Could malaria explain the global distribution of the angiotensin converting enzyme I/D polymorphism? A systematic review and ecological study [version 1; peer review: 1 approved with reservations, 1 not approved]

Chris Kenyon, Anna Rosanas
Institute of Tropical Medicine, Antwerp, Antwerp, 2000, Belgium

Abstract
Background: The D-allele of the angiotensin converting enzyme (ACE1) has been linked to an increased risk of certain diseases including hypertension and COVID-19 but a decreased risk of cerebral malaria. We hypothesized that malaria played a role in determining variations in the global distribution of ACE1 I/D polymorphism.

Methods: A systematic review was conducted to summarize the frequency of ID/DD genotypes in all countries with available data.

Results: The ID/DD genotype frequency was found to be highest in Africa (86.4%, IQR 83.6-94.7%) and Eastern Mediterranean (median 84.5%, IQR 78.3-89.8%) and lowest in South East Asia (55%, 49.5-67.8%) and Western Pacific (61.1%, IQR 55.0-67.2%). Linear regression revealed positive associations between ID/DD genotype frequency and the incidence of malaria, malaria mortality as well as hemoglobin S allele frequency (all P<0.05).

Conclusions: Our findings are compatible with the hypothesis that malaria played a role in establishing the differential frequency of the D-allele.

Keywords
Malaria, ACE, polymorphism
Summary
The D-allele of the angiotensin converting enzyme (ACE1) has been linked to a number of diseases including hypertension, obesity and most recently COVID-19. Previous research has suggested that the prevalence of this allele may vary between world regions. Because the D-allele has been shown to be protective against cerebral malaria, we hypothesized that malaria played a role in determining variations in the global distribution of ACE1 I/D polymorphism. We conducted a systematic review of published literature to estimate the prevalence of the D-allele by country and World Health Organization region. The D-allele was found to be most prevalent in Africa and least prevalent in the Western Pacific. A positive association was found between the prevalence of the D-allele and three markers of malaria burden of disease. Our findings are thus compatible with the hypothesis that malaria played a role in establishing the differential frequency of the D-allele.

Introduction
The renin-angiotensin system plays an important role in the regulation of cardiovascular, renal and immune physiology. The angiotensin converting enzyme-1 (ACE1) converts angiotensin I to angiotensin II which is a key effector molecular in this system. The ACE1 gene insertion (I) and a 287 bp Alu repeat deletion (D) polymorphism explain a large proportion of the individual variation in ACE1 levels both in serum and tissue. ACE1 DD homozygotes and ID heterozygotes have approximately 65% and 30%, respectively, higher ACE1 levels in serum ACE1 than II homozygote individuals. It is unknown why the frequency of this polymorphism varies between different populations around the world. This increased ACE1 expression is thought to partly explain the association between the D-allele and an increased risk for diseases such as hypertension, heart failure, cerebrovascular disease, diabetic nephropathy, arthritis, various cancers, asthma, acute respiratory distress syndrome (ARDS) and possibly COVID-19 incidence and mortality. Increased ACE1 in D-allele carriers generates higher angiotensin II levels, whose stimulation of the AT2 receptor then activates vasoconstriction as well as a number of pro-inflammatory and fibrotic pathways which increase the risk of these conditions. Increased ACE1 in D-allele carriers generates higher angiotensin II levels, whose stimulation of the AT2 receptor then activates vasoconstriction as well as a number of pro-inflammatory and fibrotic pathways which increase the risk of these conditions. The D-allele has been found to be associated in case-control studies with protection against progression to severe malaria. In vitro and in vivo studies have demonstrated that angiotensin II can reduce the entry of plasmodia into erythrocytes and protect the blood-brain barrier against invasion by plasmodia. These findings have led a number of authors to propose that differential intensity in malaria exposure may have driven variations in the prevalence of the D-allele in a similar way to how malaria influenced the prevalence of haemoglobin S (HbS), and other polymorphisms associated with different susceptibility to malaria.

To do this we first conducted a systemic review of the frequency of the ID/DD genotypes in different countries around the world. A prior systematic review measured D allele frequencies by ethnicity, sex and age and reported D allele frequencies of 56.2% in “whites”, 60.3% in ‘Blacks’ (60.3%) and the lowest (39.1%) in ‘Asians’. Results were similar for when DD genotype was measured. This study did not evaluate the geographical variation of the I/D polymorphism.

