

# UC Santa Barbara

## UC Santa Barbara Previously Published Works

### Title

Applying an evolutionary mismatch framework to understand disease susceptibility.

### Permalink

<https://escholarship.org/uc/item/5919v104>

### Journal

PLoS Biology, 21(9)

### Authors

Lea, Amanda

Clark, Andrew

Dahl, Andrew

et al.

### Publication Date

2023-09-01

### DOI

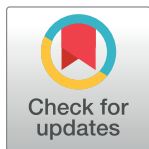
10.1371/journal.pbio.3002311

Peer reviewed

## ESSAY

# Applying an evolutionary mismatch framework to understand disease susceptibility

Amanda J. Lea<sup>1\*</sup>, Andrew G. Clark<sup>2</sup>, Andrew W. Dahl<sup>3</sup>, Orrin Devinsky<sup>4</sup>, Angela R. Garcia<sup>5</sup>, Christopher D. Golden<sup>6</sup>, Joseph Kamau<sup>7</sup>, Thomas S. Kraft<sup>8</sup>, Yvonne A. L. Lim<sup>9</sup>, Dino J. Martins<sup>10</sup>, Donald Mogoi<sup>11</sup>, Päivi Pajukanta<sup>12</sup>, George H. Perry<sup>13</sup>, Herman Pontzer<sup>14,15</sup>, Benjamin C. Trumble<sup>16,17</sup>, Samuel S. Urlacher<sup>18</sup>, Vivek V. Venkataraman<sup>19</sup>, Ian J. Wallace<sup>20</sup>, Michael Gurven<sup>21</sup>, Daniel E. Lieberman<sup>22</sup>, Julien F. Ayroles<sup>23,24\*</sup>



**1** Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee, United States of America, **2** Department of Molecular Biology & Genetics, Cornell University, Ithaca, New York, United States of America, **3** Department of Medicine, University of Chicago, Chicago, Illinois, United States of America, **4** Department of Neurology, NYU Langone Comprehensive Epilepsy Center, NYU Grossman School of Medicine, New York, New York, United States of America, **5** Department of Anthropology, Stanford University, Stanford, California, United States of America, **6** Department of Nutrition, Harvard T H Chan School of Public Health, Boston, Massachusetts, United States of America, **7** One Health Centre, Institute of Primate Research, Karen, Nairobi, Kenya, **8** Department of Anthropology, University of Utah, Salt Lake City, Utah, United States of America, **9** Department of Parasitology, Faculty of Medicine, Universiti Malaya, Kuala Lumpur, Malaysia, **10** Turkana Basin Institute, Stony Brook University, Stony Brook, New York, United States of America, **11** Department of Medical Services and Public Health, Ministry of Health Laikipia County, Nanyuki, Kenya, **12** Department of Human Genetics, David Geffen School of Medicine at UCLA, University of California Los Angeles, Los Angeles, California, United States of America, **13** Departments of Anthropology and Biology, The Pennsylvania State University, University Park, Pennsylvania, United States of America, **14** Department of Evolutionary Anthropology, Duke University, Durham, North Carolina, United States of America, **15** Duke Global Health Institute, Duke University, Durham, North Carolina, United States of America, **16** School of Human Evolution and Social Change, Arizona State University, Tempe, Arizona, United States of America, **17** Center for Evolution and Medicine, Arizona State University, Tempe, Arizona, United States of America, **18** Department of Anthropology, Baylor University, Waco, Texas, United States of America, **19** Department of Anthropology and Archaeology, University of Calgary, Calgary, Alberta, Canada, **20** Department of Anthropology, University of New Mexico, Albuquerque, New Mexico, United States of America, **21** Department of Anthropology, University of California Santa Barbara, Santa Barbara, California, United States of America, **22** Department of Human Evolutionary Biology, Harvard University, Cambridge, Massachusetts, United States of America, **23** Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, New Jersey, United States of America, **24** Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey, United States of America

## OPEN ACCESS

**Citation:** Lea AJ, Clark AG, Dahl AW, Devinsky O, Garcia AR, Golden CD, et al. (2023) Applying an evolutionary mismatch framework to understand disease susceptibility. *PLoS Biol* 21(9): e3002311. <https://doi.org/10.1371/journal.pbio.3002311>

**Published:** September 11, 2023

**Copyright:** © 2023 Lea et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** A.J.L. was funded by grants from the Canadian Institute for Advanced Research, the Searle Scholars Program, and National Institutes of Health (NIGMS - R35-GM147267). J.F.A. is funded by NIH (NIEHS R01ES029929 and NIGMS R35-GM124881). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

**Abbreviations:** CVD, cardiovascular disease; GxE, genotype by environment; LMIC, low- and middle-income country; NCD, noncommunicable disease; PRS, polygenic risk score; SNP, single nucleotide polymorphism.

☞ These authors contributed equally to this work.

\* [amanda.j.lea@vanderbilt.edu](mailto:amanda.j.lea@vanderbilt.edu) (AJL); [jayroles@princeton.edu](mailto:jayroles@princeton.edu) (JFA)

## Abstract

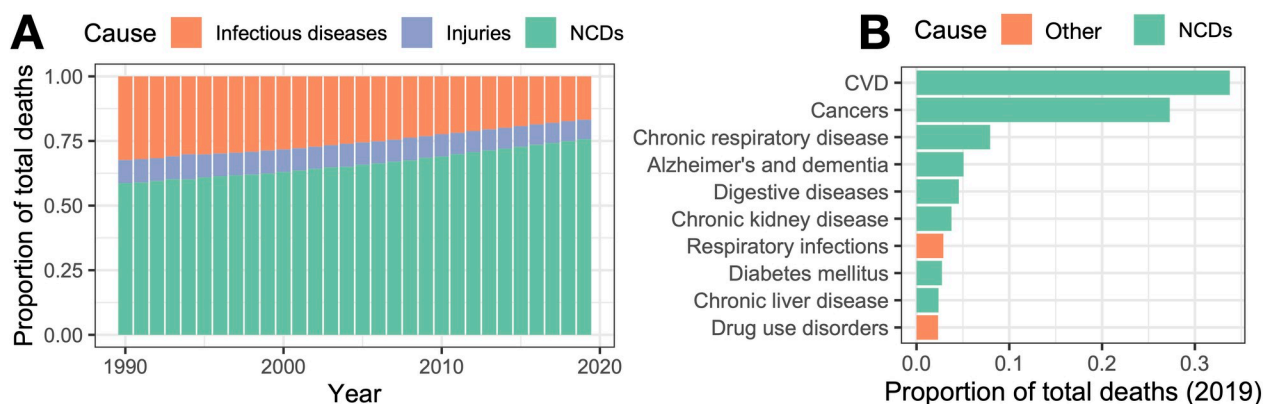
Noncommunicable diseases (NCDs) are on the rise worldwide. Obesity, cardiovascular disease, and type 2 diabetes are among a long list of “lifestyle” diseases that were rare throughout human history but are now common. The evolutionary mismatch hypothesis posits that humans evolved in environments that radically differ from those we currently experience; consequently, traits that were once advantageous may now be “mismatched” and disease causing. At the genetic level, this hypothesis predicts that loci with a history of selection will exhibit “genotype by environment” (GxE) interactions, with different health effects in “ancestral” versus “modern” environments. To identify such loci, we advocate for combining

genomic tools in partnership with subsistence-level groups experiencing rapid lifestyle change. In these populations, comparisons of individuals falling on opposite extremes of the “matched” to “mismatched” spectrum are uniquely possible. More broadly, the work we propose will inform our understanding of environmental and genetic risk factors for NCDs across diverse ancestries and cultures.

## Introduction

Noncommunicable diseases (NCDs) such as cardiovascular disease (CVD), type 2 diabetes, and Alzheimer’s disease are among the leading causes of death worldwide (Fig 1). NCDs are often difficult to prevent and treat, because they result from complex and poorly understood interactions between a person’s genetic makeup and their environment. For example, CVD has a heritability of 40% to 50%, with dozens of loci now mapped through genome-wide association studies [1–3]. However, when tallied together in an additive framework, these loci explain only a small fraction of the heritable genetic effect. This has led many to conclude that environmental risk factors, such as a diet high in processed foods and low levels of physical activity, interact with genetic variation to shape NCD risk [4,5]. In other words, genetic variation may predispose individuals toward physiological sensitivity or resilience in the face of environmental perturbations, a phenomenon known as “genotype by environment” (GxE) interactions.

Despite major interest in GxE interactions in the context of NCDs, scientists have struggled in practice to identify them. There are many reasons for this, including that the relevant environmental factors are often unknown, difficult to measure, or minimally variable within the study population (e.g., most individuals in postindustrial contexts consume processed foods). Further, large sample sizes are needed to test for interaction effects, and even more so to overcome the multiple testing burden incurred by testing for interactions between many genetic variants and many environments [6,7]. To overcome power issues, current state-of-the-art approaches have leveraged very large studies such as the UK Biobank to scan for interactions between genome-wide genetic variation and selected lifestyle factors (e.g., smoking, diet, or physical activity) [8–11]. However, these studies have not delivered as expected and have only



**Fig 1. Noncommunicable diseases are the leading cause of death worldwide.** (A) Proportion of worldwide deaths attributable to noncommunicable diseases (NCDs), communicable or infectious diseases, and injuries through time. (B) Proportion of deaths within the USA in 2019, broken down by the top 10 causes of death. NCDs are highlighted in green. For both panels, data were sourced from [ourworldindata.org](https://ourworldindata.org) and represent all ages.

<https://doi.org/10.1371/journal.pbio.3002311.g001>

uncovered a handful of GxE interactions for NCDs such as obesity, type 2 diabetes, and depression.

