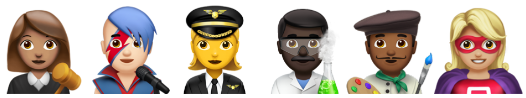
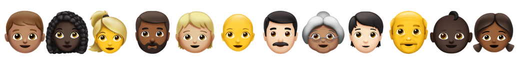
 . . .



The Emoji Guide to Human Genetic Diversity

Chapter 1: Our Deep Past

You are a member of a species that looks like this:



A long time ago, your lineage looked more like this:

What happened? Individuals vary within a species, and the sorting out of those variants over time causes a species to change. That's evolution. It's a simple concept. But these two ideas – that humans are different from each other, and that humans evolved – can make us uncomfortable. Many folks mistrust the science, misunderstand the science, or misconstrue the science. This guide will try to clear everything up.

Don't be this guy

Humans variation is correlated with geography. People living in one place might look like this:

People from Site 1

And people from another place might look like this:

People from Site 2

How to make sense of the differences? A common strategy is to divide people into “races”. But “race” isn’t a biology word and isn’t defined by what biologists can measure. The idea of race is tied up with culture and social assumptions and isn’t particularly well correlated with actual genetics.

Alan 👷🏿

Bob 👨🏼‍🔧

Chuck 👩🏿‍🏭

Alan is genetically more similar to Bob than to Chuck, but Alan and Chuck are the same race and Bob is another race

That’s why geneticists use words like “ancestry” and “population” rather than “race.” Those terms aren’t just synonyms for “race,” which is largely defined by subjective beliefs and social norms. Instead, they’re defined by objective data like who your close relatives are and where in the world they live. But even that can cause confusion. A common mistake is to think of human populations as discrete units, like islands in the sea.

Population 1 **Population 2**

Instead, it’s more realistic to think of each person as an island, and populations as archipelagos. We can find clusters of islands and call them distinct archipelagos, but it’s somewhat subjective and arbitrary. Like this:

Or alternatively:

Even accepting that the borders between populations are fuzzy, there is a tendency to overestimate the differences. It’s true that humans across the globe have been subject to slightly different evolutionary pressures and random changes, leading to distinct genetic characteristics. And it’s true that these are the same kind of adaptive processes that can eventually cause species to diverge. So far, so good. But at this point it’s easy to be misled. To see why, consider the big picture of human evolution. We’ve all seen those diagrams that show a parade of progress, from curly-tailed monkey, to knuckle-dragging ape, to hairy caveman, to erect modern human.

🐒 → 🦍 → 🧔🏾 → 👨‍💻

A common, but misleading, depiction of evolution

There are a few problems with such images, such as:

👩‍💻 ❓

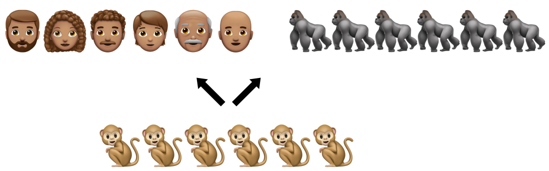
Where are the women?

and

🐒🐒🐒🐒🐒🐒 → 🦍🦍🦍🦍🦍🦍

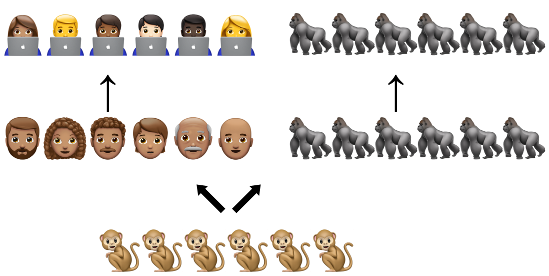
Populations, not individuals, evolve

and



Populations are continuously-evolving branches, not static steps in a linear progression

and



All living humans, and all other living organisms, are equally "modern"

Also, focus on the space between each advancing primate.

🐒 → 🦍 → 🧔 → 👨‍💻

Here it is again

You might assume from the picture that the amount of change between apes and ancient humans should be similar to the amount of change between ancient humans and today. However, a little math will show how wrong this assumption is.

🧮

Get out your calculator

The last common ancestor of all living humans is estimated to have lived only three or four thousand years ago (Rohde et al. 2004). That was about 120 generations ago. Of course, at that time there were already humans living in very different cultures all over the world. But there were always a few travelers, even across continents, sowing their wild oats.

Number of "greats" (generations) for who shares a great-great-etc. grandparent with you:

10^0👨🏽👩🏽

10^1👦🏾🧓🏽

10^2👱🏼‍♂️👩🏿‍🦱**(all humans)**

10^3

10^4

10^5

10^6🦍🦧**(all apes)**

10^7🐵🐒

10^8🐰🐯

10^9🦈🐙**(all animals)**

10^{10}🌻🌲

10^{11}

10^{12}🦠🌏**(all life)**

Picture 120 people standing in a line. The first is your mother or father, the next is one of your grandparents, then a great-grandparent, and so on. They would take up about the length of a football field.

🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂
🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂
🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂
🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂🙂
🙂🙂🙂🙂

= 🏟️

120 ancestors = 1 football field

Every person on Earth today could trace a line back to the same ancestor at the end zone. We're all family, 120th cousins or closer. However, that greatest grandparent didn't actually impart DNA to all of us. Chromosomes get shuffled with every new baby. Each parent only provides about half of your DNA, and so eventually the traces of most ancestors are lost.

👨🏻‍🦰×👩🏻‍🦱

↓

👩🏻‍🦰

The DNA behind Mom's hair might have missed you

To find someone who actually contributed genetically to all modern humans, we'd need to go back to the dawn of modern humans, between 100 and 200 thousand years ago. About 6000 generations.

In other words, a queue of ancestors the length of fifty football fields laid end to end, reaching just under three miles. A bit longer than Central Park. You could stroll past them in less than an hour.

🏟️
🏟️
= 🚶

50 football fields = 1 walk

All of these 6000 ancestors are completely human, physically indistinguishable from folks alive today. To escape our species, you'd need to go back farther. Our closest living relatives are the chimpanzees, which shared an ancestor with us around six million years ago. That's about 300,000 generations. The line now stretches over 140 miles. The width of Indiana. A trek composed of fifty park-sized walks, or a highway drive of a couple of hours.

🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶
🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶🚶
= 🚗

50 walks = 1 drive

As you drove past them, you ancestors would gradually look less human, but change would be slow. To get to actual monkeys, the kind with tails, the line would need to be even longer. Our common ancestor with baboons would stand behind more than a million other ancestors, 800 miles behind the start of the line. The distance from Canada to the Gulf of Mexico. Our ancestor with other monkeys like capuchins would need an even longer line. If you wanted to pass by all of them in less than a day, you'd want an airplane.

🚗🚗🚗🚗🚗. . . =✈️

5 (or more) drives = 1 flight

The point is that humans haven't been separated for very long. Evolution is slow. 1500 BC sounds like a long time ago, but it's a blip compared to Earth's prehistory. The amount of evolutionary change that has happened within the human species is a blip compared to the amount of change between species. So while you *could* think of the gaps among human populations as just "smaller versions" of the gaps between species, you need to remember that they are *a lot* smaller:

🐒⟶⟶**[insert thousands of arrows here]**⟶⟶👨🏼‍🔧🧑🏿‍🏭

This is not to say that big evolutionary changes can't happen on short time scales. Just look at agriculture. Most of our food comes from organisms that we have bred to be dramatically different from the way their wild ancestors were just a few thousand years ago:

🥔**toxic**

🌽**tiny**

🥕 **white**

🐖🐓🐄**bony & aggressive**

🍌**seeds!**

But that's not how evolution *usually* works. Most of the time, the main thing that natural selection does is keep things the same. If you think about how old the Earth is, the really surprising pattern is how *little* change has occurred. Let's think back to the days of the giant dinosaurs.

🦕🦖🦕🦖🦕🦖🦕🦖🦕🦖🦕🦖

100 million years ago, our ancestor was a small fuzzy creature that also gave rise to lots of other mammals.

🐭→🦇🦌🐬🦏🐆🐒🐩🐘👦🏽

One of the many differences between it and us is that our bodies are larger. So over time, body size increased in the lineage leading to us.

🐭 → 🐭

Small fuzzy creatures beget large fuzzy creatures

How fast did it evolve? Imagine that it became 1% larger every 10,000 years. That doesn't sound like a particularly fast rate of evolution, and certainly we've measured faster rates in other species. If the animal started out weighing 1 pound, then after 100 million years it would weigh 10^{43} pounds. About as much as the Milky Way Galaxy:

⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️
⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️
⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️
⭐✨🌟☀️⭐✨🌟☀️⭐✨⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨⭐✨
🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨
🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨
🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨
🌟☀️⭐✨🌟☀️⭐✨⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨🌟☀️⭐✨⭐✨🌟☀️

A very large fuzzy creature

Clearly that didn't happen. Instead, it evolved much more slowly. That's because the most common thing for natural selection to do is to constrain change, not promote it. That's not really surprising, because organisms are already really well optimized to Earth's environment. The vast majority of mutations will just screw things up and reduce the fitness of the organism. So, natural selection will weed the mutations out.

🍎🍎🍎🍎→🍎🍏🍎🍎→🍎🍎🍎🍎

The fate of most mutations

This means that from first principles, you shouldn't expect human genomes to have changed *very much* since our species first appeared. So what's a *better* way to envision human populations? Humans first evolved in Africa. The highest levels of human genetic diversity are still seen in Africa. Non-Africans represent a subset of this diversity. Both Africans and non-Africans have evolved *a little bit* since humans started to spread across the planet. Most genetic variants, though, are similarly abundant among all continents. All humans are unique combinations of this shared gene pool, which Chapter 2 will examine.

Chapter 2: Our Shallow Gene Pool

Chapter 1 explained why we shouldn't expect much evolutionary diversification since the last human common ancestor. And yet, if you look at humans from around the world, you'll see body parts in a variety of colors and shapes:

💪🏿💇🏻‍♀️👂🏾🦶🏼💋💪🏻✋🏾👃🏼💇🏼‍♂️👂🏽💅🏼🦵🏾👣🦵🏼👃🏿👄

Disembodied diversity

One often hears comments about the richness of human diversity. But interestingly enough, our species actually has relatively little genetic diversity.

👨🏽👶🏿👱🏿‍♀️👧🏾👱🏼‍♂️👩🏾👦🏽👵🏾👶🏻👩🏻👱🏻‍♀️👦🏿👧🏽👴🏿<🐰

A typical non-human species harbors more genetic diversity than humans do

To demonstrate this, imagine choosing any gene from your genome. You have two copies of it, one from your dad and one from your mom, so pick one at random. Do the same for another human with a different racial background. Compare their DNA with your own. You'll see a few differences. But you'd see even more differences between the two copies that exist side-by-side within one individual rabbit, gibbon, or sparrow. People are genetic paupers. We have less total biodiversity than most other animals [(Leffler et al. 2012)].

Genetic diversity per species:
0.0000

0.0005🦦

0.0010👦

0.0015🐼🦍🐺🐘🌵

0.0020🐸🐗🐃

0.0025🐒

0.0030🦧

0.0035🕷🦐

0.0040🐁🐋🦆🐡🌲

.

.

.

0.0060🦎

0.0065🐦

.

.

.

0.0085🐰🐝

0.0090🦗

.

0.0100🌳

.

.

.

.

.

0.0130🦂

.

0.0140🦟

.

.

.

.

.

.

.

.

.

.

.

0.0200🦋

.

.

.

0.0220🌻

The reason for our dearth of diversity lies in our history. Our species appeared quite recently, initially as only a few thousand individuals. We have since expanded into the billions, but there hasn't been enough time for new mutations to spread and reflect our new global ubiquity. Genetically, we still look like a rare species (1000 Genomes Project Consortium 2015). And if there is little genetic variation in our species as a whole, there it is little that can be partitioned out among humans in different places.

👥👥👥→🏙🏙🏙🏙🏙🏙🏙🏙🏙🏙🏙🏙

We quickly went from a few scattered people to a swarm of enormous cities

However, let's not oversimplify things. Overall, two copies of a chromosome chosen at random from two humans will be 99.9% the same and 0.1% different. But what does that mean? A single mutation could have no effect, or it could mean the difference between life and death. So, there is no straightforward connection between variation in DNA sequences and variation in traits we can actually see.

