Human Genetic Variation and Yellow Fever Mortality during 19th Century U.S. Epidemics

Lauren E. Blake,a,b,c Mariano A. Garcia-Blancoa,c,d,e

Center for RNA Biology, Duke University, Durham, North Carolina, USA; Department of Evolutionary Anthropology, Duke University, Durham, North Carolina, USA; Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, North Carolina, USA; Department of Medicine, Duke University Medical Center, Durham, North Carolina, USA; Program in Emerging Infectious Diseases, Duke-NUS Graduate School, Singapore

ABSTRACT We calculated the incidence, mortality, and case fatality rates for Caucasians and non-Caucasians during 19th century yellow fever (YF) epidemics in the United States and determined statistical significance for differences in the rates in different populations. We evaluated nongenetic host factors, including socioeconomic, environmental, cultural, demographic, and acquired immunity status that could have influenced these differences. While differences in incidence rates were not significant between Caucasians and non-Caucasians, differences in mortality and case fatality rates were statistically significant for all epidemics tested ($P<0.01$). Caucasians diagnosed with YF were 6.8 times more likely to succumb than non-Caucasians with the disease. No other major causes of death during the 19th century demonstrated a similar mortality skew toward Caucasians. Nongenetic host factors were examined and could not explain these large differences. We propose that the remarkably lower case mortality rates for individuals of non-Caucasian ancestry is the result of human genetic variation in loci encoding innate immune mediators.

IMPORTANCE Different degrees of severity of yellow fever have been observed across diverse populations, but this study is the first to demonstrate a statistically significant association between ancestry and the outcome of yellow fever (YF). With the global burden of mosquito-borne flaviviral infections, such as YF and dengue, on the rise, identifying and characterizing host factors could prove pivotal in the prevention of epidemics and the development of effective treatments.

The identification of host factors that aid or restrict pathogens has transformed our understanding of disease susceptibility, outcomes, and possible treatment responses for major global public health threats, including malaria (1) and HIV/AIDS (2). Understanding yellow fever (YF) host factors is a crucial part of a comprehensive approach to protect the public, as information about host variation, when combined with climate, geographical, and additional population data, could be used to better predict the future locations of especially severe outbreaks of YF. Most importantly, an understanding of critical host factors could help develop novel antiflaviviral treatments.

Identification of host factors for mosquito-transmitted flaviviruses, important global and reemerging threats, has begun using cell-based assays (3–5), but studies of human populations are only beginning. In fact, genome-wide association studies (GWAS) have recently been used to identify potential genetic host factors for dengue hemorrhagic fever (DHF)/dengue shock syndrome (DSS) (6) and also nonsevere dengue fever (DF) (7). Estimates typically put the global burden of YF at 200,000 cases and 30,000 deaths annually (8), although a recent study suggests that these numbers may significantly underestimate this burden (4). The reemergence of YF has been attributed to lack of vaccination in regions at risk for YF, the spread of the mosquito vector range, and increased urbanization (9). These factors, in tandem with increased globalization and climate change, contribute to a growing number of areas at risk for YF; of particular concern is the potential of an urban epidemic in a tropical or subtropical megacity (9).

The United States in the 19th century was a unique setting to observe differences in YF susceptibility between populations of different ancestry, as the country contained large populations of European and African ancestries living in close proximity. Furthermore, the United States was never an area where YF was endemic, and therefore, few people contracted YF outside of recorded epidemics, decreasing the ambiguity of diagnosis, which complicates historical evaluations of disease variance between populations. Anecdotal evidence from 19th century YF epidemics in the United States suggests that Caucasians succumbed to the disease at considerably higher rates than those of African descent (10), but this difference in mortality has never been rigorously established. Studies on historical epidemics (10, 11) and research focusing on responses to the YF vaccine (12) have used data trends to suggest that differences in mortality between races could be...
explained by genetic factors. This explanation and other explanations for the different susceptibilities (13) have not been subjected to statistics-based hypothesis testing.

In this study, we integrated data from a variety of sources, including primary-source documents created by medical personnel who treated patients during the epidemics, census data, and information about the epidemics compiled by contemporary historians. Taking advantage of unique data sets from 19th century U.S. urban YF epidemics, we utilized tests of statistical inference to establish significant differences in YF mortality and case fatality rates between Caucasians and non-Caucasians. By performing rigorous hypothesis testing on the effects of socioeconomic, environmental, cultural, and demographic factors on this association, our results strongly suggest that host genetic variation affects yellow fever virus (YFV) infection outcomes. Based on these results, a large-scale genetics study to identify host factors for YF susceptibility is warranted.

