

# Tracing evolutionary relicts of positive selection on eight malaria-related immune genes in mammals

Bing-Hong Huang and Pei-Chun Liao

Innate Immunity  
2015, Vol. 21(5) 463–476  
© The Author(s) 2014  
Reprints and permissions:  
sagepub.co.uk/journalsPermissions.nav  
DOI: 10.1177/1753425914547744  
ini.sagepub.com  


## Abstract

*Plasmodium*-induced malaria widely infects primates and other mammals. Multiple past studies have revealed that positive selection could be the main evolutionary force triggering the genetic diversity of anti-malaria resistance-associated genes in human or primates. However, researchers focused most of their attention on the infra-generic and intra-specific genome evolution rather than analyzing the complete evolutionary history of mammals. Here we extend previous research by testing the evolutionary link of natural selection on eight candidate genes associated with malaria resistance in mammals. Three of the eight genes were detected to be affected by recombination, including *TNF- $\alpha$* , *iNOS* and *DARC*. Positive selection was detected in the rest five immunogenes multiple times in different ancestral lineages of extant species throughout the mammalian evolution. Signals of positive selection were exposed in four malaria-related immunogenes in primates: *CCL2*, *IL-10*, *HO1* and *CD36*. However, selection signals of *G6PD* have only been detected in non-primate eutherians. Significantly higher evolutionary rates and more radical amino acid replacement were also detected in primate *CD36*, suggesting its functional divergence from other eutherians. Prevalent positive selection throughout the evolutionary trajectory of mammalian malaria-related genes supports the arms race evolutionary hypothesis of host genetic response of mammalian immunogenes to infectious pathogens.

## Keywords

Eutherian, evolutionary arms race, functional divergence, immunogenes, malaria, *Plasmodium*, positive selection, recombination

Date received: 20 April 2014; revised: 28 May 2014; 27 June 2014; 17 July 2014; accepted: 24 July 2014

## Introduction

*Plasmodium*-induced malaria is one of the most prevalent infectious diseases in the human population. In addition to humans, *Plasmodium* species also infect primates and non-primates;<sup>1–5</sup> primates were reported as an adapted host of *Plasmodium*.<sup>1–3</sup> Bats may have been the first mammal host of malaria syndrome.<sup>6</sup> Two lineages of *Plasmodium* genus using bats as the reservoir host are phylogenetically close to the *Plasmodium* parasites from rodents thereby generating rodent parasites as valuable experimental, medical and genetic models of human-form malaria.<sup>6,7</sup> This implies the low threshold host shift of *Plasmodium* among mammals and divergent malaria disease susceptibility by genetic components of the mammalian immune system.<sup>8–11</sup> In other words, different *Plasmodium* infections could result in different selection pressures for the evolution of the mammalian immune system.

Effective spread of adaptive alleles of a host's immunogenes is the major strategy of disease resistance

against pathogen infection.<sup>12,13</sup> Several evolutionary models were developed for evaluating the immunogenes.<sup>14–19</sup> Most of these evolutionary hypotheses for the host–parasite co-evolution involve positive selection as a main component for explaining the broad responses of hosts to non-specific pathogens (e.g. arms races hypothesis), in addition to the balancing selection model (trench warfare hypothesis).<sup>20</sup> Mechanisms for malaria resistance and the evolution model are well studied in humans.<sup>7,21</sup> However, the evolutionary genetics of the malaria-related immunogenes in mammals is still unclear. Understanding

Department of Life Science, National Taiwan Normal University, Taipei, Taiwan, Republic of China

### Corresponding author:

Pei-Chun Liao, Department of Life Science, National Taiwan Normal University, No.88, Sec. 4, Ting-Chou Road, Wunshan District, Taipei 116, Taiwan, Republic of China.  
Email: pcliao@ntnu.edu.tw

immunogene evolution is important to understanding this infectious epidemic.

Evidence of positive selection has been found in mammalian innate immunogenes. For example, the radical amino acid changes in certain structurally important domains of the TLRs that recognize the infectious agents to initiate the innate immune response were found,<sup>22,23</sup> and signatures of positive selection across primates and across mammals were detected.<sup>24,25</sup> The adaptive evolution of *TLRs* might be more episodic in nature rather than a long-term constant phenomenon along primate evolution.<sup>24</sup> Another innate immunity gene that codes for the viral-sensing RIG-I-like receptors (RLRs) also represent apparent signals of positive selection along the evolutionary history of mammals.<sup>26–28</sup> Two *RLR* genes, *RIG-I* and *LGP2*, even revealed parallel evolution at the same sites that are important positions for the antiviral response.<sup>28</sup> These studies show evidence of episodic positive selection contributing to the innate species-specific resistance/susceptibility to diverse pathogens in mammals.

