The Natural Selection on Infectious Disease Resistance and It's effect on contemporary Health

by C. Justin Cook

Sara Ariza
Laura V. León

Economic Growth and Comparative Development

June 18 2020

Table of Contents

About the author

C. Justin Cook

https://cjustincook.weebly.com/

Assistant Professor of Economics at University of California, Merced.

About the author

C. Justin Cook

Interests:

- Economic Growth.
- Health.
- Development.
- Macrogenoeconomics.

Education:

- Louisiana State University, Baton Rouge, LA USA.
  - Ph.D., Economics, 2012. Dissertation (Advisor: Areendam Chanda): Investigating the Role of Genetic Variation in Explaining Long Run Economic Outcomes
  - M.S., Economics, 2009.
  - B.S., Economics (Concentration in Empirical Studies), 2006.

Abstract

"This paper empirically tests the association between genetically determined resistance to infectious disease and cross-country health differences. A country-level measure of genetic diversity for the system of genes associated with the recognition and disposal of foreign pathogens is constructed. Genetic diversity within this system has been shown to reduce the virulence and prevalence of infectious diseases and is hypothesized to have been naturally selected from historical exposure to infectious pathogens. Base estimation shows a statistically strong, robust, and positive relationship between this constructed measure and country-level health outcomes in times prior to, but not after, the international epidemiological transition."

Introduction

Prior to the major medical discoveries associated with the international epidemiological transition, infectious diseases were a major determinant of mortality and subsequent differences in life expectancy across countries.

Introduction

Prior to the major medical discoveries associated with the international epidemiological transition, infectious diseases were a major determinant of mortality and subsequent differences in life expectancy across countries.

- The discovery and widespread use of effective medicines in the late 1940s to early 1950s is labeled by (Acemoglu&Johnson, 2007)(1) as the international epidemiological transition.

Introduction

Prior to the major medical discoveries associated with the international epidemiological transition, infectious diseases were a major determinant of mortality and subsequent differences in life expectancy across countries.

- The discovery and widespread use of effective medicines in the late 1940s to early 1950s is labeled by (Acemoglu&Johnson, 2007)(1) as the international epidemiological transition.
- In 1940, the average cross-country life expectancy at birth was 47 years with a standard deviation of 12 years; however, for 1980 the average life expectancy grew to 66 years, and the standard deviation across countries fell to 9 years.

Introduction

Question?

What were the causes of the initial cross-country disparities in the virulence of infectious diseases?

Introduction

Question?

What were the causes of the initial cross-country disparities in the virulence of infectious diseases?

Hypothesis

Innate resistance did influence country-level response to infectious disease prior to the international epidemiological transition, however, the effects of innate resistance are dissipated by more efficacious health technologies.

Introduction

The HLA system

The measure of genetic resistance is found within the human leukocyte antigen (HLA) system. The HLA is responsible for locating foreign proteins in order to direct cells of the immune system to initiate an immune response and is broken into two major classes, class I and class II, with both classes being associated with the recognition of certain pathogens (Piertney&Oliver, 2006)(2).

Introduction

The HLA system

The measure of genetic resistance is found within the human leukocyte antigen (HLA) system. The HLA is responsible for locating foreign proteins in order to direct cells of the immune system to initiate an immune response and is broken into two major classes, class I and class II, with both classes being associated with the recognition of certain pathogens (Piertney&Oliver, 2006)(2).

- Using country-level aggregations of ethnic-level genetic data, the author constructs a cross-country measure for diversity within the HLA system: HLA heterozygosity .

Introduction

The HLA system

The measure of genetic resistance is found within the human leukocyte antigen (HLA) system. The HLA is responsible for locating foreign proteins in order to direct cells of the immune system to initiate an immune response and is broken into two major classes, class I and class II, with both classes being associated with the recognition of certain pathogens (Piertney&Oliver, 2006)(2).

- Using country-level aggregations of ethnic-level genetic data, the author constructs a cross-country measure for diversity within the HLA system: HLA heterozygosity .
- **HLA heterozygosity or Expected heterozygosity** is defined as the probability that two randomly selected individuals differ in regard to genetic variants, or alleles, for a particular locus.

Background

Historic Differences in Infectious Disease Environments

The number of infectious diseases humans face increased substantially with the introduction of agriculture, commonly referred to as the **Neolithic revolution** (Wolfe, Dunavan&Diamond 2007)(3)

Background

Historic Differences in Infectious Disease Environments

The number of infectious diseases humans face increased substantially with the introduction of agriculture, commonly referred to as the **Neolithic revolution** (Wolfe, Dunavan&Diamond 2007)(3)

- Agriculture allowed the development of large, dense, and sedentary populations.

Background

Historic Differences in Infectious Disease Environments

The number of infectious diseases humans face increased substantially with the introduction of agriculture, commonly referred to as the **Neolithic revolution** (Wolfe, Dunavan&Diamond 2007)(3)

- Agriculture allowed the development of large, dense, and sedentary populations.
- The domestication of animals in the Neolithic provided closer contact between animals and humans.

Background

Historic Differences in Infectious Disease Environments

The number of infectious diseases humans face increased substantially with the introduction of agriculture, commonly referred to as the **Neolithic revolution** (Wolfe, Dunavan&Diamond 2007)(3)

- Agriculture allowed the development of large, dense, and sedentary populations.
- The domestication of animals in the Neolithic provided closer contact between animals and humans.

The timing of the Neolithic revolution is associated with contemporary differences across populations in genetic diversity within the HLA system.

Background

The Natural Selection of HLA Heterozygosity

The set of genes comprising the HLA system represents one of the most genetically diverse regions of the **genome** (Jeffery & Bangham, 2000) (4).

Background

The Natural Selection of HLA Heterozygosity

The set of genes comprising the HLA system represents one of the most genetically diverse regions of the **genome** (Jeffery & Bangham, 2000) (4).