In this Essay, we argue for a complementary approach informed by anthropological methods, genomic tools, and evolutionary theory. In particular, we believe there is much to learn by viewing GxE interactions through the lens of the “evolutionary mismatch” hypothesis and by partnering with genetically and environmentally diverse small-scale, subsistence-level populations to map them. The evolutionary mismatch hypothesis posits that traits that evolved under past selection regimes are often imperfectly or inadequately suited to modern environments, leading to “mismatches” in the form of NCDs [12–16]. At the genetic level, we would thus expect that previously neutral or beneficial alleles are now disease causing.

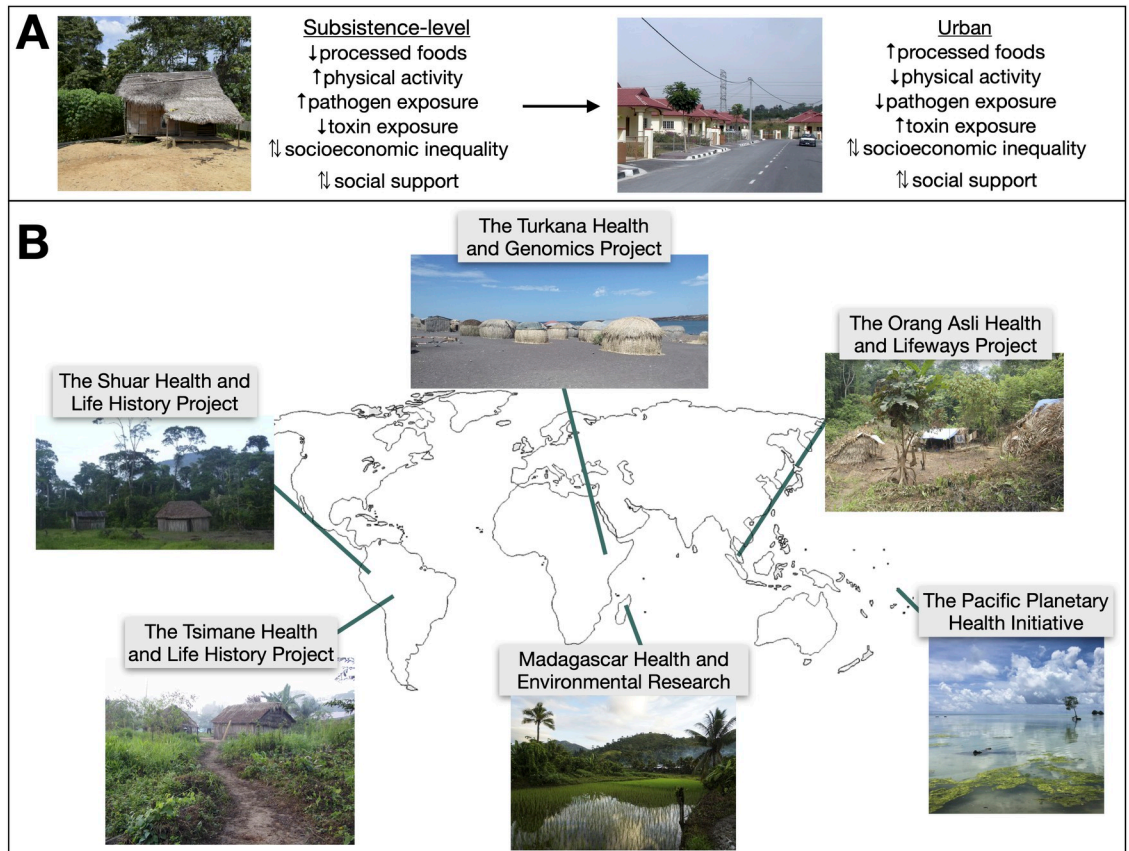
While we cannot go back in time to evaluate genotype–phenotype relationships in past environments, we can collaborate with populations that practice nonindustrial, subsistence-level lifestyles and thus fall further toward the “matched” end of the matched–mismatched spectrum than individuals in postindustrial contexts (though we caution that, of course, no modern population is perfectly representative of their evolutionary past). Further, many subsistence-level populations are currently exposed to globalizing forces causing rapid environmental shifts; this situation creates a quasi-natural experiment for studying the transition from traditional to modern lifeways within a single group [17] (Fig 2A). Additionally, the ecology and culture of many subsistence-level groups has already been well characterized through long-term work with anthropologists (Fig 2B), setting the stage for integration of genomic studies.

Uniting an evolutionary mismatch framework, long-term anthropological work with subsistence-level groups, and cutting-edge genomic tools can increase our power to identify and understand GxE interactions. Specifically, because the mismatch framework provides clear expectations for the types of loci and environments we expect to affect NCDs, we can narrow the search space considerably. Further, by focusing on populations where Western diets and lifestyles are the exception rather than the norm, we can design studies that explicitly sample environmental extremes, thereby boosting power. Finally, by studying many genetically distinct populations under a uniting intellectual framework, we can identify new loci that have so far been invisible to studies focused on individuals of European descent. With these goals in mind, we first review the evolutionary mismatch hypothesis and discuss its current support at the phenotypic and genetic levels. Second, we propose recommendations for integrating mismatch principles with molecular and genomic techniques, focusing on collaborations with subsistence-level groups. Third, we discuss the payoffs for scientists and study communities that would come from implementing these partnerships.

## Overview of the evolutionary mismatch hypothesis

An evolutionary mismatch is a condition that is more common or severe in an organism because it is imperfectly or inadequately adapted to a novel environment [27]. While mismatches are not unique to humans, their frequency may be unusually high in our species. This is because human culture can generate rapid and profound environmental change: In just a few generations, industrialization has transformed human diets, physical activity patterns, and toxin exposure landscapes, and these changes presumably contribute to the long list of NCDs that used to be rare or nonexistent [28–30].

For at least a century, a wide range of conditions have been assumed to be “diseases of civilization” or “lifestyle diseases” [31,32], but mismatches need to be explicitly and rigorously tested according to 3 criteria [33]. First, a mismatch condition should be more common or severe in the “novel” (e.g., postindustrial) relative to the “ancestral” environment (or correlated

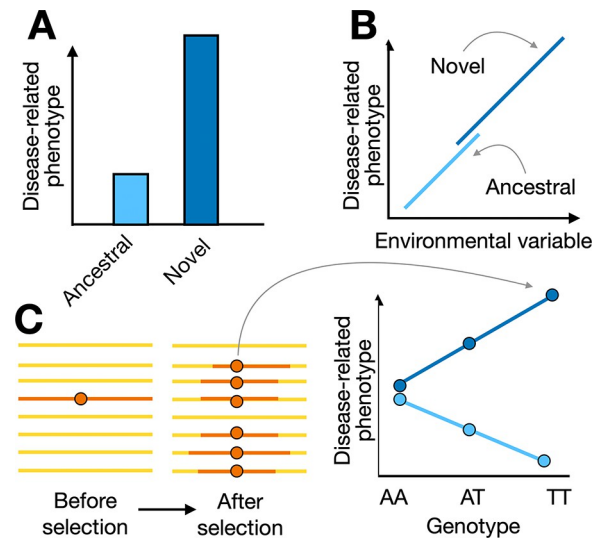


**Fig 2. Subsistence-level groups experiencing lifestyle change are a potential model for uncovering GxE interactions.** (A) Subsistence-level groups faced with urbanization, market integration, and modernization experience extreme variation in diet and physical activity levels, pathogen and toxin exposures, and social conditions. This list of environmental components for which there is extreme variation is not exhaustive and, in many cases, will also be population specific. We highlight a few broad categories that tend to change consistently during lifestyle transitions. Bidirectional arrows indicate factors that could either increase or decrease during urban transitions. (B) Studies such as The Turkana Health and Genomics Project [18,19], The Orang Asli Health and Lifeways Project [20], The Pacific Planetary Health Initiative, Madagascar Health and Environmental Research [21–23], The Tsimane Health and Life History Project [24], and The Shuar Health and Life History Project [25,26] all combine anthropological and biomedical data collection in transitioning societies and are thus poised to uncover GxE interactions in the context of evolutionary mismatch. We note that this list is meant to be illustrative and only includes projects directed by authors of this Essay; it does not by any means cover all of the rich and ongoing projects of small-scale, subsistence-level groups.

<https://doi.org/10.1371/journal.pbio.3002311.g002>

with some continuous metric of novel versus ancestral; Fig 3A). Small-scale, subsistence-level societies typically stand in as the best available proxy for the “ancestral” environment. This is because modern subsistence-level societies on average experience a closer “match” between their recent evolutionary history and their current environments relative to individuals in postindustrial contexts, though we caution they are not themselves “ancestral” populations. We also caution that modern subsistence-level groups (or any human group) will never be perfectly matched to their ancestral conditions on any time scale, given the near-constant fluctuations in human cultures, ecologies, and lifestyles. Nevertheless, these populations do all experience major environmental components consistent with the human evolutionary past, namely, they eat diets devoid of, or low in, processed foods, engage in high levels of physical activity, and never or rarely experience medical intervention.

In addition to the hypothesized mismatch condition being more prevalent in postindustrial versus subsistence-level groups, the second criterion is that it should also be tied to some environmental variable that differs between these groups (Fig 3B). One complication for achieving



**Fig 3. Mismatch diseases must be tested according to 3 criteria.** (A) Disease-related phenotypes must be more common or severe in the novel versus ancestral environment. We note that here we show mean differences in the phenotype between environments, but environmental effects could also impact trait variance. (B) These disease-related phenotypes must be attributable to an environmental variable, which will most often differ in mean and range between groups (e.g., physical activity influences cardiovascular health and is consistently higher in subsistence-level groups relative to individuals in postindustrial contexts). (C) It is necessary to establish a mechanism by which an environmental shift generates variation in disease-related phenotypes. At the genetic level, this could manifest as a locus for which a variant exhibits a past history of positive selection and is associated with health benefits in the ancestral environment but health detriments in the novel environment. A single locus with opposing effects is shown here for simplicity, but in reality, most complex traits will have highly polygenic architectures and diverse patterns of GxE interactions [34]. In panel C, horizontal lines represent haplotypes and the dark orange circle represents the selected variant. In all panels, dark blue represents the novel environment and light blue represents the ancestral environment.