In this study, we investigated the impact of natural selection on the evolution of eight genes associated with the response of malaria infection in mammals: *CCL2*, *IL-10*, *iNOS*, *HO1*, *TNF- $\alpha$* , *DARC*, *G6PD* and *CD36*. The major functions of these genes in response to *Plasmodium* infection are described in Table 1. We also extended previous studies for genetic signatures of natural selection on immunogene evolution in mammals, including primates, rodents, and other typical

eutherians along with outgroups monotremes and marsupials.<sup>1,2,29–38</sup> Based on the genetic analysis surveys, we aimed to answer questions that if these malaria-related immunogenes were positively selected on specific lineages and specific codons throughout the mammalian evolution and at what time these selective pressures drove the divergence of mammalian immunogenes. A comparison of divergence of reference genes helps to estimate when the selective pressures acted on these malaria-related immunogenes. Phylogenetic analysis by maximum likelihood (PAML) also provides statistical evidences of positive selection on specific lineages and codons, and illustrates how these genes face the disease stress under the evolutionary history.

## Materials and methods

### Data collection

In this study, the nucleotide sequences of eight malaria-related immunogenes (*CCL2*, *IL-10*, *iNOS*, *HO1*, *TNF- $\alpha$* , *DARC*, *G6PD* and *CD36*) of mammals were collected from the National Center for Biotechnology Information's (NCBI) GenBank (Washington, DC, USA). Besides humans, members from two other groups of mammals, including non-human primates and rodents, and also other typical eutherians, were investigated for the evolutionary study of immunogenes (Table 2). Sequence comparison programs that follow the principle of searching for closest matches in a large

**Table 1.** The main function of the eight immunogenes during *Plasmodium* infection.

Gene	Function
CCL2	CCL2 is a cytokine that can recruit monocytes, memory T cells and dendritic cells to tissue in inflammation during <i>Plasmodium</i> infection
IL-10	IL-10 serves as an anti-inflammatory cytokine and may inhibit the expression of pro-inflammatory response cytokines, such as <i>TNF-<math>\alpha</math></i> and CCL2. Thus, it can prevent an exacerbated inflammatory response during <i>Plasmodium</i> infection
iNOS	iNOS can be rapidly expressed in response to pro-inflammatory process and catalyze the generation of NO. The NO or its oxidative products may serve as cell toxicity for inhibit growth of <i>Plasmodium</i> during infection
TNF- $\alpha$	TNF- $\alpha$ is a pro-inflammatory cytokine that regulates various immune cells, and subsequently induces inflammatory responses during early stage of <i>Plasmodium</i> infection
DARC	DARC is a glycosylated membrane-bound protein and serves as a non-specific receptor for several chemokines. DARC can be mainly found on endothelium cell, especially erythrocyte. DARC also plays an important role in erythrocyte internalization of certain <i>Plasmodium</i> (e.g. <i>P. vivax</i> and <i>P. knowlesi</i> ). Furthermore, mutation or null alleles of DARC can serve as a species-specific barrier for <i>Plasmodium</i> infection
CD36	CD36 play an important role on <i>P. falciparum</i> -infected erythrocyte sequestration. CD36 serves as a receptor for adherence of infected erythrocyte to capillary, and has been believed to contribute to <i>P. falciparum</i> survival from phagocytic clearance in spleen. However, it can also play a role in the immune response and recruit phagocytosis cells to the infected erythrocytes
HO-1	Lysis of infected erythrocytes releases free heme into blood, and excess free heme causes strong oxidative stress for tissues. HO-1 can catalyze free heme, protecting tissue from oxidative damage
G6PD	G6PD is an enzyme that catalyzes the first step of the pentose phosphate pathway. It is one of the most important enzymes in providing energy for erythrocytes. Although there is no growth difference in <i>Plasmodium</i> between normal or G6PD-deficient erythrocytes, the phagocytosis of parasite erythrocytes is more abundant in G6PD-deficient erythrocytes