- This high level of diversity is hypothesized to have been naturally selected as a mechanism of resistance to infectious pathogens (Hughes&Yeager, 1997)(5) (Penman et al., 2013)(6) (Spurgin&Richardson, 2010)(7).

Background

The Natural Selection of HLA Heterozygosity

The set of genes comprising the HLA system represents one of the most genetically diverse regions of the **genome** (Jeffery & Bangham, 2000) (4).

- This high level of diversity is hypothesized to have been naturally selected as a mechanism of resistance to infectious pathogens (Hughes&Yeager, 1997)(5) (Penman et al., 2013)(6) (Spurgin&Richardson, 2010)(7).
- This natural selection for diversity within the HLA system is from balancing selection.

Background

The Natural Selection of HLA Heterozygosity

The set of genes comprising the HLA system represents one of the most genetically diverse regions of the **genome** (Jeffery & Bangham, 2000) (4).

- This high level of diversity is hypothesized to have been naturally selected as a mechanism of resistance to infectious pathogens (Hughes&Yeager, 1997)(5) (Penman et al., 2013)(6) (Spurgin&Richardson, 2010)(7).
- This natural selection for diversity within the HLA system is from balancing selection.
- Balancing selection results from two distinct reasons (Slade&McCallum, 1992)(8):
  - Overdominance.
  - Frequency dependence.

Background

Out-of-Africa Migration and Genetic Diversity

The overall level of genetic diversity within a population has recently been shown to be a function of the population's migratory distance from East Africa (Ashraf Galor, 2013)(9) (Ramachandran et al., 2005)(10).

Background

Out-of-Africa Migration and Genetic Diversity

The overall level of genetic diversity within a population has recently been shown to be a function of the population's migratory distance from East Africa (Ashraf Galor, 2013)(9) (Ramachandran et al., 2005)(10).

- Modern human populations originated within East Africa (roughly Ethiopia) and subsequently migrated to all other continents, excluding Antarctica.

Background

Out-of-Africa Migration and Genetic Diversity

The overall level of genetic diversity within a population has recently been shown to be a function of the population's migratory distance from East Africa (Ashraf Galor, 2013)(9) (Ramachandran et al., 2005)(10).

- Modern human populations originated within East Africa (roughly Ethiopia) and subsequently migrated to all other continents, excluding Antarctica.
- Given that the entire set of genetic diversity was contained within the initial East African population and that migrating populations contain only a subset of this diversity, a strong, negative, and linear association exists between the distance along migration routes a country is from East Africa and the genomic diversity of populations within a country.

Data

HLA Heterozygosity

Ashraf and Galor (2013; hereafter AG)(9) explore the role of genetic variation in explaining historical and contemporary levels of development. In order to measure genetic diversity, they use the expected heterozygosity, roughly defined as: "the probability that two randomly selected individuals differ with respect to the gene in question" (AG, p. 3).

Data

HLA Heterozygosity

Ashraf and Galor (2013; hereafter AG)(9) explore the role of genetic variation in explaining historical and contemporary levels of development. In order to measure genetic diversity, they use the expected heterozygosity, roughly defined as: "the probability that two randomly selected individuals differ with respect to the gene in question" (AG, p. 3).

- Expected heterozygosity is calculated with the frequency of gene variants, or alleles, at a particular site on the genome, or locus:

Data

HLA Heterozygosity

Ashraf and Galor (2013; hereafter AG)(9) explore the role of genetic variation in explaining historical and contemporary levels of development. In order to measure genetic diversity, they use the expected heterozygosity, roughly defined as: "the probability that two randomly selected individuals differ with respect to the gene in question" (AG, p. 3).

- Expected heterozygosity is calculated with the frequency of gene variants, or alleles, at a particular site on the genome, or locus:

$$H_{exp} = 1 - \frac{1}{m} \sum_{l=1}^{m} \sum_{l=1}^{kl} p_i^2 \tag{4.1}$$

Data

HLA Heterozygosity

Ashraf and Galor (2013; hereafter AG)(9) explore the role of genetic variation in explaining historical and contemporary levels of development. In order to measure genetic diversity, they use the expected heterozygosity, roughly defined as: "the probability that two randomly selected individuals differ with respect to the gene in question" (AG, p. 3).

- Expected heterozygosity is calculated with the frequency of gene variants, or alleles, at a particular site on the genome, or locus:

$$H_{exp} = 1 - \frac{1}{m} \sum_{l=1}^{m} \sum_{l=1}^{kl} p_i^2 \tag{4.1}$$

Where:

pi represents the fraction of allele *i* within each population and expected heterozygosity is found by the average across *m* loci.

Data

The author measures of interest is referred to as HLA heterozygosity.

- This measure intends to capture balancing selection within the HLA system from historical exposure to infectious pathogens and is not representative of the entire genoma.

Data

The author measures of interest is referred to as HLA heterozygosity.

- This measure intends to capture balancing selection within the HLA system from historical exposure to infectious pathogens and is not representative of the entire genoma.
- HLA heterozygosity is constructed with data on SNPs from the Allele Frequency Database at Yale University, referred to as ALFRED (Kidd et al., 2003)(11).

Data

The author measures of interest is referred to as HLA heterozygosity.

- This measure intends to capture balancing selection within the HLA system from historical exposure to infectious pathogens and is not representative of the entire genoma.
- HLA heterozygosity is constructed with data on SNPs from the Allele Frequency Database at Yale University, referred to as ALFRED (Kidd et al., 2003)(11).
- A SNP (single-nucleotide polymorphism) is a single change along a strand of DNA.

Data

The author measures of interest is referred to as HLA heterozygosity.

