit together to give us a coherent and internally consistent view of our history. As we shall see, genomes are providing us with a powerful new tool that we can use to bring the puzzle much more clearly into view.

Ancestors in Our Genome

CHAPTER 1

Looks Can Be Deceiving

When I was in graduate school in physical anthropology in the 1990s, there was a war around me—well, let's say a battle—being waged between those of us who studied the anatomy of bones and fossils and those who studied genetics. We even occupied different floors in our department, which only further heightened the divide. So who had the high ground? Well, from my perspective, it was the ninth-floor morphologists like myself who were studying anatomy. By studying the anatomy of fossils from extinct primates, some of them our close relatives, we could determine if they ate fruits, leaves, or grass seeds. We could tell if they climbed in trees or walked on the ground, how much they weighed, and how big their brains were. We also felt we had the upper hand in determining the evolutionary relationships of primates—the details of the skeleton could tell us more about the kinship among different species than any single gene. Plus, on the ninth floor we had nice sunny views of lower Manhattan, which added a little to our sense of self-importance!

Genetics, on the other hand, could tell us nothing about how a primate moved around in its habitat, how much it weighed, or what it ate. So what was it good for? What could it reveal about our evolutionary history? The debate between morphology and genetics played out on a larger stage at the annual meetings of the American Association of Physical Anthropologists and in our professional journals. The main contentious issue centered on whether genetics or morphology was better for determining the evolutionary relationships of the primates. A prominent book published at that time, *Molecules versus Morphology: Conflict or Compromise?* by Colin Patterson,[1] polarized the debate and helped fuel the arguments.

For relief from the battle, I decided to go fishing. Not fishing in the literal sense, though I did escape from reading and measuring bones occasionally to visit the shore. I went fishing for a dissertation project. My main interests at the time were in studying the morphological diversity of primates and humans to reconstruct how extinct primates moved, using evidence gleaned from their fossilized remains. I became especially interested in studying anatomy in order to reveal the evolutionary relationships among one group of our African relatives, the papionin monkeys, which includes the long-faced and terrestrial baboons, mandrills, and their relatives. An accurate evolutionary tree for this group was in doubt because morphological evidence was pointing in one direction and some new genetic evidence was pointing in another. Since bones and muscles were all I knew at the time, I planned to do a thorough analysis of the anatomy of this group of monkeys. The way in which my research played out, and my personal ideological evolution during that time, revealed several very important lessons directly applicable to our understanding of the evolution of great apes and humans.

ENTER THE MONKEYS

Monkeys and apes from Africa and Asia are grouped together due to the fact that they share derived and newly evolved details of their ear anatomy, the evolutionary loss of a premolar tooth as well as many genetic aspects. But Old World monkeys and apes separated from one another about twenty-four million years ago. The papionin monkeys are a subgroup of these monkeys that began to differentiate into different species at about the same time that the hominoid apes— chimpanzees, gorillas, and humans—were splitting into different species. No doubt the two groups of primates came into contact on the African plains, and very likely sometimes on less than friendly terms. There is archeological evidence from the site of Olorgesailie in Kenya that our human ancestors, most likely *Homo erectus*, butchered and ate (somewhere in the time frame of 700,000 and 400,000 years ago) a now extinct giant relative of the living papionin monkey called the gelada.[2]

The African papionin monkeys include three large-bodied, long-faced, and very terrestrial members (the geladas, mandrills, and baboons). It also includes two smaller-bodied and shorter-faced monkeys, known as the mangabeys, which are partially terrestrial to arboreal in their habits. Up until the early 1990s, morphological studies had concluded that the larger-bodied and long-faced monkeys were the closest evolutionary