**Table 2.** Accession numbers of sequences used in this study.

	CCL2	IL-10	iNOS	HOI	TNF- $\alpha$	DARC	G6PD	CD36	FGG	RAGI
Primate										
<i>Homo sapiens</i>	AK311960.1	NM_000572.2	NP_000616.3	NP_002124.1	NM_000594.3	NM_002036.3	NM_000402.3	NM_001001547.2	NM_000509.4	NM_000448.2
<i>Callithrix jacchus</i>	XM_002748333.2	XM_002760733.1	XM_002806845.2	XM_00274372.1	NM_00125726.1	XM_002760156.1	XM_002807924.2	XM_002751659	XM_002806643.2	XM_002755187.1
<i>Macaca mulatta</i>	NM_00103282.1	NM_001044727.1	XM_001106245.1	XM_00111324.1	NM_001047149.1	XM_001117200.2	NM_001266470	NM_001032913.1	AC191446.4	NM_001266772.1
<i>Nonascus leucogenys</i>	XM_003278400.2	XM_003272964.1	XM_003277085.1	XM_003264685	XM_003272094.1	XM_003258679.1	XM_003279364.1	XM_003252176.1	XM_003257865.2	XM_003254437.2
<i>Pan troglodytes</i>	XM_001174551.2	NM_001135620.2	XM_001148238.1	XM_525579.4	NM_001045511.1	XM_001170629.3	XM_001146640.3	XM_003318555.1	XM_001138780.3	XM_001154240.3
<i>Pongo abelii</i>	XM_002827247.2	XM_002809541.1	XM_002827149.1	NM_001132886.1	XM_002816720.2	XM_002809955.2	XM_002832311.1	XM_002818297.2	XM_002815235.2	XM_002821865.2
Rodent										
<i>Cavia porcellus</i>	NM_001172926.1	NM_001260485.1	NM_001172984.1	XM_003462326.1	NM_001173025.1	XM_003466665.1	XM_003462133.1	XM_003469814.1	XM_003476822.1	XM_003463833.1
<i>Cricetulus griseus</i>	XM_003495792.1	XM_003501219.1	XM_003495779.1	XM_003511957.1	XM_003508487.1	XM_003500262.1	NM_001246727.1	XM_003504836.1	XM_003499470.1	XM_003497460.1
<i>Mus musculus</i>	NM_011333.3	NM_010548.2	AY090567.1	NM_010442.2	NM_013693.2	NM_010045.2	NM_019468.2	NM_001159557.1	BC019828.1	NM_009019.2
<i>Rattus norvegicus</i>	NW_001084656.1	NM_012854.2	NM_012611.3	NM_012580.2	NM_012675.3	XM_002724893.1	NM_017006.2	NM_031561.2	NM_012559.2	NM_053468.1
Monotremes										
<i>Ornithorhynchus anatinus</i>	—	—	XM_001506837.2	XM_001520223.2	XM_001515885.1	—	XM_001505586.1	XM_001506533.1	XM_001512228.1	NM_001242754.1
Marsupials										
<i>Monodelphis melanaleuca</i>	—	XM_003340167.1	XM_001375545.2	XM_001375607.2	XM_001368652.1	XM_001379684.1	XM_001362790.1	XM_001364338.1	XM_001375602.2	XM_001364998.2
Other mammals										
<i>Aluropoda melanaleuca</i>	XM_002912354.1	XM_002919274.1	XM_002912338.1	XM_002925580.1	XM_002930032.1	XM_002927274.1	XM_002929683.1	XM_002918262.1	XM_002913851.1	XM_002926498.1
<i>Bos taurus</i>	NM_174006.2	NM_174088.1	NM_001076799.1	NM_001014912.1	NM_173966.2	NM_001015634.1	NM_001244135.1	NM_174010.2	BC102629.1	XM_867321.1
<i>Canis lupus</i>	NM_001003297.1	NM_001003077.1	NM_001003186.1	NM_001194969.1	NM_001003244.4	XM_003640206.1	XM_538209.3	NM_001177734.1	XM_532698.3	XM_540538.1
<i>Equus caballus</i>	NM_001081931.1	NM_001082490.1	NM_001081769.1	XM_001498885.2	NM_001081819.1	XM_001490641.2	XM_001492232.3	XM_001487907.1	XM_001914798.2	NM_001256901.1
<i>Oryctolagus cuniculus</i>	NM_001082294.1	NM_001082045.1	XM_002718780.1	XM_002711415.1	NM_001082263.1	XM_002715308.1	NM_001171382.1	XM_002712016.1	XM_002716891.1	NM_001171140.1
<i>Ovis aries</i>	XM_004012469.1	NM_001009327.1	XM_004012488.1	XM_004023352.1	NM_001024860.1	XM_004002645.1	NM_001093780.1	XM_004007809.1	XM_004017182.1	XM_004016411.1
<i>Sus crofa</i>	NM_001164515.1	NM_214041.1	NM_001143690.1	NM_001004027.1	NM_214022.1	NM_001244095.1	XM_00360515.2	NM_001044622.1	NM_001244524.1	AB091392.1