- This measure intends to capture balancing selection within the HLA system from historical exposure to infectious pathogens and is not representative of the entire genoma.
- HLA heterozygosity is constructed with data on SNPs from the Allele Frequency Database at Yale University, referred to as ALFRED (Kidd et al., 2003)(11).
- A SNP (single-nucleotide polymorphism) is a single change along a strand of DNA.
- ALFRED provides allele frequencies for anthropologically defined ethnicities, providing genetic data for 156 SNPs within 19 HLA genes for 51 distinct ethnic groups. These ethnic data are then aggregated to the country level by matching ethnic compositions within (Alesina et al., 2003)(12)

Data

HLA Heterozygosity

- The country-level measure of HLA heterozygosity is the weighted average of ethnic-level HLA heterozygosities, where weights are determined by the fraction of the contemporary population associated with each ethnicity.

Data

HLA Heterozygosity

- The country-level measure of HLA heterozygosity is the weighted average of ethnic-level HLA heterozygosities, where weights are determined by the fraction of the contemporary population associated with each ethnicity.
- Through this method, the author constructed genetic diversity scores for 175 countries, of which 131 are used in the baseline regression model.

Data

HLA Heterozygosity

Table A1. List of Ethnicities by Region in ALFRED

Region	Ethnicities
Africa	Bantu, Biaka, Mandenka, Mbuti, San, Yoruba
Americas	Amerindians, Karitiana, Maya, Pima, Surui
East Asia	Cambodian, Dai, Daur, Han, Hezhe, Japanese, Korean, Lahu, Miao, Naxi, She, Tu, Tujia, Uyghur, Xibe, Yi
Europe	Adygei, Basque, Estonian, French, Italian, Orcadian, Russian, Sardinian
Middle East	Balochi, Bedouin, Brahui, Burusho, Druze, Hazara, Kalash, Mongolian, Mozabite, Oroqen, Palestenian, Pashtun, Sindhi, Yakut
Oceania	Melanesian, Papuan New Guinean

Notes: This table provides the ethnic groups by region for which ALFRED has data. Our country-level measure of HLA heterozygosity is based on genetic data for these ethnicities.

Data

Health Outcomes prior to the International Epidemiological Transition

- The author consider a number of country-level health outcomes prior to the discovery and diffusion of medical technologies associated with the international epidemiological transition.

Data

Health Outcomes prior to the International Epidemiological Transition

- The author consider a number of country-level health outcomes prior to the discovery and diffusion of medical technologies associated with the international epidemiological transition.
- These include both predicted mortality from infectious disease and life expectancy in 1940, as well as life expectancy in 1960.

Data

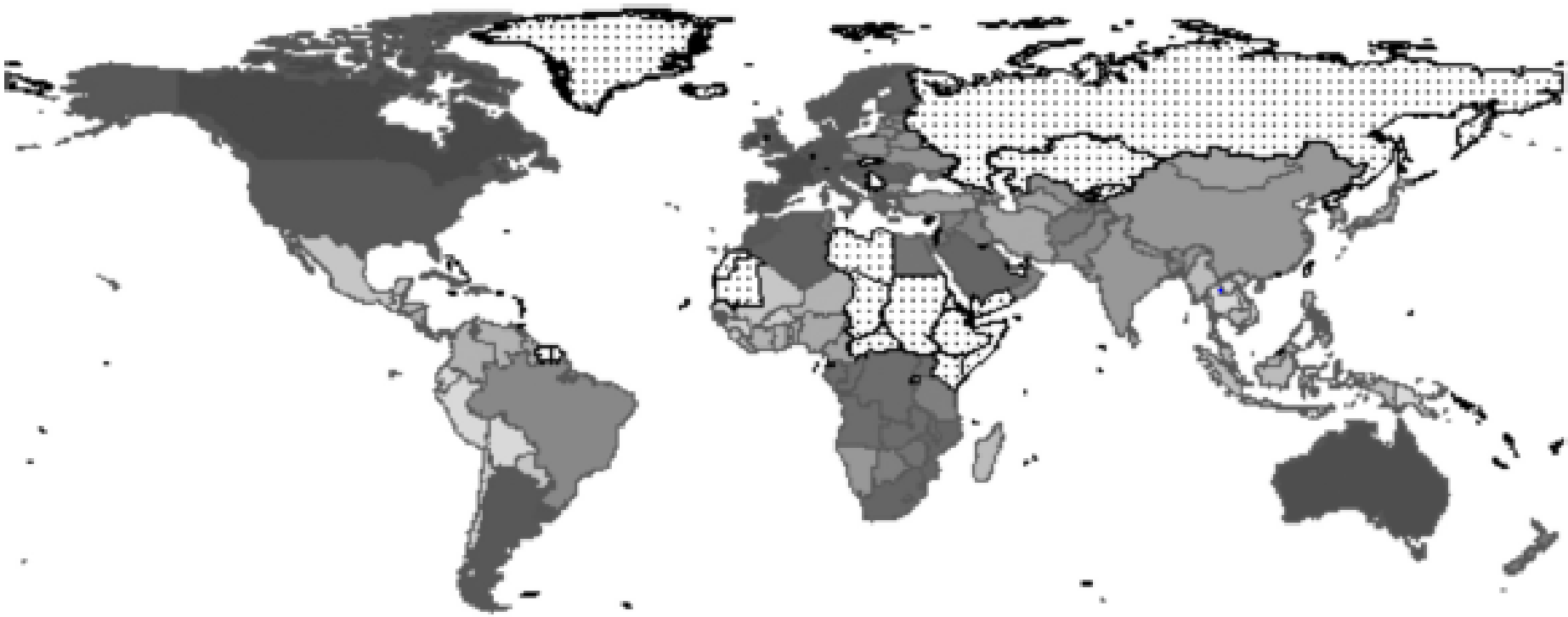
Health Outcomes prior to the International Epidemiological Transition

- The author consider a number of country-level health outcomes prior to the discovery and diffusion of medical technologies associated with the international epidemiological transition.
- These include both predicted mortality from infectious disease and life expectancy in 1940, as well as life expectancy in 1960.
- The use of 1940s data, while truly before the epidemiological transition, is problematic due to a lack of data in relatively poor countries, leading to possible selection bias. Therefore, the primary dependent variable is country level life expectancy at birth in 1960

Data

HLA Heterozygosity

FIGURE 1.—CROSS-COUNTRY HLA HETEROZYGOSITY



Darker areas represent increased HLA heterozygosity. HLA heterozygosity is calculated using contemporary populations.

