Out of Africa: what skin color tells us about human evolution

Educator guide

PAPER DETAILS

Original title: Loci associated with skin pigmentation identified in African populations

Reference: Vol. 358, Issue 6365, eaan8433

Authors: Nicholas G. Crawford, Sarah A. Tishkoff

Issue name: *Science*

Original publication date: 11/17/2017

DOI: 10.1126/science.aan8433

Annotator(s): Allison Mayle, Beth Ruedi

TABLE OF CONTENTS

Science in the Classroom AAAS

LEARNING STANDARDS ALIGNMENT

Learning Performance: Students will examine genetic analyses and compare data sets that inform the genetic basis and evolution of skin pigmentation in African populations, exploring methods for determining the age of genetic variants and assessing which variants cause differences in skin color.

The following tables provide an overview of the learning standards covered by this article, including the A Framework for K-12 Science Education (Framework), Common Core State Standards English Language Arts-Literacy (CCSS ELA), Common Core State Standards Statistics and Probability (CCSS HSS), AP Science Practices, and Vision and Change for Undergraduate Education. Where applicable, activities and information will be marked with specific standards to which they are linked.

A Framework for K-12 Science Education		
Science and Engineering Practices	Disciplinary Core Ideas	Crosscutting Concepts
Developing and Using Models (SEP2) Develop and/or use multiple types of models to provide mechanistic accounts and/or predict phenomena, and move flexibly between model types based on merits and limitations. **Analyzing and Interpreting Data (SEP4)** Analyze data using tools, technologies, and/or models (e.g., computational, mathematical) in order to make valid and reliable scientific claims. Compare and contrast various types of data sets (e.g., self-generated, archival) to examine consistency of measurements and observations. Evaluate the impact of new data on a working explanation and/or model of a proposed process or system.	**LS4.A: Evidence of Common Ancestry and Diversity** Genetic information provides evidence of evolution. DNA sequences vary among species, but there are many overlaps; in fact, the ongoing branching that produces multiple lines of descent can be inferred by comparing the DNA sequences of different organisms. Such information is also derivable from the similarities and differences in amino acid sequences and from anatomical and embryological evidence.	**Cause and Effect** Empirical evidence is required to differentiate between cause and correlation and make claims about specific causes and effects. **Stability and Change** Much of science deals with constructing explanations of how things change and how they remain stable. Change and rates of change can be quantified and modeled over very short or very long periods of time. Some system changes are irreversible.

Science in the Classroom AAAS

Common Core State Standards English Language Arts-Literacy		
Key Ideas and Details	Craft and Structure	Integration of Knowledge and Ideas
RST.9-10.1 Cite specific textual evidence to support analysis of science and technical texts, attending to the precise details of explanations or descriptions. **RST.9-10.2** Determine the central ideas or conclusions of a text; trace the text's explanation or depiction of a complex process, phenomenon, or concept; provide an accurate summary of the text. **RST.11-12.1** Cite specific textual evidence to support analysis of science and technical texts, attending to important distinctions the author makes and to any gaps or inconsistencies in the account. **RST.11-12.2** Determine the central ideas or conclusions of a text; summarize complex concepts, processes, or information presented in a text by paraphrasing them in simpler but still accurate terms.	**RST.9-10.4** Determine the meaning of symbols, key terms, and other domain-specific words and phrases as they are used in a specific scientific or technical context relevant to grades 9-10 texts and topics. **RST.9-10.5** Analyze the structure of the relationships among concepts in a text, including relationships among key terms (e.g., force, friction, reaction force, energy). **RST.9-10.6** Analyze the author's purpose in providing an explanation, describing a procedure, or discussing an experiment in a text, defining the question the author seeks to address. **RST.11-12.4** Determine the meaning of symbols, key terms, and other domain-specific words and phrases as they are used in a specific scientific or technical context relevant to grades 11-12 texts and topics. **RST.11-12.5** Analyze how the text structures information or ideas into categories or hierarchies, demonstrating understanding of the information or ideas. **RST.11-12.6** Analyze the author's purpose in providing an explanation, describing a procedure, or discussing an experiment in a text, identifying important issues that remain unresolved.	**RST.9-10.8** Assess the extent to which the reasoning and evidence in a text support the author's claim or a recommendation for solving a scientific or technical problem. **RST.9-10.9** Compare and contrast findings presented in a text to those from other sources (including their own experiments), noting when the findings support or contradict previous explanations or accounts. **RST.11-12.8** Evaluate the hypotheses, data, analyses, and conclusions in a science or technical text, verifying the data when possible and corroborating or challenging conclusions with other sources of information. **RST.11-12.9** Synthesize information from a range of sources (e.g., texts, experiments, simulations) into a coherent understanding of a process, phenomenon, or concept, resolving conflicting information when possible.