<https://doi.org/10.1371/journal.pbio.3002311.g003>

this is that NCDs arise from complex multifactorial causes, and thus, while between-population comparisons are necessary, they can be confounded by many covariates that must also be taken into account (e.g., sanitation, access to medical care, or age structure, given that risk for most NCDs increases with age in postindustrial contexts [35]).

The third criterion is that it is necessary to establish a molecular or physiological mechanism by which the environmental shift generates the proposed mismatch condition. At the genetic level, this should manifest as a locus for which a variant exhibits a past history of positive selection and is associated with health benefits in the ancestral environment but health detriments in the novel environment, or one for which past stabilizing selection has created a situation where 2 intermediate alleles have similar fitness in the ancestral environment, but one allele becomes associated with health detriments in the novel environment (Fig 3C and Box 1).

### Box 1. GxE interactions in population genetics: Definitions and related concepts

In population genetics, the simplest conceptualization of a GxE interaction involves 3 genotypes for a single biallelic locus, with each of the 3 genotypes found in 2 different environments and with fitnesses varying across these 6 conditions (Fig 3C). At equilibrium, this population will harbor, among other types of genetic variation, alleles that have been selected to high frequency as a consequence of directional selection (i.e.,

selection on a trait value in a particular direction), and alleles that are at intermediate frequency as a consequence of stabilizing selection (i.e., selection to keep trait values near an optimum). If the environment changes quickly, previously selected alleles may now be associated with a trait that is no longer beneficial, and even disease causing, but will remain at a high frequency for some time before selection is able to purge them.

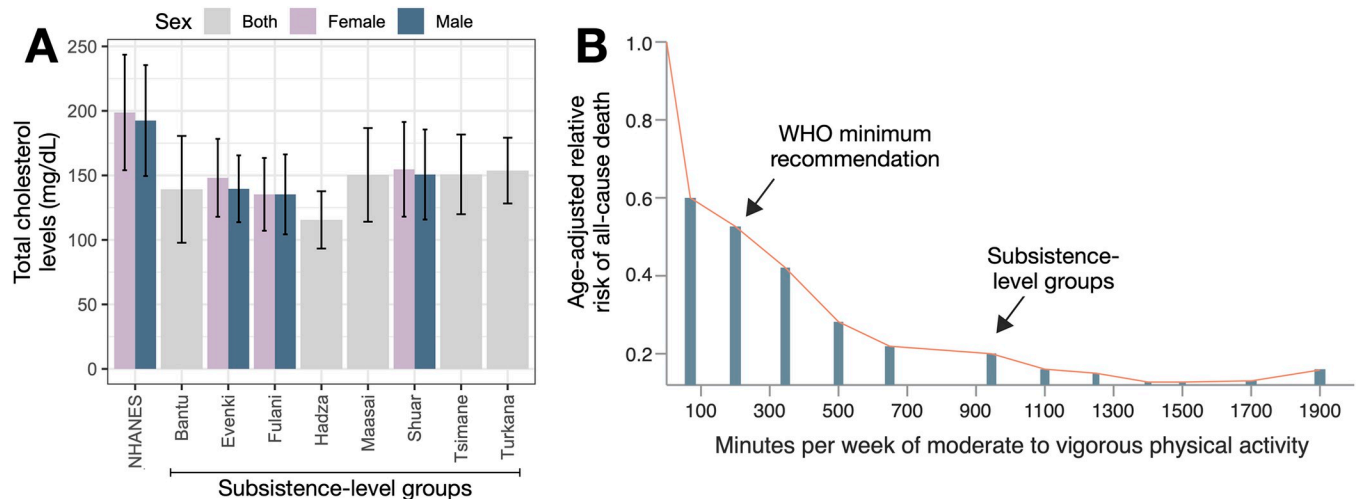
A few notes are important on this simple thought example. First, loci with no genetic variation (e.g., fixed beneficial mutations) could still be involved in mismatches in the new environment, but in the absence of genetic variation, we will be unable to identify them. Second, most complex traits have highly polygenic architectures, and while our simple examples (here and throughout) have focused on a single biallelic locus, the same logic applies under polygenicity [36]. Third, stabilizing selection is thought to be the most common mode of evolution shaping complex traits [37], and, thus, mismatch scenarios involving alleles that have previously undergone stabilizing selection may be the most common.

In addition to GxE interactions, a quantitative genetic concept relevant to evolutionary mismatch is “decanalization” [16,38]. Canalization refers to the process of stabilizing selection that selects for trait values that closely track some optimum in a given environment. However, in the presence of rapid environmental change or other strong perturbations, the optimum can shift and lead to decanalization [39]. While canalization acts to decrease genetic and phenotypic variance in a trait over time, decanalization involves an increase in the trait’s variance that is generally thought to be associated with the unmasking of loci that only impact the trait in the new environment [40]. Decanalization can thus be thought of as a specific form of evolutionary mismatch. Evolutionary mismatch can occur without having a previously canalized trait and is a more general term not necessarily linked to stabilizing selection. A final term that is distinct from all of these is “robustness.” Robustness refers to a property of individual genotypes, wherein they are able to retain an advantageous phenotype despite genetic or environmental hazards [39]. In contrast, evolutionary mismatch and decanalization are population-level phenomena.

## Current evidence for evolutionary mismatch at the phenotypic level

Scientists have been relatively successful at testing the first 2 criteria for mismatch, especially in the context of CVD, the single largest cause of mortality worldwide [41]. In support of the first criteria, subsistence-level groups experience remarkably low rates of CVD [30,42,43] relative to individuals in postindustrial contexts, as well as minimal age-associated increases in CVD or its biomarkers (e.g., hypertension, cholesterol) [44–46] (Fig 4A). Studies of small-scale societies in the midst of socioeconomic transition have demonstrated within-population effects of industrialization [18,47,48], strengthening the findings from between-population comparisons.

In support of the second criteria, recent work has also isolated salient environmental changes by which industrialization promotes CVD. People in subsistence-level communities are generally very physically active, accruing 5 to 10 times more daily physical activity than adults in postindustrial contexts [52,53]. Moderate to vigorous physical activity increases nitric oxide production and arterial elasticity [54,55] and reduces inflammation, all of which are protective against CVD [56]. Within industrialized populations, individuals accruing daily physical activity similar to those of subsistence-level individuals experience similarly low rates of



**Fig 4. Evidence for evolutionary mismatch at the phenotypic level.** (A) Mean levels of total cholesterol are much lower in selected subsistence-level populations relative to adults in the USA (>18 years old) profiled as part of the US National Health and Nutrition Examination Survey (NHANES) [49] (subsistence-level data from [17]). (B) Evidence that, within industrialized populations, individuals accruing daily physical activity similar to those of men and women in subsistence-level societies experience similarly low rates of CVD as well as all-cause mortality from NCDs. Dose-response relationship between minutes/week of moderate to vigorous leisure time physical activity and age-adjusted relative risk of death from a sample of 661,137 adults from the USA and Europe [50]. The arrow for physical activity estimates in subsistence-level groups is based on studies of Hadza individuals (estimated at  $x = 944$  minutes [43]) and Tsimane individuals ( $x = 924$  minutes [51]).

<https://doi.org/10.1371/journal.pbio.3002311.g004>

CVD, as well as NCD-related mortality [57] (Fig 4B). Although physical activity has a critical role in averting CVD, it is not a panacea and several other factors are surely important. For example, relative to individuals in postindustrial contexts, subsistence-level groups subsist on diets dominated by unprocessed or minimally processed foods and experience different types and degrees of social integration and inequality, all of which can impact CVD risk [58–60].

We note that while we have focused this section on CVD as an illustrative example of the type of comprehensive evidence required for fulfilling the first 2 criteria of mismatch, several other conditions also have relatively clear evidence. For example, inflammatory and autoimmune disorders have increased during the 20th century, which has been linked to a reduced exposure to parasites and microorganisms (a phenomenon attributed to the “hygiene hypothesis” or “old friends hypothesis”) [61–63].

### Current evidence for evolutionary mismatch at the genetic level

As mentioned above, to fulfill the third criteria for mismatch, we would need to identify a locus for which there is evidence of past selection (positive or stabilizing), and for which performance of at least 1 allele varies across environments and confers inflated risk of an NCD in the novel environment (Fig 1B and Box 1). One would think this would be easy to find, but in fact, there are only a handful of clear cases, despite good evidence for the existence of GxE interactions in general [64–67]. One clear example of mismatch involves variants in *APOLI1*, which provide resistance to trypanosome infections. Given the prevalence of trypanosomes across Africa, beneficial alleles are found at high frequency in African populations, as well as in African Americans. However, these same variants confer an increased risk of kidney disease in African American individuals living in the USA [68,69].

Another example is related to the “thrifty genotype” hypothesis [68], which suggests that individuals living in environments where food is unpredictably and periodically scarce should experience selection to store body fat in times of plenty. Recently, an intriguing variant was



found in Samoans, who are also susceptible to extreme obesity when eating a Western diet: A single amino acid variant (p.Arg475Gln) in *CREBRF* exhibits signatures of past selection and is currently associated with a 1.3-fold increased risk of obesity (though puzzlingly, also a 1.6-fold decreased risk of type 2 diabetes). Subsequent functional work in cell culture models demonstrated that p.Arg475Gln has direct effects on metabolism, reducing energy use while increasing lipid storage [70].

In addition to these well-characterized examples (see also Fig 2 of [71]), recent genomic work has shown that, in aggregate, variants that serve as modern-day risk alleles for particular NCDs (namely, CVD and autoimmune diseases) are more likely to show signatures of past selection relative to nonrisk alleles [72–74]. More broadly, there is now ample evidence that human populations can adapt to their local ecologies quite quickly (e.g., in thousands of years) [75], setting the stage for mismatches when local conditions shift. For example, the high *Plasmodium vivax* malaria risk experienced by West Africans has selected for changes to a key chemokine receptor encoded by *DARC* [76,77], whereas the spread of dairying in Europe has selected for lactase persistence through changes in the regulation of *LCT* [78,79]. Both of these changes have occurred within the last 10,000 years. As pathogen environments and diets inevitably change, local adaptation sets the stage for mismatches to occur.

