Mangano, VD; Clark, TG; Auburn, S; Campino, S; Diakite, M; Fry, AE; Green, A; Richardson, A; Jallow, M; Sisay-Joof, F; Pinder, M; Griffiths, MJ; Newton, C; Peshu, N; Williams, TN; Marsh, K; Molyneux, ME; Taylor, TE; Modiano, D; Kwiatkowski, DP; Rocke-ett, KA (2009) Lack of association of interferon regulatory factor 1 with severe malaria in affected child-parental trio studies across three African populations. PLoS One, 4 (1). e4206. ISSN 1932-6203 DOI: 10.1371/journal.pone.0004206

Downloaded from: http://researchonline.lshtm.ac.uk/4455/

DOI: 10.1371/journal.pone.0004206

Usage Guidelines

Please refer to usage guidelines at http://researchonline.lshtm.ac.uk/policies.html or alternatively contact researchonline@lshtm.ac.uk.

Available under license: http://creativecommons.org/licenses/by/2.5/
Lack of Association of Interferon Regulatory Factor 1 with Severe Malaria in Affected Child-Parental Trio Studies across Three African Populations

Valentina D. Mangano1,2*, Taane G. Clark1,3, Sarah Auburn1,3,9, Susana Campino1,3,9, Mahamadou Diakite1,9, Andrew E. Fry1,3, Angela Green1, Anna Richardson1, Muminatou Jallow4, Fatou Sisay-Joof4, Margaret Pinder4, Michael J. Griffiths5, Charles Newton5, Norbert Peshu5, Thomas N. Williams5,6, Kevin Marsh5,6, Malcolm E. Molyneux7,8, Terrie E. Taylor7,9, David Modiano2, Dominic P. Kwiatkowski1,3, Kirk A. Rockett1

1 Childhood Infection Group, The Wellcome Trust Centre for Human Genetics, Oxford, United Kingdom, 2 Department of Public Health Sciences, Section of Parasitology, University of Rome La Sapienza, Rome, Italy, 3 Wellcome Trust Sanger Institute, Hinxton, United Kingdom, 4 Medical Research Council, Fajara, The Gambia, 5 Kenya Medical Research Institute Centre for Geographical Medicine Research (Coast), Kilifi, Kenya, 6 Nuffield Department of Medicine, John Radcliffe Hospital, Oxford, United Kingdom, 7 The Malawi–Liverpool–Wellcome Trust Programme of Clinical Tropical Research, Blantyre Malaria Project, College of Medicine, Blantyre, Malawi, 8 Liverpool School of Tropical Medicine, Liverpool, United Kingdom, 9 Department of Internal Medicine, College of Osteopathic Medicine, Michigan State University, East Lansing, Michigan, United Kingdom

Abstract

Interferon Regulatory Factor 1 (IRF-1) is a member of the IRF family of transcription factors, which have key and diverse roles in the gene-regulatory networks of the immune system. IRF-1 has been described as a critical mediator of IFN-gamma signalling and as the major player in driving TH1 type responses. It is therefore likely to be crucial in both innate and adaptive responses against intracellular pathogens such as Plasmodium falciparum. Polymorphisms at the human IRF1 locus have been previously found to be associated with the ability to control P. falciparum infection in populations naturally exposed to malaria. In order to test whether genetic variation at the IRF1 locus also affects the risk of developing severe malaria, we performed a family-based test of association for 18 Single Nucleotide Polymorphisms (SNPs) across the gene in three African populations, using genotype data from 961 trios consisting of one affected child and his/her two parents (555 from The Gambia, 204 from Kenya and 202 from Malawi). No significant association with severe malaria or severe malaria subphenotypes (cerebral malaria and severe malaria anaemia) was observed for any of the SNPs/haplotypes tested in any of the study populations. Our results offer no evidence that the molecular pathways regulated by the transcription factor IRF-1 are involved in the immune-based pathogenesis of severe malaria.

Introduction

Severe malaria is a life-threatening disease that kills over a million individuals every year, with 90% of the deaths occurring in Sub-Saharan African children below the age of five [1]. People living in an endemic area often become infected with Plasmodium falciparum malaria during childhood, but a small proportion of children experience severe complications, the clinical outcome depending on many factors, including the genetic make-up of the human host. Malaria is recognised as a very strong selective force in the recent evolutionary history of the human genome [2] and therefore is a candidate of great interest for genetic association analysis. Association studies can inform us on mechanisms of protective immunity against malaria at every stage of infection, the understanding of which is crucial for the development of effective vaccines.