🦟🦟🦟🦟🦟🦟🦟🦟🦟🦟🦟🦟🦟🦟🦟🦟

Even with high genetic diversity, every member of a species could look the same

What, then, does it mean for a species to have a lot of biodiversity? Consider, for example, the common side-blotched lizard from western North America. This is a single species, but among these lizards are several "morphs" the look and act differently (Corl et al. 2010). Orange-throated males guard large harems of several females. Blue-throated males are monogamous and guard a single mate. Yellow-throated males sneak into the territories of orange-throated males and mate with the females there, but forego such trysts with the ever-guarded mates of the blue throats. Meanwhile, orange-throated females produce many small eggs, yellow-throated females produce a few large eggs, and there are no blue-throated females. Most populations include all three colors, but some have just one or two. These ratios vary with geography, and along with other genetic differences they divide the lizards into several "subspecies."

🟠🦎🆚🟡🦎🆚🔵🦎

There are many other species with similarly mind-blowing diversity. The Numata longwing is a South American butterfly that comes in seven forms, each with a completely different pattern of orange, yellow, and black pigment on its wings. If you didn't know better you'd assume they were separate species, and in fact they have each evolved to mimic a different poisonous butterfly (Joron et al. 2011). Steelhead and rainbow trout are actually the same fish species. The former is more than twice as large and migrates to the sea to hunt, while the latter maintains a more laid-back lifestyle in the watershed of its birth. Woodland strawberries have distinct subspecies with totally different sexual systems: one includes females, while the other consists entirely of hermaphrodites (Staudt 1989). In these animals and plants, the different morphs or subspecies each have their own ecological roles and adaptive strategies.

🦎🦋🐟🍓🚷

Species with outstanding biodiversity. Humans need not apply.

It should be obvious that human biodiversity pales in comparison. People practice different strategies for finding mates, but these are influenced culturally, not genetically. Unlike the lizards, a person of any genotype could grow up to be a celibate, a monogamous spouse, or a polygamist.

👨🏽 🆚 🤵🏽 👰🏻 🆚 👯‍♀️🕺🏽👯‍♀️

Your ancestry doesn't define your lifestyle

Furthermore, even if prehistoric humans like *Homo erectus* were still around, you would have no trouble sorting out *Homo sapiens* from any other species, as you might with the longwing butterflies. Human populations may vary, but we don't differ in what kind of fluid we are able to breathe, like freshwater versus saltwater fish. Nor in the presence or absence of an entire sexual organ, as in strawberries. For these reasons, humans are not classified into morphs, subspecies, or breeds[Norton et al. 2019]. There is one standard type of human, with minor customizable flourishes.

🚙🚙🚗🚕🚕🚙🚙🚗🚗🚗🚕

🆚

🚙🚕🏎️🚓🚜🛺🚗🚐🚑🚚🚛

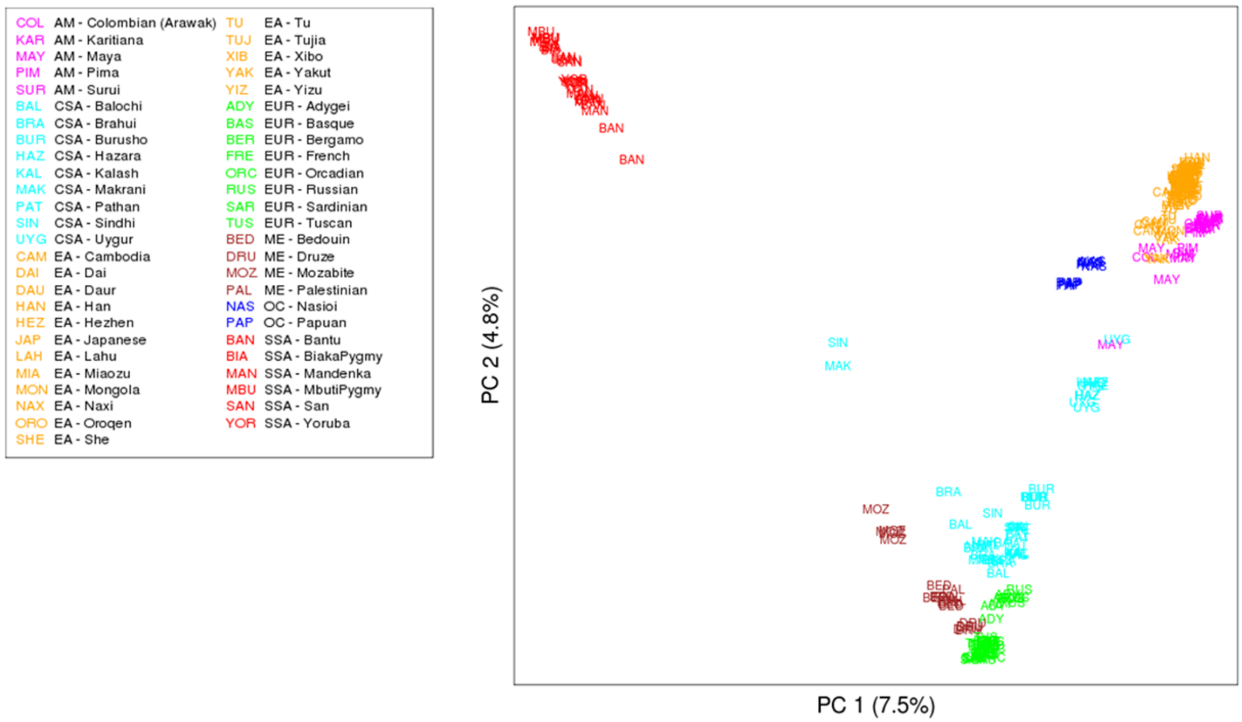
Some variety (top), but not as much as there could be (bottom)

To be clear, the variation we do possess is not distributed equally around the world. Humanity is not homogeneous. If you send a geneticist your DNA, they could get a pretty good estimate of where your recent ancestors lived. If you showed those ancestry estimates to someone else, they could make a few general predictions about your overall physical appearance.

