Praise for the book

'If you really want to know who we are, and how we got here, this is the book for you. A thrilling account of our extraordinary past – I couldn't put it down.' **Gurcharan Das**

'An amazing book, written in an engaging style . . . It grabs your interest from the first sentence.' **Bibek Debroy**

'At a time when the issue of the peopling of the planet by anatomically modern humans is becoming hotter and hotter, Tony Joseph has admirably cooled it down. He goes deep into recent developments in ancient DNA studies on human remains and delves into the issue of the "Aryan Migration". This is a vexing issue and not so easy to conclude. However, Joseph has marshalled multidisciplinary data from archaeology, linguistics, genetics and literature to support his stand on the issue. This is perhaps the most scientific way of presenting the Aryan debate. Lucidity is the hallmark of this book.' **Ravi Korisettar**

'Intellectual omnivore Tony Joseph offers an enjoyable meander through a minefield of how our ancestors got here.' **Pranay Lal**

'Masterful and unbiased reconstruction of human presence in India using evidence from archaeology, ancient and modern history, linguistics, geography and genetics, with a tilt on genetic evidence.' **Partha P. Majumder**

'Joseph deftly and brilliantly summarizes new findings of genetics that definitively solve old problems in South Asian history, and show we are all migrants and, ultimately, kin. A timely, fascinating and courageous book.' **Sheldon Pollock**

'There has been a lot of controversy about the origins of various populations, and in India, much of this is driven by a quasi-religious ideology. It is therefore refreshing to see how recent advances in DNA sequencing from people of various ethnicities as well as remains of ancient people is shedding light on the origins, migration and intermixing of people throughout history. In this very readable account, Tony Joseph has distilled the results of recent research and his book should be of interest to anyone curious about the waves of migration and intermixing that resulted in the rich tapestry that makes up the people of today's India.' **Venki Ramakrishnan**

'DNA studies of Indians dating to the millennia BCE confirm that they were a mixed population, and at a particular time included migrants from Central Asia. Given that these are initial studies their readings require circumspection when equating DNA identities with those from other sources. Tony Joseph's perceptive summary suggests how this new information might help clarify some of our understanding of the early past.' **Romila Thapar**

'Tony Joseph's book provides a remarkably accessible overview of the early stages of ancient Indian history, starting with the immigration from Africa of current humans to the age of the Vedas. He provides evidence from several fields of scientific enquiry, notably archaeology, linguistics, ancient texts and the very recent study of ancient genes (aDNA). The latter is currently revolutionizing ancient history not just of India but also of Europe, Africa and South America. Accordingly, T. Joseph lays to rest the question about the origins of the so-called (Indo-)Aryans and their settlement in ancient India – which has basically been politically motivated, especially for the past 40 years. As common in scholarship, not all individual scholars may agree on *all* questions and conclusions (such as the nature of the Indus civilization and its relation with the origin of the Dravidian speakers). However, finally, a firm basis for writing the history of ancient India is laid. The various sciences, in the end, lead us from darkness to the light of insight.' **Michael Witzel**

'Joseph has told an extremely complex tale very lucidly. It's quite remarkable . . . India could have faced a drought of accessible writing on academic questions after the generation represented by Jayant Narlikar, Irfan Habib and Romila Thapar, but new arrivals like . . . Joseph's exploration of the humanscape, offer considerable promise. Read *Early Indians* soon.' **Pratik Kanjilal, *Indian Express***

'A lucid and enthralling narrative that combines forensic examination with the compelling twists and turns of an unfolding mystery.' **Sunil Sethi, *Business Standard***

'An astonishing tale, difficult to put down . . . This book is excellent science journalism, the kind that we need more of in other disciplines.' **Sujatha Byravan, *The Hindu***

'A book of such national importance . . . Joseph deserves our gratitude for writing a masterpiece with such clarity and lucidity, despite numerous opportunities for

confusion, and a flair that renders the stories of our roots exciting . . . every Indian should read the book.' **Jay Desai, Wire**

'The question "Who are we?" has been answered brilliantly.' **Swaminathan A. Aiyar, *Economic Times***

'Tony Joseph has given us a book that, with its racy style, is easy to read. By proposing puzzles, he makes dull evidence lively for the reader. He simultaneously maintains a level of academic accuracy that adds weight to his hypotheses and conclusions. This is a book which all who are interested in both history and historical method should read – and enjoy.' **Irfan Habib, *Studies in People's History***

'Remarkably accessible to the reader, dense as it is with evidence from multiple branches of knowledge such as archaeology, linguistics, ancient texts and, most notably, the recent study of ancient genes (aDNA) . . . here is a firm basis on which the study of Indian history can begin.' **Kesavan Veluthat, *Frontline***

'Joseph has presented new findings from the fields of archaeology, anthropology, linguistics and genetics (including ancient DNA) and very precisely describes the complete story in four chapters.' **K. Thangaraj and L. Kumar, *Current Science***

'[A] fascinating history of Indians as a people . . . Joseph writes with an infectious enthusiasm . . . a readable overview of India's early history as well as how advances in human genetics – but also linguistics and other sciences – have revolutionized our understanding of pre-history.' **Peter Gordon, Asian Review of Books**

'There is so much going on in DNA studies . . . that having a solid guide to stitch it all together, including papers that landed with a giant thud in 2018, would be so very, very nice . . . Tony Joseph's *Early Indians* fits the bill mightily. It brings together many lines of research centering around ancient and modern DNA studies in a lucid, engaging way that makes complex concepts easier to understand . . . Tony Joseph has provided a wonderfully engaging and sophisticated lens to start understanding and absorbing all this information through.' **Harappa.com**

'Does an excellent job presenting a clear and concise narrative of the origins of early Indians.' **Razib Khan, *India Today***

'Important, accessible and thought-provoking . . . Joseph has done an excellent job . . . opens the doorways of prehistory to whole new audiences . . . groundbreaking.' **Anirudh Kanisetti,** ***Pragati***

'Brilliantly sums up the recent genetic research and rightly characterizes the Indian Civilization as a "multi-source civilization".' **Ashish Kumar,** ***Telegraph***

Early Indians

Early Indians

The Story of Our Ancestors and Where We Came From

Tony Joseph

juggernaut

JUGGERNAUT BOOKS
C-I-128, First Floor, Sangam Vihar, Near Holi Chowk,
New Delhi 110080, India

First published in hardback by Juggernaut Books 2018
Published in paperback 2021

10 9 8 7 6 5 4 3 2 1

P-ISBN: 9789391165956
E-ISBN: 9789391165260

The international boundaries on the maps of India are neither purported to be correct nor authentic by Survey of India directives.

Passages quoted from *Who We Are and How We Got Here: Ancient DNA and the New Science of the Human Past* (2018) by David Reich are reprinted by permission of Oxford University Press.

Passages quoted from *Greater Magadha: Studies in the Culture of Early India* (2007) by Johannes Bronkhorst are reprinted by permission of the author.

While every effort has been made to trace copyright holders and obtain permission, this has not been possible in all cases; any omissions brought to our attention will be remedied in future editions.

The views and opinions expressed in this book are the author's own. The facts contained herein were reported to be true as on the date of publication by the author to the publishers of the book, and the publishers are not in any way liable for their accuracy or veracity.

Typeset in Adobe Caslon Pro by R. Ajith Kumar, Noida

Printed and bound at Nutech Print Services - India

To
My parents, for everything
My wife, who made this book possible
My daughter, who is the reason I wrote it

Contents

A Short Chronology of the Modern Human in Indian Prehistory xiii

Introduction 1
1. The First Indians 16
2. The First Farmers 64
3. The First Urbanites: The Harappans 104
4. The Last Migrants: The 'Arya' 167
Epilogue 215

Appendix 235
Summary 243
Frequently Asked Questions 264
Bibliography 277
Acknowledgements 288
A Note on the Author 294
Index 295

A Short Chronology of the Modern Human in Indian Prehistory

~ 300,000 years: The age of the earliest remains of a modern human, *Homo sapiens*, ever found – in a cave in Jebel Irhoud, about fifty kilometres from the city of Safi in Morocco.

~ 210,000 years: The age of the earliest modern human fossil found outside of Africa – in the Apidima Cave of southern Greece.

~ 70,000 years ago: Geneticists calculate that the earliest successful Out of Africa (OoA) migration happened around this time. This migration was termed 'successful' because these migrants are the ancestors of all of today's non-African populations. (Earlier modern humans outside of Africa have not left a lineage that is detectable today.) The OoA migrants 70,000 years ago are likely to have taken the Southern Route that would have brought them from Africa (specifically, from modern-day Eritrea and Djibouti) into Asia (modern-day Yemen) through Bab el Mandeb at the southern tip of the Red Sea.