Members of the Interferon Regulatory Factor (IRF) family of transcription factors have key and diverse roles in the gene-regulatory networks of the immune system [3,4,5]. In particular IRF-1 has been described as a critical mediator of IFN-gamma signalling and as the major player in driving TH1 type responses [3]. It is therefore likely to be crucial in both innate and adaptive responses against intracellular pathogens such as P. falciparum.

In two recent complementary association studies conducted in rural communities and malaria patients of Burkina Faso, we found that IRF1 polymorphisms were associated with the ability to
control *P. falciparum* infection: in adult asymptomatic individuals living in rural villages *IRF-1* alleles were associated with a higher prevalence of *P. falciparum* infection; the same alleles were associated with a higher parasite density in children with uncomplicated and severe malaria [6]. However, our work did not provide conclusive evidence for a role of this locus contributing to or providing protection from the severe manifestations of malaria disease [6], as may be found for a well known genetic factor such as haemoglobin S (HbS) [7].

In order to investigate whether genetic variation at the *IRF1* locus contributes to a child's risk of developing severe malaria, we conducted a large multi-centre association study. As human populations in Africa show a great genetic diversity [8] population structure must be taken into account when conducting population based genetic association studies since it might lead to false positive results [9]. To avoid this drawback we applied the complementary Transmission Disequilibrium Test (TDT) approach [10]. Genotype data were collected from family trios consisting of one affected child and his/her two parents, and the genotype distribution observed in the severe malaria cases was compared to its expected distribution derived on Mendel's law of segregation. A comparison of such family-based association methods versus case-control approaches has been carried out using the well-established association of the sickle cell trait with protection against severe malaria. The two methods were found to have similar power and to give similar estimates of the level of protection [7]. Epidemiological studies have suggested that features of severe malaria, with regards to both age-incidence profiles and clinical spectrum, vary with levels of *P. falciparum* transmission intensity [11]. We therefore conducted our association analysis in trios from three sub-Saharan African countries covering a range of malaria ecologies, The Gambia, Kenya and Malawi.

Here we assess the role of *IRF1* genetic variation in severe malaria, and by considering multiple populations, we able to investigate regional differences, whilst standardizing phenotype definition and study design.

**Results**

We have genotyped 18 Single Nucleotide Polymorphisms (SNPs) in the *IRF1* genetic region (Figure 1) in 961 nuclear trios comprising one child affected by severe malaria and his/her two parents. Severe malaria patient samples were collected as part of ongoing epidemiological studies at the Royal Victoria Hospital, Banjul, The Gambia (555 trios); the Queen Elizabeth Central Hospital, Blantyre, Malawi (202 trios); and Kilifi District Hospital, Kilifi, Kenya (204 trios). The family trios were assessed for pedigree mis specification using a panel of 48 additional markers spread across the genome and the Nucl3ar software package [12].

The allele frequencies calculated among the trio founders (Table 1) differ for many markers between The Gambia and either Kenya or Malawi [significant difference for 14 and 12 SNPs respectively], while they show a greater similarity between the two East-African countries [significant difference for 3 SNPs only], as we could expect on the basis of the geographical distribution and history of the three populations.

The power of a TDT analysis to detect transmission distortion is a function of the number of heterozygous parents in the sample and the level of distortion at the locus [13]. Trios in which at least one parent is heterozygous are informative in TDT analysis, and therefore the heterozygosity of a locus can be used as a measure of its information content in family-based association analysis. The observed heterozygosity (H) in the study populations is relatively high: The Gambia, HG = 40.8%; Kenya, HK = 38.7%; Malawi, HM = 38.0%. For comparison, in the HapMap populations Yoruba (Ibadan, Nigeria) and CEPH (Utah residents with ancestry from northern and western Europe) the heterozygosity of the *IRF1* gene is the highest among the Th2 cluster (HY = 38% and HG = 45%) while the heterozygosity of the neighbouring *IL5* gene

![Figure 1. The *IRF1* locus and Single Nucleotide Polymorphisms (SNPs) tested for association with severe malaria in child-parental trios.](image-url)
is the lowest among the Th2 cluster (H1 = 25% and H2 = 21%) (HapMap data release 3.1).