**RESULTS AND DISCUSSION**

Caucasians were more likely to succumb to yellow fever than non-Caucasians during 19th century epidemics. There was no consistent difference in the incidence of YF between Caucasians and non-Caucasians (see Text S1 and Table S1 in the supplemental material). On the other hand, mortality rates, which varied widely between epidemics, were 1.86- to 39.42-fold higher for Caucasians in all 16 epidemics for which data are available (Table S2). The differences in mortality rates between populations were statistically significant \( P < 0.01 \), indicating that population-based differences in YF susceptibility existed.

Since the set of clinical features of severe YF cases is distinctive, the potential of mistaking the cause of death as YF was unlikely. Nonetheless, it is possible that mortality rates could have been significantly affected by differences in socioeconomic status, particularly access to health care providers. Cultural bias could also have affected mortality rates, since until the 1840s many physicians believed that those of non-Caucasian descent could not contract YF (14), a misconception which has been refuted. To minimize these confounding variables, we focused on epidemics where non-Caucasians were poorly followed up by medical personnel. Six epidemics met this additional inclusion criterion.

In these six epidemics, the case fatality rate for Caucasians varied from 25.0 to 72.5%, while the case fatality rate for non-Caucasians varied from 1.1 to 14.1% (Table 1). This remarkable skew in case fatality rates between races was observed in all six epidemics. When we constructed regression models for each of the six epidemics, the increased case fatality rates for Caucasians relative to non-Caucasians were all significant \( P < 0.01 \) [Table 1]. When the data sets were combined, the case fatality rate was 6.8 times greater for Caucasians than non-Caucasians, and the odds of Caucasian patients succumbing to YF were 14.6 times greater than those of non-Caucasian patients (with a 95% confidence interval of 13.59 to 15.76 [Table 1]).

We considered recording bias as an explanation for this dramatic difference in case fatality rates. For instance, it is possible that non-Caucasians were poorly followed up by medical personnel. This confounding variable was minimized for the epidemic among the U.S. troops in New Orleans in 1867, given the War Department’s emphasis on record-keeping (15). The Caucasian troop case fatality rate was 1.99-fold higher \( (P < 0.01) \) and the odds of Caucasian troops dying from YF were 2.56 times greater than those of non-Caucasian soldiers (Table 1). Given the more-homogeneous treatment among troops, these data also argue against the possibility that YF “treatments” such as quinine, castor oil, and calomel (15, 16) or some other aspect of the care (17) could have increased case fatality among Caucasians, a population that was more likely to have access to care, particularly earlier in the disease cycle or for milder cases (see Text S2 in the supplemental material). We conclude from these data that there was a remarkably higher and statistically significant case fatality rate

**TABLE 1** Case fatality rates stratified by race during yellow fever epidemics in the United States (1808 to 1878)*

<table>
<thead>
<tr>
<th>Yellow fever epidemic by location and year</th>
<th>No. of deaths/100 YF cases in Caucasians (no. of cases in Caucasians)</th>
<th>No. of deaths/100 YF cases in non-Caucasians (no. of cases in non-Caucasians)</th>
<th>Fold change in case fatality rate ((C/nC))</th>
<th>Odds (95% confidence interval) of Caucasians versus non-Caucasians succumbing to YF</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>St. Mary’s, GA (1808)</td>
<td>48.3 (87)</td>
<td>6.7 (45)</td>
<td>7.24***</td>
<td>13.07 (4.33–56.78)</td>
<td>28</td>
</tr>
<tr>
<td>Norfolk, VA (1855)</td>
<td>46.2 (143)</td>
<td>6.0 (50)</td>
<td>7.69***</td>
<td>13.43 (4.64–57.05)</td>
<td>29</td>
</tr>
<tr>
<td>Portsmouth, VA (1855)</td>
<td>42.0 (2,264)</td>
<td>5.0 (1,980)</td>
<td>8.40***</td>
<td>13.76 (11.11–17.22)</td>
<td>26</td>
</tr>
<tr>
<td>U.S. troops in New Orleans, LA (1867)</td>
<td>29.6 (659)</td>
<td>14.1 (163)</td>
<td>2.10***</td>
<td>2.56 (1.63–4.19)</td>
<td>15</td>
</tr>
<tr>
<td>Port Royal and Sea Islands, SC (1877)</td>
<td>25.0 (96)</td>
<td>1.1 (87)</td>
<td>21.75**</td>
<td>28.67 (5.84–518.70)</td>
<td>31</td>
</tr>
<tr>
<td>Memphis, TN (1878)</td>
<td>72.5 (5,800)</td>
<td>8.6 (11,000)</td>
<td>7.44***</td>
<td>20.53 (18.87–22.34)</td>
<td>30</td>
</tr>
<tr>
<td>All epidemics</td>
<td>58.6 (9,049)</td>
<td>8.4 (13,325)</td>
<td>6.80***</td>
<td>14.63 (13.59–15.76)</td>
<td></td>
</tr>
</tbody>
</table>