~ 65,000 years ago: The OoA migrants reach India and are faced with a robust population of archaic humans. They perhaps take both an inland sub-Himalayan route and a coastal route, to keep themselves out of the way of other *Homo* species in the subcontinent who dominated central

and southern India, and then move across the Indian subcontinent into south-east Asia, east Asia and Australia.

60,000–40,000 years ago: The descendants of the OoA migrants populate central Asia and Europe over this period.

~ 40,000 years ago: Neanderthals go extinct in Europe, with the Iberian peninsula in south-western Europe (modern-day Portugal and Spain) being their last refuge and stand.

45,000–20,000 years ago: The First Indians, the descendants of the OoA migrants in the subcontinent, start using Microlithic technology, and their population increases dramatically in central and eastern India. South Asia becomes the place where 'most of humanity' lives. Modern humans move into what would have been long-established refuges of other *Homo* species in southern and central India.

~ 16,000 years ago (14,000 BCE): Modern humans reach the Americas, the last major continent to be settled in by modern humans, after crossing Beringia, the land bridge between Siberia and Alaska.

~ 7000 BCE: In a village that is today called Mehrgarh, at the foot of the Bolan Hills in Balochistan, a new agricultural settlement begins that would ultimately become one of the largest habitations of its period between the Indus and the Mediterranean.

7000–2600 BCE: The Mehrgarh site shows evidence for cultivation of barley and wheat, and increasing consumption of domesticated animals. The site was abandoned somewhere between 2600 BCE and 2000 BCE. By then agricultural settlements had spread all across north-western India – in the Indus and Ghaggar–Hakra river valleys and in Gujarat.

7000 BCE: From around this period there is evidence for rice harvesting and sedentary settlement at Lahuradewa in the Sant Kabir Nagar district of Uttar Pradesh in the Upper Ganga plain. The chronology of transition from harvesting wild rice to cultivating domesticated

rice is not yet certain, but there is no doubt that Lahuradewa indicates experiments in agriculture were taking place at several places in south Asia around the same time and that Mehrgarh was not an isolated case.

5400–3700 BCE: Genetic studies suggest that a population related to the early farmers and herders of Iran and its Zagros mountain region mixed with the descendants of the First Indians at least by this period.

5500–2600 BCE: The Early Harappan era, which witnesses early agricultural settlements growing into towns with their own unique styles, such as Kalibangan and Rakhigarhi in India and Rahman Dheri in Pakistan.

3700–1500 BCE: Evidence of early agriculture starts to appear in different parts of India – eastern Rajasthan, southern India, the Vindhya region of central India, eastern India and the Swat valley of Kashmir.

2600–1900 BCE: The Mature Harappan period, which sees many sites being newly built or rebuilt, and many existing sites being abandoned. There is also a visible and higher level of standardization across the region, with a common script, seals, motifs and weights. The transition from the Early Harappan to the Mature Harappan phase happened over four or five generations, or 100 to 150 years.

2300–1700 BCE: The period of the Bactria–Margiana Archaeological Complex (BMAC), a civilization centred on the Oxus river (also called Amu Darya) and covering today's northern Afghanistan, southern Uzbekistan and western Tajikistan. The BMAC had close trade and cultural relations with the Harappan Civilization.

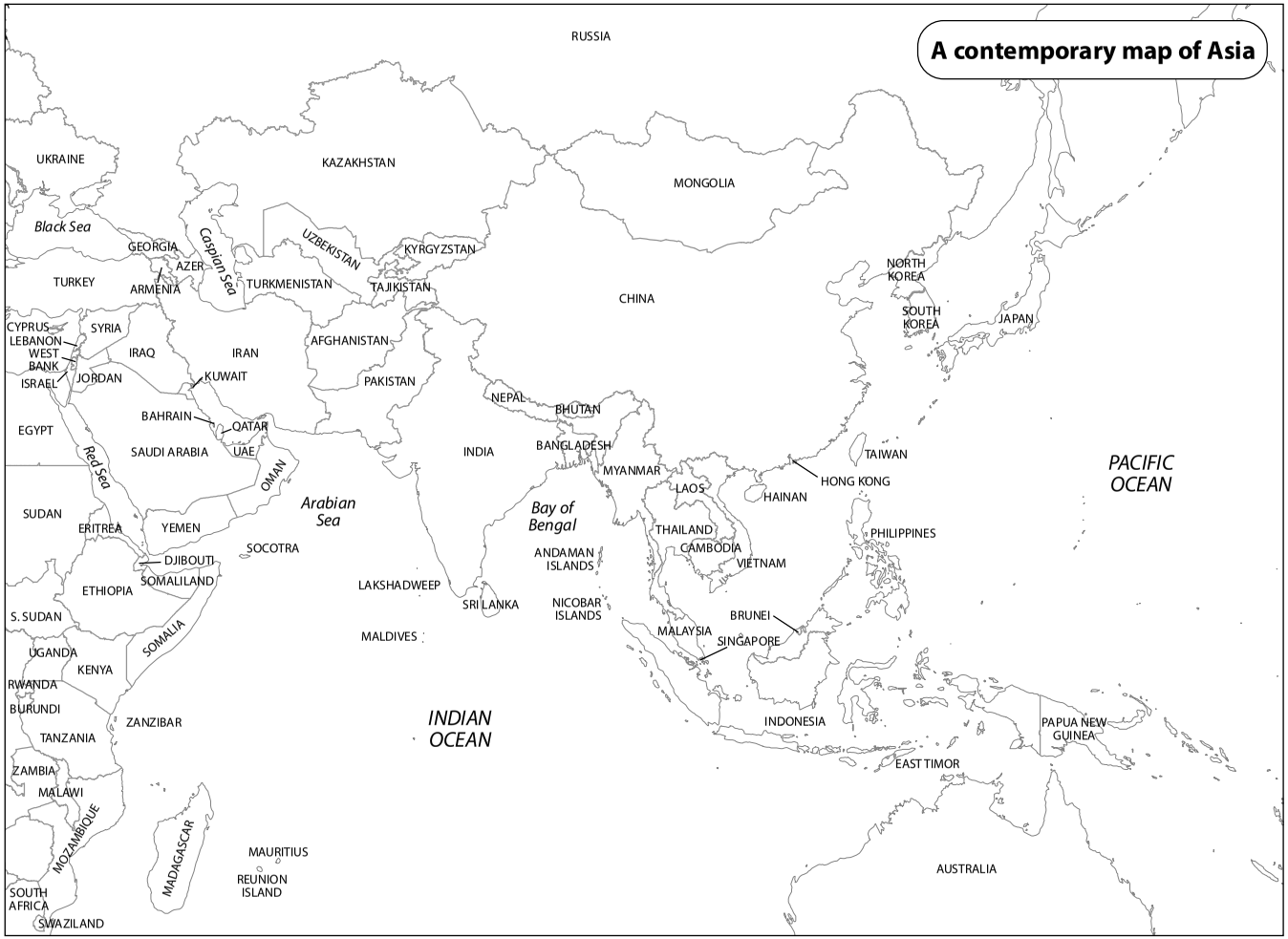
2100 BCE: A southward migration of pastoralists from the Kazakh Steppe, towards the southern central Asian regions that would today be called Turkmenistan, Uzbekistan and Tajikistan. The migrants make an impact on the BMAC, but mostly bypass it and move towards

south Asia during the first half of the second millennium BCE, as listed below (2000–1500 BCE).

2000 BCE: Two major waves of migrations with their origin in China – after it had gone through the farming revolution and the resultant population surge – reshape south-east Asia. One of these reach India, bringing Austroasiatic languages, new plants and a new variety of rice to India after 2000 BCE.

2000–1500 BCE: Multiple waves of Steppe pastoralist migrants from central Asia reach south Asia, bringing Indo-European languages and new religious and cultural practices.

1900–1300 BCE: The Late Harappan period that sees the decline and eventual disappearance of the Harappan Civilization, primarily due to the effects of a long drought that affected civilizations in west Asia, Egypt and China as well.