## A new path forward: Integrating genomic tools and partnerships with transitioning populations

In principle, GxE interactions are most simply identifiable using a mismatch framework by testing for environmentally dependent genetic effects in transitioning populations. However, in practice, this would be difficult because most NCDs arise from many small genetic effects distributed across the genome. Further, the standard approach to resolve this needle-in-a-haystack problem—using a massive sample size—is difficult in small-scale groups who typically have modest population sizes. Sample sizes in the thousands, but not hundreds of thousands (e.g., biobank scale), are currently feasible; however, many anthropological studies have invested in long-term relationships with particular communities and are thus able to generate highly longitudinal datasets [24], where repeated samples and within-individual study designs could boost power. With these limitations in mind, we discuss how advanced genomic methods can be combined with the mismatch framework in a principled way to quantify the role of GxE interactions in NCDs.

First, we can improve GxE test power by focusing on genetic loci with already demonstrated evidence for phenotypic relevance, for example, those with evidence for recent selection in the study group or those that have already been discovered in urban/industrialized environments. For example, recent work on the *APOE* locus found that the *E4* variant—a well-known risk factor for CVD and Alzheimer's disease in individuals in postindustrial contexts—is associated with lower innate inflammation and may have beneficial effects on lipid moderation and cognition in a high pathogen/low obesity environment [80–82]. We might expect similar successes in elucidating GxE mismatches at other well-known risk loci that replicate across postindustrial contexts (e.g., *FTO*, *ADCY3*, *BRCA1/2*), though we caution that candidate gene studies should always be undertaken with care due to potential bias and replication issues [83,84]. A related approach is to test for GxE enrichment at the level of known genes or pathways with evolutionary or phenotypic relevance in the study population. These set-based approaches (i.e., that target predefined genes, genomic regions, or single nucleotide polymorphisms (SNPs)) may also perform well, even in cases where the specific causal variants are not shared between the focal population and the dataset in which they were identified.

Second, polygenic approaches that integrate GxE signals across the genome can improve power when studying complex traits such as NCDs. For example, recent methodological

developments have extended the popular polygenic risk score (PRS) framework to allow for PRS–environment interaction tests, thus providing a polygenic GxE test [85–87]. This approach has so far been used to show how diet and other lifestyle factors modulate the genetic risk of obesity, metabolic traits, and type 2 diabetes [40,88–90]. While polygenic approaches such as the PRS sacrifice variant-level resolution, they yield much greater power to detect GxE interactions, an invaluable exchange for quantifying evolutionary mismatch in transitioning populations. Three downsides to PRS–environment interaction tests, however, are that compared to single, large-effect allele results, one can be left with no suggestion of underlying mechanism; power depends on the predictive power of the PRS as well as its portability, which is a clear problem, given that most PRS work has focused on European ancestry individuals in postindustrial contexts, and, thus, this is where the summary statistics to build a PRS in other groups will have to come from (for the time being); and an underlying assumption is that risk effects are systematically stronger in one environment than another [91]. Any work in this area will consequently require replication across populations and will dramatically benefit from biobank-scale datasets that are currently being built in underrepresented, non-European ancestry contexts (e.g., [92,93]); these datasets will surely catalyze better multi-ancestry PRS methods.

Third, and perhaps most feasibly with current sample sizes, we can add power and interpretability for GxE interactions using intermediate molecular phenotypes such as gene expression, DNA methylation, and chromatin accessibility. One approach is to impute these functional genomic features from genotype data and then test them for environmental interaction (e.g., akin to a GxE version of transcriptome-wide association studies) [94,95]. The imputation step can use large, publicly available functional genomic datasets from US and European cohorts but will improve when similar datasets are available for the study populations. A second approach is to directly measure gene expression, DNA methylation, or other molecular features and identify variants that impact these features in different ways across different environmental contexts; this “molecular QTL” framework has so far proven very powerful and could be extended to transitioning populations [64,96–98]. Moreover, GxE molecular QTLs can be validated experimentally by exposing cell lines or model organisms to stimuli that mimic aspects of the environmental gradients experienced by transitioning populations; indeed, this can pinpoint key components of the incredibly complex environmental shifts that drive GxE interactions. A third option is to use functional genomic experiments to narrow the search space, by first identifying regulatory elements that respond to mismatch-relevant environments. For example, Garske and colleagues recently identified chromatin elements that respond to dietary fatty acids in adipocytes and then focused GxE follow-up work on variants in these responsive elements. By doing so, they were able to gain power to search for interaction effects between genotype and dietary saturated fat intake on body mass index [99]. Similar *in vitro* functional genomic experiments (using field-collected samples) could be leveraged to target regions of the genome that may be most important for responding to key aspects of lifestyle transitions.

## Payoffs for NCD prevention and treatment

Testing the degree to which GxE interactions arise from evolutionary mismatch would answer mechanistic questions about how GxE interactions manifest. For example, are loci that were involved in adaptation to a population’s past environment more likely to exhibit GxE effects when the environment shifts? To what degree does the nature of GxE interactions vary across ancestries with distinct evolutionary histories? What is the envelope of “optimal” human environmental conditions that do not provoke mismatch? Molecular insights into evolutionary

mismatch would allow us to prioritize the study of genetic variants that may adversely affect health outcomes in novel environments (i.e., those that have historically been under stabilizing or positive selection). It would also enable prediction of potential future adverse environments that could accelerate the onset of disease (i.e., those that represent strong deviations from the human evolutionary past). Furthermore, it could help us refine explanations for already observed ancestry-related differences in disease susceptibility. We emphasize that these are potential outcomes if mismatch is rigorously tested according to the criteria we lay out and subsequently supported; currently, its generalizability to the study of many complex traits and NCDs remains unclear due to a need for more empirical data.

The studies we recommend would more broadly advance our understanding of health issues in minority, Indigenous, and other underrepresented groups. Most subsistence-level populations in low- and middle-income countries (LMICs) are facing rapid rises in NCD risk, and the limited reports from these countries suggest that population responses to urbanization and market integration are highly variable. Studies of European ancestry individuals in postindustrial contexts are not well suited to explain why. Partnering with transitioning groups to conduct evolutionarily and culturally informed studies is needed to better serve their health concerns ([Box 2](#)).

### Box 2. Ethical considerations of conducting genomic work with subsistence-level populations

Community engagement and ethical research is fundamental to achieving the broader vision of this Essay. There is widespread consensus that broader population representation in biomedical research is critical for reducing health disparities [100], but moving forward on this agenda requires that we simultaneously acknowledge and learn from past mistakes and abuses.

At the heart of ethical considerations in genetics research is a situation in which diverse populations are dually underrepresented and underconsulted [101]. Recent work has outlined best practices for overcoming these issues [101–108]. For example, Claw and colleagues [102] suggest 6 principles of research ethics: understand community sovereignty and research regulations; engage and collaborate; build cultural competencies; improve transparency; build local research capacity; and disseminate research in accessible formats. The common thread behind these principles is the importance of building trustful and long-term relationships based on principles of dynamic consent, reciprocity, beneficence, and sovereignty. In our own experience, building these sorts of relationships takes time (typically years) but is essential to do before engaging in research.

Basic research with populations in LMICs can lead to important insights, yet the value-added benefits from basic research (e.g., shaping health policy based on epidemiological trends, and/or the development of novel treatment strategies) often can take decades to materialize. Mechanisms for participant community involvement in these longer-term benefits should be explicitly embedded in initial plans [100]. It is also important to recognize that community benefits can extend beyond the research itself. The needs and desires of local communities will vary widely, but populations in LMICs may face problems that are deeply interconnected and often stem from systemic discrimination: poor nutrition and sanitation (often due to environmental degradation), minimal access to

education, few economic opportunities, and loss of land rights. The priorities of communities will seldom match perfectly with the aims of scientists, especially when participant communities lack basic infrastructure and face discrimination. Prioritizing solutions to these problems is an opportunity to have great impact that will require cooperation between researchers, study participants, universities, nongovernmental organizations, governments, and funding bodies.

## Conclusions and future directions

The basic argument of this Essay is that we can further our understanding of evolution as well as the genetic architecture of human disease by combining genomic tools with studies of transitioning populations (as has been discussed previously [6,12,13,15,105,106], though not in the context of genomics). This recommended path improves on current approaches, which typically rely on “brute forcing” GxE scans across many SNPs and many environments. Instead, we advocate for using evolutionary theory to parse a priori which genotypes and environments we expect to interact. More specifically, under a mismatch framework, we expect genomic regions under positive or stabilizing selection in past environments to be enriched for GxE interactions revealed in postindustrial environments. If this framework proves true, leveraging its predictions could boost power and better position us to understand and predict GxE interactions in the etiology of NCDs. More generally, the work we propose would provide much needed insight into urgent health issues affecting vulnerable populations around the world.

Because the interdisciplinary perspective we take here necessarily touches on several fields, we did not attempt an exhaustive review of research on either evolutionary mismatch or GxE interactions (instead, we refer readers to excellent existing work [6,12,13,15,109,110]). However, there are several interesting new directions in these fields that are worth highlighting. For example, a growing body of work has begun to conceptualize the human microbiome as an evolved trait that is currently “mismatched” to its environment, often with serious health implications [111]. Given that the microbiome is under host genetic control and can therefore be a target of natural selection [112] and that industrialization can induce large scale changes in gut microbial communities [113–115], this is an exciting area in which to investigate GxE interactions that generate mismatch diseases. Another emerging research topic is sex differences in the response to lifestyle change: Several recent studies have found that women experience greater NCD risk following economic and nutritional transitions than men [18,25,116,117], yet how sex-specific genetic, physiological, or environmental variation interact to produce this phenomenon is still unknown [34]. Moreover, it is well established that early life experiences are important for predicting NCD risk later in life [118–120], and the timing of lifestyle change, as well as the degree to which individuals experience environmental mismatches within their lifetimes, may therefore be important to consider and to intersect with GxE frameworks (Box 3). In many cases, long-term partnerships with focal communities have already led to the creation of longitudinal datasets well positioned to take a life course approach. Moving forward, we expect that longitudinal perspectives on environmental change, NCD risk, and GxE interactions will be especially fruitful.