We conducted single marker and haplotype TDT association analysis using a custom script implemented in R statistical software and based on Transmit [14], and the equivalent FBAT approach [15] for comparison of the results. Considering the differences in allele frequencies observed, along with possible differences in clinical spectrum, we conducted the association analysis separately within each LD block for association with severe malaria.

The application of the single locus TDT to the 18 IRF1 SNPs revealed no evidence of association (all P-values > 0.05) across all populations, neither with severe malaria or with the sub-phenotypes of cerebral malaria (CM) and severe malaria anaemia (SA) under any of the models tested (additive, dominant, recessive and heterozygous advantage). The results for the test of association with severe malaria under the additive model are shown in Table 2. There was no evidence of heterogeneity or differences in the allelic odds ratios between the CM and SA phenotypes for each SNP (all P-values > 0.05), supporting the interpretation above that there were no strong SNP effects being masked by sub-clinical phenotype.

In comparison, the HbS allele shows, as expected, a very strong protective effect against severe malaria, with a reduction in risk of 81% and 87% in the Gambian (P = 6.6 × 10−19) and Kenyan (P = 9.1 × 10−18) populations, respectively. The effect of the HbS allele could not be tested in the trios from Malawi, due to its low allele frequency in this population and therefore the number of informative trios is not sufficient to compute the statistic.

Linkage Disequilibrium (LD) analysis was performed using the HAPLOVIEW application [16] and revealed a similar haplotype architecture in the three populations (Figure 2). Two blocks of high LD have been identified: the first corresponding to the downstream and coding region (SNPs rs7719499 to rs2070721; average r2 = 0.70) and the second to the proximal upstream region (<5 kb from the first exon, SNPs rs2706384 to rs2549002; average r2 = 0.63). The LD was observed in the intergenic upstream region (>5 kb from the first exon, SNPs rs7719499 to rs2549002; average r2 = 0.56)–while very low levels of LD were observed in the second exon (P = 9.1 × 10−18) and Kenyan (P = 9.1 × 10−18) populations, respectively. The results are consistent with previous findings in two other West-African populations from Burkina Faso [6].

We tested the common haplotypes (frequency higher than 5%) within each LD block for association with severe malaria phenotypes. Similarly to the single locus TDT, no significant association was observed (all P-values > 0.05) in any of the study populations. Table 3 shows the results for the test of association with severe malaria.