A contemporary map of Asia
RUSSIA
UKRAINE
KAZAKHSTAN
MONGOLIA
Black Sea
Caspian Sea
GEORGIA
AZER
ARMENIA
UZBEKISTAN
KYRGYZSTAN
TURKMENISTAN
TAJIKISTAN
TURKEY
CHINA
NORTH KOREA
SOUTH KOREA
JAPAN
CYPRUS
LEBANON
WEST BANK
ISRAEL
SYRIA
IRAQ
IRAN
AFGHANISTAN
JORDAN
KUWAIT
PAKISTAN
NEPAL
BHUTAN
BAHRAIN
QATAR
UAE
EGYPT
SAUDI ARABIA
OMAN
Red Sea
INDIA
BANGLADESH
MYANMAR
TAIWAN
HONG KONG
HAINAN
LAOS
PACIFIC OCEAN
Arabian Sea
Bay of Bengal
SUDAN
ERITREA
YEMEN
SOCOTRA
DJIBOUTI
SOMALILAND
ETHIOPIA
THAILAND
CAMBODIA
VIETNAM
PHILIPPINES
ANDAMAN ISLANDS
LAKSHADWEEP
SRI LANKA
NICOBAR ISLANDS
S. SUDAN
SOMALIA
MALDIVES
BRUNEI
MALAYSIA
SINGAPORE
UGANDA
KENYA
RWANDA
BURUNDI
ZANZIBAR
INDIAN OCEAN
INDONESIA
PAPUA NEW GUINEA
TANZANIA
EAST TIMOR
ZAMBIA
MALAWI
MOZAMBIQUE
MADAGASCAR
MAURITIUS
REUNION ISLAND
AUSTRALIA
SOUTH AFRICA
SWAZILAND

Introduction

How We, the Indians, Came to Be

The story of our ancestors, the early Indians, who came from Africa, west Asia, east Asia and central Asia and made this land theirs over the last 65,000 years.

'But have you ever considered how fast you are really moving when it seems you are not moving at all?'

Professor Andrew Fraknoi, astronomer

Things are often not what they seem. As you read this sentence, perhaps sitting in a comfortable chair in your study, you would probably consider yourself at rest. But you are really not, because the Milky Way galaxy of which you are a part is moving through space at 2.1 million kilometres an hour. And that is without taking into account the effects of the earth's rotation on its own axis (1600 kilometres an hour at the equator and zero at the poles), its orbiting around the sun (107,000 kilometres an hour) and the sun's journey around the Milky Way (792,000 kilometres an hour).[1]

[1] Andrew Fraknoi, 'How Fast Are You Moving When Sitting Still', in *Universe in the Classroom* (Astronomical Society of the Pacific, 2007).

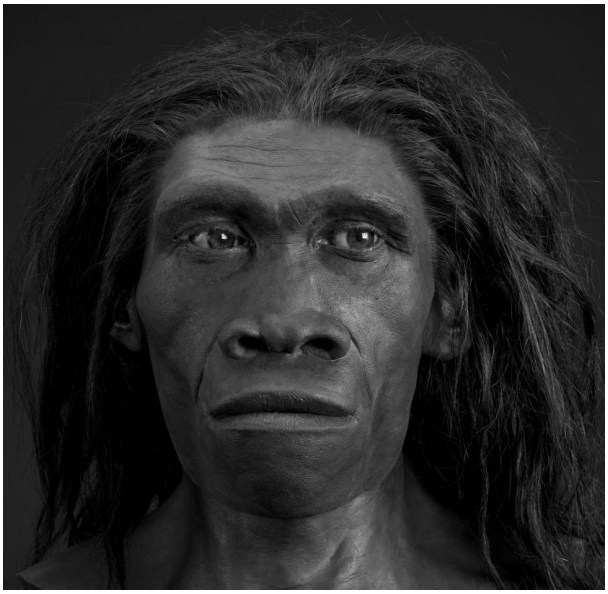
So in the roughly twenty seconds that it would have taken you to read the paragraph above, you have already moved thousands of kilometres without even knowing it!

Each successive discovery that led to the calculations above – that the earth is just one of many planets circling the sun; that the sun is just an average, middle-aged star in the Milky Way galaxy; that the galaxy itself is just one of at least a hundred billion galaxies – made some humans feel a little smaller while the wiser ones felt a new sense of awe at the size and majesty of what we are all a part of.

And this is true not just in a cosmic sense, but in a biological sense as well. Ever since Darwin shocked humanity a century and a half ago by formulating the theory of evolution and suggesting that our closest living relatives could be chimpanzees, every subsequent discovery has gone on to destroy the special status we had generously given ourselves previously. First, we thought that when we, the modern humans or *Homo sapiens*,[2] arrived on the scene, there was a sudden and appreciable difference in the kind of tools that were being made, as well as an efflorescence of artistry and abstract thought. Now we know that all that was conceit, and that the tools made by us and those made by our closest evolutionary cousins – *Homo erectus*, *Homo neanderthalensis*, Denisovans – were often indistinguishable from each other and that there was no watershed moment. All these extinct members of the *Homo* species (*Homo sapiens* being the only surviving member of the *Homo* family today) also had large brains like us.

In the past decade, we even learned that they were close enough to *Homo sapiens* genetically for us to have mated with them and

[2] Modern humans and *Homo sapiens* are used synonymously throughout this book. Humans, by contrast, could mean any member of the *Homo* species, such as *Homo habilis*, *Homo erectus*, *Homo neanderthalensis* or *Homo sapiens*. Archaic humans refers to those members of the *Homo* species that are extinct – such as *Homo habilis*, *Homo erectus* or *Homo neanderthalensis*. However, in the Holocene (from ~9700 BCE onward), 'humans' will mean modern humans since archaic humans are believed to have gone extinct by then.

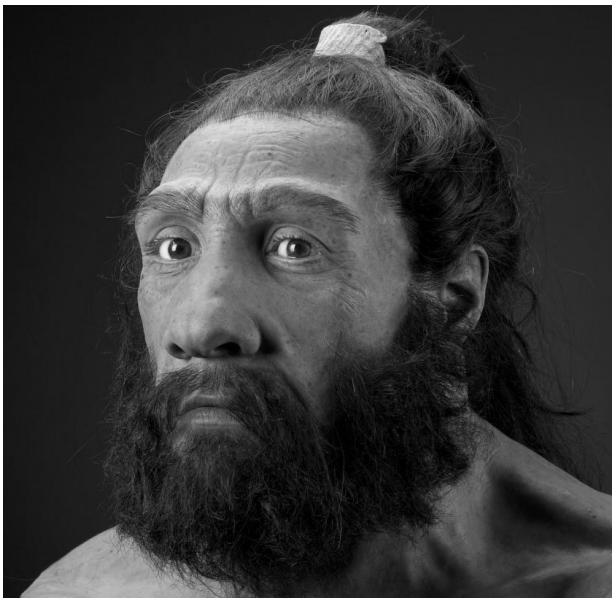


Smithsonian National Museum of Natural History / Wikimedia Commons

Homo erectus (lived approximately 1.89 million to 143,000 years ago in parts of Africa and Asia)

Smithsonian National Museum of Natural History / Wikimedia Commons

Homo heidelbergensis (lived approximately 700,000 to 200,000 years ago in parts of Africa, Asia and Europe)



Smithsonian National Museum of Natural History / Wikimedia Commons

Homo neanderthalensis (lived approximately 400,000 to 40,000 years ago in Europe and south-western and central Asia)

produced children who grew up to be fertile. We know this because all non-African *Homo sapiens* today carry about 2 per cent Neanderthal genes in their DNA. Some of us – like the Melanesians, Papuans and Aboriginal Australians – also carry 3 to 6 per cent Denisovan DNA. Because of this genetic inheritance, we may call them our ancestors, but it is perhaps more reasonable to see them as our evolutionary cousins with whom *Homo sapiens* did dally. Biologically, we are just a part of a gradual continuum of evolution, with chimpanzees sharing 96 per cent of *Homo sapiens* DNA. And the emergence of *Homo sapiens* itself was not a single, dramatic episode. It was a slow process, involving several beginnings and intermixing of various members of the *Homo* species, all of them now extinct. For the small-minded among us, this would be a forgettable, if not an unacceptable, fact. For the rest, this would be yet another reason to appreciate the life around us, and wonder at the unity that binds all life together so tightly.

What applies to cosmology and biology applies to our history as well. In the rather short history of *Homo sapiens* (just around 300,000 years, compared to the 3.8 billion years that there has been life on earth), each of our tribes, clans, kingdoms, empires and nations have considered themselves to be of superior status. Some thought that they were the children of a special God, others that they were the chosen people, and still others that they were divinely ordained to rule over everyone else. People also thought that the spot of earth they occupied was at the very centre of it all – for example, the Middle Kingdom of the Chinese or the 'Midgard' (Middle Enclosure) of Norse mythology. The new nationalisms of the eighteenth and nineteenth centuries built on all these ideas to make everyone believe that the newly created 'nations' they belonged to were infinitely superior to all other nations, and that they had always existed, from 'time immemorial'! In fact, 'time immemorial' is the phrase we hear most often when we try to grasp our deep history.