🇯🇵🇲🇽🇿🇼🇵🇼🇪🇸🇮🇳🇲🇬🇧🇷🇮🇩🇮🇸🇦🇹🇬🇦🇸🇦🇻🇳

There's no perfect match between your DNA and your ancestors' nationalities, but there are correlations

But DNA is not cleanly partitioned by homeland. Remember the archipelagos. Populations geneticists often make plots of individuals that actually look a lot like islands in an ocean. However, these plots are easy to misinterpret. There is no literal ocean or any other geographical space. Positions on the plot are defined by multivariate statistics. Genetically similar individuals cluster together. Consider the following classic principal component analysis (PCA), for example [López Herráez et al. 2009].



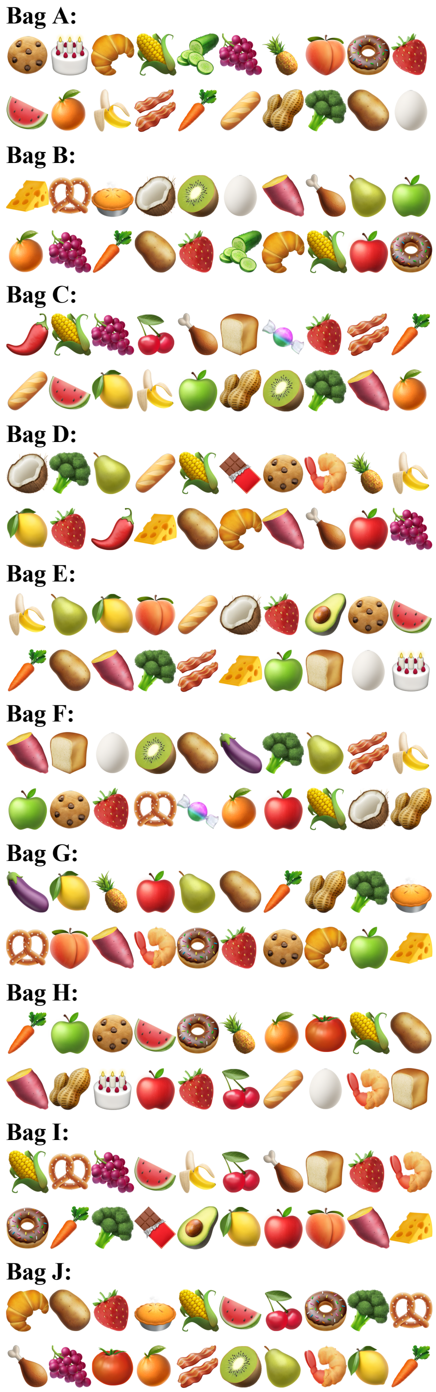
Studies like this don't randomly sample humans across the globe. They typically target people belonging to several pre-defined groups from different geographic locations (as in Pima, Miaozu, Yoruba, etc. above). This isn't necessarily a fault of the study, but it can make the borders between groups look really stark, more so than they are in reality. A common misconception is that there are "pure" populations like the ones in the figure, and then some individuals like Barack Obama are "hybrids" between these populations.

A hybrid of two pure populations: not a useful way to think about human genetics

In reality, human populations have been mixing and remixing since the dawn of our species. Every "pure" population is really descended from two or more other populations that used to be separate, which themselves have a mixed ancestry, and so on[Reich et al. 2018]. We are all admixed.

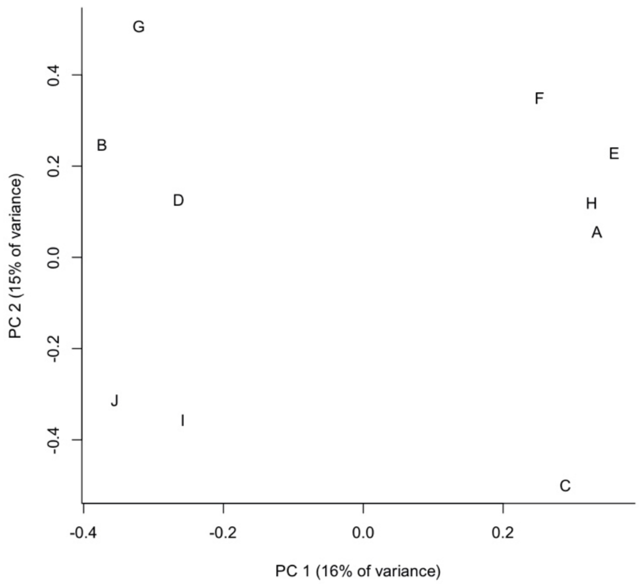
A stew of mixed ingredients plus another stew of mixed ingredients yields an especially mixed stew

More importantly, images like the PCA above are meant to highlight subtle dissimilarities by turning the contrast up to eleven. They doesn't say anything about the *size* of the differences among those ancestry groups. As an analogy, consider the following ten bags of groceries, each with twenty items:

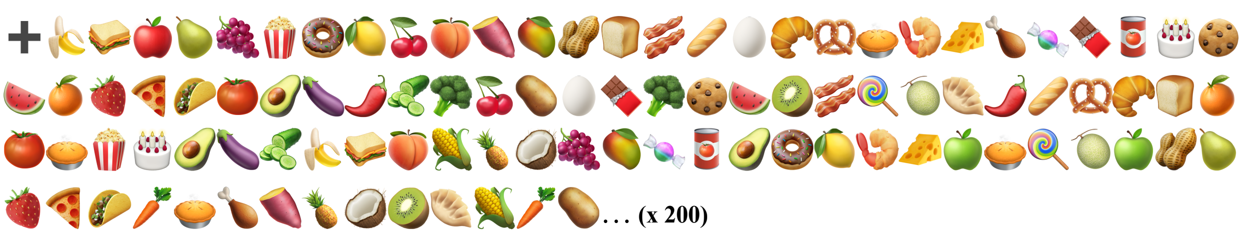


Ten grocery bags that were randomly packed. Or were they?