Science in the Classroom AAAS

AP Science Standards	
AP Science Practices	AP Biology Content Standards
Science Practice 1 (SP1) The student can use representations and models to communicate scientific phenomena and solve scientific problems. **Science Practice 6 (SP6)** The student can work with scientific explanations and theories, construct explanations of phenomena based on evidence, articulate the reasons that scientific explanations and theories are refined or replaced, and evaluate alternative scientific explanations.	**Essential knowledge 1.B.1 (EK1.B.1)** Organisms share many conserved core processes and features that evolved and are widely distributed among organisms today. Structural and functional evidence supports the relatedness of all domains. **Essential knowledge 1.C.3 (EK1.C.3)** Populations of organisms continue to evolve. Scientific evidence supports the idea that evolution continues to occur. **Essential knowledge 3.A.4 (EK3.A.4)** The inheritance pattern of many traits cannot be explained by simple Mendelian genetics. Many traits are the product of multiple genes and/or physiological processes.

Connections to the Nature of Science	
Vision and Change for Undergraduate Biology Education Core Competencies and Disciplinary Practices	A Framework for K-12 Science Education Understandings About the Nature of Science
Ability to apply the process of science Understand that science is evidence based and grounded in the formal practices of observation, experimentation, and hypothesis testing **Ability to use modeling and simulation** Understand how mathematical and computational tools to describe complex living systems	**Scientific Investigations Use a Variety of Methods** Science investigations use diverse methods and do not always use the same set of procedures to obtain data. Scientific investigations use a variety of methods, tools, and techniques to revise and produce new knowledge. **Science Models, Laws, Mechanisms, and Theories Explain Natural Phenomena** Models, mechanisms, and explanations collectively serve as tools in the development of a scientific theory.

ARTICLE OVERVIEW

Article summary (recommended for educator use only)

Skin pigmentation in humans is highly variable across populations. This variation may be due to adaptation to local environments. Using a diverse array of genetic techniques, the authors examined skin pigmentation variation across African populations. They could identify four major regions within the genome that were associated with differences in skin pigmentation. Using model systems, they were also able to specify the functionality of several of these regions, as well as determine the age of specific genetic variants. Using these tools, they reveal some of the evolutionary history of skin pigmentation in humans.

Importance of this research

Despite the overwhelming evidence that our common ancestors were from Africa, prior research examining the genetic basis for differences in human skin pigmentation has focused mostly on European populations. The authors of this article wanted to specifically examine genetic variation in skin pigmentation between African populations to better understand the evolution of skin pigmentation in humans. The results of this study provide a more complete picture of the genetic basis of adaptation in humans, something that will help inform future work. In addition, this is a reminder that scientists must examine *all* populations in order to make informed inferences about the evolutionary history of our species.

Experimental methods

- **Phenotyping** (**skin pigmentation assessment**): Researchers used a handheld device, the DSM II ColorMeter, to quantify skin pigmentation across populations in Ethiopia, Tanzania, and Botswana. They also took blood samples from many of these individuals for DNA analysis, and made note of the age, gender, and familial relationships of all subjects.

- **Genome-wide association study** (**GWAS**): DNA was extracted from blood samples and genotyped—in other words, examined for single nucleotide polymorphisms (SNPs), which are changes to a single base at specific locations along the genome. The authors did this using whole-genome genotyping, and examined if any differences in skin pigmentation were associated with particular SNP variants. To ensure that there weren't other factors confusing the associations, they accounted for differences due to age, sex, and familial relationships across the data set. They identified four regions that had highly significant association of DNA variants with skin color. To narrow down the specific loci in the regions and verify which SNPs had the strongest association, they performed genome sequencing on a smaller subset of DNA samples, as well as comparing their results with a different whole-genome sequencing data set available to the public.