* Data from references 15 and 26 to 31. The data for all six epidemics are lumped together and shown in boldface type.

* The case fatality rate in Caucasians \((C)\) to the case fatality rate in non-Caucasians \((nC)\) is shown. Race was found to be a significant predictor of outcome in the logistic regression model as follows: **, \( P < 0.01 \); ***, \( P < 0.001 \).
among Caucasians compared to individuals of African descent during 19th century YF epidemics in the United States.

Testing potential explanations for population-based differences in YF mortality. Biological, demographic, cultural, socioeconomic, or environmental factors may have contributed to the population-based differences in YF severity, so existing data were analyzed to assess whether each of these factors could have significantly affected these rates.

(i) It is unlikely that differences in viral genotypes explain differential severity within epidemics. As can be noted from Table 1, and consistent with what is known about dengue (5), severity can vary between epidemics, and this could be attributed in part to the differential virulence of viral genotypes. Since the addresses and workplaces of Caucasians and non-Caucasians within a city were often segregated (14), we considered the possibility that the differences in severity within epidemics were because the two races were being affected by two different strains of YFV. Given the unavailability of YFV isolates from these epidemics, we examined historical data to ascertain whether there was evidence for differences between outbreaks among Caucasians and non-Caucasians. We investigated the kinetics of the YF epidemic in Memphis, TN, in the two populations by charting deaths from YF over time and found similar kinetics (Fig. 1). The consistency in disease kinetics between populations was mirrored when examining cases diagnosed and deaths over time in the 1867 YF epidemic in New Orleans, LA, among U.S. Army troops (see Fig. S1A and S1B in the supplemental material). The time courses for YF-associated deaths were similar in both populations, which is most consistent with one YF genotype spreading in both populations. Moreover, both the Memphis 1878 and the U.S. Army 1867 data indicate that at all times during an epidemic, the case fatality rate for Caucasians greatly exceeded that for non-Caucasians.

(ii) Acquired immunity is unlikely to explain differences in yellow fever case fatality rates. Population-based differences in YF outcome have been attributed to differing rates of acquired immunity. Immigration from Africa early in the 19th century, immigration of Caucasians from regions where YF is not endemic later in the century, and different behavioral or cultural patterns between the two groups could contribute to differing levels of acquired immunity between the two populations (13). It is important to note, however, that individuals who had previously acquired immunity were not included in the calculation of case fatality rates, because multiple clinical infections of YF in one individual have not been documented in the historical literature (15, 18) and are considered exceedingly rare (http://www.cdc.gov/travel-training/local/HistoryEpidemiologyandVaccination/page24492.html). Therefore, for this important reason, and others discussed in detail in Text S3 in the supplemental material, we conclude that acquired immunity was not a determinant of differential YF case fatality between Caucasians and non-Caucasians. The confounding issue of cross-reacting partial immunity due to prior infection with dengue viruses cannot be completely ruled out. Nevertheless, we argue against a major role for cross-reacting immunity for two reasons. First, it is likely that incidence of dengue fever, which was also epidemic in the 19th century United States and was transmitted by the same mosquito vector, was similar for Caucasians and non-Caucasians, as we note for YF (Table S1). Therefore, immunity levels should have been comparable for the two populations. Second, given what is known about cross protection between different dengue serotypes, it is unlikely that these viruses could provide lasting protection against YFV.

(iii) Demographic factors are unlikely to explain differences in case fatality rates for YF. Both age and gender could affect susceptibility to YF, indeed males may be more likely to succumb to YF than females (12) and children may have milder YF cases than adults (17). We examined whether demographic differences between Caucasian and non-Caucasian populations could explain the difference in YF case fatality. Case fatality rates stratified by age and population were analyzed (Table 2; see Table S3 in the supplemental material). Even when controlling for age and gender (there were only males in the data set), the YF case fatality rates were at least 1.86-fold higher for Caucasians ($P < 0.01$ [Table 2]). In contrast, these population-based trends were not seen with cholera, as the differences in case fatality rates were the same or minimally lower in Caucasian than non-Caucasian soldiers (Table 2).