None of these beliefs are true, of course. No human community is of exceptional status relative to others. None are children of God, or chosen people, unless all are. And none of us live upon the centre of the earth any more than we live on its periphery, since we live on the surface of a globe. Nations as we understand them today are no older than a few centuries, and we are all interconnected – genetically, culturally and historically – far more than we imagine. And even 'time immemorial', it turns out, can increasingly be pinned down, dated, analysed and grasped. And when we do that, we get a far better understanding of our society and culture, and what went into their making.

Would this be upsetting to some? You bet. It is like being told the secrets of the magician who held you spellbound in your childhood. When you learn his secrets, you can either bemoan your lost innocence and the ruined charm of the magic, or you can revel in your new knowledge, the clarity it brings to many things and the possibilities of

what you could do with it. This book is betting that you, dear reader, are of the second kind.

In the chapters that follow, we will be looking at how and when modern humans, or *Homo sapiens*, first arrived in India; what evidence they left behind for us to see; who their descendants are today; who else followed them as migrants to this land of ours; how and when we started farming and building the world's largest civilization of its time; when and why this civilization declined; and what happened next.

This book is about prehistory, and prehistory is about the period that comes before history. History begins when writing begins and places and individuals come alive before us, with their own names and, sometimes, recognizable stories. In prehistory there are no written records and hence we cannot know for sure the names of people and places or the stories of individuals. But, to some extent, we can work out what the life of people might have been like back then, using other kinds of evidence. The evidence in prehistory comes from fossils, archaeological excavations of ancient human settlements, various objects made by humans, like tools, and, increasingly importantly, the DNA of both ancient and present-day individuals.

So where does India's prehistory end and history begin? This is a tricky question because we have written records on seals and tablets from the Harappan Civilization that thrived between 2600 BCE and 1900 BCE, so in a sense we can say that is when our history begins and prehistory ends. But we have not yet deciphered the Harappan script and, therefore, have no knowledge of what is written in those records, so that period falls outside of history and within prehistory. But then, we do have some references to the Harappan Civilization in the contemporary records of the Mesopotamian Civilization in west Asia, so that makes it part of history again.

It is this ambiguity that prompted some historians to use the label 'proto-history' to describe the period between prehistory and history. In this book, we will come right up to the tail end of the Harappan

Civilization and a few centuries after that, leaving the rest for another book, perhaps!

Why this book now

There is a reason why this book could have been written only now, and not earlier. It is because our understanding of deep history has changed dramatically in the last one decade or so. Large stretches of our prehistory are being rewritten as we speak, based on analysis of DNA extracted from individuals who lived thousands or tens of thousands of years ago. Many 'facts' that we took for granted have been proved wrong, and many questions left dangling in the air as historians, archaeologists and anthropologists argued it out among themselves have been given convincing new answers – thanks to the recently acquired ability of genetic scientists to successfully extract DNA from ancient fossils and then sequence it to understand all that bound people together, or distinguished them from each other. If technology had not matured to the level it has, scientists would not have been able to make the discoveries they are making today. And if it were not for their latest findings, our prehistory would have remained as vague and contentious as earlier and this book would not have been written.

Just to get a sense of the speed at which things have moved, consider this: when work on this book began a decade ago, we did not know who were the people of the Harappan Civilization or where their descendants had gone, but now we do. A decade ago, we did not know how much of our ancestry we owed to the original Out of Africa migrants who reached India about 65,000 years ago, but now we do. A decade ago, we did not know when the caste system began, but now we can zero in on the period with a fair degree of genetic accuracy. These are just a few examples that demonstrate our rapidly improving understanding of prehistory, and not only with regard to India.

Here's a short list of things that have changed about human prehistory in other parts of the world because of ancient DNA: we now know that large portions of European populations were replaced not once but twice within the last 10,000 years. First, a mass migration of farmers from west Asia around 9000 years ago mixed with or replaced already established hunter-gatherers in Europe. And then a mass migration from the Eurasian Steppes about 5000 years ago mixed with or replaced the then existing population of European farmers. In the Americas, we now know that native American populations, before the arrival of Europeans, owed their ancestry to not one but at least three migrations from Asia. In east Asia, we know that much of the ancestry of people in the region derives from two or more major expansions of populations from the Chinese agricultural heartland. In 2010 we learned that modern humans had interbred with Neanderthals and in 2014 we learned that our ancestors had interbred with Denisovans (a member of the *Homo* species that was identified only because of ancient DNA sequencing) as well.

When this journey began in 2012, though, I did not know that the field I was getting into, prehistory, was just about to experience an explosion of new knowledge. That is something that happened serendipitously. When I started, I was fascinated by the Harappan Civilization and the questions that were still unsettled: who were the people who built the largest civilization of their time, and where did they go? I visited Harappan sites from Dholavira and Lothal in Gujarat to Rakhigarhi in Haryana, which led me on to many meetings and email discussions with leading historians, archaeologists, epigraphists, linguists and geneticists both in India and from around the world – Romila Thapar and B.B. Lal in New Delhi; Sheldon Pollock in New York; Michael Witzel, David Reich and Vagheesh Narasimhan in Harvard; Iravatham Mahadevan in Chennai; Martin B. Richards in Huddersfield, UK; Peter Underhill in Stanford; M.K. Dhavalikar, V.N. Misra, Vasant Shinde and K. Paddayya in Pune; Shereen Ratnagar in Mumbai; Ravi Korisettar

in Dharwad; Partha Majumder in Kolkata; K. Thangaraj in Hyderabad; Lalji Singh in Varanasi; Niraj Rai in Lucknow; Michael Petraglia in Jena, Germany; Madhav M. Deshpande in California . . . the list is long.

Not all of them agreed with each other, and while every discussion answered some of my questions, it left me with even more questions, not just about the Harappan Civilization but also about the periods preceding it. Before I knew it, the question I was dealing with had morphed from who were the Harappans to how we, the Indians, came to be.

Somewhere along this route, it became clear that the most important revelations were coming from the new field of population genetics. This led me on a search for population genetics papers dealing with the peopling of south Asia – and there were dozens of them – often followed by meetings or discussions with the authors. I met K. Thangaraj, principal scientist at the Centre for Cellular and Molecular Biology (CCMB), Hyderabad, and Lalji Singh, former head of CCMB and, later, vice chancellor of the Banaras Hindu University, Varanasi. This was in 2015 and I was hoping to publish my first story about the Harappan Civilization and the issue of 'Arya migrations'[3] based on these conversations and the research. But there was a problem. I could not complete my article because what Singh and Thangaraj told me did not match up with what I read in the paper they had authored along with other scientists from around the world in 2009.[4] I, therefore,

[3] 'Arya migration' refers to the theory that Indo-European languages, including an early version of Sanskrit, were brought to India by migrants from the Eurasian Steppes, who called themselves Arya, sometime after 2000 BCE. 'Arya' is the self-description of this group of people speaking the same family of languages. Wherever the phrase 'Arya migration' is used in this book, it has to be read as the short version of 'migration of Indo-European-language-speaking people who called themselves Arya'. And wherever the word 'Arya' is used, it has to be read as 'people who called themselves Arya'.

[4] David Reich, et al., 'Reconstructing Indian Population History', *Nature* 461: 489–94 (September 2009).

decided to put the story on hold and gain a better understanding of population genetics before writing anything on it. Then, two years later, in 2017, I came across a paper titled 'A Genetic Chronology of the Indian Subcontinent Points to Heavily Sex-biased Dispersals', co-authored by Professor Martin B. Richards of the University of Huddersfield in the UK along with his team.[5] I read this paper again and again till things started slowly falling in place. I finally got a grip on the issue and could zero in on what was causing the disconnect.

The confusion arose because when I met the scientists in 2015, they had put forward a new hypothesis to me that did not figure in their 2009 paper. This hypothesis was that there were no large-scale migrations to India during the last 40,000 years or so. They also said that there were two very ancient populations, one located in north India and the other in south India and that all of today's populations had descended from the mixing of these two groups, technically given the tags Ancestral North Indian (ANI) and Ancestral South Indian (ASI).