Just by looking at their contents, could you easily classify the bags into distinct categories? Probably not. Or if you did it would be arbitrary. You wouldn't expect other people to make the same decisions. All the bags seem to contain a fairly random assortment of items. However, a computer algorithm could readily detect nuanced patterns and place the bags into clear groups. Here's a PCA. The first axis explains 16% of the variation among bags. A, E, F, H, and (to a lesser extent) C all seemingly share something.



If you only saw the PCA result, you might assume an absolute gulf between the left and right sides. However, you'd be wrong. The right-side bags often (but not always) have certain groceries, like bananas and eggs, that are usually (but not always) absent from the left-side bags. And vice versa. That's it. Those gradients are enough to separate them, but in the overall chaos of snacks and entrées, it's hardly a noteworthy difference. Same with people. We can cluster people based on traits, but that doesn't mean we're revealing some essential distinctiveness. In fact, people are even more similar than that, because 99.9% of DNA is the same. It would be like we also added an identical set of 20,000 groceries to each bag, and the only thing unique about each bag were the 20 groceries shown above.



. . . (x 200)

That's a really big grocery bag

For most genes, diversity within our species and divergence among populations is neither large nor meaningful. But is that true for every gene? Chapter 3 will explore whether and how some genes may impact important traits that vary among populations.

Chapter 3: Ancestry and Ability

Despite the paltry human diversity outlined in Chapter 2, are there some genes that still explain the differences in important traits that we care about? Traits that underlie how we look, think, feel, act, heal, perform, and so on? Among individuals, definitely. There's a genetic component to most aspects of both brawn and brains. Often a substantial one. *Between* populations, though, genetic effects are typically weak if they exist at all. Most of the human genome looks pretty much the same on average all over the world, even if it varies within each population. But there are a few genes that have diverged dramatically among continents.

👖👖👖👖👖👖👖👖👖👖👖🩳👖👖👖
Most jeans are the same, but not all

These genes are rare exceptions that stand out against the background of more typical genes. Many of them have to do with infectious disease, diet, or climate. Factors that clearly aren't the same in different environments. There are no known effects of these genes on personality or cognitive ability. And even for the most extreme genes, there are no absolute differences among populations, just average differences. One version of a gene is more common in some populations, while an alternate version is more common elsewhere.

Genes with adaptive differences among human populations:

🩸🦟**malaria resistance: *ACKR1***

🥛🧀**digest dairy: *LCT***

👂💦**ear wax & sweat odor: *ABCC11***

🆎🅾️**blood type: *ABO***

🏔️🗻**breathe at elevation: *EPAS1***

✋🏿🖐🏻**skin tone: *SLC24A5***

☃️❄️**cold tolerance: *TRPM8***

Those are the outliers. They are few and far between in a genome of over 20,000 genes (the exact number is still debated). And while you sometimes hear that these variants have major overall effects on how bodies function, this is wrong.

How to actually eat for your 🩸 type:

🅰️:🥕🥗🌯🥜🍣🍠🥖🍅🧀
🅱️:🥕🥗🌯🥜🍣🍠🥖🍅🧀
🆎:🥕🥗🌯🥜🍣🍠🥖🍅🧀
🅾️:🥕🥗🌯🥜🍣🍠🥖🍅🧀

There are other genetic variants with *moderate* levels of divergence among populations. Such moderate differences are more numerous, and it's hard to say what all of them do. Some are associated with personality traits, for example in the gene *MAOA*, sometimes called the "warrior gene" because it's linked to aggressive behavior (McDermott et al. 2009). It's an overly exaggerated moniker: its effect is minor, and as Chapter 4 will explain, *MAOA* isn't even the main gene we should associate with aggression. And regardless, one cannot conclude from any specific gene that aggressive behavior differs innately among populations. Complex traits are influenced by many genes, as well as by the environment. *MAOA* only explains a small proportion of the variance in violent behavior, and even then it depends a great deal on a person's upbringing. Unless we can tally up the effects of all of the many genes and other factors that influence aggression — and we have nowhere near the ability to do that — we shouldn't expect populations to differ in any particular way.

Relative abundances of genetic variants in people with ancestry in western Africa and northern Europe:

Variant near *ACKR1* 🩸🦟:
🔷🔷🔷🔷🔷🔷🔷🔷🔷🔷🔷🔷**W Africa**
🔶🔶🔶🔶🔶🔶🔶🔶🔶🔶🔶🔶**N Europe**

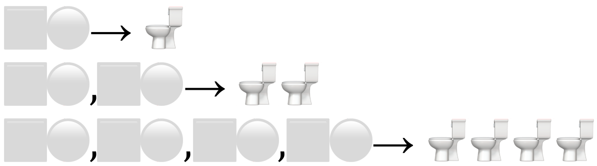
Variant near *LCT* 🥛🧀:
🔶🔶🔶🔶🔶🔶🔶🔶🔶🔶🔶🔶**W Africa**
🔷🔷🔷🔶🔶🔶🔶🔶🔶🔶🔶🔶**N Europe**

Variant near *MAOA*🤬😡:
🔷🔶🔶🔶🔶🔶🔶🔶🔶🔶🔶🔶**W Africa**
🔷🔷🔷🔶🔶🔶🔶🔶🔶🔶🔶🔶**N Europe**

A more typical variant😐😑:
🔷🔷🔷🔶🔶🔶🔶🔶🔶🔶🔶🔶**W Africa**
🔷🔷🔶🔶🔶🔶🔶🔶🔶🔶🔶🔶**N Europe**

Another typical variant😐😑:
🔷🔶🔶🔶🔶🔶🔶🔶🔶🔶🔶🔶**W Africa**
🔶🔶🔶🔶🔶🔶🔶🔶🔶🔶🔶🔶**N Europe**

Part of the problem comes from thinking of DNA like a blueprint. This is a common metaphor, even among scientists. It's a nice shorthand for the idea that DNA records information that is later manifest in the physical construction of the body. But it's not a perfect metaphor, and it's easy to misinterpret. In an actual blueprint, as used by architects, every symbol has a precise meaning. Toilets are symbolized by little shapes that look like a toilet as viewed from above. The number of these symbols on the blueprint will exactly equal the number of toilets in the finished building. This is true whether the building will be constructed in Paris or Bangkok.