Cells containing a '—' indicate that sequences were not obtained from GenBank.

database such as BLAST-n and BLAST-x searches were performed by NCBI. Two nuclear genes, *FGG* and *RAG1*, were taken as the reference markers (phylogenetic marker) for comparison with malaria defense-related genes. Nucleotide sequences were aligned based on the translated protein sequences by ClustalW implemented in BioEdit v.7.<sup>39</sup>

### Phylogenetic reconstruction

Species trees were reconstructed based on reference genes *FGG* and *RAG1* as the input tree for positive selection analysis. We chose this reference phylogeny as the input tree because it could better reflect the evolutionary relationship of selected taxa and was independent from the tested immunogenes to prevent circular arguments using the immunogene trees.<sup>40</sup> We used the method of Heled and Drummond to reconstruct a species tree under Yule's pure-birth model.<sup>41,42</sup> The general time-reversible substitution model with gamma and invariant site heterogeneity model by BEAST v. 1.7.3 were applied based on the substitution model test implemented in MEGA v. 5.05.<sup>43,44</sup> Multiple-times independent pre-runs of 10 million generations of the length of Markov chain Monte Carlo (MCMC) were performed for obtaining better parameter priors to the last independent 10 million generations of the MCMC process. Genealogies were sampled every 1000 generations with the first 70% discarded as burning according to the simulation trajectory. Statistics of all the output values were summarized by TRACER v. 1.5.<sup>45</sup> Sampled trees were summarized by Tree Annotator v. 1.6.1,<sup>43</sup> and Fig Tree v. 1.3.1<sup>46</sup> was used to display the consensus tree. The divergent times of mammals estimated by O'Leary et al. were consulted for inferring the times of positive selection events.<sup>47</sup>

### Recombination analyses

Recombination could cause excessively long terminal branches in phylogeny owing to the introduction of apparent substitution rate heterogeneity among sites,<sup>48</sup> and therefore resulted in type-I error (false positive) when detecting signals of positive selection using phylogenetic-based approaches.<sup>49</sup> Probable recombinant fragments were then identified by GeneConv (<http://www.math.wustl.edu/~sawyer/geneconv/>), in which the segments that shared identical 5' and 3' ends are thought as consequences of recombination events.<sup>50</sup> Recombination between ancestors of two sequences (i.e. global inner fragments) was assessed by 10,000 random permutations to indicate the significance of putative recombinant regions.

### Detecting genetic signatures of positive selection

The evolution of specific taxa and codons at a specific locus was tested for deviation from neutrality. We used

the maximum likelihood approach to get the ratio of non-synonymous mutation rate ( $Ka$ ) over the synonymous mutation rate ( $Ks$ ) with higher statistical power by codon-based analysis (codeml) implemented in the PAML v. 4 package.<sup>51</sup>  $Ka/Ks$  ( $= \omega$ ) is regarded as the common approach to test the signature of molecular adaptation in response to natural selection. We used three selective models to test for positive selection in mammalian malaria-related genes: the branch model (free-ratio model), site model and branch site model.

The free-ratio model assumes independent  $\omega$  in all lineages and average  $\omega$  over sites. We used this model to test whether the positive selection acted on specific mammal taxa by comparison with the null model (one-ratio, M0). The likelihood ratio test (LRT) statistic, 2 $\Delta$ L, against the  $\chi^2$  distribution with degrees of freedom equal to the difference in the number of parameters was used to compare data fit for testing whether the free-ratio model provides a suitable explanation for the evolution of malaria-related genes.<sup>52</sup>