But the paper that Lalji Singh and Thangaraj had co-authored with scientists from the Harvard Medical School in 2009 (titled 'Reconstructing Indian Population History') had made no claims about there having been no large migrations to India in the last 40,000 years. The paper had clearly stated that ANI, unlike ASI, were related to west Eurasians (west Asians, Europeans, central Asians and people of the Caucasus region). This would have given strong support to the theory that Indo-European-language speakers who called themselves the Arya had migrated to India within the last 4000 years or so, after the Harappan Civilization started declining. The issue of 'Arya migration' has been a political hot button for decades, with many opposing the suggestion that the 'Arya' were late migrants to the country, not part of the earliest Indian population. There was the additional problem of the Harappan Civilization: if this mighty civilization which has left an indelible imprint on India preceded 'Arya migrations', then that

[5] Marina Silva, et al., 'A Genetic Chronology . . .', *BMC Evolutionary Biology* (2017).

cuts at the root of the right-wing position that the 'Arya', Sanskrit and the Vedas are the fundamental wellspring of Indian culture. (See also the section 'The second method: Whole genome data' in chapter 2, p. 89.)

The paper co-authored by Martin B. Richards was published on 23 March 2017 and I found it a week later. I spent the following two months reading and rereading tough-to-understand genetics papers from different time periods dealing with the formation of the Indian population; trying to correlate their often contradictory findings with the state of development of population genetics when each of these papers was written; getting in touch with the authors of these papers, many of them doyens of their field with many path-breaking discoveries to their credit; and checking and double-checking the conclusions I was arriving at; and reading more and more papers.

On 17 June 2017, *The Hindu* published my article 'How Genetics Is Settling the Aryan Migration Debate'. Here, I explained how DNA evidence supported the theory that Indo-European-language speakers who called themselves Arya had migrated to India from central Asia around 4000 years ago. The statements made in that story were reconfirmed in March 2018 by a path-breaking paper written by ninety-two scientists from around the world, 'The Genomic Formation of South and Central Asia', and posted in the preprint server for biology, bioRxiv. Reich and Thangaraj were among the co-directors of the study. The scale of the study and the fact that it was based on ancient DNA made the findings far more robust and the chronology of migrations far more accurate. An updated version of the paper with even stronger DNA evidence for a migration from the central Asian Steppe was finally published on 16 September 2019 in the peer-reviewed journal *Science* with the title 'The Formation of Human Populations in South and Central Asia'. This is what the summary of the paper said in its conclusion: 'Earlier work recorded massive population movement from the Eurasian Steppe into Europe early in the third millennium BCE, likely spreading Indo-European languages.

We reveal a parallel series of events leading to the spread of Steppe ancestry to South Asia, thereby documenting movements of people that were likely conduits for the spread of Indo-European languages.'

On the same day, another paper based on the ancient DNA of a woman who lived in the Harappan site of Rakhigarhi about 4600 years ago was published in the peer-reviewed journal *Cell*, co-authored by twenty-eight scientists including many co-authors of the *Science* report such as Thangaraj, Reich, Vagheesh Narasimhan of the Harvard Medical School, Niraj Rai of the Birbal Sahni Institute of Palaeosciences in Lucknow and Vasant Shinde, then vice chancellor of the Deccan College, Pune. With so many common authors, it is no surprise that the two studies published on the same date were in conformity with each other and had arrived at the same conclusions.

The *Cell* paper's title seemed straightforward enough: 'An Ancient Harappan Genome Lacks Ancestry from Steppe Pastoralists and Iranian Farmers'. But, in fact, it required an understanding of the subject to avoid misinterpretations. Many newsapers, not surprisingly, ran with the headline that the Rakhigarhi study had disproved 'Arya' migrations from the Steppe into India, since the 'Harappa genome lacked the Steppe ancestry'. But this interpretation was utterly wrong, if not disingenuous. Steppe ancestry is widespread among Indian populations today and, therefore, the fact that there was no such ancestry in the Rakhigarhi DNA strongly supported the argument that the 'Arya' were not present during the Harappan Civilization and that their migration happened later. In fact, the paper went on to say: 'However, a natural route for Indo-European languages to have spread into south Asia is from Eastern Europe via central Asia in the first half of the second millennium BCE, a chain of transmission that did occur as has been documented in detail with ancient DNA.'

The statement that the Harappan genome lacked ancestry from Iranian 'farmers' had a different implication: it meant that while the Rakhigarhi DNA carried an ancestry 'related to' Iranian farmers, this ancestry had separated from the 'Iranian farmers' before farming itself had begun anywhere in the world. In other words, the paper

was saying that it was not Iranian 'farmers' whose lineage was visible in Rakhigarhi, but the lineage of a population that split from the 'Iranian farmers' before agriculture was invented and, therefore, were not farmers themselves *yet*. They were hunter-gatherers then. We will talk more about this in chapter 2.

There is one more recent finding that could provide additional evidence of a population movement into south Asia soon after 2100 BCE, the date of the first genetic evidence of a Steppe migration that was recorded in Turan, in the region of the Bactria–Margiana Archaeological Complex (BMAC). This is the discovery of three 'chariots' – carts, according to some – buried in the Sinauli necropolis (large cemetery of an ancient city) in the Baghpat district of Uttar Pradesh, about seventy kilometres from Delhi.[6] The vehicles have been dated to around 1900 BCE and since this fits within the period of Steppe migrations into India (~ 2000 BCE to 1500 BCE), this could point to the early arrival of Indo-European-language speakers in the subcontinent.

However, a Discovery Plus documentary released in early 2021 titled *Secrets of Sinauli* argued against this possibility by stating that the Archaeological Survey of India (ASI) had 'obtained a DNA report of a skeleton found in the burial site' and then adding that this report showed no links to central Asian or European ancestry. But

[6] The dispute exists because of the fact that the three buried vehicles have solid wheels, not spoked wheels that normally separate chariots pulled by horses from carts pulled by oxen or other animals. There has been no evidence for horses in Sinauli either. But those who call these 'chariots', or even 'war chariots', argue that the height of the wheel and the lengths of the pole and the yoke suggest that these vehicles were pulled by horses, not oxen. They also rely on the fact that these vehicles were buried along with weaponry – such as copper-decorated shields, antennae swords with hilts, whip, torch, bows and helmets – to argue that these vehicles existed in a war context and that some of the graves belonged to the 'warrior class'. Many of these findings – from the chariots to the shields to the whip to the torch to the helmet – are the first of their kind to be discovered in India. This puts the uniqueness of the site in stark relief and needs an explanation.

this statement is unsupported. No such report based on Sinauli DNA had been published when the documentary was released. And no such report has been published since then. We will discuss more about the Sinauli findings on p. 210.

My experiences during the writing of this book have taught me that even in the most professional of settings, personal preferences can play a part in how research findings are interpreted. And often it may not be a question of bias, but a genuine belief that the truth might cause harmful side effects and, therefore, needs to be treated cautiously. For instance, there could be a fear that the fact of Steppe migrations may reinvigorate old divisions of language and region, just as there might have been a fear among some Indian historians over half a century ago that details of medieval atrocities might cause enmity between different religions. But in reality, holding back the truth cannot heal divisions. It can only cause them to fester underground with even more vigour. Also, no scientist or writer can accurately predict the consequences of a particular truth being withheld: history is made up almost entirely of unintended consequences. So the only reasonable position for any scientist, or any writer for that matter, to take is to let the facts speak, but make sure that no unsupported conclusions are drawn from them.

In this case, it is true that there was large-scale migration of Indo-European-language speakers to south Asia in the second millennium BCE (you will read more about this in chapter 4), but it is also true that all of today's population groups in India draw their genes from several migrations to India: there is no such thing as a 'pure' group, race or caste that has existed since 'time immemorial' (and this holds for all of the world). Of course, the degree to which the mixing between different populations has occurred differs across regions and communities. So the fact of Indo-European migrations has to be told along with the truth of multiple migrations and large-scale population mixing that happened over millennia. We are today a uniquely Indian civilization that has drawn together many population groups with

different migration histories, and its impulses, culture, traditions and practices come from multiple sources, not just one singular source.

In the pages that follow, we will use the new findings made possible by ancient DNA as well as the latest fascinating discoveries made by archaeologists, anthropologists, epigraphists (people who study ancient inscriptions), linguists, palaeoscientists (scholars of the geologic past) and historians to peel the layers of our ancient past one by one. It is a fascinating story and one that is rarely told. Come along.