Results

Explaining HLA Heterozygosity

Table 1 shows the summary statistics for : *i)* HLA heterozygosity and *ii)* the overall level of genetic diversity from (Ashraf&Galor, 2013)(9).

TABLE 1.—SUMMARY STATISTICS: HLA HETEROZYGOSITY VERSES AGGREGATE HETEROZYGOSITY

Variable	N	Mean	SD	Minimum	Maximum
HLA heterozygosity by continent	131	0.3183	0.0207	0.2347	0.3529
Europe	32	0.3343	0.010	0.3184	0.3529
Africa	37	0.3158	0.0176	0.2844	0.3352
Asia	35	0.3149	0.0146	0.2711	0.3298
Americas	24	0.3078	0.0253	0.2588	0.3503
Oceania	3	0.3022	0.0590	0.2347	0.3439
Overall heterozygosity by continent (Ashraf & Galor, 2013)	131	0.7248	0.0272	0.6279	0.7653
Europe	32	0.7342	0.0057	0.7217	0.7429
Africa	37	0.7444	0.0097	0.7282	0.7653
Asia	35	0.7208	0.023	0.682	0.7519
Americas	24	0.6912	0.0324	0.6279	0.7446
Oceania	3	0.6983	0.0360	0.6573	0.7248

HLA heterozygosity is a measure for genetic variation with the HLA system, which is responsible for the recognition of foreign pathogens. Overall heterozygosity is a measure for genetic variation across the entire genome. This measure is calculated with neutral genetic markers, or markers due to random genetic drift between isolated populations. Due to the origination of modern human populations within Africa, aggregate heterozygosity is a declining, linear function of the migratory distance from East Africa. Using genetic diversity data for 53 ethnic groups and the migratory distance from East Africa, Ashraf and Galor (2013) predict a country-level measure of heterozygosity. These variables are explained in greater detail in the appendix of variables.

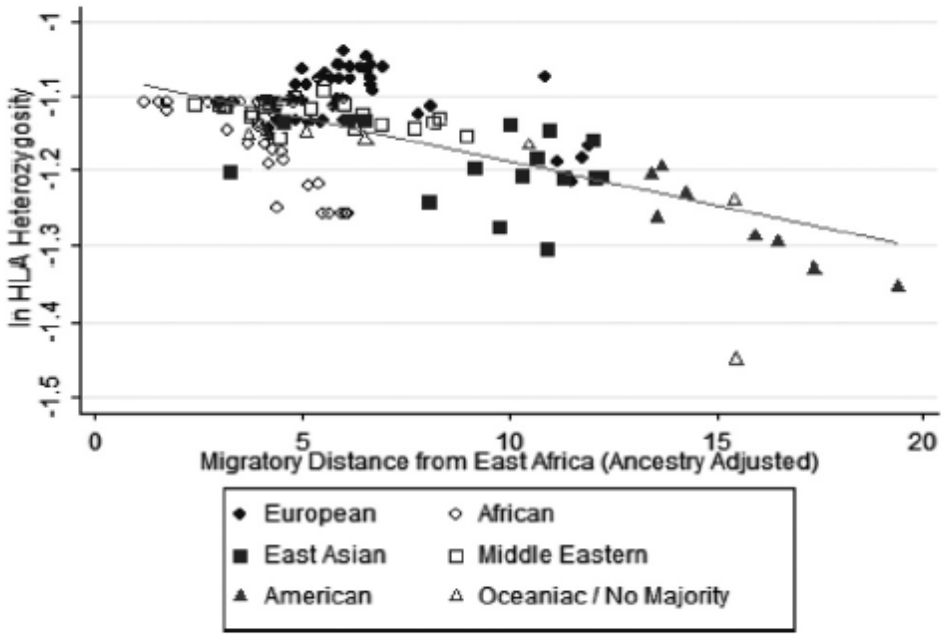
Acording to the results, African countries are found to have the highest levels of overall diversity, while countries from Europe contain the greatest amount of diversity within the HLA system.

Results

Explaining HLA Heterozygosity

Figure 2 plots the measure of diversity, HLA heterozygosity, as a linear function of migratory distance from East Africa.

FIGURE 2.—HLA HETEROZYGOSITY AND MIGRATORY DISTANCE FROM EAST AFRICA



This figure displays the cross-country relationship between HLA heterozygosity and an ancestry-adjusted measure of migratory distance from East Africa. Countries are shaded by the majority of the population being from a noted region within the ALFRED data. Note that countries with populations from Europe and Middle East contain greater levels of HLA heterozygosity than is predicted by the linear trend.

Results

Explaining HLA Heterozygosity

From *figure 2*, countries with a majority population from Europe and the Middle East tend to break from the linear association between HLA heterozygosity and migratory distance from East Africa, containing higher-than-predicted levels of HLA heterozygosity.

The role of both the migratory distance from East Africa and the Neolithic revolution in explaining HLA heterozygosity is tested in *table 2*.