However, the  $\omega$  ratio averaged over sites may have low power to detect positive selection owing to the assumption that all sites in a sequence are under equal selective pressure, a potentially unrealistic assumption, as most amino acids in proteins are under functional constraints.<sup>53</sup> Therefore, site models that allow no positive selection of codon sites including M1a (the nearly neutral model) and M7 [the model allowing  $\omega$  variation assumes a beta distribution over the interval (0, 1)] were tested against the nested models that allow positive selection of codon sites, including M2a (the positive selection model) and M8 (the model adds a discrete  $\omega$  class to the beta model with positive selection  $\omega > 1$ ), respectively. The M8 model was also compared to the M8a model, which has similar assumptions as the M8 model except limiting  $\omega = 1$ . To perform the PAML analysis, non-eutherian mammals were removed from the analyses to prevent the high saturation in nucleotide replacement to bias the selection analysis. The input prior tree is shown as the left tree topology in Figure 1. The LRT was used to evaluate the fit of nested models to the data for positive selection.<sup>52</sup> We accepted the presence of positive selection when a significantly better fit of the selective model was found, and calculated the posterior probabilities of codon with  $\omega > 1$  according to Bayes empirical Bayes (BEB) analysis.<sup>54</sup> In order to confirm the robustness of the results, we also used gene trees of each loci as inputs to repeat site model analyses to ensure the inference of positive selection.

In addition, the positive selection of malaria-related genes was focused on primates and Glires (rodents + rabbit). We applied branch-site model A to the eight malaria-related genes and designated primates, Glires or primates + Glires as the foreground branches, respectively, in a series of independent tests.<sup>55</sup> The LRT was used to evaluate whether branch-site model



A had a significantly better fit for the codon site with  $\omega > 1$  in comparison with branch-site model A1, which fixes  $\omega$  to 1.0 on the branches of interest.<sup>55</sup> If the LRT significantly rejected model A1, further tests for functional divergence between the foreground and background clades were performed by DIVERGE v.3.0 using 1000 bootstrap type-II divergence analysis.<sup>56</sup> For understanding the evolutionary trajectory of the change of positively selected codons in foreground clades, the ancestral states of these codons were inferred using the probabilistic-based maximum likelihood method under the M5 substitution model with the assistance of FastML.<sup>52,57</sup>

## Results and discussion

### Phylogenetic tree reconstruction and divergent time estimation

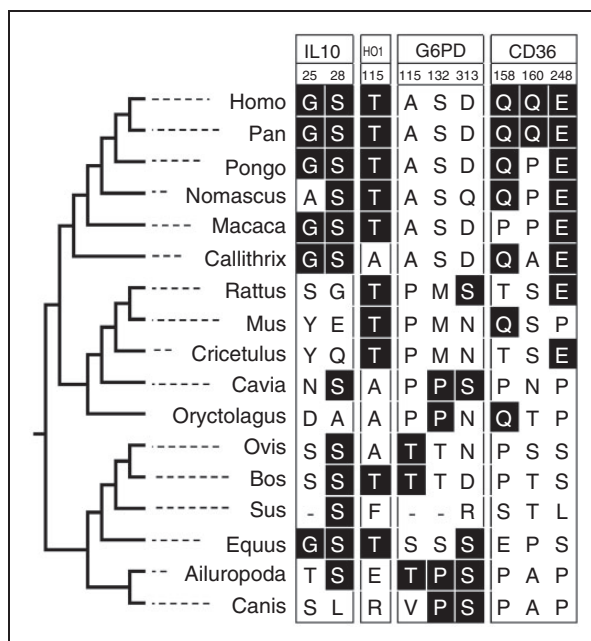
The Bayesian inference of evolutionary relationships of the studied species by two nuclear reference genes (left panel of Figure 1) is similar to the parsimonious analysis based on combined nuclear and phenomic data.<sup>47</sup> Posterior probabilities for all nodes are  $>0.90$ , indicating a reliable tree topology, except a relatively low support for the grouping of rabbits (*Oryctolagus cuniculus*) and rodents (posterior probability = 0.683). O'Leary et al. dated the divergent times of mammals by phenotypic and molecular data, and estimated the coalescent times of mammals, Theria, eutherians, primates, Glires

and rodents as 166.2 Mya, 127.5 Mya, 64.85 Mya, 53.1 Mya, 63.4 Mya and 56.8 Mya, respectively.<sup>47</sup> The splitting times between old world monkey, gibbon, orang utan, chimpanzee and human are 23.3 Mya (between Oligocene and Miocene),  $14.9 \pm 2.0$  Mya,  $11.3 \pm 1.3$  Mya,  $6.4 \pm 1.5$  Mya and  $5.4 \pm 1.1$  Mya (Miocene), respectively. Comprehensive surveys on nearly 3000 fecal samples from wild chimpanzees, bonobos and gorillas in Central Africa suggested that ancestors of human-infected *P. falciparum* could be relatives of gorilla and chimpanzee parasites (e.g. *P. falciparum* and *P. reichenowi*, respectively).<sup>8,58</sup> Several newly discovered *Plasmodium* isolates (*P. gaboni*, *P. billbrayi* and *P. billcollinsi*, and two unpublished isolates, *P. GorA* and *P. GorB*) were also shown to be able to infect great apes.<sup>10,11,59</sup> The infectious history of hominid could be traced to the divergence of the *P. guboni* and *P. falciparum*/*P. reichenowi* lineages since 21 Mya,<sup>9</sup> which is roughly during the divergence period of old world monkeys and gibbons (23.3–14.9 Mya). The estimated splitting times are helpful for inferring the time of positive-selection events.