- **Functional assessment of regions of interest**: Once specific SNPs were identified as having a significant association with skin pigmentation, the authors needed to verify that the SNPs were indeed involved processes leading to skin pigmentation. They used various methods to explore how each gene functioned.
    - ***Gene expression analysis***: The authors wanted to identify which variants were "causal"—in other words, which SNPs caused a change, or were near a locus that caused a change, in gene expression or protein function. They did this both by statistical assessment of the whole genome data set, and by conducting further experiments using human melanocyte cell lines. Melanocytes are cells that produce melanin.

- ***Functional characterizations***: The authors used cell lines and model organisms to examine the function of one of the candidate regions: *MFSD12*
  - *Mfsd12* gene function was knocked down in human melanocyte cell lines; melanin content in the cells was then quantified to examine changes in pigment
  - CRISPR/Cas-9, a method for editing DNA, was used to target *mfsd12a* in zebrafish and "knockdown" its production of protein. Then, fish were stained and examined by microscope for any differences in wild-type and knockdown strains.
  - CRISPR/Cas-9 was also used to delete a region of *Mfsd12* in mice; this is sometimes referred to as gene "knockout." Hairs were plucked from wild-type and knockout mice and examined for pigment differences under the microscope.

- **Evolution and population genetics**: Using the newly discovered variants associated with differences in skin pigmentation coupled with the rest of the whole-genome data set, the authors used several statistical methods to examine evolutionary history and human population dynamics. These methods included:

  - ***Tajima's D***: A statistical method to distinguish between a DNA sequence having evolved randomly (under neutral selection), or having evolved nonrandomly (undergoing directional selection or balancing selection).
    - Directional selection occurs when one trait is favored over others, leading to a shift in allele frequency toward the allele associated with the favored trait. This also leads to an excess of low-frequency alleles, and a negative value of Tajima's D.
    - Balancing selection is a process by which multiple alleles are maintained in the population, for example when heterozygotes are at an advantage (which keeps both of their alleles at a relatively high frequency). This leads to a lack of rare alleles, and a positive Tajima's D.
    - Tajima's D statistic of approximately zero indicates that the population is evolving with little or no selection.

  - ***Coalescent analysis***: A mathematical model used to determine how long ago two variants of the same gene may have arisen from a common ancestor. This model also allows researchers to determine which variant is most likely the ancestral variant, and which is newer, or "derived."

  - ***Haplotype analysis***: DNA variants can be inherited together more frequently than we would expect by chance; this is known as linkage disequilibrium. This phenomenon is useful for helping us understand population mixing, ancestry, and evolutionary history. Haplotype analysis uses sets of variants in linkage disequilibrium across the genome to investigate a range of factors that have led to evolution (ex: natural selection, genetic drift, population bottlenecks, inbreeding) and to help put newly discovered genetic variants (like the SNPs from this paper) in context. Humans have distinct sets of "haplotype blocks," meaning groups of nonoverlapping loci in strong linkage disequilibrium. That has been extremely helpful in human genetic analyses.

  - ***Allele frequency patterns***: Examining the different frequencies, or occurrences, of specific variants within subpopulations of humans helped the authors understand which alleles were prevalent in which geographic locations, leading to a better understanding of why certain variants might be more common in specific areas.

Conclusions

- Four regions of the human genome associated significantly with differences in skin pigmentation. Within these regions, eight different single nucleotide polymorphisms were identified as causal candidate variants. Together, these eight SNPs accounted for nearly 30% of the heritable variation in skin pigmentation in African populations. Skin pigmentation is considered a "complex" trait because it is not controlled by a single gene. However, the genetic basis of skin pigmentation seems to be simpler than other complex traits, like height.

- The regions identified were near the following genes, listed by magnitude of variation in skin pigmentation that each region was associated with: *SLC24A5*, *MFSD12*, *DDB1/TMEM138*, and *OCA2/HERC2*. These all seem to affect multiple phenotypes in humans. In terms of molecular function, many of the DNA variants in these regions appear to be "regulatory," or controlling expression of other genes in the genome. The authors demonstrated causal effects of the variants on skin pigmentation. Some were involved in production of yellow pigmentation; others were involved in production of dark-brown/black pigmentation.