Results

Explaining HLA Heterozygosity

TABLE 2.—EXPLAINING HLA HETEROZYGOSITY

	Dependent Variable: ln HLA Heterozygosity					
	(1)	(2)	(3)	(4)	(5)	(6)
ln Years since Neolithic Revolution	0.0211**	0.0300***				
(Ancestry Adjusted, 1500–1960)	(0.0098)	(0.0086)				
ln No. of Potential Domesticate Animals			0.0266***	0.0361***		
Ancestry Adjusted, 1500–1960)			(0.0064)	(0.0051)		
ln Population Density in 1 CE					0.0144***	0.0158***
(Ancestry Adjusted, 1500–1960)					(0.0043)	(0.0033)
Migratory Distance from East Africa		−0.0121***		−0.0133***		−0.0124***
(Ancestry Adjusted, 1500–1960)		(0.0014)		(0.0015)		(0.0015)
N	131	131	89	89	113	113
R^2	0.0217	0.4265	0.1176	0.5968	0.0626	0.4578

This table displays the relationship of factors associated with HLA heterozygosity. I consider four main determinants: the years a country has practiced agriculture, the availability of domesticate animals, the density of historic populations, and the migratory distance from East Africa. The Neolithic revolution is the ultimate cause of differential disease environments by providing dense populations and close contact with domesticate animals. Domesticate animals are associated with the initiation of many infectious diseases, whereas historically dense populations provide necessary hosts. Due to the serial founder effect, migratory patterns from East Africa have strong associations with overall genetic diversity. All coefficients are statistically significant with the expected sign.
HLA heterozygosity is a measure for genetic variation with the HLA system, which is responsible for the recognition of foreign pathogens. Sources and definitions of all variables are given within the appendix of variables. Ancestry-adjusted measures use the matrix of migration between 1500 and 1960 from Chanda, Cook, and Putterman (2014). OLS coefficients are reported in each column. Significant *10%, **5%, ***1%. Robust standard errors are in parentheses.

Results

HLA Heterozygosity, Country-Level Health Outcomes, and the International Epidemiological Transition

Equation used to test that genetic resistance to infectious disease, measured by HLA heterozygosity, is positively associated with country-level health outcomes prior to the international epidemiological transition.,

Results

HLA Heterozygosity, Country-Level Health Outcomes, and the International Epidemiological Transition

Equation used to test that genetic resistance to infectious disease, measured by HLA heterozygosity, is positively associated with country-level health outcomes prior to the international epidemiological transition.,

$$\ln y_i^{t<i.e.t.} = \alpha + \beta_1(\ln HLA_i) + \boldsymbol{\beta}_2' \mathbf{X_i} + \boldsymbol{\beta}_3' \mathbf{I_i^c} + \epsilon_i \qquad (5.1)$$

Results

HLA Heterozygosity, Country-Level Health Outcomes, and the International Epidemiological Transition

$$\ln y_i^{t<i.e.t.} = \alpha + \beta_1(\ln HLA_i) + \boldsymbol{\beta}_2' \mathbf{X_i} + \boldsymbol{\beta}_3' \mathbf{I_i^c} + \epsilon_i$$

Results

HLA Heterozygosity, Country-Level Health Outcomes, and the International Epidemiological Transition

$$\ln y_i^{t<i.e.t.} = \alpha + \beta_1(\ln HLA_i) + \boldsymbol{\beta}_2' \mathbf{X_i} + \boldsymbol{\beta}_3' \mathbf{I_i^c} + \epsilon_i$$

- i: is a country indicator.
- $yt \leq i.e.t.$: i represents aggregate health outcomes prior to the international epidemiological transition.
- β_1: measures the effect of HLA heterozygosity.
- X_i: is a vector of country-level controls, including ethnic fractionalization, agricultural productivity, geography, and an ancestry-adjusted measure for the agricultural transition.
- I_i^c: is an indicator variable as to whether country i is within continent c.
- ϵ_i: is the cross-country error term.

Results

Explaining HLA Heterozygosity

TABLE 3.—THE EFFECT OF HLA HETEROZYGOSITY PRIOR TO THE INTERNATIONAL EPIDEMIOLOGICAL TRANSITION

	ln Predicted Mortality in 1940		ln Life Expectancy in 1940		ln Life Expectancy in 1960	
Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)
ln HLA Heterozygosity	−5.0080***	−3.8610***	2.1436***	1.0771***	1.6200***	1.0151***
	(0.7117)	(1.1282)	(0.3310)	(0.3771)	(0.2455)	(0.1899)
ln Ethnic Fractionalization		−0.3970		0.0724		−0.1210
		(0.3961)		(0.1338)		(0.0927)
ln Years since Neolithic Revolution		0.5975**		−0.0031		0.0370
(Ancestry Adjusted, 1500–1960)		(0.2358)		(0.0549)		(0.0397)
ln Fraction of Arable Land		0.0449		−0.1060***		−0.0087
		(0.1220)		(0.0366)		(0.0160)
ln Suitability of Agriculture		0.0362		0.0530**		0.0080
		(0.0844)		(0.0238)		(0.0128)
ln Abs. Latitude		−0.3260**		0.0980**		0.0199
		(0.1297)		(0.0410)		(0.0156)
Continent fixed effects	No	Yes	No	Yes	No	Yes
N	73	73	71	71	131	131
R^2	0.3488	0.5133	0.3329	0.7192	0.2271	0.7446

This table displays the relationship between HLA heterozygosity and Acemoglu and Johnson's (2007) measures of health before the epidemiological transition as well as the baseline estimation, which uses life expectancy at birth in 1960 as the main dependent variable. The use of 1960 data is intended to increase the sample in terms of both size and global representation. Odd-numbered columns provide the simple bivariate relationship between HLA heterozygosity and the alternative measures of premedicinal health, and even-numbered columns include baseline controls and continent fixed effects.

HLA heterozygosity is a measure for genetic variation with the HLA system, which is responsible for the recognition of foreign pathogens. Sources and definitions of all variables are given within the appendix of variables. Continent fixed effects consist of dummies for Africa, the Americas, Asia, and Europe. Ancestry-adjusted measures use the matrix of migration between 1500 and 1960 from Chanda et al. (2014). OLS coefficients are reported in each column. Significant *10%, **5%, ***1%. Robust standard errors are in parentheses.