### Recombination analyses

Among the sampled sequences, *TNF- $\alpha$*  is the only gene that was detected to represent signals of recombination in both inner and outer fragments ( $P = 0.0034$  and  $P = 0.0031$ , respectively; Table 3). The inner recombination fragments inferred with simulation  $P < 0.05$  are nucleotide sites 258–398 in the pairs marmoset–human, marmoset–chimpanzee and marmoset–orang utan. For preventing the false positive inference of positive selection due to recombination, *TNF- $\alpha$*  was removed for further analyses of positive selection.

In addition to *TNF- $\alpha$* , the panda *CD36* ( $P = 0.0007$ ), sheep *DARC* ( $P = 0.0467$ ), orang utan *G6PD* ( $P = 0.0001$ ) and marmoset *iNOS* ( $P < 0.0001$ ) also represent signals of recombination between ancestral sequences outside the alignment sequences (Table 3). However, because only short recombination fragments were detected in these four genes (Table 3), and the estimation of recombination by outer fragments could be caused by biased sampling, it is difficult to determine whether recombination really affects the tests of positive selection. Therefore, we repeated the analysis assuming that sequences were linked by a 'star' phylogeny to remove the affect of recombination events, and if consistent LRTs were obtained, the recombination effect could be ignored.<sup>60</sup> In the reanalyzed results, *CD36* and *G6PD* showed consistent LRTs in all model tests, suggesting none or less interference of recombination for the positive selection test (Table 4). In contrast, the *DARC* and *iNOS* showed inconsistent LRTs in site- and free-ratio models and branch-site model when using different input trees, respectively (Table 4), indicating that the recombination could



**Figure 1.** The positively selected amino acids estimated under the site model with a posterior probability  $>0.8$ . Amino acids marked in dark background are estimated evolving with  $\omega > 1$  under the M8 model.

**Table 3.** Summary of recombination (Rec.) events identified by GeneConv of eight malaria-related immunogenes.

	Global inner fragments <sup>a</sup>				Global outer fragments <sup>b</sup>			
	Simulation	P-Value	Rec. <sup>c</sup>	Breakpoint Species pair	Simulation	P-Value	Rec. <sup>c</sup>	Breakpoint Species
CCL2	0.3540	0	—		0.4949	0	—	
IL-10	0.5464	0	—		0.1120	0	—	
iNOS	0.0986	0	—		<b>&lt;0.0001</b>	2	485–542, 466–483	<i>Callithrix jacchus</i>
HO1	0.3993	0	—		1.0000	0	—	
TNF- $\alpha$	<b>0.0034</b>	1	258–398	<i>Callithrix jacchus</i> – <i>Pan troglodytes</i> <i>Callithrix jacchus</i> – <i>Pongoa belii</i> <i>Homo sapiens</i> – <i>Callithrix jacchus</i>	<b>0.0031</b>	1	233–245	<i>Pan troglodytes</i>
DARC	0.1087	0	—		<b>0.0467</b>	1	534–536	<i>Ovis aries</i>
G6PD	0.4838	0	—		<b>0.0001</b>	1	1303–1307	<i>Pongoa belii</i>
CD36	0.5459	0	—		<b>0.0007</b>	1	433–437	<i>Ailuropodame lanoleuca</i>

Bold values denote  $P < 0.05$ .

<sup>a</sup>Gene conversion events between ancestors of two sequences in the alignment.<sup>50</sup>

<sup>b</sup>Gene conversion events between ancestors outside the alignment or gene conversion events that have subsequently been obscured by mutation or recombination events.<sup>50</sup>

<sup>c</sup>Recombination events: fragments with the same 5' or 3' breakpoints identified by GeneConv were considered to represent one recombination event.<sup>50</sup>

cause a false positive selection analysis. Therefore, these two genes were also removed in the further analyses of positive selection to prevent the false positive result.