Methods
Systematic review of ACE1 DD and ID/DD prevalence
In April 2020, we used PubMed and Google Scholar to conduct a systematic review with the objective of obtaining country-level ACE1 estimates of DD and ID+DD frequency. We used the following MeSH terms: ‘ACE’ OR ‘angiotensin-converting enzyme’ AND (‘polymorphism’ OR ‘deletion’ OR ‘insertion’). The following inclusion criteria were used: sample size of at least 50; published in English and the population could be considered representative of the general population (the studies were typically the control groups of case control studies evaluating the associations between I/D polymorphisms and specific diseases). Studies from previous systematic reviews of associations between D-allele frequency and various diseases could be included if they met the entry criteria. No date restrictions were used. Duplicate studies were removed (Figure 1). The data was extracted by a single author (CK), and manually checked for accuracy. Risk of bias was not assessed. The principle summary measure used was the prevalence of the genotype of interest expressed as a percentage. The PRISMA guidelines were followed for the manuscript (see Extended data, STable 1).

HbS: The HbS allele frequency (expressed as a percentage) was obtained from a study that estimated national HbS allele frequencies for 190 countries.

Malaria incidence: The national incidence of malaria per 1,000 population at risk for the year 2000 (the earliest year with available data), was obtained from the World Health Organization (WHO), Global Health Observatory Data Repository. A value of zero cases per 1,000 was used for countries classified as having eliminated malaria in 2000.

Malaria mortality. The age-standardized national number of deaths from malaria per 100,000 individuals. The data was obtained from the International Health Metrics and Evaluation Global Burden of Diseases study for the year 1990 – the first year with available data.

Geographical grouping. For the global analyses, countries were classified into 6 regions according to the WHO system.

A previous study limited to European countries found an ecological association between COVID-19 incidence and D-allele frequency, with both having higher values in Southern European countries. This provided the motivation to compare the ID/DD prevalence between European regions. We used the four European Regions as defined by the United Nations Geoscheme for Europe for this purpose.

Statistical analysis. Linear regression was used to evaluate the country-level association between the prevalence of ID/DD genotypes and the three measures of intensity of malaria...
exposure. Differences in ID/DD genotype prevalence between world regions were assessed via the Wilcoxon rank-sum test. The analyses were performed in STATA version 16 (Stata Corp, College Station, Tx). The maps were produced with Data Wrapper v2.

Results
Systematic review of ACE1 DD and ID/DD prevalence
Searches of the literature identified 5099 references. The application of the selection criteria restricted the total to 248 informative references published between 1993 and 2020, providing 253 prevalence estimates from 71 countries (Figure 1; Extended data, STable 2). Of these, 7 were located in the Americas, 15 in Eastern Mediterranean, 26 in Europe, 8 in sub Saharan Africa, 5 in South East Asia and 10 in Western Pacific (Table 2).

The ID/DD frequency was found to be highest in Africa (86.4%, IQR 83.6-94.7%) and Eastern Mediterranean (84.5%, IQR 78.3-89.8%) and lowest in South East Asia (55%, 49.5-67.8%) and Western Pacific (61.1%, IQR 55.0-67.2%; Figure 2 & Figure 5; Table 1). Within Europe, the prevalence was highest in Southern Europe (85.3%, IQR 84.3-88.2%) and lowest in Northern Europe (75.7%, IQR 73.9-76.7%; Table 2).

Linear regression revealed positive associations between ID/DD frequency and the incidence of malaria (coef. 4.96, 95% confidence interval [CI] 1.73-8.20; N=63; Table 3 and Figure 3), malaria mortality (coef. 0.69, 95% CI 0.20-1.19; N=70) as well as HbS allele frequency (coef. 0.12, 95% CI 0.04-0.19; N=68). Results were similar when these regressions were run with DD frequency as the outcome variable (Extended data, STable 4).

Discussion
Differential exposure to malaria has been found to play a role in driving the large geographical variations in the prevalence of various haemoglobinopathies, thalassemias, glucose 6-phosphate dehydrogenase deficiencies, erythrocyte membranopathies and various mediators of inflammation and immunity. We found country- and regional-level differences in the prevalence of the ACE1 D-allele that were associated with malaria incidence and mortality and HbS allele. These findings add into previous individual-level studies showing protection against severe malaria by the D allele, suggesting that malaria may have played a role in the global variations in D-allele frequency.

An important weakness of the study is our use of malaria incidence and mortality data from 1990 and 2000. Although

Figure 1. PRISMA flowchart.
Table 1. Prevalence of the ID/DD genotypes according to World Health Organization regions.