Results

HLA Heterozygosity, Country-Level Health Outcomes, and the International Epidemiological Transition

The Estimates in *table 3* consider three alternatives for measuring country-level health outcomes prior to the epidemiological transition:

Results

HLA Heterozygosity, Country-Level Health Outcomes, and the International Epidemiological Transition

The Estimates in *table 3* consider three alternatives for measuring country-level health outcomes prior to the epidemiological transition:

- The mortality rate from fifteen infectious diseases in 1940.

Results

HLA Heterozygosity, Country-Level Health Outcomes, and the International Epidemiological Transition

The Estimates in *table 3* consider three alternatives for measuring country-level health outcomes prior to the epidemiological transition:

- The mortality rate from fifteen infectious diseases in 1940.
- Life expectancy at birth in 1940.

Results

HLA Heterozygosity, Country-Level Health Outcomes, and the International Epidemiological Transition

The Estimates in *table 3* consider three alternatives for measuring country-level health outcomes prior to the epidemiological transition:

- The mortality rate from fifteen infectious diseases in 1940.
- Life expectancy at birth in 1940.
- Life expectancy at birth in 1960.

Results

HLA Heterozygosity, Country-Level Health Outcomes, and the International Epidemiological Transition

Table 4 explores the effect of HLA heterozygosity on life expectancy from 1960 to 2010 and tests the effect of HLA heterozygosity, or innate resistance, in post epidemiological transition environments. It shows that 1960 is an early enough period to capture the effects of innate resistance.

TABLE 4.—THE EFFECT OF HLA HETEROZYGOSITY AFTER THE INTERNATIONAL EPIDEMIOLOGICAL TRANSITION
Dependent Variable: ln Life Expectancy

	1960 (1)	1970 (2)	1980 (3)	1990 (4)	2000 (5)	2010 (6)
ln HLA Heterozygosity	1.0151***	1.0405***	0.8421***	0.5490***	0.2512	0.1788
	(0.1899)	(0.1696)	(0.1700)	(0.1798)	(0.1541)	(0.1335)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Continent fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	131	131	131	131	131	131
R^2	0.7446	0.7340	0.6995	0.6886	0.7683	0.7480

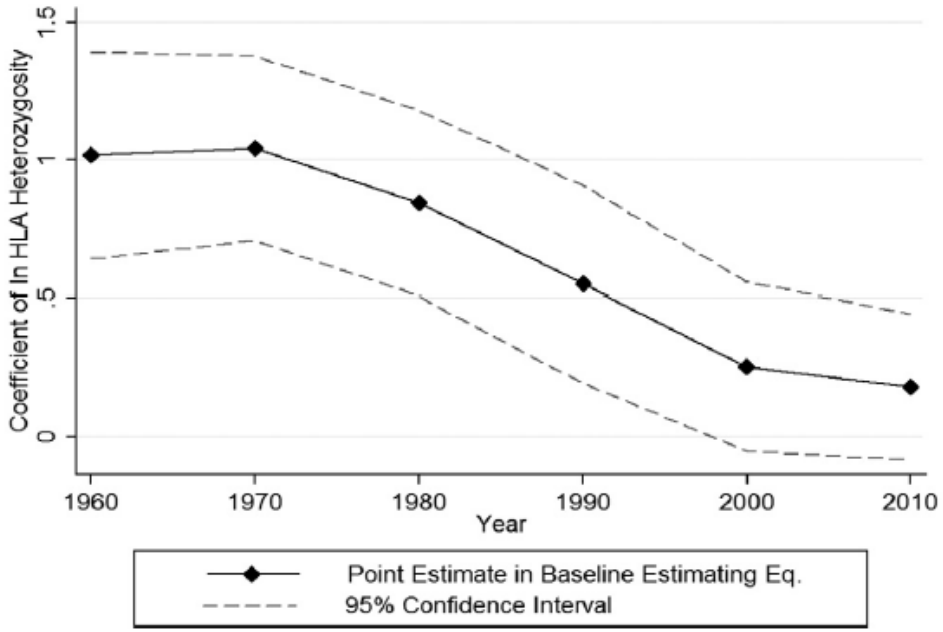
This table provides support for the lessened benefit of genetically determined resistance following the international epidemiological transition and for the use of 1960 data as a valid proxy for health outcomes prior to the diffusion of health technologies associated with epidemiological transition. More contemporary periods are associated with a greater prevalence and use of medical technologies of the epidemiological transition; therefore, the effect of inherent resistance on life expectancy lessens in magnitude, becoming insignificantly different from zero in 2000 CE.

HLA heterozygosity is a measure for genetic variation with the HLA system, which is responsible for the recognition of foreign pathogens. Baseline controls include ethnic fractionalization, an ancestry-adjusted measure for the years a country has practiced agriculture, the fraction of arable land within a country, the suitability of agriculture within a country, and absolute latitude. Continent fixed effects consist of dummies for Africa, the Americas, Asia, and Europe. OLS coefficients are reported in each column. Significant *10%, **5%, ***1%. Robust standard errors are in parentheses.

Results

HLA Heterozygosity, Country-Level Health Outcomes, and the International Epidemiological Transition

FIGURE 3.—THE EFFECT OF HLA HETEROZYGOSITY FOLLOWING THE DIFFUSION OF HEALTH TECHNOLOGIES FROM THE INTERNATIONAL EPIDEMIOLOGICAL TRANSITION



This figure displays the point estimates for the coefficient of HLA heterozygosity in table 4. Note that the effect of HLA heterozygosity declines over time. A key hypothesis is that this decline is due to the diffusion of medical technologies associated with the international epidemiological transition.

Results

Robustness

Controlling for regional ethnic differences.

The relationship between HLA heterozygosity and life expectancy in 1960 may reflect some underlying role of European populations in promoting greater health outcomes. In addition, other populations may have unobserved effects that are also correlated with both HLA heterozygosity and life expectancy in 1960.

Results

Robustness

Controlling for regional ethnic differences.