### Discussion

Rodent malaria studies have shown that IFN-gamma is a key inducer of the immune effector mechanisms that are responsible for the control of both pre-erythrocytic and blood-stage malaria
Table 2. Single marker Transmission Disequilibrium Test (TDT) of association with severe malaria in Gambian, Kenyan and Malawian child-parental trios.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Inform The Gambia</th>
<th>Inform Kenya</th>
<th>Inform Malawi</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR    LCL LCL</td>
<td>OR    LCL LCL</td>
<td>OR    LCL LCL</td>
</tr>
<tr>
<td>HbS (rs334)</td>
<td>170 0.19 0.13 0.28</td>
<td>26 0.13 0.04 0.43</td>
<td>- - - -</td>
</tr>
<tr>
<td>rs7719499</td>
<td>513 0.98 0.82 1.17</td>
<td>169 0.99 0.75 1.34</td>
<td>195 0.89 0.67 1.18</td>
</tr>
<tr>
<td>rs10065633</td>
<td>499 0.96 0.80 1.14</td>
<td>168 1.00 0.74 1.35</td>
<td>100 0.91 0.68 1.21</td>
</tr>
<tr>
<td>rs2070729</td>
<td>496 1.00 0.84 1.19</td>
<td>100 1.00 0.73 1.37</td>
<td>131 0.77 0.55 1.09</td>
</tr>
<tr>
<td>rs2070727</td>
<td>504 0.93 0.78 1.11</td>
<td>169 0.97 0.71 1.30</td>
<td>194 0.94 0.71 1.25</td>
</tr>
<tr>
<td>rs2070726</td>
<td>494 0.96 0.81 1.15</td>
<td>163 1.04 0.76 1.41</td>
<td>178 0.98 0.73 1.31</td>
</tr>
<tr>
<td>rs2070724</td>
<td>460 0.97 0.81 1.17</td>
<td>167 1.06 0.78 1.44</td>
<td>193 1.05 0.79 1.40</td>
</tr>
<tr>
<td>rs10213701</td>
<td>452 0.94 0.78 1.13</td>
<td>172 1.02 0.76 1.38</td>
<td>156 0.88 0.64 1.20</td>
</tr>
<tr>
<td>rs2070722</td>
<td>493 0.95 0.80 1.13</td>
<td>161 1.01 0.74 1.38</td>
<td>183 0.89 0.66 1.19</td>
</tr>
<tr>
<td>rs2070721</td>
<td>468 1.00 0.83 1.20</td>
<td>103 1.24 0.84 1.83</td>
<td>111 0.79 0.54 1.15</td>
</tr>
<tr>
<td>rs2706384</td>
<td>489 0.97 0.81 1.16</td>
<td>131 0.87 0.62 1.23</td>
<td>178 0.89 0.67 1.20</td>
</tr>
<tr>
<td>rs2549005</td>
<td>484 1.03 0.86 1.23</td>
<td>156 0.93 0.68 1.27</td>
<td>150 0.97 0.71 1.34</td>
</tr>
<tr>
<td>rs41525648</td>
<td>412 0.98 0.81 1.19</td>
<td>138 1.00 0.72 1.40</td>
<td>114 0.84 0.58 1.21</td>
</tr>
<tr>
<td>rs2549004</td>
<td>490 0.98 0.82 1.16</td>
<td>166 0.93 0.69 1.26</td>
<td>164 0.97 0.73 1.29</td>
</tr>
<tr>
<td>rs2549002</td>
<td>483 0.96 0.80 1.14</td>
<td>166 0.95 0.70 1.29</td>
<td>180 0.96 0.71 1.28</td>
</tr>
<tr>
<td>rs736801</td>
<td>85 0.73 0.48 1.13</td>
<td>9 0.80 0.21 2.98</td>
<td>4 0.33 0.03 3.20</td>
</tr>
<tr>
<td>rs4705952</td>
<td>462 1.00 0.83 1.20</td>
<td>191 1.20 0.90 1.59</td>
<td>177 1.08 0.81 1.45</td>
</tr>
<tr>
<td>rs2706390</td>
<td>313 0.97 0.78 1.21</td>
<td>83 0.93 0.60 1.43</td>
<td>63 1.42 0.86 2.35</td>
</tr>
<tr>
<td>rs2706391</td>
<td>95 0.73 0.48 1.09</td>
<td>26 1.00 0.46 2.16</td>
<td>28 1.55 0.72 3.30</td>
</tr>
</tbody>
</table>

The table shows the SNPs tested in association analysis with severe malaria and results of the TDT in the three populations of affected child-parental trios. Inform is the number of informative trios. OR: Odds Ratio comparing the risk of the minor vs major allele. LCL and UCL: Lower and Upper Confidence Limits of a 95% Confidence Interval respectively. TDT P: P-value for the family-based test of association. P-values < 0.05 are shown in bold.

doi:10.1371/journal.pone.0004206.t002

Figure 2. Linkage Disequilibrium (LD) architecture of the IRF1 locus in The Gambia, Kenya and Malawi. Typed single nucleotide polymorphisms (SNPs) are represented on the vertical axis and ordered by chromosome position as in Table 1. Diagonal lines in two directions link each SNP to each of the other 17 SNPs. The LD (\(D^2\)) between each pair of markers is represented by diamonds of different colours (see legend). \(D^2\) is a simple measure of correlation between markers and is calculated as follows: \((f_{AB} - f_A f_B)^2/(f_A f_B)\), where \(f\) stands for frequency, \(A\) and \(a\) denote the alleles at the first marker, \(B\) and \(b\) denote the alleles at the second marker, \(A\) and \(b\) is the haplotype carrying the major allele for the first marker and the minor allele for the second marker and so on. Blocks 1 and 2 are set of markers in high LD identified within the gene using the HAPLOVIEW application [16].

doi:10.1371/journal.pone.0004206.g002
infection [reviewed in 17]. Consistent with the mouse model, studies conducted in human populations have shown that malaria induced IFN-gamma production is associated with lower infection rates and resistance to re-infection [18] and with reduced risk of clinical malaria [19]. For these reasons IFN-gamma is considered as a correlate of cellular immunity to malaria and thus used as a marker for vaccines immunogenicity and efficacy [17]. On the other hand, both in mice and humans, an excessive inflammatory response with high systemic levels of IFN-gamma, TNF-alpha, IL-6 and IL-1, has long been correlated with immune-based pathogenesis of severe disease [reviewed in 20]. The tight regulation of IFN-gamma production and signalling appears therefore of great importance in determining susceptibility to malaria.