<table>
<thead>
<tr>
<th>Region</th>
<th>Number of studies(^b)</th>
<th>Median Prevalence, % (IQR)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Europe</td>
<td>26</td>
<td>79.2% (76.1-82.0)</td>
<td>Ref(^a)</td>
</tr>
<tr>
<td>Africa</td>
<td>8</td>
<td>86.4% (83.6-94.7%)</td>
<td>0.0005</td>
</tr>
<tr>
<td>Americas</td>
<td>7</td>
<td>74.0% (64.1-79.2%)</td>
<td>0.0798</td>
</tr>
<tr>
<td>Eastern Mediterranean</td>
<td>15</td>
<td>84.5% (78.3-89.8%)</td>
<td>0.1093</td>
</tr>
<tr>
<td>South East Asia</td>
<td>5</td>
<td>55% (49.5-67.8%)</td>
<td>&gt;0.0001</td>
</tr>
<tr>
<td>Western Pacific</td>
<td>10</td>
<td>61.1% (55.0-67.2%)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

\(^a\) Europe was used as the reference category as it had the largest sample size.

\(^b\) The number of studies used to generate the median prevalence estimates.

Figure 2. Boxplots of the prevalence of ACE1 ID/DD genotypes according to World Health Organization regions.

Table 2. Prevalence of the ID/DD genotypes in four European sub-regions.

<table>
<thead>
<tr>
<th>Sub-region</th>
<th>Number of studies(^b)</th>
<th>Median Prevalence, % (IQR)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern Europe</td>
<td>7</td>
<td>75.7% (73.9-76.7%)</td>
<td>Ref(^a)</td>
</tr>
<tr>
<td>Eastern Europe</td>
<td>6</td>
<td>77.6% (76.1-79.7%)</td>
<td>0.0478</td>
</tr>
<tr>
<td>Southern Europe</td>
<td>4</td>
<td>85.3% (84.3-88.2%)</td>
<td>0.0061</td>
</tr>
<tr>
<td>Western Europe</td>
<td>6</td>
<td>80.4% (78.6-81.0%)</td>
<td>0.0478</td>
</tr>
</tbody>
</table>

\(^a\) Northern Europe was used as the reference category as it had the largest sample size.

\(^b\) The number of studies used to generate the median prevalence estimates.
this was the earliest data we had available, it underestimates the historical burden of malaria for much of the world’s population up to the beginning/mid-20th century as shown in Figure 4. This figure of malaria prevalence from the preintervention period (around 1900) was produced by Lysenko in 1968. Unfortunately it has not been translated into estimated numerical values of national prevalence that could be used to test numerical associations. Although we did not test this statistically, it does appear that a number of the countries with high D-allele prevalence but low malaria incidence in 2000 in our analysis had a high malaria prevalence in 1900 (Figure 3 and Figure 4). For example, all 10 countries with ID/DD prevalence of above 81% and malaria incidence of 0 in 2000 were classified as being at least hypoendemic for malaria transmission in 1900 according to Lysenko. Our finding of higher prevalence of the D-allele in Southern than Northern Europe is also

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Coef.</th>
<th>95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malaria incidence</td>
<td>63</td>
<td>4.96</td>
<td>1.73-8.20</td>
<td>0.003</td>
</tr>
<tr>
<td>Malaria mortality</td>
<td>71</td>
<td>0.69</td>
<td>0.20-1.19</td>
<td>0.006</td>
</tr>
<tr>
<td>HbS</td>
<td>68</td>
<td>0.12</td>
<td>0.04-0.19</td>
<td>0.002</td>
</tr>
</tbody>
</table>

**Table 3.** Linear regression assessing the country-level association between the frequency of the ACE1 ID/DD genotypes and malaria incidence per 1,000, malaria mortality per 100,000 and the allele frequency of hemoglobin S (HbS).
Figure 4. The map global malaria endemicity originally produced by Lysenko et al. and reproduced by Hay et al. Endemicity as defined by Lysenko is according to the parasite rate in the 2–10-year age cohort (hypoendemic <0.1; mesoendemic 0.11–0.5; hyperendemic 0.51–0.75) except the holoendemic class (0.75) where the PR refers to the 1-year age group. The black line was added by Hay et al. to delineate the 2002 limit of malaria risk.