### Box 3. Life course perspectives on NCD risk

Development is a period of heightened environmental sensitivity, and challenging experiences early in life increase lifelong risk of most NCDs [118,120,121]. Subsistence-level societies are an underutilized yet potentially powerful model for studying early life influences on NCD risk. Many of these groups are currently experiencing rapid lifestyle changes leading to (1) extreme variation in early life conditions within a single population, and (2) frequent mismatch between early life and adult environments—a situation that is thought to put individuals at risk for later life health issues [118–120]. Point 1 provides a clear opportunity to leverage the distributional extremes to study early life effects on health [26,122]. Further, point 2 affords us the opportunity to compare outcomes when individuals experience within-lifetime environmental “matches” versus “mismatches.” To date, studies of industrial transitions have come to mixed conclusions about the importance of within-lifetime mismatches [18,47,123,124]. More work in this area is urgently needed to understand when, why, and how early life experiences shape adult health in these groups.

Genomic tools applied to populations undergoing lifestyle change could also provide valuable insight into how early life experiences become “embedded” into lifelong physiology. At the molecular level, this process is thought to be mediated by stable changes in gene regulation (e.g., DNA methylation, chromatin accessibility, and gene expression). However, many gene regulatory elements are also dynamic and responsive to environmental perturbations throughout life. This fact leads to challenges in disentangling the effects of early versus later life environments, especially when the two are highly correlated (as is common in postindustrial contexts). By contrast, subsistence-level groups in transition often experience decoupled early life and adult experiences, which could be leveraged to disentangle early versus later life influences. Genotype data collected for the same individuals could also be used to identify rarely studied GxE interactions where the “E” encompasses early life experiences. Overall, integrative studies of transitioning populations are primed to reveal which individuals will be most susceptible to NCDs during lifestyle transitions as well as when in the life course these exposures matter most.

### Acknowledgments

We thank all participants from the “Evolutionary Mismatch Hypothesis in the Genomics Era” symposium, which generated many of the ideas discussed here. We also thank all of the communities, staff, and scientists that have participated in the long-term studies mentioned in Fig 2.

### Author Contributions

**Conceptualization:** Amanda J. Lea, Andrew G. Clark, Andrew W. Dahl, Orrin Devinsky, Angela R. Garcia, Christopher D. Golden, Joseph Kamau, Thomas S. Kraft, Yvonne A. L. Lim, Dino J. Martins, Donald Mogoï, Päivi Pajukanta, George H. Perry, Herman Pontzer, Benjamin C. Trumble, Samuel S. Urlacher, Vivek V. Venkataraman, Ian J. Wallace, Michael Gurven, Daniel E. Lieberman, Julien F. Ayroles.

**Project administration:** Amanda J. Lea, Julien F. Ayroles.

**Supervision:** Amanda J. Lea, Julien F. Ayroles.

**Writing – original draft:** Amanda J. Lea, Andrew G. Clark, Andrew W. Dahl, Orrin Devinsky, Angela R. Garcia, Christopher D. Golden, Joseph Kamau, Thomas S. Kraft, Yvonne A. L. Lim, Dino J. Martins, Päivi Pajukanta, George H. Perry, Herman Pontzer, Benjamin C. Trumble, Samuel S. Urlacher, Vivek V. Venkataraman, Ian J. Wallace, Michael Gurven, Daniel E. Lieberman, Julien F. Ayroles.

**Writing – review & editing:** Amanda J. Lea, Andrew G. Clark, Andrew W. Dahl, Orrin Devinsky, Angela R. Garcia, Christopher D. Golden, Joseph Kamau, Thomas S. Kraft, Yvonne A. L. Lim, Dino J. Martins, Päivi Pajukanta, George H. Perry, Herman Pontzer, Benjamin C. Trumble, Samuel S. Urlacher, Vivek V. Venkataraman, Ian J. Wallace, Michael Gurven, Daniel E. Lieberman, Julien F. Ayroles.

## References

1. Kathiresan S, Srivastava D. Genetics of human cardiovascular disease. *Cell*. 2012 Mar 16; 148(6):1242–1257. <https://doi.org/10.1016/j.cell.2012.03.001> PMID: 22424232
2. Khera AV, Kathiresan S. Genetics of coronary artery disease: discovery, biology and clinical translation. *Nat Rev Genet*. 2017 Jun; 18(6):331–344. <https://doi.org/10.1038/nrg.2016.160> PMID: 28286336
3. McPherson R, Tybjaerg-Hansen A. Genetics of Coronary Artery Disease. *Circ Res*. 2016 Feb 19; 118(4):564–578. <https://doi.org/10.1161/CIRCRESAHA.115.306566> PMID: 26892958
4. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature*. 2009 Oct 8; 461(7265):747–753. <https://doi.org/10.1038/nature08494> PMID: 19812666
5. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. *Am J Hum Genet*. 2012 Jan 13; 90(1):7–24. <https://doi.org/10.1016/j.ajhg.2011.11.029> PMID: 22243964
6. McAllister K, Mechanic LE, Amos C, Aschard H, Blair IA, Chatterjee N, et al. Current Challenges and New Opportunities for Gene-Environment Interaction Studies of Complex Diseases. *Am J Epidemiol*. 2017 Oct 1; 186(7):753–761. <https://doi.org/10.1093/aje/kwx227> PMID: 28978193
7. Thomas D. Gene–environment-wide association studies: emerging approaches. *Nat Rev Genet*. 2010 Mar 9; 11(4):259–272. <https://doi.org/10.1038/nrg2764> PMID: 20212493
8. Wang H, Zhang F, Zeng J, Wu Y, Kemper KE, Xue A, et al. Genotype-by-environment interactions inferred from genetic effects on phenotypic variability in the UK Biobank. *Sci Adv*. 2019 Aug; 5(8):eaaw3538. <https://doi.org/10.1126/sciadv.aaw3538> PMID: 31453325
9. Tyrrell J, Wood AR, Ames RM, Yaghootkar H, Beaumont RN, Jones SE, et al. Gene–obesogenic environment interactions in the UK Biobank study [Internet]. *Int J Epidemiol*. 2017:dyw337. <https://doi.org/10.1093/ije/dyw337> PMID: 28073954
10. Nag A, Mahajan A, Mccarthy M. 306-OR: Gene-Environment Interactions for Type 2 Diabetes in UK Biobank. *Diabetes*. 2019; 68:306–OR. <https://doi.org/10.2337/db19-306-or>
11. Arnau-Soler A, Macdonald-Dunlop E, Adams MJ, Clarke T-K, MacIntyre DJ, Milburn K, et al. Genome-wide by environment interaction studies of depressive symptoms and psychosocial stress in UK Biobank and Generation Scotland. *Transl Psychiatry*. 2019 Feb 4; 9(1):14. <https://doi.org/10.1038/s41398-018-0360-y> PMID: 30718454
12. Lieberman D. *The Story of the Human Body: Evolution, Health, and Disease*. Vintage; 2014. 460 p.
13. Nesse RM, Williams GC. *Why We Get Sick: The New Science of Darwinian Medicine*. Vintage; 2012. 304 p.
14. Neel JV. Diabetes Mellitus: A “Thrifty” Genotype Rendered Detrimental by “Progress”? *Am J Hum Genet*. 1962; 1949:353–62.
15. Corbett S, Courtiol A, Lummaa V, Moorad J, Stearns S. The transition to modernity and chronic disease: mismatch and natural selection. *Nat Rev Genet*. 2018 Jul; 19(7):419–430. <https://doi.org/10.1038/s41576-018-0012-3> PMID: 29743650
16. Gibson G. *It Takes a Genome: How a Clash Between Our Genes and Modern Life Is Making Us Sick*. Upper Saddle River, New Jersey: FT Press Science; 2009.
17. Garruto RM, Little MA, James GD, Brown DE. Natural experimental models: the global search for biomedical paradigms among traditional, modernizing, and modern populations. *Proc Natl Acad Sci U S A*. 1999 Aug 31; 96(18):10536–10543. <https://doi.org/10.1073/pnas.96.18.10536> PMID: 10468644