To examine whether distortions in the underlying population structure could have affected skews within or between populations, a case study on the 1878 YF epidemic in Memphis, TN, was performed. Although there were 13% more Caucasian males than females versus 18% more non-Caucasian females than males residing in Memphis in 1870 (see Fig. S2A and S2B and Table S4 in the supplemental material), these small distortions are not sufficient to explain a 7.44-fold difference in case fatality rate for the 1878 Memphis YF epidemic (Table 1). Indeed, the U.S. Army data presented above (Table 2 and Table S3) demonstrated that among age-matched men, YF case fatality was greater for Caucasians. Therefore, while age and gender may have played a small role in the differences in YF case fatality rate between the two populations, these factors could not account for the remarkable differences observed.

(iv) Socioeconomic, environmental, and cultural factors are unlikely to account for differences observed. Socioeconomic status could have had widespread ramifications for disease response, as it could have affected overall health, access to health care, and diet. The absence of any evidence for nutritional deprivation lead-


TABLE 2 Case fatality rates from yellow fever and cholera among U.S. Army troops in 1867

<table>
<thead>
<tr>
<th>Disease and age of soldier</th>
<th>No. of deaths/100 cases of disease in Caucasians (total no. of cases)</th>
<th>No. of deaths/100 cases of disease in non-Caucasians (total no. of cases)</th>
<th>Fold change in case fatality rates (C/nC)</th>
<th>Odds (95% confidence interval) of Caucasians versus non-Caucasians succumbing to disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yellow fever</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 20 yr</td>
<td>30.0 (237)</td>
<td>12.1 (33)</td>
<td>5.67***</td>
<td>15.97 (6.02–55.30)</td>
</tr>
<tr>
<td>20–29 yr</td>
<td>30.2 (770)</td>
<td>16.3 (123)</td>
<td>1.86**</td>
<td>2.23 (1.38–3.79)</td>
</tr>
<tr>
<td>Cholera</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 20 yr</td>
<td>22.2 (45)</td>
<td>50.0 (6)</td>
<td>0.44</td>
<td>0.29 (0.05–1.75)</td>
</tr>
<tr>
<td>20–29 yr</td>
<td>37.8 (185)</td>
<td>39.1 (87)</td>
<td>0.97</td>
<td>0.95 (0.56–1.61)</td>
</tr>
</tbody>
</table>

* Data from reference 15.

** The average age of soldiers less than 20 years old was not given in the primary source.

† The case fatality rate in Caucasians (C) to the case fatality rate in non-Caucasians (nC) is shown. Race was found to be a significant predictor of outcome in the logistic regression model as follows: **, P < 0.01; ***, P < 0.001.

The weakening of alternative hypotheses, which invoke non-genetic factors, led us to conclude that genetic factors explain the different YF responses in Caucasians and non-Caucasians. As a result, genetic differences, particularly those contributing to intrinsic and innate immunity, should be directly explored. This is fully consistent with proposals that populations whose ancestors lived in areas where YF is endemic would be more likely to have undergone selection for resistance that increase YF survival (10), similar to the malarial resistance afforded to those of African descent (1). While the research presented cannot formally confirm the possibility of host genetic differences significantly influencing YF severity, it is an important first step toward this goal. It should be noted that before allelic variation in *IL28B* was implicated in treatment-induced hepatitis C virus clearance, genetic variation was inferred based on strong differences between individuals of dissimilar ancestry (19). By analogy, we posit that our work strongly suggests the need for future experiments that compare genotypes of individuals with different YF outcomes.

**MATERIALS AND METHODS**

**Study design and definitions.** Reviews listing 19th century YF epidemics (11, 20, 21) were used to identify periods labeled at the time as “YF epidemics” by medical personnel or organizations (20) and to locate primary-source documents about these epidemics. The case numbers or numbers of deaths from YF by population were sought in reading these primary sources. If these data were available, then the epidemic was subjected to rigorous selection criteria, which are defined immediately below, to determine whether it would be included in this research.