- The genes and variants identified in this paper have diverse functions, including UV damage repair and melanocyte (cells that produce melanin) biology.

- Both dark and light pigmentation variants arose before the origin of modern humans, and light and dark pigmented skin is still evolving. Variants associated with dark pigmentation in Africa are identical by descent (due to common ancestry) to variants in South Asia and indigenous Malaysian, Philippine, Indonesian, and Australian populations, giving us a better idea of the migratory patterns of humans.

Science in the Classroom AAAS

ACTIVITIES FOR INTERACTIVE ENGAGEMENT

Learning Performance: Students will examine genetic analyses and compare data sets that inform the genetic basis and evolution of skin pigmentation in African populations, exploring methods for determining the age of genetic variants and assessing which variants cause differences in skin color.

Writing an abstract Students write a new abstract for the article at a grade-appropriate reading level.	**RST.9-10.2** **RST.11-12.2** **Nature of Science**
Locating this study in the larger field Students use the annotated list of references to explain how this research builds on the published work of at least one other independent group of scientists. Students will evaluate whether data from this research supports or contradicts previous conclusions, and reflect on the statement that scientific knowledge is a "community effort."	**RST.9-10.8** **RST.11-12.8** **Nature of Science**
Science in the news Students explore news stories in the Related Resources tab and evaluate the stories for tone, accuracy, missing information, etc. They may then write their own news stories on the article.	**RST.9-10.5** **RST.11-12.5** **RST.9-10.6** **RST.11-12.6** **RST.9-10.8** **RST.11-12.8**
Mapping human migration Students explore theories about the patterns of human migration by searching literature reviews. In groups or as a class, students decide which migration map(s) seems to have the strongest consensus within the scientific community. Students then add in the migratory patterns that the current paper seems to indicate (or reinforce). The authors describe the directionality of gene flow of different genetic variants throughout the paper (ex: one light pigmentation variant first occurred outside of Africa, and then appears to have been reintroduced through population mixing).	**SEP4** **SP6** **LS4.A** **EK1.B.1** **EK1.C.3**
Results and conclusions Students diagram each of the experiments presented in the study (divided up by figure, if appropriate). They then consider the results depicted in each figure, and how these results support the conclusions of the study.	**SEP2** **SEP4** **SP1** **LS4.A** **Nature of Science**
The next steps Students design a follow-on experiment to this study that either addresses flaws or unanswered questions in the research at hand, or builds on it to explore a new question.	**SEP4** **SP6** **Nature of Science**

DISCUSSION QUESTIONS

1. What were the model systems used to explore the function of genetic variants? Why did the authors decide to use multiple model organisms? Could this work be done in humans?	**SEP2** **LS4.A** **SP1** **EK1.B.1**
2. What genetic analyses and statistical tests did the authors use in this study? Why were so many different methods used to explore the genetic basis for African skin pigmentation? How did each new result work to enhance or further develop the initial understanding of genetic variation leading to differences across populations?	**SEP6** **SP4** **Cause and Effect** **EK3.E.2** **Nature of Science**
3. How was it possible for the authors to use whole-genome data to determine how long ago specific genetic variants most likely arose? How could they identify the "ancestral" and "derived" variants?	**LS4.A** **Stability and Change** **EK1.B.1**
4. With the age of the different genetic variants the authors discovered, what did they determine about the evolution of ancestral populations and those of modern human populations? Do you think that human populations are still evolving? Would it be possible to witness adaptation in humans?	**Stability and Change** **EK1.C.3**
5. Geneticists often refer to traits as "simple" or "complex." What are they referring to—do these words mean the same thing in genetics as they do in other situations? How did the authors identify skin pigmentation as either a "simple" or "complex" trait? Think about how you were originally taught the principles of genetics—for example, with Mendel's peas. How does that compare with skin pigmentation? How many traits do you think are simple vs. complex?	**SP6** **EK3.A.4**
6. Why do you think that previous studies investigating skin pigmentation in humans were limited primarily to European populations? Do you think there was a scientific reason for the focus on those populations? What made the current authors decide to investigate African populations?	**Nature of Science**