We have previously investigated the role of the transcription factor IRF-1, a critical effector molecule in IFN-gamma signalling, in susceptibility to malaria by conducting genetic association studies in Burkina Faso based on cross-sectional parasitological surveys and case-control clinical studies. We have shown that IRF1 polymorphisms are associated with the control of P. falciparum infection (parasite prevalence and density), both in healthy adult subjects and in children with uncomplicated and severe malaria. In those studies we did not find conclusive evidence for the involvement of this locus in the severe manifestations of malaria disease. Indeed, only a trend of association with small differential risk of severe disease was suggested [6].

In the present study we investigated whether genetic variation at the IRF1 locus affects susceptibility to severe malaria by performing a family based association test in nuclear trios consisting of an affected child and his/her two parents from The Gambia, Kenya and Malawi. The trio design is robust to population stratification. We did not observe any significant associations between the SNPs or haplotypes analysed in severe malaria in any of the three study populations.

There are various possible interpretations of these findings. The effects of IRF1 polymorphisms on severe malaria may be too small to be detected with this sample size, or may affect only a sub-phenotype of severe malaria. It is likely that the genetic basis of susceptibility to malaria is built upon many different protective genes, each individually resulting in small effects [21]. This study lacks power to detect a small effect [13]. For example, consider a situation where we have 555 trios and wish to detect (assuming a false positive rate of 5%) a multiplicative genotypic effect at a SNP with minor allele frequency of 0.2. If the magnitude of the odds ratio to detect is 1.5 (50% increased risk) we would have >90% power. However, if we wish to detect an odds ratio of only 1.2 (20% increased risk), we would have only 42% power, and would require ~1400 trios to achieve 80% power. These estimates also assume that the SNP genotyped is in perfect linkage disequilibrium (A^2=1) with the causal polymorphism. Overall, we find no evidence that IRF1 is a strong determinant of susceptibility to severe disease in our study populations.

When considering the exceptionally high diversity of the human genome in African populations [8] we must also take into account that different malaria-resistant alleles may have arisen in different populations due to differences in selective pressure and demographic history. This is the case, for example, of HbC, whose prevalence varies greatly between neighbouring countries and even villages [22,23]. It is therefore possible that IRF1 functional variants with an effect on malaria susceptibility/resistance exist in the Fula and Mossi populations of Burkina Faso but not in the study populations presented here. Moreover, we cannot rule out that a different pattern of association could be observed in other African populations as a consequence of a different genetic background. Further disease-association studies across Africa and including Burkina Faso will be needed to answer this question.

Finally, it might be possible that the association of IRF1 polymorphisms with the ability to control P. falciparum infection and the lack of association with susceptibility to severe disease actually reflect differences in the molecular mechanisms underlying protective/pathological immune responses at different stages of infection. The gene regulatory networks of the immune system that are activated during malaria infection and that are responsible for parasite killing may differ from those leading to an over-reaction to the parasite that is harmful to the host, depending for example on the molecular triggers from the parasite or on the tissues involved. That this could be the case has been recently suggested by a genome-wide linkage analysis of parasite density and mild malaria in two Senegalese villages. Regions of linkage showed little if any overlap with genes previously described to be

---

**Table 3.** Haplotype Transmission Disequilibrium Test (TDT) of association with severe malaria in Gambian, Kenyan and Malawian child-parental trios.