The relationship between HLA heterozygosity and life expectancy in 1960 may reflect some underlying role of European populations in promoting greater health outcomes. In addition, other populations may have unobserved effects that are also correlated with both HLA heterozygosity and life expectancy in 1960.

TABLE 5.—ROBUSTNESS TO THE INFLUENCE OF REGIONAL POPULATIONS
Dependent Variable: ln Life Expectancy in 1960

	% European = 0 (1)	% European ∈ (0,1) (2)	% European = 1 (3)	Full (4)
ln HLA Heterozygosity	0.7241*	0.8049***	0.7281**	0.8870***
	(0.3623)	(0.2923)	(0.3111)	(0.2872)
Baseline controls	Yes	Yes	Yes	Yes
Continent fixed effects	Yes	Yes	No	Yes
Population fixed effects	No	No	No	Yes
N	55	52	24	131
R^2	0.4183	0.5891	0.5162	0.7891

This table performs both a sample truncation and the inclusion of population fixed effect to control for any potential health benefits associated with populations from a particular region. Easterly and Levine (2012) argue that European populations contained both human capital and institutional advantages in accumulating wealth. I therefore restrict the sample based on the fraction of the contemporary population that is from Europe in columns 1 to 3. Column 4 includes population fixed effects; that is, we control for the contemporary fraction of each regional population within each country. Controlling for regional populations, with special attention given to populations from Europe, does not alter the main finding.

HLA heterozygosity is a measure for genetic variation with the HLA system, which is responsible for the recognition of foreign pathogens. ALFRED data provide six regions: Africa, the Americas, East Asia, Europe, the Middle East, and Oceania. The country-level population fraction from each region is calculated using ethnic compositions from Alesina et al. (2003) matched to the ethnic groups of ALFRED. Baseline controls include ethnic fractionalization, an ancestry-adjusted measure for the years a country has practiced agriculture, the fraction of arable land within a country, the suitability of agriculture within a country, and absolute latitude. Continent fixed effects consist of dummies for Africa, the Americas, Asia, and Europe. OLS coefficients are reported in each column. Significant *10%, **5%, ***1%. Robust standard errors are in parentheses.

Results

Robustness

Omitted variables.
The author includes omitted variables that may be associated with either the measure of HLA heterozygosity or life expectancy in 1960.

Results

Robustness

Omitted variables.
The author includes omitted variables that may be associated with either the measure of HLA heterozygosity or life expectancy in 1960. The additional variables are broken into two classes:

Results

Robustness

Omitted variables.

The author includes omitted variables that may be associated with either the measure of HLA heterozygosity or life expectancy in 1960. The additional variables are broken into two classes:

- Exogenous: Include genetic, geographic, and historic population controls *(Table 6)*.

Results

Robustness

Omitted variables.
The author includes omitted variables that may be associated with either the measure of HLA heterozygosity or life expectancy in 1960. The additional variables are broken into two classes:

- Exogenous: Include genetic, geographic, and historic population controls *(Table 6)*.
- Endogenous: Consist of income, human capital, and demographics in 1960 *(Table 7)*.

Results

Robustness

TABLE 6.—ROBUSTNESS TO EXOGENOUS OMITTED VARIABLES
Dependent Variable: ln Life Expectancy in 1960

	(1)	(2)	(3)	(4)
ln HLA Heterozygosity	1.0771***	0.7713***	0.8699***	0.5373*
	(0.3401)	(0.2040)	(0.2239)	(0.3093)
Aggregate heterozygosity	−0.2784			0.1385
(Ancestry adjusted, 1500–1960)	(1.1998)			(1.0280)
Fraction of population at risk of contracting malaria		−0.1613***		−0.1730***
		(0.0481)		(0.0487)
% within tropics		−0.0003		−0.0003
		(0.0006)		(0.0006)
% within desert		−0.0016		−0.0013
		(0.0015)		(0.0014)
Mean distance to coast or river		0.0116		0.0299
		(0.0385)		(0.0381)
Fraction of nonindigenous population			0.0982	0.1203
			(0.0632)	(0.0734)
Baseline controls	Yes	Yes	Yes	Yes
Continent fixed effects	Yes	Yes	Yes	Yes
N	131	131	131	131
R^2	0.7449	0.7741	0.7504	0.7824

A number of potentially confounding exogenous omitted variables are included in the table: aggregate genetic diversity, a measure for the suitability of malaria, the fraction of a country within the tropics, the fraction of a country that is desert, the average distance within a country to a coast or river, and the fraction of the population that has migrated into the country as of 1960. The inclusion of these controls, both piecemeal and jointly, does not alter the positive, statistically significant effect of HLA heterozygosity, although in the joint estimation of column 5, the coefficient is lessened in magnitude and significance falls to the 10% level.

HLA heterozygosity is a measure for genetic variation with the HLA system, which is responsible for the recognition of foreign pathogens. Sources and definitions of all variables are given within the appendix of variables. Baseline controls include ethnic fractionalization, an ancestry-adjusted measure for the years a country has practiced agriculture, the fraction of arable land within a country, the suitability of agriculture within a country, and absolute latitude. Continent fixed effects consist of dummies for Africa, the Americas, Asia, and Europe. OLS coefficients are reported in each column. Significant *10%, **5%, ***1%. Robust standard errors are in parentheses.