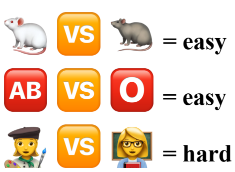
Blueprint: the number of toilet symbols equals the number of actual toilets

A genome is a product of billions of years of evolution. A creationist might believe that is was designed by an Architect. But reality is not nearly so coherent. Genomes are messy. In genetics, there is usually no one-to-one association between a gene and a trait. Many genes and other factors contribute to each part of the final product. Most functional genetic variants are more like a person's comment at a town meeting when a building is being designed. It might have an eventual effect on the number of toilets, but only in an indirect and contingent way.

Not a blueprint

So can we take genetic data and directly predict the effects on bodily traits? Yes, but only after controlling for other factors at play. This is relatively easy if you are working with, say, mice in a laboratory setting. It's much harder if you are looking at humans across different societies. It's still doable for simple traits like blood type. But it's especially hard for complex traits like behavior that are affected by many genes as well as culture.

Predicting traits from 🧬:



Even if you don't know the specific genes involved, you can tell if a trait is heritable if it runs in families. Can we use heritability to estimate innate differences among populations? Not really. Heritability is easy to misinterpret. It has to do with how much variation *in one particular population at one particular time* is due to genetics. A trait like "number of legs", while obviously influenced by DNA, actually has really low heritability in humans. Almost all of the variation is due to wars and accidents, not genetics.

Not a genetic trait

Even a highly heritable trait can be influenced by the environment, often in subtle ways. Skin color is determined genetically, but it can also vary because of tanning. Analogously, a trait with high heritability could differ between two populations, without the *between*-population differences having anything to do with genetics.

December:
👶🏼🧕🏻👱🏻🧔🏿👵🏻👲🏻👧🏼🕵🏻‍♀️👴🏽🎅🏻

→

July:
👶🏽🧕🏽👱🏽🧔🏿👵🏽👲🏽👧🏼🕵🏾‍♀️👴🏾🎅🏽

Not an example of evolution in action

Thus, heritability isn't some fixed feature of a trait, but depends on the environment. Imagine if all bicycles were exactly the same style and of a uniform, non-adjustable size. In that case, ability to ride a bike would be highly heritable because you would need to have just the right leg and arm length, etc. In the real world, almost anyone can find a bike that fits their body, and whether you learn to ride or not depends mostly on personal choice. So, simply by revising a piece of technology, it's possible to change heritability.

🚴🏽🚴🏽🚴🏽🚴🏽🚴🏽🚴🏽🚴🏽🚴🏽
🆚
🚴🏽🚴🏽🚴🏽🚴🏽🚴🏽🚴🏽🚴🏽🚴🏽

Heritability depends on the tools available

Suppose two genetically-distinguishable groups of humans differ in average SAT score. That does not imply that they have different brain wiring because of their DNA. Consider the following analogy. There are two populations of flowers. The western population has more pink pigment than the eastern one, and also produces more seeds, on average.

Natural situation:
🌺🌺🌺 🌼🌼🌼
🌰🌰🌰 🌰🌰🌰
🌰🌰🌰

A botanist wants to know why this is so. It could be that some aspect of the environment, like the soil, is simply different between the habitats. If so, growing the flowers together in a common garden should eliminate the difference

Poor soil:
🌺🌺🌺 🌼🌼🌼
🌰🌰🌰 🌰🌰🌰

Rich soil:
🌺🌺🌺 🌼🌼🌼
🌰🌰🌰 🌰🌰🌰
🌰🌰🌰 🌰🌰🌰

Or, it could be that the difference is innate, and doesn't depend on the soil:

Any habitat:

🌺🌺🌺   🌼🌼🌼
🌰🌰🌰   🌰🌰🌰
🌰🌰🌰

Or, it could be that the soil matters, but it has an effect early on. For example, plants grown in poor soil produce nutrient-deficient seeds which themselves will grow up to produce only a few seeds. To test this hypothesis, you'd need to raise a few generations of plants in the same garden first.

Mom grown in poor soil:

🌺🌺🌺   🌼🌼🌼
🌰🌰🌰   🌰🌰🌰

Mom grown in rich soil:

🌺🌺🌺   🌼🌼🌼
🌰🌰🌰   🌰🌰🌰
🌰🌰🌰   🌰🌰🌰

Or, there might be something else going on. What if bees favor pink flowers and are more likely to pollinate them? Then the question is really about the genetic basis of flower color, which only indirectly influences seed count. Flower color could be genetic, or due to soil, or anything else.

Indirect effect of pollinators:

🌺🌺🌺   🌼🌼🌼
🐝🐝🐝   🐝
🌰🌰🌰   🌰🌰🌰
🌰🌰🌰

It's not that pink flowers are innately better at seed production. Their seed production machinery has nothing to do with it. It's an arbitrary preference of the bees. Release some butterflies that enjoy all flowers equally, and the difference disappears. Change the environment, change the trait. Even if the trait depends on genes.

Equal-opportunity pollinators:

🌺🌺🌺   🌼🌼🌼
🦋🦋🦋   🦋🦋🦋
🌰🌰🌰   🌰🌰🌰
🌰🌰🌰   🌰🌰🌰

An analogous thing happens with humans. Is success in life due to nature or nurture? Any serious scientist would agree that both genes and environment play important roles. But they can interact in complex ways. Genes influence physical appearance. In a society that judges people by appearances, these genes influence professional achievement. Even a seemingly objective test like the SAT is heavily affected by how a person has been treated in the past. But just like the flowers, this is an indirect and arbitrary effect. It could be changed.

🧠🧠🧠🧠🧠🧠🧠🧠🧠🧠🧠🧠🧠🧠🧠

These aren’t that different after all

Different populations are one thing. But what about different sexes? Chapter 4 will address the most striking genetic variant of them all.

Chapter 4: Sex, Gender, and Beyond

In humans, there is one gene that impacts observable variation more than any other, by far. It is equally common everywhere, so it doesn't cause differences *between* populations, just *within* populations. Shared across mammals, it represents the oldest genetic difference in our lineage and one of the oldest in the entire biosphere. Unlike typical DNA variants, it is associated with a major rearrangement of genome's basic structure. This gene is called *SRY*, and it is responsible for the sex chromosomes, X and Y.