1

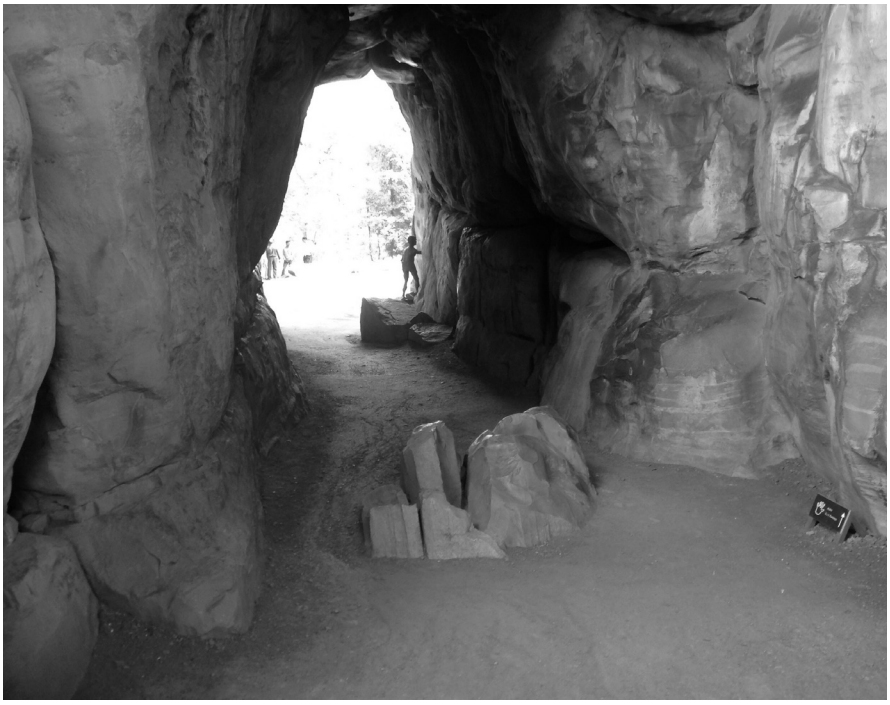
The First Indians

How a band of Out of Africa migrants found their way to India, dealt with their evolutionary cousins and a range of environmental challenges, mastered new technology, made this land their own and became the largest modern human population on earth.

If you want to get as close as possible to the lives of the first modern humans in India, one of the best places to go to is Bhimbetka in Madhya Pradesh's Raisen district, about forty-five kilometres from the state capital, Bhopal. It is an enchanting place spread over seven hills and full of naturally occurring rock shelters that are perhaps more imposing and majestic than most man-made residences of the twenty-first century. There are perennial springs, creeks and streams filled with fish; plenty of fruits, tubers and roots; deer, boar and hare; and, of course, as many quartzite rocks as you need to make all the tools you want. Moreover, the elevation of the hills makes it possible for the residents to keep track of who is approaching them: food or predator, nilgai or leopard!

In the world of early humans, this must have been the equivalent of a much sought-after luxury resort. Ever since it was first occupied

some 100,000 years ago, it has never lain vacant for too long, and it is easy to imagine there having been a long waiting list to get in. A place so well liked that millennia after millennia, one or the other *Homo* species, including our own ancestors, the *Homo sapiens*, lived and hunted and painted and partied there. Yes, the rock shelters are full of paintings, including some that depict people dancing to drumbeats. The paintings are not well-dated, so it is quite likely that most of them, though not all, were made within the last few thousand years, rather than many tens of thousands of years ago. But there are a few petroglyphs, or rock carvings or markings, that could be the earliest evidence of art created by members of the *Homo* species anywhere in the world – a few perfect cupules (small cup-like depressions) with lines beside them.



Tony Joseph

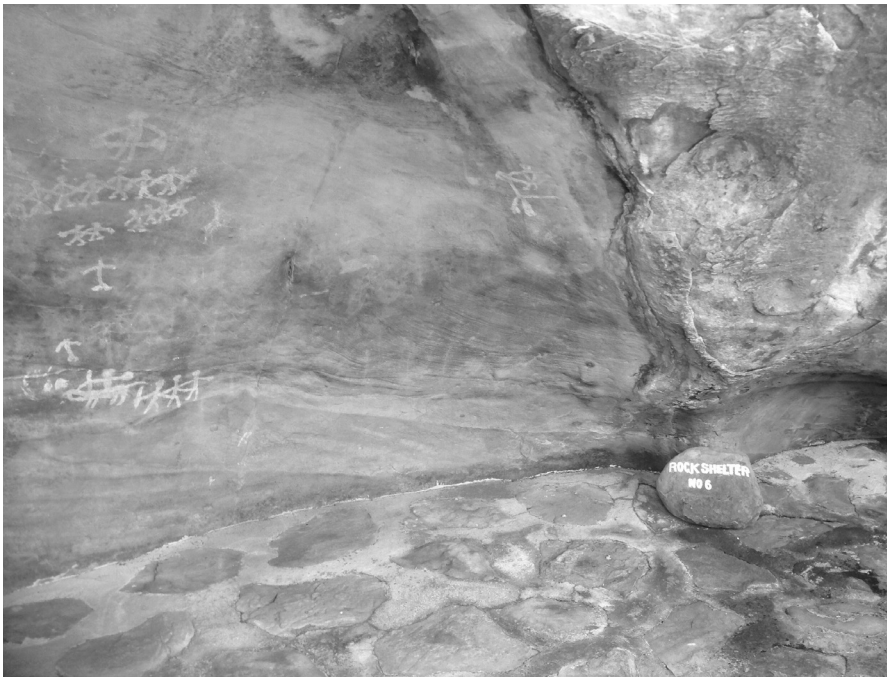
A rock shelter in Bhimbetka

Tony Joseph

Perfect cup-like depressions made on the walls of a rock shelter in Bhimbetka. This is perhaps the earliest evidence of art made by members of the *Homo* species anywhere in the world.

But do we know exactly when the first modern humans set foot in Bhimbetka or, for that matter, in India? The answer to that is a bit complex. First we need to define what we mean when we say 'first modern humans in India'. The technical meaning of the phrase would be any individual belonging to the *Homo sapiens* species who set foot in India first. However, when we say 'first modern humans in India' we also often mean to say the earliest direct ancestors of people living in India today. It is important to know that there is a difference between the two.

For example, let us say the first *Homo sapiens* in India were a group of thirty people in Bhimbetka 80,000 years ago. Let us also say that some calamity – like the huge Toba supervolcanic eruption that occurred in Sumatra, Indonesia, 74,000 years ago and impacted



Tony Joseph

Paintings on the wall of a rock shelter in Bhimbetka, perhaps a few thousand years old

the entire region from east Asia to east Africa – directly or indirectly killed off every one of this first group of modern humans, leaving behind no one to populate the subcontinent.[1] Let us then imagine a second group of modern humans in Bhimbetka around 50,000 years ago, who successfully settle down and leave behind a lineage of people still found in India. Are we referring to the second group when we say the 'first modern humans in India'? This may look like a matter of semantics and it is so, in a way, but it has meaningful implications for us when we interpret archaeological or other evidence to understand the history of early Indians.

[1] This is a hypothetical scenario. Recent research suggests the impact of the volcanic eruption on life in the region was not as severe as earlier understood.

If you ask Indian archaeologists when the first modern humans arrived in India, at least some of them are likely to put a date that is perhaps as early as 120,000 years ago. But if you ask a population geneticist, that is, a geneticist studying genetic variations within and between population groups, the answer is likely to be around 65,000 years or so ago. This seemingly irreconcilable difference between the two sciences is not necessarily contradictory. When geneticists talk about the first modern humans in India, they mean the first group of modern humans who have successfully left behind a lineage that is still around. But when archaeologists talk about the first modern humans in India, they are talking about the first group of modern humans who could have left behind archaeological evidence that can be examined today, irrespective of whether or not they have a surviving lineage.[2]

Did we really have to come from elsewhere?

But why do we assume that modern humans arrived in India from elsewhere at all? Why couldn't they have originated right here? Until a few decades ago, this would have been considered a reasonable question, because the theory that modern humans evolved in different parts of the world separately, from archaic or extinct members of the *Homo* species such as *Homo erectus* that had spread out all over Eurasia by about 1.9 million years ago, was still prevalent – even though Charles Darwin had suggested the African origin of modern humans as early as in 1871. The theory was that the later intermingling of very differently evolved populations kept us together as one species, thus preventing us from branching off into different species in different continents.