Figure 5. The country level prevalence of the ACE1 ID/DD genotypes (in percentages) according to the data produced by the current systematic review.
commensurate with the higher historical prevalence of malaria in this part of Europe\textsuperscript{26,27}. Another weakness of our study is the small number of countries with data for D-allele frequency and malaria burden of disease. We also acknowledge the fact that some of the differences in ID/DD frequency between studies and countries may be attributed to differences in how control groups were selected, differences in age distribution and in the methodology used to genotype individuals\textsuperscript{8}. It is also important to note that recent genome wide association studies have not found the ACE1 I/D polymorphism to be associated with severe malaria\textsuperscript{22,23}. The lack of an association in these studies may however be related to the high prevalence of the D-allele in malaria endemic areas, as well as the high level of genetic diversity and weak linkage disequilibrium in Africans compared to elsewhere\textsuperscript{24,25}. The association between malaria and the D-allele could be explained by confounding. It is possible that the actual driver of the D-allele is an environmental factor such as another infection that is associated with malaria. Finally, not all studies we drew our data from reported if the I- and D-alleles were in Hardy-Weinberg equilibrium. Including data from studies where these alleles were not in Hardy-Weinberg equilibrium would, however, be expected to introduce a non-differential misclassification bias into the analysis\textsuperscript{26}. This would be expected to dilute the strength of association between the prevalence of the D-allele and malaria severity. Given these concerns, we can only conclude that our data is compatible with the hypothesis that malaria played a role in establishing frequency distribution of the D-allele. This association could provide a potential evolutionary explanation for the commonly observed higher prevalence of hypertension in individuals with African genetic background\textsuperscript{22,26}. If the association between the D-allele and COVID-19 incidence/morbidity is confirmed, our mapping of D-allele prevalence could assist with estimating morbidity and mortality in different populations.

Data availability

Underlying data

All data underlying the results are available as part of the article and no additional source data are required.

Extended data

Figshare: Could malaria explain the global distribution of the angiotensin converting enzyme I/D polymorphism? A systematic review and ecological study. https://doi.org/10.6084/m9.figshare.12982112.v2\textsuperscript{17}.

This project contains the following extended data:

- STable 1. PRISMA checklist.
- STable 2. Study characteristics of studies included in systematic review, including frequency of II, ID and DD genotypes of ACE1, arranged according to country and year of study.
- STable 3. Study references for STable 2.
- STable 4. Linear regression assessing the country-level association between the prevalence of the DD genotypes of the ACE1 gene and malaria incidence per 1,000, malaria mortality per 100,000 and the allele frequency of hemoglobin S.

Extended data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Authors contributions

CK conceptualized the study, was responsible for the acquisition, analysis and interpretation of data including the systematic review and wrote the analysis up as a manuscript. AR assisted with the study design and analysis. Both authors contributed to and approved the final draft.

References


Alfred Amambua-Ngwa
Medical Research Council Unit, London School of Hygiene and Tropical Medicine, London, UK

The authors present a somewhat systematic review of publications to determine the relationship between ACE1 ID/DD alleles and malaria. The authors used malaria incidence rather than prevalence information as well as the distribution of Hbs and malaria mortality. The review has a strong rationale given the suspected role of ACE1 in COVID-19 and associations with high blood pressure in populations of African descent. However, there are several limitations with the data gathering and analysis process that does not render full credibility to the conclusions of malaria being a contributing factor to the prevalence of the ID/DD alleles:

- Data curation by a single author is not in line with standards for systematic reviews.

- The authors have not clearly stated the number of different publication types that were included, short communications, letters, posters, reviews, and conference abstracts. The method seemed limited by the number of databases or sources used for the search and the numbers in the flow chat need further clarification on how some publications were eliminated. For example, why were only 312 texts assessed for eligibility out of the 402 texts obtained after screening? Why were the final 248 texts included after excluding 90 out of the 312 assessed?

- Both first (Title/abstract) and second (full text) screening were done by one author which might have biased the screening strategy. Moreover, the author CK seems to have done everything in the manuscript, and this questions the contribution(s) of the second author.

- It is not clear how the authors assessed the criteria of “the population being a representative of the general population” in the selected 248 articles. This is important because allele frequencies depend on census sizes from the population. Yet the authors did not provide country-level information on the number of samples. This reflects a general flaw in the statistics as results are presented per sub-region. Country-level malaria incidence for example can be widely different within sub-regions and so country-level data would have been more informative. Given that the frequency of the D allele is almost fixed in most of Africa and equally high in malaria-free regions, the regression in my opinion...
cannot be robust. No explanation for example was given for the lower frequency in South East Asia where there is historically a higher prevalence of malaria than in Europe.