To be included, an epidemic must have occurred during the 19th century and labeled as such by medical personnel at the time. We restricted the analysis to epidemics consisting of at least 50 documented cases of YF within a 5-month period (midsummer to fall of the same year) in the same geographical area. The case number cutoff was determined to mitigate the effects of misdiagnosis, and the time period reflected when the climate would most likely support the mosquitoes, which transmitted YF. When clinical descriptions from an epidemic were available, the signs and symptoms were compared to the clinical features of YF recognized today (22). While some features of YF are nonspecific, such as fever, headache, myalgia and weakness, lumbar sacral pain, anorexia, nausea, vomiting, reflecting the fact that non-Caucasians were deployed in battle less frequently than Caucasians (Fig. 2). We conclude from these comparisons that overall health or health care disparities are very unlikely to explain the skew toward higher case fatality rate for Caucasians. Furthermore, our analysis indicates that cultural factors did not significantly impact this skew (Text S2).

The weakening of alternative hypotheses, which invoke non-genetic factors, led us to conclude that genetic factors explain the different YF responses in Caucasians and non-Caucasians. As a result, genetic differences, particularly those contributing to intrinsic and innate immunity, should be directly explored. This is fully consistent with proposals that populations whose ancestors lived in areas where YF is endemic would be more likely to have undergone selection for resistance that increase YF survival (10), similar to the malarial resistance afforded to those of African descent (1). While the research presented cannot formally confirm the possibility of host genetic differences significantly influencing YF severity, it is an important first step toward this goal. It should be noted that before allelic variation in *IL28B* was implicated in treatment-induced hepatitis C virus clearance, genetic variation was inferred based on strong differences between individuals of dissimilar ancestry (19). By analogy, we posit that our work strongly suggests the need for future experiments that compare genotypes of individuals with different YF outcomes.

**FIG 2** The yellow fever case fatality rates among U.S. Union troops in 1864 demonstrate a skew toward Caucasians. The ratios of case fatality rates for Caucasians over the case fatality rate for non-Caucasians for 81 selected cases of death among U.S. Union troops in 1864 are shown. The total numbers of deaths from different causes follow: 223 from typhus fever, 64 from mumps, 1,445 from remittent fever, 1,922 from measles, 1,917 from typho-malarial fever, 223 from typhoid fever, 3,031 from smallpox and varioloid, and 336 from yellow fever.
ing, and a decreased pulse, the clinical signs of severe cases, jaundice, hematemesis (sometimes referred to as "black vomit" in the historical record), oliguria, bleeding from orifices, and delirium, when observed together and in the context of an epidemic, are considered pathognomonic for YF (22–24). Sixteen out of 333 19th century U.S. YF epidemics (20) met these criteria.

**Study population.** In each epidemic, an individual's race was used as a proxy for ancestry, Caucasian or non-Caucasian. Indeed, U.S. Census data reveal that over 97% of non-Caucasians living in the United States during the late 19th century were of African descent (25) (see supplemental material).

**Statistical analysis.** Data analysis was performed in R Studio using the contingency tables for the raw data (http://www.rstudio.com/). Fold differences between Caucasians and non-Caucasians were found by dividing the Caucasian case fatality rate by the non-Caucasian case fatality rate.

For each epidemic, a logistic regression of race on outcome was found in R Studio using Caucasians as the reference level. The odds ratios and the corresponding 95% confidence intervals were then found from the model. All the data sets for which an incidence rate, case mortality rate, or case fatality rate could be found were then combined, and three regression models were fit for the three types of rates. These models were analyzed in the same way as the individual epidemics.

**SUPPLEMENTAL MATERIAL**

Supplemental material for this article may be found at http://mbio.asm.orglookup/suppdoi:10.1128/mBio.01253-14-JDSupplemental.

**ACKNOWLEDGMENTS**

This work was supported by Public Health Service grants U54 AI057157 and RO1 AI089526 from NIAID. The funding sources had no role in the writing of the manuscript or in the decision to submit it.

We thank Margaret Humphries and Lee Baker for their perspectives about yellow fever and race in the United States. We are grateful to the members of M. A. Garcia-Blanco’s laboratory (particularly Shelton Bradf...


30. Keating JM, Howard Association. 1879. A history of the yellow fever. The yellow fever epidemic of 1878, in Memphis, Tenn., embracing a complete list of the dead, the names of the doctors and nurses employed, names of all who contributed money or means, and the names and history of the Howards, together with other data, and lists of the dead elsewhere. Howard Association of Memphis, Memphis, TN.