[2] Not all archaeologists agree with this distinction, though. 'This idea of successful and failed dispersal is also under scrutiny,' says the archaeologist Ravi Korisettar, adding, 'All dispersal events are successful.'

But this theory has now gone into the dustbin and no serious scientist anywhere puts this forward as a possibility any more (though there may be some isolated holdouts especially in China which, till very recently at least, was wedded to the idea of indigenous, independent evolution of the Chinese people from archaic humans). The reasons why this theory went into disuse are both archaeological and genetic. The fossil record of Africa is rich with the remains of our closest relatives – *Sahelanthropus tchadensis* 7 million years ago, *Ardipithecus ramidus* 4 million years ago, *Kenyanthropus platyops* 3.5 million years ago, *Homo habilis* 2.4 million years ago and *Homo heidelbergensis* 700,000 to 200,000 years ago – and there is no other region in the world that comes anywhere close to it. But the clinching argument against multiple origins of humans on different continents is genetic. The DNA evidence has been conclusive that modern humans outside of Africa are all descendants of a single population of Out of Africa (OoA) migrants who moved into Asia sometime after 70,000 years ago and then spread around the world, perhaps replacing their genetic cousins such as *Homo neanderthalensis* along the way. All recent discoveries have gone on to reaffirm the African origins of all modern humans. As recently as in June 2017 came the news that an ancient skull from a cave in Jebel Irhoud, about fifty kilometres from the city of Safi in Morocco, has been classified as belonging to the *Homo sapiens* species and was dated to about 300,000 years ago.

Until the Jebel Irhoud fossil was dated and classified, the oldest discovered modern human fossils were two skullcaps dated to about 195,000 years ago, found at the archaeological site of Omo Kibish in Ethiopia. So the Jebel Irhoud discovery takes back modern human origins by about 100,000 years and also removes any remaining doubt about where we came from. Though the skull from Jebel Irhoud looks quite like us in its facial traits, the back of the skull is elongated like that of archaic humans and it also has 'very large' teeth, suggesting

that the modern human didn't emerge suddenly and fully formed, but was a work in progress as early as 300,000 years ago.

The logic of genetics

But even if you accept that modern humans arose in Africa, how did the geneticists arrive at the conclusion that all non-African populations descend from a single Out of Africa migration that happened less than 70,000 years ago? One needs to know a little bit of genetics to follow their argument. Genetics can sound somewhat complex to anyone who hasn't paid attention to it earlier, but it is worth investing a few minutes to get familiar with it. You will be able to follow the story even without a perfect understanding of the mechanics of the science described here, so don't get hassled if the explanations given here are not clear enough. Once you get more familiar with the vocabulary, you can come back and read this part again. So here we go.

Almost all the genetic code that humans need is packed into twenty-three pairs of chromosomes that we all carry inside the nuclei of our cells. There is one exception and that is the mitochondrial DNA, or mtDNA, which stays outside the cell nuclei. Each person inherits his or her mtDNA exclusively from his or her mother (the father also carries mtDNA passed on by his own mother, but he doesn't pass it on to any of his children, male or female). The twenty-three chromosomes together with the mtDNA comprise a person's genome.

Unlike the mtDNA, each of the twenty-three pairs of chromosomes in the cell nuclei has one half contributed by the mother and the other by the father. The two chromosomes that make up each pair are similar to each other, carrying similar codes at similar locations. But they are only similar, not identical. The differences between the chromosomes contributed by each of our parents usually amount to about 0.1 per cent. This is the same as the difference between the genomes of any two individuals, on average. These differences arise because of mutations, or random errors that happen especially during

cell division – a necessary part of reproduction in living things. These mutations are then passed down through generations – assuming that, on balance, they are not harmful and, therefore, not weeded out by natural selection.

You could look at a genome as a genetic code written using an 'alphabet' of just four chemicals – A (adenine), C (cytosine), G (guanine) and T (thymine) – and if you do that, then each genome is made up of about three billion individual letters.[3] A 0.1 per cent difference between the genomes of two people translates to about three million differences between the two genomes. If the two genomes came from people who shared a recent ancestor, then the differences would be smaller (which also means that genetic differences can be used as a measure of how close or distant two individuals are genetically).

Notice that although each person carries twenty-three pairs of chromosomes inherited from their parents, they pass on only twenty-three chromosomes (not twenty-three pairs of chromosomes each) to their children. How does this happen? The genetic term for this is recombination and what this means is that each parent randomly shuffles and divides the twenty-three pairs of chromosomes they inherited from their own parents and then passes on only one set of twenty-three chromosomes to their child. In other words, each parent does not pass on all of the genetic material they inherited from their own parents. They pass only twenty-three chromosomes each, thus together giving their offspring a complete set of twenty-three pairs of chromosomes.

But there is one exception to this rule: the twenty-third chromosome pair, or the sex chromosomes. Sex chromosomes are what makes a

[3] According to the National Human Genome Research Institute, 'A chromosome is the structure housing DNA in a cell . . . DNA is a remarkably simple structure. It's a polymer of four bases – A, C, T, and G – but it allows enormous complexity to be encoded by the pattern of those bases, one after another.' Pieces of DNA, or strings of code, that lead to observable traits such as height or eye colour are called genes.

person male or female. If a person carries two sex chromosomes of the type XX, the person will be female, and if the person carries two sex chromosomes of the type XY, the person will be male. For a series of complex reasons, the Y part of the sex chromosome that every male carries comes directly from his own father, with no recombination. In other words, in the case of a male, the Y-chromosome he carries in his sex chromosome comes exclusively through the paternal line going back hundreds of thousands of years.

So we could say, up to an extent, that the Y-chromosome – or Y-DNA, as it is sometimes called – is a mirror image of the mtDNA, which is inherited exclusively through the maternal line, going back hundreds of thousands of years. If the Y-chromosome comes from your father and his father and his father and so on, the mtDNA comes to you from your mother and her mother and her mother and so on. Where the parallel breaks is in the fact that while both men and women carry mtDNA, only men carry the Y-chromosome. Since women's sex chromosomes are of the XX type, they do not have the Y-chromosome at all. There's a reason for this apparent lack of symmetry. Within every cell, mtDNA performs an extremely critical function – it has the code to convert chemical energy from food into a form that cells can use. No wonder mtDNA is often called 'the powerhouse of the cell'. So to put it plainly, no man can do without the mtDNA, but every woman can do without the Y-chromosome.

This nature of the Y-chromosome and mtDNA – that they are inherited without recombination and trace the exclusively paternal and exclusively maternal lines of a person – has proved to be of enormous help, especially in the early stages of population genetics, in understanding the migration history of individuals and populations. What made this possible were mutations, or copying errors, as we discussed earlier. If the mtDNA of a person were exactly the same as her mother's, grandmother's and so on, or if the Y-chromosome of a man were exactly the same as his father's, grandfather's and so on, there would be no substantive information or insight to be had by

analysing anyone's mtDNA or Y-chromosome. But mutations that accumulate over time ensure that the Y-chromosome or mtDNA of a person carries the genetic track record of all that happened in the exclusively paternal or maternal lineage of that person.

For example, if Great-Grandmother had a mutation called PCX on her mtDNA, then she would have passed that on to all her daughters and all her granddaughters born to her daughters and so on. And if you are doing genetic testing of a population in a particular area and come across multiple cases of PCX on the mtDNA, you would be able to create a genetic tree for people with that mutation – and all other mutations that followed since then, if any. In other words, if you have the mtDNA or Y-chromosome of a person, you will be able to locate that person's maternal or paternal lineage over time. Since global human genetic databases exist for both the Y-chromosome and the mtDNA, it is now possible to locate where in the world people who belong to the same group or mutation are currently widely present.

But that is not all either. Scientists have long noticed that there is a certain pattern or regularity in mutations. This is not an exact science but still, they have worked out mutation rates with large confidence margins for the whole genome, as well as for specific regions of the genome such as the Y-chromosome and mtDNA.

While the track record of mutations as reflected in the mtDNA and Y-chromosome allows us to create genetic family trees, the mutation rate allows us to work out the approximate time that has passed since two branches or sub-branches of a tree diverged.

Population geneticists have given names to the branches of the global mtDNA and Y-chromosome family trees that they have created using extensive genetic studies. The equivalent word in population genetics for a branch is haplogroup – haplo means single in Greek, so haplogroup means single group.[4] While a parent branch is called macro-haplogroup,

[4] This is a reference to the fact that the Y-chromosome and mtDNA are haploid – inherited from a single parent, without mixing with the DNA of the other parent.