○ From figure 3 there is a higher prevalence of the D allele in higher malaria prevalence regions, but the relationship is far from linear. Hence the coefficients presented cannot be accurate.

○ The authors should consider adding a reference for the WHO system mentioned in the geographical grouping section of the methods.

○ In the first paragraph of the results, the authors are referring to table 2 instead of table 1 and in the second paragraph, they refer the reader to table 1 instead of table 2.

○ The author should consider rearrangement of figures for easy flow. Changing figure 5 to figure 3 will enable a better understanding of the second paragraph of the results section by the reader.

○ The authors might consider capturing the P-values in figure 2.

○ The limitation of drawing data from studies without a report that I and D alleles are in Hardy-Weinberg equilibrium can and should be corrected. Withdrawing data from such studies in a review on polymorphism can create big bias in the analysis and conclusions drawn. I herein suggest the authors go back to all included 248 papers, sort out those without reports that the I and D alleles are in Hardy-Weinberg equilibrium, and then do additional analysis on data from only studies with reports on Hardy-Weinberg equilibrium. This will enable the authors to present results of data drawn from studies with and without reports from Hardy-Weinberg equilibrium as well as results with data drawn from only studies with reports on I and D alleles being in Hardy-Weinberg equilibrium, hence reducing bias in the analysis and final conclusion(s).

Are the rationale for, and objectives of, the Systematic Review clearly stated?
Yes

Are sufficient details of the methods and analysis provided to allow replication by others?
No

Is the statistical analysis and its interpretation appropriate?
Partly

Are the conclusions drawn adequately supported by the results presented in the review?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: parasitology, genomics, genetics, epidemiology, transmission, drug resistance

I confirm that I have read this submission and believe that I have an appropriate level of
expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 24 November 2020

https://doi.org/10.5256/f1000research.29501.r74480

© 2020 Pei B. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Bin Pei
Evidence-Based Medicine Center and The Third Ward of Orthopedic, Xiangyang No. 1 People’s Hospital, Hubei University of Medicine, Xiangyang, China

In this study, the authors performed a systematic review to analyze the association between ACE1 I/D polymorphism and malaria. The topic is novel, but there are some serious defects that cannot be ignored.

1. This study included previous systematic reviews regarding ACE1 I/D, which is inconsistent with the scientificity of a systematic review. Besides, the search strategy and databases could not ensure the comprehensiveness of the researches included in this study, and the included 248 literatures were neither listed in this paper nor in the supplement. The data was extracted by only a single author, this is inconsistent with the basic requirement of data extracting method in systematic reviews. Moreover, the author did not evaluate the publication bias and method quality of the included studies, which may influence the reliability of the conclusion.

2. For the systematic review of gene polymorphism, Hardy-Weinberg test was not performed in this study. Whether the included researches meets Hardy-Weinberg test is important to reflect the representativeness of the population.

3. There are defects in statistical analysis. Linear regression was used to evaluate the country-level association between the prevalence of ID/DD genotypes and the three measures of intensity of malaria. However, the scatter plots (Figure 3) could not reflect the feasibility of linear regression, and the authors did not explain whether the data meet the assumption of linear regression. Besides, the prevalence of the ACE1 ID/DD genotypes should be obtained only from the healthy population, but the authors did not display the data resource in the figures or tables. Additionally, in Table 1 and Table 2, there is a methodological error to analyze the association between gene polymorphisms and malaria in other regions based on the region with the largest sample size, which may obtain fundamentally wrong results and conclusions.

4. The results of ID/DD genotype were analyzed in the study. Although it was mentioned in the last sentence of the results section that “Results were similar when these regressions were run with DD frequency as the outcome variable”, it could not be concluded that the D allele is associated with malaria. It is suggested that on the basis of overcoming the above
problems, the authors should separately study the correlation between I allele and malaria and the correlation between D allele and malaria.

5. The research on gene polymorphism is always divided by race rather than region.

**Are the rationale for, and objectives of, the Systematic Review clearly stated?**
Yes

**Are sufficient details of the methods and analysis provided to allow replication by others?**
Partly

**Is the statistical analysis and its interpretation appropriate?**
No

**Are the conclusions drawn adequately supported by the results presented in the review?**
Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Evidence-based medicine, clinical epidemiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.