🕺🐂🐏🆚🐑🐄💃
Vive la différence

The Y chromosome is basically a shriveled version of the X chromosome, and it *usually* contains *SRY*. A body with *SRY usually* develops male traits like testicles and facial hair. A person with testicles and facial hair *usually* identifies as male gender.

How it *usually* works:
🇾=***SRY***=🥎🥎🍆=🕺
🚫🇾=🚫***SRY***=🌷=💃

Notice all of the *usuallys*. That's because none of these steps are guaranteed. Sometimes *SRY* switches from the Y to the X. Sometimes bodies with *SRY* develop female features, or intersex features than are not unambiguously male or female. Sometimes a person's gender doesn't match what you might guess from looking at their body.

How it *sometimes* works:
🚫🇾=***SRY***=🥎🥎🍆=🕺
=De la Chapelle syndrome
🇾=🚫***SRY***=🌷=💃
=Swyer syndrome
🇾=***SRY***=🌷=💃
=Androgen insensitivity syndrome
🇾=***SRY***=🥎🥎,🚫🍆=🕺
=Aphallia
🇾=***SRY***=🥎🥎🍆=💃
=Trans woman
🚫🇾=🚫***SRY***=🌷=🕺
=Trans man
etc.

It gets even more complicated, because some people have combinations of chromosomes other than the typical XX or XY. Since gender isn't perfectly coupled to sex, any of these could occur in any gender:

👧:
🅇🅇 or 🅇🅈 or 🅇 or 🅇🅇🅈 or 🅇🅈🅈 or 🅇🅇🅇 or 🅇🅇🅈🅈
👦:
🅇🅇 or 🅇🅈 or 🅇 or 🅇🅇🅈 or 🅇🅈🅈 or 🅇🅇🅇 or 🅇🅇🅈🅈
🧒:
🅇🅇 or 🅇🅈 or 🅇 or 🅇🅇🅈 or 🅇🅈🅈 or 🅇🅇🅇 or 🅇🅇🅈🅈

And there is nothing special about the Y chromosome or *SRY* across the tree of life. Most species have completely unrelated sex-determining mechanisms, or don't even have two primary sexes at all.

Sexual systems
🐱🦟: 🅇 🅈 ♂ | 🅇 🅇 ♀
🐦🦋: 🅉 🅉 ♂ | 🅉 🅆 ♀
🟩(moss): 🅅 ♂ | 🅄 ♀
🐝: 🄰 ♂ | 🄰 🄰 ♀
🐊: 🔥 ♀ | 🌥️ ♂ | ❄️ ♀
🐌: ⚥
🍌: 🚫

Sex and gender aren't always cut and dry. But in the big picture, *most* people are male or female, and there seem to be really big differences between these two categories. Not only in which organs are present, but differences in size, strength, personality, behavior, etc. Of course these are merely trends, not absolutes. Clearly the world's weakest man is no match for the world's strongest woman. But are the *average* differences biologically real?

🕺 🆚 💃
Genetic or synthetic?

The answer is complicated. There are a lot of baseless stereotypes out there. But while myths about race tend to lack any biological justification at all, there are genuine sex differences grounded in biology. Sex hormones permeate the body and influence the development of tissues from head to toe.

❤️🧠🦴👁️💪💀👄
Some organs affected by sex hormones

But, remember the bicycles in Chapter 3. The impact of a genetic difference depends on the environment, which can change.

🚳🆚🚲
Whether you can ride depends on the bike not your genes

A century or two ago, it was almost unheard of for women to be successful doctors, scientists, engineers, or judges. Society didn't permit it. Many assumed that women were innately incapable of such tasks. These assumptions were wrong.

👩🏼‍⚕️👩🏾‍🔬👩🏽‍🔧👩🏻‍⚖️
You didn't see this in the olden days

Still today we see average differences between men and women. Some of these differences are influenced by *SRY* or other sex-linked genes. But, it would be foolish to assume that we have now removed all barriers to equality and levelled the playing field. The sensible prediction is that gaps between the sexes will continue to shrink. Still, some differences may be relatively immune to environmental adjustments. These could include differences in athletic achievement as well as brain differences, like what kind of person you are sexually attracted to. The real question is, are these differences large and robust enough to matter in any practical sense? Should we expect men and women to perform differently on average in academic or professional realms, solely because of biology? The answer is no.

👨🏾‍🔧👨🏼‍🔧👨🏿‍🔧👨🏽‍🔧👨🏾‍🔧👨🏻‍🔧👨🏿‍🔧👨🏻‍🔧👨🏿‍🔧👨🏽‍🔧👨🏼‍🔧👨🏻‍🔧👨🏿‍🔧👨🏽‍🔧
Don't blame biology if your organization is dominated by one gender

Here's why. Perhaps the strongest behavioral effect of *SRY* not directly tied to reproduction is its influence on aggression. Violent behavior is strongly correlated with *SRY*, much more than so-called "warrior gene" *MAOA*, or any other gene that isn't sex-linked. Now, imagine a society where most leadership positions were held by women.

👩🏿‍💼👮🏾‍♂️👩🏼‍💼👸🏽👩🏻‍🎓👩🏽‍💼👸🏿🦸🏾‍♀️👮🏼‍♂️👩🏻‍🏫
What if "The Man" weren't The Man?

From that vantagepoint, one could readily argue that men were not naturally suited to hold power, because *SRY* condemns them to angry irrationality. This argument would seem valid, but it would be wrong. We all know that men are fully capable of controlling their behavior. *SRY* is not an excuse to attack people. Even our legal system agrees.

🔪💣🔫🧨💥☠️
***SRY* didn't make you do it**

The effects of *SRY* on violence can be culturally and willfully overcome. Every law-abiding carrier of *SRY* knows this, and it should also be obvious from the vastly different crime rates seen within the same society but at different points in history. And if *that* major genetic effect can be toppled, what does that say for any weaker trends in behavior? Surely they are even less robust.

♂🧠⚖️♀🧠
Hard to find any meaningful imbalance in brain function between the sexes

And now, to wrap things up.