18. Lea AJ, Martins D, Kamau J, Gurven M, Ayroles JF. Urbanization and market integration have strong, nonlinear effects on cardiometabolic health in the Turkana. *Sci Adv*. 2020 Oct; 6(43). <https://doi.org/10.1126/sciadv.abb1430> PMID: 33087362
19. Lea AJ, Waigwa C, Muhoya B, Lotukoi F, Peng J, Henry LP, et al. Socioeconomic status effects on health vary between rural and urban Turkana. *Evol Med Public Health*. 2021 Nov 25; 9(1):406–419. <https://doi.org/10.1093/emph/eoab039> PMID: 34987823
20. Wallace IJ, Lea AJ, Lim YAL, Chow SKW, bin Mohd Sayed I, Ngui R, et al. Orang Asli Health and Life-ways Project (OA HeLP): Study Protocol. *bioRxiv*. 2021. Available from: <http://medrxiv.org/lookup/doi/10.1101/2021.10.24.21265442>.
21. Golden CD, Anjaranirina EJG, Fernald LCH, Hartl DL, Kremen C, Milner DA Jr, et al. Cohort Profile: The Madagascar Health and Environmental Research (MAHERY) study in north-eastern Madagascar. *Int J Epidemiol*. 2017 Dec 1; 46(6):1747–1748d. <https://doi.org/10.1093/ije/dyx071> PMID: 29040632
22. Golden CD, Borgerson C, Rice BL, Allen LH, Anjaranirina EJG, Barrett CB, et al. Cohort Description of the Madagascar Health and Environmental Research–Antongil (MAHERY–Antongil) Study in Madagascar. *Front Nutr*. 2019; 6. Available from: <https://www.frontiersin.org/article/10.3389/fnut.2019.00109>. <https://doi.org/10.3389/fnut.2019.00109> PMID: 31428615
23. Golden CD, Rice BL, Randriamady HJ, Vonona AM, Randrianasolo JF, Tafangy AN, et al. Study Protocol: A Cross-Sectional Examination of Socio-Demographic and Ecological Determinants of Nutrition and Disease Across Madagascar. *Front Public Health*. 2020 Sep 17; 8:500. <https://doi.org/10.3389/fpubh.2020.00500> PMID: 33042943
24. Gurven M, Stieglitz J, Trumble B, Blackwell AD, Beheim B, Davis H, et al. The Tsimane Health and Life History Project: Integrating anthropology and biomedicine. *Evol Anthropol*. 2017 Apr; 26(2):54–73. <https://doi.org/10.1002/evan.21515> PMID: 28429567
25. Liebert MA, Snodgrass JJ, Madimenos FC, Cepon TJ, Blackwell AD, Sugiyama LS. Implications of market integration for cardiovascular and metabolic health among an indigenous Amazonian Ecuadorian population. *Ann Hum Biol*. 2013 May; 40(3):228–242. <https://doi.org/10.3109/03014460.2012.759621> PMID: 23388068
26. Urlacher SS, Liebert MA, Josh Snodgrass J, Blackwell AD, Cepon-Robins TJ, Gildner TE, et al. Heterogeneous effects of market integration on sub-adult body size and nutritional status among the Shuar of Amazonian Ecuador. *Ann Hum Biol*. 2016 Jul; 43(4):316–329. <https://doi.org/10.1080/03014460.2016.1192219> PMID: 27230632
27. Lieberman DE. *The Story of the Human Body: Evolution, Health, and Disease*. New York: Pantheon; 2013;
28. Holowka NB, Wallace IJ, Lieberman DE. Foot strength and stiffness are related to footwear use in a comparison of minimally- vs. conventionally-shod populations. *Sci Rep*. 2018 Feb 27; 8(1):3679.
29. Trumble BC, Finch CE. THE EXPOSOME IN HUMAN EVOLUTION: FROM DUST TO DIESEL. *Q Rev Biol*. 2019 Dec; 94(4):333–394. <https://doi.org/10.1086/706768> PMID: 32269391
30. Pontzer H, Wood BM, Raichlen DA. Hunter-gatherers as models in public health. *Obes Rev*. 2018 Dec; 19(Suppl 1):24–35. <https://doi.org/10.1111/obr.12785> PMID: 30511505
31. Price WA, Nguyen T. Nutrition and Physical Degeneration: A Comparison of Primitive and Modern Diets and Their Effects. *EnCognitive.com*; 2016. p. 550.
32. Rousseau J-J. *Discourse on the Origin of Inequality*. Oxford University Press; 1999. p. 127.
33. Gurven MD, Lieberman DE. WEIRD bodies: mismatch, medicine and missing diversity. *Evol Hum Behav*. 2020 Sep; 41(5):330–340. <https://doi.org/10.1016/j.evolhumbehav.2020.04.001> PMID: 33100820
34. Zhu C, Ming MJ, Cole JM, Edge MD, Kirkpatrick M, Harpak A. Amplification is the primary mode of gene-by-sex interaction in complex human traits. *Cell Genom*. 2023 May 10; 3(5):100297. <https://doi.org/10.1016/j.xgen.2023.100297> PMID: 37228747
35. NCD Countdown 2030 collaborators. NCD Countdown 2030: worldwide trends in non-communicable disease mortality and progress towards Sustainable Development Goal target 3.4. *Lancet*. 2018 Sep 22; 392(10152):1072–1088.
36. Boyle EA, Li Yi, Pritchard JK. Leading Edge Perspective An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell*. 2017; 169(7):1177–1186.
37. Harpak A, Przeworski M. The evolution of group differences in changing environments. *PLoS Biol*. 2021 Jan; 19(1):e3001072. <https://doi.org/10.1371/journal.pbio.3001072> PMID: 33493148
38. Gibson G. Decanalization and the origin of complex disease. *Nat Rev Genet*. 2013; 10:134–140.
39. Gibson G, Lacey KA. Canalization and Robustness in Human Genetics and Disease. *Annu Rev Genet*. 2020 Nov 23; 54:189–211. <https://doi.org/10.1146/annurev-genet-022020-022327> PMID: 32867542

40. Nagpal S, Tandon R, Gibson G. Canalization of the polygenic risk for common diseases and traits in the UK Biobank cohort. *Mol Biol Evol.* 2022 Apr 11; 39(4):msac053. <https://doi.org/10.1093/molbev/msac053> PMID: 35275999
41. Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet.* 2020 Oct 17; 396(10258):1204–1222. [https://doi.org/10.1016/S0140-6736\(20\)30925-9](https://doi.org/10.1016/S0140-6736(20)30925-9) PMID: 33069326
42. Kaplan H, Thompson RC, Trumble BC, Wann LS, Allam AH, Beheim B, et al. Coronary atherosclerosis in indigenous South American Tsimane: a cross-sectional cohort study. *Lancet.* 2017 Apr 29; 389(10080):1730–1739. [https://doi.org/10.1016/S0140-6736\(17\)30752-3](https://doi.org/10.1016/S0140-6736(17)30752-3) PMID: 28320601
43. Raichlen DA, Pontzer H, Harris JA, Mabulla AZP, Marlowe FW, Josh Snodgrass J, et al. Physical activity patterns and biomarkers of cardiovascular disease risk in hunter-gatherers. *Am J Hum Biol.* 2017 Mar; 29(2):e22919. <https://doi.org/10.1002/ajhb.22919> PMID: 27723159
44. Truswell AS, Kennelly BM, Hansen JD, Lee RB. Blood pressures of Kung bushmen in Northern Botswana. *Am Heart J.* 1972 Jul; 84(1):5–12. [https://doi.org/10.1016/0002-8703\(72\)90299-2](https://doi.org/10.1016/0002-8703(72)90299-2) PMID: 5080283
45. Shave RE, Lieberman DE, Drane AL, Brown MG, Batterham AM, Worthington S, et al. Selection of endurance capabilities and the trade-off between pressure and volume in the evolution of the human heart. *Proc Natl Acad Sci U S A.* 2019 Oct 1; 116(40):19905–19910. <https://doi.org/10.1073/pnas.1906902116> PMID: 31527253
46. Gurven M, Blackwell AD, Rodríguez DE, Stieglitz J, Kaplan H. Does blood pressure inevitably rise with age?: longitudinal evidence among forager-horticulturalists. *Hypertension.* 2012 Jul; 60(1):25–33. <https://doi.org/10.1161/HYPERTENSIONAHA.111.189100> PMID: 22700319
47. Wallace IJ, Felson DT, Worthington S, Duryea J, Clancy M, Aliabadi P, et al. Knee osteoarthritis risk in non-industrial societies undergoing an energy balance transition: evidence from the indigenous Tarahumara of Mexico. *Ann Rheum Dis.* 2019; 78:1693–1698. <https://doi.org/10.1136/annrheumdis-2019-215886> PMID: 31519654
48. Urlacher SS, Snodgrass JJ, Dugas LR, Madimenos FC, Sugiyama LS, Liebert MA, et al. Childhood Daily Energy Expenditure Does Not Decrease with Market Integration and Is Not Related to Adiposity in Amazonia. *J Nutr.* 2021 Mar 11; 151(3):695–704. <https://doi.org/10.1093/jn/nxaa361> PMID: 33454748
49. National Health and Nutrition Examination Survey (NHANES). [cited 2022 Feb 16]. Available from: <https://methods.sagepub.com/reference/encyclopedia-of-survey-research-methods/n318.xml>.
50. Arem H, Moore SC, Patel A, Hartge P, Berrington de Gonzalez A, Viswanathan K, et al. Leisure time physical activity and mortality: a detailed pooled analysis of the dose-response relationship. *JAMA Intern Med.* 2015 Jun; 175(6):959–967. <https://doi.org/10.1001/jamainternmed.2015.0533> PMID: 25844730
51. Gurven M, Jaeggi AV, Kaplan H, Cummings D. Physical activity and modernization among Bolivian Amerindians. *PLoS ONE.* 2013 Jan 31; 8(1):e55679. <https://doi.org/10.1371/journal.pone.0055679> PMID: 23383262
52. Burn PH. *New Research Blows the Lid Off How We Really Burn Calories, Lose Weight, and Stay Healthy.* Penguin; 2021. p. 384.
53. Lieberman D. *Exercised: Why Something We Never Evolved to Do Is Healthy and Rewarding.* Pantheon Books; 2021. p. 464.
54. Nosarev AV, Smagliy LV, Anfinogenova Y, Popov SV, Kapilevich LV. Exercise and NO production: relevance and implications in the cardiopulmonary system. *Front Cell Dev Biol.* 2014; 2:73. <https://doi.org/10.3389/fcell.2014.00073> PMID: 25610830
55. Green DJ, Hopman MTE, Padilla J, Laughlin MH, Thijssen DHJ. Vascular Adaptation to Exercise in Humans: Role of Hemodynamic Stimuli. *Physiol Rev.* 2017 Apr; 97(2):495–528. <https://doi.org/10.1152/physrev.00014.2016> PMID: 28151424
56. Petersen AMW, Pedersen BK. The anti-inflammatory effect of exercise. *J Appl Physiol.* 2005 Apr; 98(4):1154–1162. <https://doi.org/10.1152/jappphysiol.00164.2004> PMID: 15772055
57. Tigbe WW, Granat MH, Sattar N, Lean MEJ. Time spent in sedentary posture is associated with waist circumference and cardiovascular risk. *Int J Obes (Lond).* 2017 May; 41(5):689–696. <https://doi.org/10.1038/ijo.2017.30> PMID: 28138134
58. Jaeggi AV, Blackwell AD, von Rueden C, Trumble BC, Stieglitz J, Garcia AR, et al. Do wealth and inequality associate with health in a small-scale subsistence society? *Elife.* 2021 May 14; 10. <https://doi.org/10.7554/eLife.59437> PMID: 33988506