<table>
<thead>
<tr>
<th>Haplotypes</th>
<th>The Gambia Freq</th>
<th>Inform</th>
<th>OR LCL UCL TDT P</th>
<th>Kenya Freq</th>
<th>Inform</th>
<th>OR LCL UCL TDT P</th>
<th>Malawi Freq</th>
<th>Inform</th>
<th>OR LCL UCL TDT P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Block 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GTCCCGATA</td>
<td>0.38</td>
<td>470</td>
<td>0.96</td>
<td>0.82</td>
<td>1.18</td>
<td>0.64</td>
<td>0.27</td>
<td>151.60</td>
<td>0.97</td>
</tr>
<tr>
<td>CCAAAATGC</td>
<td>0.36</td>
<td>438</td>
<td>1.09</td>
<td>0.86</td>
<td>1.25</td>
<td>0.35</td>
<td>0.45</td>
<td>186.80</td>
<td>0.98</td>
</tr>
<tr>
<td>CCAAAAGGC</td>
<td>0.14</td>
<td>213</td>
<td>0.84</td>
<td>0.71</td>
<td>1.21</td>
<td>0.20</td>
<td>0.16</td>
<td>110.80</td>
<td>1.06</td>
</tr>
<tr>
<td>GTACCGATC</td>
<td>0.11</td>
<td>208</td>
<td>1.06</td>
<td>0.78</td>
<td>1.35</td>
<td>0.67</td>
<td>0.11</td>
<td>79.70</td>
<td>1.03</td>
</tr>
<tr>
<td><strong>Block 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CATGG</td>
<td>0.51</td>
<td>498</td>
<td>1.00</td>
<td>0.84</td>
<td>1.19</td>
<td>0.96</td>
<td>0.52</td>
<td>183.60</td>
<td>0.90</td>
</tr>
<tr>
<td>AGGCT</td>
<td>0.28</td>
<td>409</td>
<td>1.03</td>
<td>0.84</td>
<td>1.23</td>
<td>0.74</td>
<td>0.24</td>
<td>136.90</td>
<td>1.01</td>
</tr>
<tr>
<td>AGTCT</td>
<td>0.10</td>
<td>185</td>
<td>1.09</td>
<td>0.78</td>
<td>1.38</td>
<td>0.58</td>
<td>0.09</td>
<td>68.80</td>
<td>1.28</td>
</tr>
<tr>
<td>CGTGG</td>
<td>0.09</td>
<td>147</td>
<td>0.91</td>
<td>0.69</td>
<td>1.33</td>
<td>0.57</td>
<td>0.11</td>
<td>83.00</td>
<td>1.08</td>
</tr>
</tbody>
</table>

The table shows the common haplotypes (frequency higher than 5%) within each LD block identified in the gene, their frequency and the results of family-based test of association with severe malaria in the three populations of affected child-parental trios. Freq: frequency of the haplotype in trio parents. These are not unbiased estimates of the population frequencies. OR: Odds Ratio comparing the risk of the untransmitted versus the transmitted haplotype. All other abbreviations as in Table 2.

doi:10.1371/journal.pone.0004206.t003
associated with severe malaria, while showing an overlap with genes involved in asthma and atopy related traits [24]. Further insights could be gained by conducting fine mapping studies in multiple populations of larger sample size, where both phenotypes, parasite load and disease severity, could be inspected against the same genetic background. However, the data so far available offer no evidence that the molecular pathways regulated by the transcription factor IRF-1 are involved in immune-based pathogenesis of severe malaria.

**Methods**

**Child-parental trios**

Severe malaria patient samples were collected as part of ongoing epidemiological studies at the Royal Victoria Hospital, Banjul, The Gambia (555 trios); the Kilifi District Hospital, Kilifi, Kenya (204 trios); and the Queen Elizabeth Central Hospital, Blantyre, Malawi (202 trios). Nuclear family trios comprised one affected child and his two parents.

The research protocols used in this study have obtained approval from the relevant ethics committees: the United Kingdom approval was given by the Oxfordshire Clinical Research Ethics Committee (OxREC); the Gambian approval was given by the Gambia Government/Medical Research Council (MRC) Joint Ethics Committee; the Kenyan approval was given by Kenya Medical Research Institute (KEMRI)/National Ethical Review Committee; the Malawi approval was given by the Health Sciences Research Committee and College of Medicine Research Committee (COMRC).

All DNA samples were collected and genotyped following informed consent from participants. Some of the samples for this study were collected with written consent, others with verbal consent. In all cases the advice of the local ethics committee about the appropriate ways of recording consent in the local research context was followed.

The family trios were assessed for pedigree mis specification using a panel of 48 additional markers spread across the genome and the Nuclear software package [12].