### Gene tree vs. species tree

The gene tree of every malaria-related immunogene reconstructed by the Neighbor-Joining method revealed inconsistent tree topology with the species tree reconstructed by two reference genes (online Supplementary Figure S1). Heterogeneous tree topologies suggest different evolutionary tracks of these functionally important genes throughout the evolutionary history of eutherians. Such phylogenetic inconsistency may be signatures of natural selection, gene duplication or horizontal gene transfer.<sup>61</sup> To ensure the robustness of the analytical results, we used species tree and gene tree as inputs to conduct the PAML analyses.<sup>62,63</sup> Concordant results of LRTs calculated by both analytical strategies (Table 4 and online Supplementary Table S2) are consistent with the simulation of Yang et al. that tree topology uncertainty may only be a mild effect for LRT.<sup>64</sup>

### Multiple episodes of positive selection on malaria-related genes on specific lineages and species codons

Evolutionary rates of non-synonymous sites to the synonymous sites are not constant among lineages through the evolutionary history of all the genes of interest. The one-ratio model was rejected in LRTs in *CCL2*, *HO1*, *G6PD* and *CD36* ( $P < 0.05$ ), and a fuzzy  $P$ -value ( $P = 0.05$ ) in *IL-10* provides moderate but not

overwhelming evidence against the null one-ratio model (Table 4), but the free-ratio model was found to be a better fit for the evolution of most malaria-related immunogenes. Lineages that evolved with  $\omega > 1$  are shown in Figure 2. Parameter estimates of site model and free-ratio model are listed in online Supplementary Table S1.

Because the M7–M8 comparison is less robust or more powerful than the M1a–M2a comparison, relatively abundant codons of a gene are able to be estimated to evolve with  $\omega > 1$ .<sup>65</sup> *IL-10*, *HO1*, *G6PD* and *CD36* were suggested to be a better fit in the M8 model than in the null M7 model, but the nearly neutral model (M1a) cannot be rejected by the positive selection model (M2a). Among the four genes that are better fit in M8 than M7, only *IL-10* could reject the null model of  $\omega = 1$  fixed (M8a) by M8, which is also consistent with the phenomenon of relatively less robust evaluation for positive selection by the M7–M8 comparison.

Cross verification of results of the molecular dating and free-ratio model analysis revealed that the timing of positive selection on malaria-related immunogenes in mammals could be divided into four phases.<sup>47</sup> The first phase is during Palaeocene, in which the primates + Glires and other eutherian mammals diverged at roughly 64.85 Mya (Figure 2).<sup>47</sup> Selective signals during this phase were detected in all genes, but these genes suffered different selective pressures on different lineages: *CCL2*, *IL-10* and *CD36* in common ancestors of primates and Glires, and *CCL2*, *HO1* and *G6PD* in ancestors of other eutherians. Continued selective pressure is acted on common ancestors of Glires' *CCL2* and

**Table 4.** Summaries of LRT results of PAML analyses using branch model, site models and branch site model with species tree.