🔚
Coming soon

Reflect back across all four chapters. The guide has explained how genetic differences between human groups, be they races or genders, are not as stark as they are often portrayed. Many people find this conclusion hard to accept. There seems to be a strong human drive to create strict categories. However, a scientific mindset seeks to avoid such biases. Just think about the popularity of astrology. It's a pseudoscience. Hopefully you already know that a person's Zodiac sign is not a major determinant of their personality or potential. Don't then fall into the trap of thinking that race or sex is a good predictor of your brain's abilities.

♈♉♊♋♌♍♎♏♐♑♒♓
What's your sign? Humans love to categorize each other for no good reason

Here's another example. Which square below is darker? It probably looks like the left square is. But in fact, they are exactly the same. Whether it's shades of geometry or shades of humanity, our minds really like to magnify differences or event invent them out of thin air.

Nothing in biology is identical. So of course there are real differences between human populations, or between human genders. It would be weird if there weren't. But these differences are typically so tiny that they are hard to even measure accurately. Stereotypes aren't based on science, and if you look closely enough to detect a difference, it will often be in the opposite direction from what a stereotype would predict. In any case, the differences are too small to matter in the vast majority of situations. Consider your right and left legs. One must be slightly longer than the other, but do you even know which? And who cares? The important thing is that they work together.

🦵🆚🦵
Technically there's a real biological difference, but so what?

We don't yet fully understand every way that DNA shapes human lives. For complex traits like behavior, it is especially challenging. Suppose you hear about a study testing for a genetic difference in some ability between two groups of humans. The data will invariably be confounded by factors that can't be controlled. We can't do proper experiments on humans like we can with flowers, after all. So always ask yourself: what are the caveats?

👩🏽‍🔬🔎🗒️🌸🌺🌼**normal science (legal)**

🦹‍♂️🧪⚡👶🏽👶🏿👶🏻**mad science (illegal)**

This brings up a final, more philosophical point. With incomplete data, we make educated guesses. Precisely how different are two particular groups in some particular aspect? Whenever there is scientific uncertainty, it's best err on the side of caution. Overestimation of differences between human groups has led to some of the most horrific tragedies of history. Meanwhile, any negative consequences of underestimating these differences are slim. If you are driving through the fog and the GPS estimates that the edge of the cliff is either 50 or 60 feet away, both of those estimates are equally consistent with the data. But where do you stop the car?

🌫️🌫️🌫️🌫️🌫️🌫️🌫️🌫️🌫️🌫️🌫️🌫️🌫️🌫️🌫️

🌫️🌫️🌫️🌫️🌫️🌫️🌫️🚗🌫️🌫️🌫️🌫️🌫️🌫️🌫️

Human evolutionary genetics is an enormous field, and this guide is just the tip of the iceberg. Hopefully you've learned something. More importantly, hopefully you are inspired to learn more. Check out the References and Acknowledgments section for where to go next.

References and Acknowledgments

The inspiration for this guide grew out of conversations with journalist Amy Harmon, as well as her call for scientists to debunk myths about human genetics: https://www.nytimes.com/2018/10/18/insider/science-genetics-white-supremacy.html. The emoji format was pioneered on my Twitter account @biolojical, as well as my personal Twitter account @JacobPhD, and many of the emoji diagrams are lifted directly from tweets. I have had many productive conversations with my fellow biologists, too many to name, who have clarified my thinking on these topics. Specific ideas in this guide have been inspired or guided by Holly Dunsworth, April Wei, Josef Uyeda, and Melissa Wilson.

If you want to learn more, there have been several good books published recently about human genetics. *A Brief History of Everyone Who Ever Lived* by Adam Rutherford and *She Has Her Mother's Laugh* by Carl Zimmer are particularly interesting and accessible. Several professional evolutionary biologists regularly write fascinating articles for non-scientists on these topics which are freely available online. Check out the works of Jennifer Raff, Graham Coop, Holly Dunsworth, Jeremy Yoder, and others. Finally, reading the primary scientific literature can be challenging for people outside the field, but if you want to delve in, here are some papers supporting the statements made in this guide:

1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR. 2015. A global reference for human genetic variation. Nature. 526:68-74.

Corl A, Davis AR, Kuchta SR, Sinervo B. Selective loss of polymorphic mating types is associated with rapid phenotypic evolution during morphic speciation. 2010. Proc Natl Acad Sci USA. 107:4254-4259.

López Herráez D, Bauchet M, Tang K, Theunert C, Pugach I, Li J, Nandineni MR, Gross A, Scholz M, Stoneking M. 2009. Genetic variation and recent positive selection in worldwide human populations: evidence from nearly 1 million SNPs. PLoS One. 4:e7888.

Joron M, Frezal L, Jones RT, Chamberlain NL, Lee SF, Haag CR, Whibley A, Becuwe M, Baxter SW, Ferguson L, Wilkinson PA, Salazar C, Davidson C, Clark R, Quail MA, Beasley H, Glithero R, Lloyd C, Sims S, Jones MC, Rogers J, Jiggins CD, ffrench-Constant RH. 2011. Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. Nature 477:203-206.

Leffler EM, Bullaughey K, Matute DR, Meyer WK, Ségurel L, Venkat A, Andolfatto P, Przeworski M. 2012. Revisiting an old riddle: what determines genetic diversity levels within species? PLoS Biol. 10:e1001388.

McDermott R, Tingley D, Cowden J, Frazzetto G, Johnson DD. Monoamine oxidase A gene (MAOA) predicts behavioral aggression following provocation. 2009. Proc Natl Acad Sci USA. 106:2118-2123.

Norton HL, Quillen EE, Bigham AW, Pearson LN, Dunsworth H. 2019. Human races are not like dog breeds: refuting a racist analogy. Evolution: Education and Outreach 12:17.

Reich, D. 2018. Who We Are and How We Got Here: Ancient DNA and the New Science of the Human Past. Oxford University Press.

Rohde DL, Olson S, Chang JT. 2004. Modelling the recent common ancestry of all living humans. Nature. 431:562-566.

Staudt G. 1989. The species of *Fragaria*, their taxonomy and geographical distribution. Acta Hortic. 265:23–24.