**Severe malaria cases and phenotypes definition**

All cases were children admitted to hospital with evidence of *P. falciparum* on blood film and clinical features of severe malaria [25,26]. We used a Blantyre coma score of ≤2 as a criterion of cerebral malaria (CM), and haemoglobin <5 g/dl or packed cell volume <15% as a criterion of severe malaria anaemia (SA). Some individuals had both CM and SA. Of the severe malaria cases that were not CM or SA by these criteria, most had lesser degrees of coma (Blantyre coma score 3) or anaemia (Hb 5–6 g/dl), or other complications such as respiratory distress.

The sub-clinical phenotype frequencies in the probands were: (i) The Gambia (443 CM, 239 SA, 141 both CM/SA, 14 other); (ii) Kenya (109 CM, 77 SA, 31 CM/SA, 49 other) and (iii) Malawi (195 CM, 52 SA, 49 CM/SA, 4 other).

The median (range) ages of the Gambian, Kenyan and Malawian cases expressed in months were 41 (0-156), 26 (0-121) and 31.5 (2-156), respectively.

**Markers selection, DNA samples preparation and genotyping**

Single Nucleotide Polymorphisms (SNPs) in the region 10 kb upstream to 10 kb downstream of the *IRF1* gene were identified from public databases (dbSNP build 126, Ensemble release 40) and the literature. 18 polymorphisms were selected on the basis of the following criteria: validated status by frequency, haplotype tagging SNP (htSNP) in the HapMap Yoruba population, overall coverage of the locus with regular spread (density of 1 SNP every 2 Kb on average), encompassing exonic, intronic, regulatory regions of the gene. Database identifiers and location with respect to the *IRF1* gene of the 18 SNPs are displayed in Figure 1 and listed in Table 1.

Genomic DNA samples underwent whole genome amplification through either Primer Extension Pre-amplification (PEP) [27] or Multiple Displacement Amplification (MDA) [28] before genotyping.

The genotypes for the selected polymorphisms were determined through the SEQUENOM® MassARRAY™ System [29]. One assay (rs839) did not pass quality filters following initial testing (call rate<90%) and evidence that the genotypes in the founders deviated from Hardy-Weinberg equilibrium [30] (P<0.001). This SNP has therefore been excluded from any further analysis.

**Haplotypes and Linkage Disequilibrium architecture**

Haplotypes have been constructed using family data and the Stephens-Donnelly method [31]. Haplotypes blocks across the genes have been defined using the HAPLOVIEW application [16], based on the method originally described by Gabriel and colleagues [32].

LD metrics and graphs showing LD patterns have been generated using the MARKER [http://www.gmap.net/perl/marker/marker_entry] application. Typed single nucleotide polymorphisms (SNPs) are represented on the vertical axis and ordered by chromosome position as in Figure 1. The linkage disequilibrium between each pair of markers is represented by diamonds of different colours.

**Statistical methods**

Allele frequencies have been calculated among pedigree founders and Hardy-Weinberg equilibrium has been tested, using the dedicated software PEDSTATS [33]. These are not unbiased estimates of the population frequencies. Allele distributions between populations have been compared by Yates-corrected $\chi^2$ test. If contingency tables had less than 5 expected events per cell, Fisher’s exact test was used.

Trio association analysis was performed using the transmission disequilibrium test (TDT). This tests for family-based association in the presence of genetic linkage between a SNP and severe malaria [10], and is robust to the presence of population structure. Single locus and haplotype versions of the TDT were implemented using R statistical software based on Transmit [14]. The FBAT application [15] (software available at http://biosun1.harvard.edu/~flat/hbat.htm) has been used for comparison of the results.

Haemoglobin genotyping was previously performed on all samples included in the study [34] and association of HbS with severe malaria has been used as positive control for the TDT results.

A P-value of 0.05 has been considered as the threshold for statistical significance.

**Acknowledgments**

We are gratefully indebted to the children and their parents for their participation in this study. This manuscript is published with the permission of the director of the Kenya Medical Research Institute.

**Author Contributions**

Conceived and designed the experiments: VDM DM DPK KR. Performed the experiments: VDM SA SC MD AEF AG AR. Analyzed the data: VDM DM DPK KR. Contributed reagents/materials/analysis tools: DPK KR. Wrote the paper: VDM TGC DM DPK KR. Conducted the clinical studies: MJ FSJ MP MG CRJCN NP TW KM MEM TET.
References