Results

Robustness

TABLE 7.—ROBUSTNESS TO ENDOGENOUS OMITTED VARIABLES
Dependent Variable: ln Life Expectancy in 1960

	(1)	(2)	(3)	(4)	(5)
ln HLA Heterozygosity	1.0000***	0.6575***	0.6127***	0.7529***	0.5296***
	(0.2118)	(0.2124)	(0.1633)	(0.1944)	(0.1673)
ln GDP per Capita in 1960		0.0977***			0.0280*
		(0.0221)			(0.0167)
ln Avg. Years of School in 1960			0.1320***		0.1037***
			(0.0195)		(0.0190)
ln Population Density in 1960				0.0002	0.0063
				(0.0152)	(0.0126)
ln Urbanization Rate in 1960				0.1137***	0.0343*
				(0.0223)	(0.0195)
ln Fraction of Population under 15 Years in 1960				−0.0614	−0.0148
				(0.0887)	(0.0685)
Baseline controls	Yes	Yes	Yes	Yes	Yes
Continent fixed effects	Yes	Yes	Yes	Yes	Yes
N	109	109	109	109	109
R^2	0.7378	0.8062	0.8727	0.8256	0.8897

This table controls for a number of potentially endogenous controls: income, human capital, and demographics in 1960. Given the relationship between income and genetic diversity (Ashraf & Galor, 2013), controlling for income and related variables is intended to dispel the effect of HLA heterozygosity working through an income channel. When controlling for these additional measures, the effect of HLA heterozygosity remains both positive and significant at the 1% level; although, as with joint inclusion of all endogenous omitted variables, the coefficient on HLA heterozygosity is reduced in magnitude in column 5.

HLA heterozygosity is a measure for genetic variation with the HLA system, which is responsible for the recognition of foreign pathogens. Sources and definitions of all variables are given within the appendix of variables. Baseline controls include ethnic fractionalization, an ancestry-adjusted measure for the years a country has practiced agriculture, the fraction of arable land within a country, the suitability of agriculture within a country, and absolute latitude. Continent fixed effects consist of dummies for Africa, the Americas, Asia, and Europe. OLS coefficients are reported in each column. Significant *10%, **5%, ***1%. Robust standard errors are in parentheses.

Conclusion

- HLA heterozygosity has a positive, statistically significant, and robust relationship with life expectancy at birth in 1960, a period argued to be before the diffusion of health technologies associated with the international epidemiological transition.

Conclusion

- HLA heterozygosity has a positive, statistically significant, and robust relationship with life expectancy at birth in 1960, a period argued to be before the diffusion of health technologies associated with the international epidemiological transition.
- The strong statistical relationship between HLA heterozygosity and life expectancy is substantially lessened by the introduction of medicines and vaccines, which dissipate any benefits from genetically determined resistance.

Conclusion

- HLA heterozygosity has a positive, statistically significant, and robust relationship with life expectancy at birth in 1960, a period argued to be before the diffusion of health technologies associated with the international epidemiological transition.
- The strong statistical relationship between HLA heterozygosity and life expectancy is substantially lessened by the introduction of medicines and vaccines, which dissipate any benefits from genetically determined resistance.
- An important source of the variation in HLA heterozygosity is the differential timing date of the Neolithic revolution, which provided the means of development for a much more severe infectious disease environment.

Bibliography

- Cook, C.J. (2015). The natural selection of infectious disease resistance and its effect on contemporary health. The Review of Economics and Statistics, 97, 742-757.

References

[1] D. Acemoglu and S. Johnson, "Disease and development: the effect of life expectancy on economic growth," *Journal of political Economy*, vol. 115, no. 6, pp. 925–985, 2007.

[2] S. Piertney and M. Oliver, "The evolutionary ecology of the major histocompatibility complex," *Heredity*, vol. 96, no. 1, pp. 7–21, 2006.

[3] N. D. Wolfe, C. P. Dunavan, and J. Diamond, "Origins of major human infectious diseases," *Nature*, vol. 447, no. 7142, pp. 279–283, 2007.

[4] K. J. Jeffery and C. R. Bangham, "Do infectious diseases drive mhc diversity?," *Microbes and infection*, vol. 2, no. 11, pp. 1335–1341, 2000.

[5] A. L. Hughes and M. Yeager, "Natural selection at major histocompatibility complex loci of vertebrates," *Annual review of genetics*, vol. 32, no. 1, pp. 415–435, 1998.

[6] B. S. Penman, B. Ashby, C. O. Buckee, and S. Gupta, "Pathogen selection drives nonoverlapping associations between hla loci," *Proceedings of the National Academy of Sciences*, vol. 110, no. 48, pp. 19645–19650, 2013.

[7] L. G. Spurgin and D. S. Richardson, "How pathogens drive genetic diversity: Mhc, mechanisms and misunderstandings," *Proceedings of the Royal Society B: Biological Sciences*, vol. 277, no. 1684, pp. 979–988, 2010.

[8] R. Slade and H. McCallum, "Overdominant vs. frequency-dependent selection at mhc loci," *Genetics*, vol. 132, no. 3, p. 861, 1992.

[9] Q. Ashraf and O. Galor, "The'out of africa'hypothesis, human genetic diversity, and comparative economic development," *American Economic Review*, vol. 103, no. 1, pp. 1–46, 2013.

[10] S. Ramachandran, O. Deshpande, C. C. Roseman, N. A. Rosenberg, M. W. Feldman, and L. L. Cavalli-Sforza, "Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in africa," *Proceedings of the National Academy of Sciences*, vol. 102, no. 44, pp. 15942–15947, 2005.

[11] K. Kidd, H. Rajeevan, M. Osier, K. Cheung, H. Deng, L. Druskin, R. Heinzen, J. Kidd, S. Stein, A. Pakstis, *et al.*, "Alfred-the allele frequency database-an update.," in *American Journal of Physical Anthropology*, pp. 128–128, WILEY-LISS DIV JOHN WILEY & SONS INC, 605 THIRD AVE, NEW YORK, NY 10158-0012 USA, 2003.

[12] A. Alesina, A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg, "Fractionalization," *Journal of Economic growth*, vol. 8, no. 2, pp. 155–194, 2003.

Contact information

Sara Ariza Murillo

Economist, Universidad del Rosario.

email: saraariza97@gmail.com

Laura Viviana León Díaz

Finance and International Trade, Universidad del Rosario.

email: leond.laura@gmail.com