59. Snyder-Mackler N, Burger JR, Gaydos L, Belsky D, Noppert GA, Campos FA, et al. Social determinants of health and survival in humans and other animals. *Science*. 2020; <https://doi.org/10.1126/science.aax9553> PMID: 32439765
60. Lea AJ, Waigwa C, Muhoya B, Lotukoi F, Peng J, Henry LP, et al. Social gradients in health vary between rural and urban Turkana [Internet]. *bioRxiv. medRxiv*; 2021. Available from: <http://medrxiv.org/lookup/doi/10.1101/2021.07.23.21260771>.
61. Strachan DP. Hay fever, hygiene, and household size. *BMJ*. 1989 Nov 18; 299(6710):1259–1260. <https://doi.org/10.1136/bmj.299.6710.1259> PMID: 2513902
62. Bach J-F. The hygiene hypothesis in autoimmunity: the role of pathogens and commensals. *Nat Rev Immunol*. 2018 Feb; 18(2):105–120. <https://doi.org/10.1038/nri.2017.111> PMID: 29034905
63. Rook G. Darwinian medicine and the 'hygiene' or 'old friends' hypothesis. *Clin Exp Immunol*. 2010; 160:70–79.
64. Nédélec Y, Sanz J, Baharian G, Tung J, Yotova V, Barreiro Correspondence LB, et al. Genetic Ancestry and Natural Selection Drive Population Differences in Immune Responses to Pathogens. *Cell*. 2016; 167:113. <https://doi.org/10.1016/j.cell.2016.09.025> PMID: 27768889
65. Kim-Hellmuth S, Bechheim M, Pütz B, Mohammadi P, Nédélec Y, Giangreco N, et al. Genetic regulatory effects modified by immune activation contribute to autoimmune disease associations. *Nat Commun*. 2017; 8(1). <https://doi.org/10.1038/s41467-017-00366-1> PMID: 28814792
66. Jelenkovic A, Hur Y-M, Sund R, Yokoyama Y, Siribaddana SH, Hotopf M, et al. Genetic and environmental influences on adult human height across birth cohorts from 1886 to 1994. *Elife*. 2016 Dec 14; 5. <https://doi.org/10.7554/eLife.20320> PMID: 27964777
67. Athanasiadis G, Meijssen JJ, Helenius D, Schork AJ, Ingason A, Thompson WK, et al. A comprehensive map of genetic relationships among diagnostic categories based on 48.6 million relative pairs from the Danish genealogy. *Proc Natl Acad Sci U S A*. 2022 Feb 8; 119(6). <https://doi.org/10.1073/pnas.2118688119> PMID: 35131856
68. Genovese G, Friedman DJ, Ross MD, Lecordier L, Uzureau P, Freedman BI, et al. Association of Trypanolytic ApoL1 Variants with Kidney Disease in African Americans. *Science*. 2010; 329:841–845. <https://doi.org/10.1126/science.1193032> PMID: 20647424
69. Friedman DJ, Pollak MR. APOL1 and Kidney Disease: From Genetics to Biology. *Annu Rev Physiol*. 2020 Feb 10; 82:323–342. <https://doi.org/10.1146/annurev-physiol-021119-034345> PMID: 31710572
70. Loos RJJ. CREBRF variant increases obesity risk and protects against diabetes in Samoans. *Nat Genet*. 2016 Aug 30; 48(9):976–978. <https://doi.org/10.1038/ng.3653> PMID: 27573685
71. Benton ML, Abraham A, LaBella AL, Abbot P, Rokas A, Capra JA. The influence of evolutionary history on human health and disease. *Nat Rev Genet*. 2021 Jan 6. <https://doi.org/10.1038/s41576-020-00305-9> PMID: 33408383
72. Byars SG, Huang QQ, Gray L-A, Bakshi A, Ripatti S, Abraham G, et al. Genetic loci associated with coronary artery disease harbor evidence of selection and antagonistic pleiotropy. *PLoS Genet*. 2017 Jun; 13(6):e1006328. <https://doi.org/10.1371/journal.pgen.1006328> PMID: 28640878
73. Brinkworth JF, Barreiro LB. The contribution of natural selection to present-day susceptibility to chronic inflammatory and autoimmune disease. *Curr Opin Immunol*. 2014 Dec; 31:66–78. <https://doi.org/10.1016/j.coi.2014.09.008> PMID: 25458997
74. Raj T, Kuchroo M, Replogle JM, Raychaudhuri S, Stranger BE, De Jager PL. Common risk alleles for inflammatory diseases are targets of recent positive selection. *Am J Hum Genet*. 2013 Apr 4; 92(4):517–529. <https://doi.org/10.1016/j.ajhg.2013.03.001> PMID: 23522783
75. Fan S, Hansen MEB, Lo Y, Tishkoff SA. Going global by adapting local: A review of recent human adaptation. *Science*. 2016 Oct 7; 354(6308):54–59. <https://doi.org/10.1126/science.aaf5098> PMID: 27846491
76. Hamblin MT, Di Rienzo A. Detection of the signature of natural selection in humans: evidence from the Duffy blood group locus. *Am J Hum Genet*. 2000 May; 66(5):1669–1679. <https://doi.org/10.1086/302879> PMID: 10762551
77. Hamid I, Korunes KL, Beleza S, Goldberg A. Rapid adaptation to malaria facilitated by admixture in the human population of Cabo Verde. *Elife*. 2021 Jan 4; 10. <https://doi.org/10.7554/eLife.63177> PMID: 33393457
78. Sabeti PC, Schaffner SF, Fry B, Lohmueller J, Varilly P, Shamovsky O, et al. Positive natural selection in the human lineage. *Science*. 2006 Jun 16; 312(5780):1614–1620. <https://doi.org/10.1126/science.1124309> PMID: 16778047
79. Evershed RP, Davey Smith G, Roffet-Salque M, Timpson A, Diekmann Y, Lyon MS, et al. Dairying, diseases and the evolution of lactase persistence in Europe. *Nature*. 2022 Aug; 608(7922):336–345. <https://doi.org/10.1038/s41586-022-05010-7> PMID: 35896751

80. Belloy ME, Napolioni V, Greicius MD. A Quarter Century of APOE and Alzheimer's Disease: Progress to Date and the Path Forward. *Neuron*. 2019 Mar 6; 101(5):820–838. <https://doi.org/10.1016/j.neuron.2019.01.056> PMID: 30844401
81. Trumble BC, Stieglitz J, Blackwell AD, Allayee H, Beheim B, Finch CE, et al. Apolipoprotein E4 is associated with improved cognitive function in Amazonian forager-horticulturalists with a high parasite burden. *FASEB J*. 2017 Apr; 31(4):1508–1515. <https://doi.org/10.1096/fj.201601084R> PMID: 28031319
82. Garcia AR, Finch C, Gatz M, Kraft T, Eid Rodriguez D, Cummings D, et al. APOE4 is associated with elevated blood lipids and lower levels of innate immune biomarkers in a tropical Amerindian subsistence population. *Elife*. 2021 Sep 29; 10. <https://doi.org/10.7554/eLife.68231> PMID: 34586066
83. Risch N, Herrell R, Lehner T. Interaction Between the Serotonin Transporter Gene (5-HTTLPR), Stressful Life Events, and Risk of Depression. *JAMA*. 2009; 301(23):2462–2472.
84. Border R, Johnson EC, Evans LM, Smolen A, Berley N, Sullivan PF, et al. No support for historical candidate gene or candidate gene-by-interaction hypotheses for major depression across multiple large samples. *Am J Psychiatry*. 2019 May 1; 176(5):376–387. <https://doi.org/10.1176/appi.ajp.2018.18070881> PMID: 30845820
85. Peyrot WJ, Milaneschi Y, Abdellaoui A, Sullivan PF, Hottenga JJ, Boomsma DI, et al. Effect of polygenic risk scores on depression in childhood trauma. *Br J Psychiatry*. 2014 Aug; 205(2):113–119. <https://doi.org/10.1192/bjp.bp.113.143081> PMID: 24925986
86. Mullins N, Power RA, Fisher HL, Hanscombe KB, Euesden J, Iniesta R, et al. Polygenic interactions with environmental adversity in the aetiology of major depressive disorder. *Psychol Med*. 2016 Mar; 46(4):759–770. <https://doi.org/10.1017/S0033291715002172> PMID: 26526099
87. Peterson RE, Cai N, Dahl AW, Bigdeli TB, Edwards AC, Webb BT, et al. Molecular Genetic Analysis Subdivided by Adversity Exposure Suggests Etiologic Heterogeneity in Major Depression. *Am J Psychiatry*. 2018 Jun 1; 175(6):545–554. <https://doi.org/10.1176/appi.ajp.2017.17060621> PMID: 29495898
88. Nagpal S, Gibson G, Marigorta UM. Pervasive Modulation of Obesity Risk by the Environment and Genomic Background. *Genes*. 2018 Aug 14; 9(8). <https://doi.org/10.3390/genes9080411> PMID: 30110940
89. Abadi A, Alyass A, Robiou du Pont S, Bolker B, Singh P, Mohan V, et al. Penetrance of Polygenic Obesity Susceptibility Loci across the Body Mass Index Distribution. *Am J Hum Genet*. 2017 Dec 7; 101(6):925–938. <https://doi.org/10.1016/j.ajhg.2017.10.007> PMID: 29220676
90. Rask-Andersen M, Karlsson T, Ek WE, Johansson Å. Gene-environment interaction study for BMI reveals interactions between genetic factors and physical activity, alcohol consumption and socioeconomic status. *PLoS Genet*. 2017 Sep; 13(9):e1006977. <https://doi.org/10.1371/journal.pgen.1006977> PMID: 28873402
91. Sheppard B, Rappoport N, Loh P-R, Sanders SJ, Zaitlen N, Dahl A. A model and test for coordinated polygenic epistasis in complex traits. *Proc Natl Acad Sci U S A*. 2021 Apr 13; 118(15). <https://doi.org/10.1073/pnas.1922305118> PMID: 33833052
92. Nagai A, Hirata M, Kamatani Y, Muto K, Matsuda K, Kiyohara Y, et al. Overview of the BioBank Japan Project: Study design and profile. *J Epidemiol*. 2017 Mar; 27(3S):S2–S8. <https://doi.org/10.1016/j.je.2016.12.005> PMID: 28189464
93. Malaria Genomic Epidemiology Network. A global network for investigating the genomic epidemiology of malaria. *Nature*. 2008 Dec 11; 456(7223):732–737. <https://doi.org/10.1038/nature07632> PMID: 19079050
94. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BWJH, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet*. 2016 Mar; 48(3):245–252. <https://doi.org/10.1038/ng.3506> PMID: 26854917
95. Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, et al. A gene-based association method for mapping traits using reference transcriptome data. *Nat Genet*. 2015 Sep; 47(9):1091–1098. <https://doi.org/10.1038/ng.3367> PMID: 26258848
96. Fairfax BP, Humburg P, Makino S, Naranbhai V, Wong D, Lau E, et al. Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science*. 2014; 343(6175):1246949. <https://doi.org/10.1126/science.1246949> PMID: 24604202
97. Lee MNHMN, Ye C, Villani A-CA-C, Raj T, Li W, Eisenhaure TM, et al. Common Genetic Variants Modulate Pathogen-Sensing Responses in Human Dendritic Cells. *Science*. 2014; 343(6175):1246980. <https://doi.org/10.1126/science.1246980> PMID: 24604203
98. Idaghdour Y, Czika W, Shianna KV, Lee SH, Visscher PM, Martin HC, et al. Geographical Genomics of Human Leukocyte Gene Expression Variation in Southern Morocco. *Nat Genet*. 2010; 42(1):62–67. <https://doi.org/10.1038/ng.495> PMID: 19966804