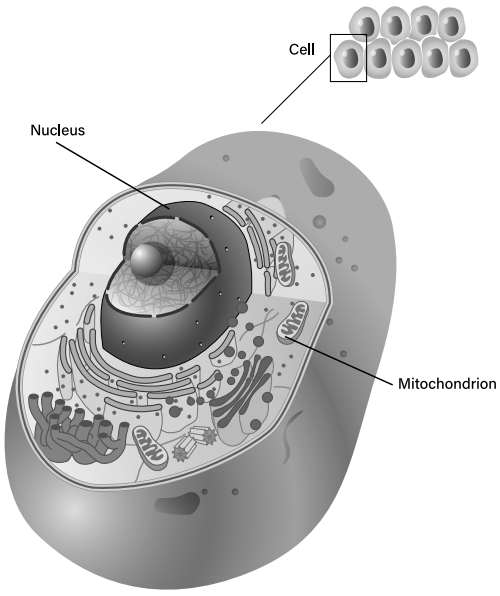
subhaplogroup or clades refers to sub-branches. Some of the oldest branches in the mtDNA genetic tree are haplogroups L0, L1, L2 and M7, while some of the oldest Y-chromosome branches are A, B, CT and D. So by identifying the mtDNA or Y-chromosome haplogroup of a person, you can broadly work out his or her long-term paternal or maternal lineage, and how close or far other lineages are from this. If two people belong to the same mtDNA haplogroup, it means they have a common female ancestor dating from the time that haplogroup originated. And if two men belong to the same Y-chromosome haplogroup, it means they share a common male ancestor dating from the time that haplogroup originated.

A caveat is in order here. Remember that the Y-chromosome or mtDNA that you carry is only a small, less than twenty-third part of your entire genome. So just figuring out your Y-chromosome or mtDNA doesn't say much about what your entire genetic make-up is: it just tells you who your entirely paternal or entirely maternal ancestors are. And they are just a small part of the people you can legitimately call your ancestors. Your mother's father, or your father's mother, or your father's mother's father, for example, are all left out in the cold if you go only by Y-chromosome or mtDNA lineages. If you go back ten generations, you will have 1024 people of that generation whom you can call your ancestors. If you count all your ancestors in each succeeding generation, the number would come to 2046. However, your mtDNA (or Y-chromosome) would have any connection with only ten of them. If you go back fifteen generations, the number of your ancestors in that generation goes up exponentially to 32,768, or 65,534 if you include all ancestors in each of the fifteen generations. But your mtDNA would be connected to only fifteen of them! This could sometimes lead to odd results.

This is unlike the rest of the chromosomes, which are diploid, or inherited from both the parents.

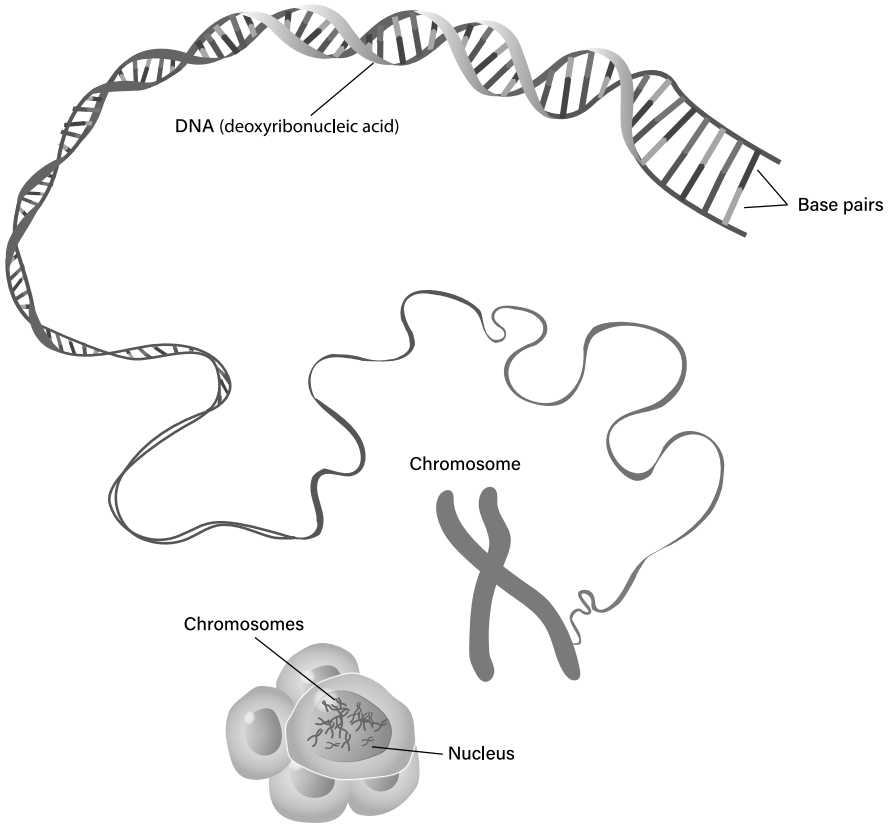
For instance, it is possible for a person to be almost entirely of Chinese ancestry, but to belong to a Y-chromosome haplogroup that is common only in India. All that would have been necessary for this to happen is for an Indian man to have left behind a son in China, say, ten centuries ago and for this son in turn to have founded a lineage with every generation having at least one son, all of whom lived in China and had children with Chinese women. A male descendant of this lineage today – the son of the son of the son . . . of the Indian – could still carry the Indian man's Y-chromosome, but he would be of Chinese ancestry for all practical purposes, because there is only one tenuous, centuries-old link that connects him to India.

So while the mtDNA and Y-chromosome are helpful ways to understand population movements or histories of individuals or groups, they may not be sufficient to grasp a person's or a population's entire genetic make-up or its relationship to other populations. For that, we need whole genome sequencing, which studies a person's entire



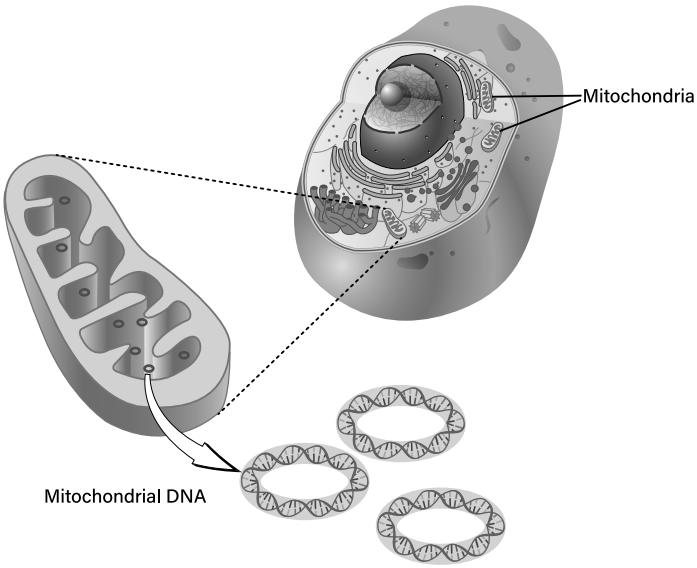
Cell

Source: National Human Genome Research Institute, Bethesda



Chromosome
Source: National Human Genome Research Institute, Bethesda

genome, not just the Y-chromosome or the mtDNA. We cannot create genetic trees out of the twenty-two non-sex chromosomes – which are called autosomes – because recombination, or the shuffling and division of genes, makes that impossible. But whole genome sequencing can clearly help measure the degree of affinity between different population groups. Whole genome sequencing used to be a very costly and time-consuming affair earlier, but with improving technology, it is becoming increasingly common in genetic studies.



Mitochondrial DNA
Source: National Human Genome Research Institute, Bethesda

Dating 'Out of Africa'

Now that the basic mechanics of genetics is out of the way, let's tackle the next question: why do geneticists say that all modern humans outside of Africa come from a single group that migrated out of that continent, and why do they put the time of the exodus to 70,000 years ago or later? The reason is straightforward. When you look at the mtDNA of people outside of Africa all around the world, you will find they all descend from a single haplogroup with deep lineage in Africa, namely, L3. Think about what this means: that all people outside of Africa are descended from a single African woman who originated the L3 mtDNA haplogroup! Africa has about fifteen other, much older, lineages with names such as L0, L1, L1a and L1c, but none of them were part of the group that went on to populate the rest of the world. L3 has two immediate descendant lineages or subhaplogroups today, M and N, with N having its own major subhaplogroup, R. Thus