99. Garske KM, Pan DZ, Miao Z, Bhagat YV, Comenho C, Robles CR, et al. Reverse gene-environment interaction approach to identify variants influencing body-mass index in humans. *Nat Metab.* 2019 Jun; 1(6):630–642. <https://doi.org/10.1038/s42255-019-0071-6> PMID: 31538139
100. Fox K. The illusion of inclusion—the “all of us” research program and indigenous peoples’ DNA. *N Engl J Med.* 2020 Jul 30; 383(5):411–413. <https://doi.org/10.1056/NEJMp1915987> PMID: 32726527
101. Jackson L, Kuhlman C, Jackson F, Keolu Fox P. Including Vulnerable Populations in the Assessment of Data From Vulnerable Populations. *Front Big Data.* 2019; 2:19. <https://doi.org/10.3389/fdata.2019.00019> PMID: 33693342
102. Claw KG, Anderson MZ, Begay RL, Tsosie KS, Fox K, Garrison NA, et al. A framework for enhancing ethical genomic research with Indigenous communities. *Nat Commun.* 2018 Jul 27; 9(1):2957. <https://doi.org/10.1038/s41467-018-05188-3> PMID: 30054469
103. Dignan MB, Burhansstipanov L, Bemis L. Successful implementation of genetic Education for Native Americans workshops at national conferences. *Genetics.* 2005 Feb; 169(2):517–521. <https://doi.org/10.1534/genetics.104.030015> PMID: 15731516
104. Tiffin N. Tiered informed consent: respecting autonomy, agency and individuality in Africa. *BMJ Glob Health.* 2018 Dec 17; 3(6):e001249. <https://doi.org/10.1136/bmjgh-2018-001249> PMID: 30613430
105. Hudson M, Garrison NA, Sterling R, Caron NR, Fox K, Yracheta J, et al. Rights, interests and expectations: Indigenous perspectives on unrestricted access to genomic data. *Nat Rev Genet.* 2020 Jun; 21(6):377–384. <https://doi.org/10.1038/s41576-020-0228-x> PMID: 32251390
106. Hernandez M, Perry GH. Scanning the human genome for “signatures” of positive selection: Transformative opportunities and ethical obligations. *Evol Anthropol.* 2021 Mar; 30(2):113–121. <https://doi.org/10.1002/evan.21893> PMID: 33788352
107. Garrison NA, Hudson M, Ballantyne LL, Garba I, Martinez A, Taulii M, et al. Genomic Research Through an Indigenous Lens: Understanding the Expectations. *Annu Rev Genomics Hum Genet.* 2019 Aug 31; 20:495–517. <https://doi.org/10.1146/annurev-genom-083118-015434> PMID: 30892943
108. Guglielmi G. Facing up to injustice in genome science. *Nature.* 2019 Apr; 568(7752):290–293. <https://doi.org/10.1038/d41586-019-01166-x> PMID: 30992587
109. Ober C, Vercelli D. Gene–environment interactions in human disease: nuisance or opportunity? *Trends Genet.* 2011 Mar 1; 27(3):107–115. <https://doi.org/10.1016/j.tig.2010.12.004> PMID: 21216485
110. Hunter DJ. Gene–environment interactions in human diseases. *Nat Rev Genet.* 2005 Apr; 6(4):287–298. <https://doi.org/10.1038/nrg1578> PMID: 15803198
111. Carmody RN, Sarkar A, Reese AT. Gut microbiota through an evolutionary lens. *Science.* 2021 Apr 30; 372(6541):462–463. <https://doi.org/10.1126/science.abf0590> PMID: 33926939
112. Grieneisen L, Dasari M, Gould TJ, Björk JR, Grenier J-C, Yotova V, et al. Gut microbiome heritability is nearly universal but environmentally contingent. *Science.* 2021 Jul 9; 373(6551):181–186. <https://doi.org/10.1126/science.aba5483> PMID: 34244407
113. Groussin M, Poyet M, Sistiaga A, Kearney SM, Moniz K, Noel M, et al. Elevated rates of horizontal gene transfer in the industrialized human microbiome. *Cell.* 2021 Apr 15; 184(8):2053–2067.e18. <https://doi.org/10.1016/j.cell.2021.02.052> PMID: 33794144
114. Schnorr SL, Candela M, Rampelli S, Centanni M, Consolandi C, Basaglia G, et al. Gut microbiome of the Hadza hunter-gatherers. *Nat Commun.* 2014 Apr 15; 5:3654. <https://doi.org/10.1038/ncomms4654> PMID: 24736369
115. Vangay P, Johnson AJ, Ward TL, Al-Ghalith GA, Shields-Cutler RR, Hillmann BM, et al. US Immigration Westernizes the Human Gut Microbiome. *Cell.* 2018 Nov 1; 175(4):962–972.e10. <https://doi.org/10.1016/j.cell.2018.10.029> PMID: 30388453
116. Snodgrass JJ, Leonard WR, Sorensen MV, Tarskaia LA, Alekseev VP, Krivoschapkin V. The emergence of obesity among indigenous Siberians. *J Physiol Anthropol.* 2006 Jan; 25(1):75–84. <https://doi.org/10.2114/jpa2.25.75> PMID: 16617212
117. NCD Risk Factor Collaboration (NCD-RisC). Rising rural body-mass index is the main driver of the global obesity epidemic in adults. *Nature.* 2019 May; 569(7755):260–264. <https://doi.org/10.1038/s41586-019-1171-x> PMID: 31068725
118. Felitti VJ, Anda RF, Nordenberg D, Williamson DF, Spitz AM, Edwards V, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The Adverse Childhood Experiences (ACE) Study. *Am J Prev Med.* 1998; 14:245–258.
119. Bellis MA, Hughes K, Ford K, Ramos Rodriguez G, Sethi D, Passmore J. Life course health consequences and associated annual costs of adverse childhood experiences across Europe and North America: a systematic review and meta-analysis. *Lancet Public Health.* 2019 Oct 1; 4(10):e517–e528. [https://doi.org/10.1016/S2468-2667\(19\)30145-8](https://doi.org/10.1016/S2468-2667(19)30145-8) PMID: 31492648

120. Hoffman DJ, Reynolds RM, Hardy DB. Developmental origins of health and disease: current knowledge and potential mechanisms. *Nutr Rev.* 2017 Dec 1; 75(12):951–970. <https://doi.org/10.1093/nutrit/nux053> PMID: [29186623](https://pubmed.ncbi.nlm.nih.gov/29186623/)
121. Cohen S, Janicki-Deverts D, Chen E, Matthews K. Childhood socioeconomic status and adult health. *Ann N Y Acad Sci.* 2010 Feb; 1186:37–55. <https://doi.org/10.1111/j.1749-6632.2009.05334.x> PMID: [20201867](https://pubmed.ncbi.nlm.nih.gov/20201867/)
122. Muehlenbein MP. *Human Evolutionary Biology.* Cambridge University Press; 2010. p. 624.
123. Chan JCN, Malik V, Jia W, Kadowaki T, Yajnik CS, Yoon K-H, et al. Diabetes in Asia: epidemiology, risk factors, and pathophysiology. *JAMA.* 2009 May 27; 301(20):2129–2140. <https://doi.org/10.1001/jama.2009.726> PMID: [19470990](https://pubmed.ncbi.nlm.nih.gov/19470990/)
124. Huxley R, James WPT, Barzi F, Patel JV, Lear SA, Suriyawongpaisal P, et al. Ethnic comparisons of the cross-sectional relationships between measures of body size with diabetes and hypertension. *Obes Rev.* 2008 Mar; 9 Suppl 1(s1):53–61. <https://doi.org/10.1111/j.1467-789X.2007.00439.x> PMID: [18307700](https://pubmed.ncbi.nlm.nih.gov/18307700/)