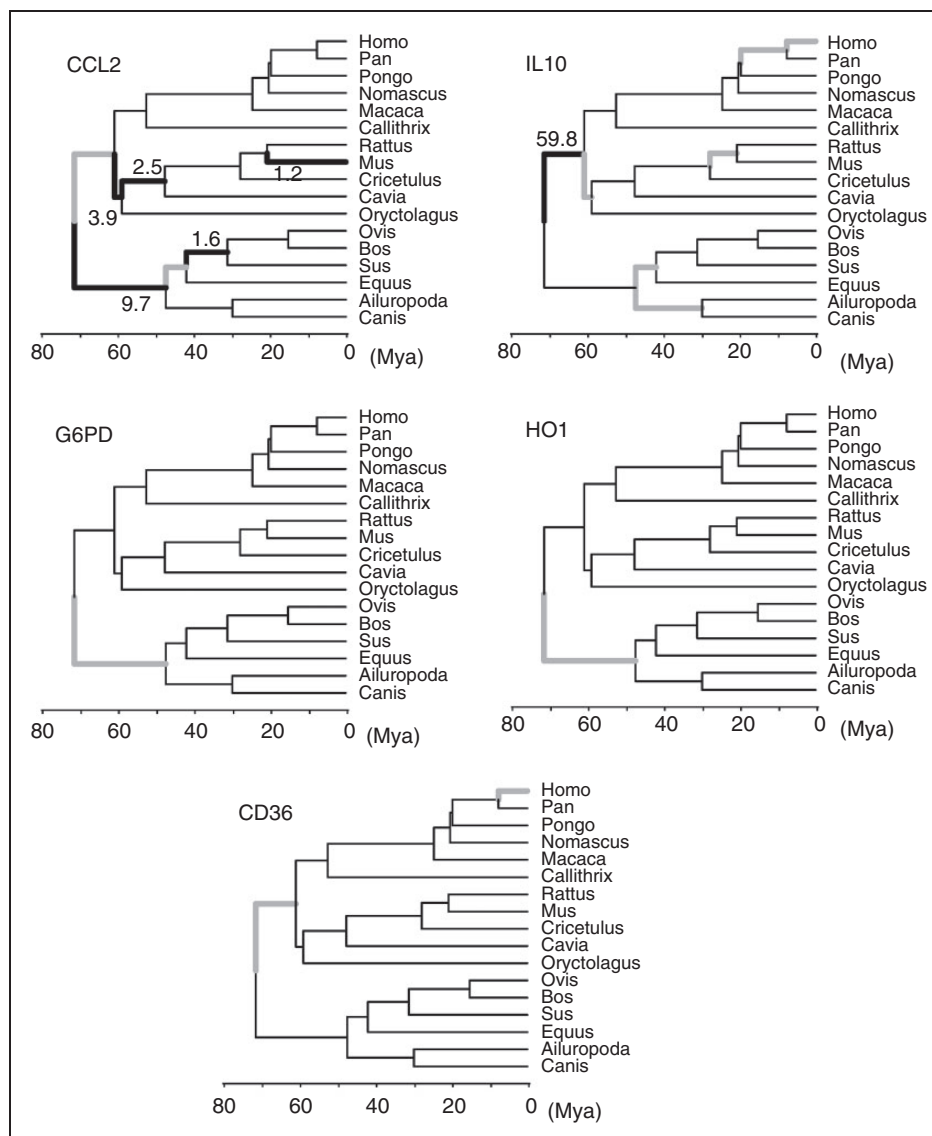
	CCL2	IL-10	HOI	G6PD	G6PD <sup>a</sup>	CD36	CD36 <sup>a</sup>	iNOS	iNOS <sup>a</sup>	DARC	DARC <sup>a</sup>
<b>Branch model</b>											
One ratio	InL	-2219.1	-3112.52	-3982.11	-6028.51	-8051.82	-8362.54	-13688.43	-13796.30	-6716.27	-7124.10
Free ratio	InL	-2192.3	-3097.48	-3972.49	-6002.72	-7997.29	-7997.29	-13653.18	-13761.80	-6693.03	-7106.79
	2ΔL	53.60	30.10	19.24	51.59	109.07	730.50	70.50	69.00	46.47	34.63
	P	<b>0.001</b>	0.051	<b>0.020</b>	<b>0.001</b>	<b>0.006</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>0.006</b>	0.120
<b>Site model</b>											
M1a	InL	-2159.97	-3057.2	-3907.37	-5943.54	-7865.33	-8143.69	-13516.91	-13607.60	-6587.29	-6957.15
M2a	InL	-2157.5	-3057.2	-3907.37	-5943.54	-7865.33	-8143.69	-13516.91	-13607.60	-6581.24	-6946.21
	2ΔL	4.95	0	<-0.001	0	0	0	<-0.001	0	12.12	21.89
	P	0.084	1	N.C.	1	1	1	N.C.	1	<b>0.001</b>	<b>&lt;0.001</b>
M7	InL	-2158.58	-3051.9	-3872.94	-5936.08	-7861.78	-8129.98	-13439.88	-13530.96	-6592.01	-6952.01
M8	InL	-2156.11	-3047.74	-3867.88	-5917.83	-7856.42	-8127.44	-13429.25	-13518.42	-6577.37	-6934.57
	2ΔL	4.92	8.32	10.11	36.5	10.71	5.08	21.26	25.08	29.27	-685.11
	P	0.085	<b>0.008</b>	<b>0.003</b>	<b>&lt;0.001</b>	<b>0.002</b>	<b>0.039</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>&lt;0.001</b>	N.C.
M8a	InL	-2156.3	-3050.16	-3868.56	-5917.83	-7856.42	-8127.44	-13429.9	-13519.17	-6584.2	-6947.61
M8	InL	-2156.11	-3047.74	-3867.88	-5917.83	-7856.42	-8127.44	-13429.25	-13518.42	-6577.37	-6934.57
	2ΔL	0.37	4.84	1.34	0	0	0	1.13	1.49	13.65	26.08
	P <sup>b</sup>	0.272	<b>0.016</b>	0.176	0.5	0.5	0.5	0.182	0.111	<b>&lt;0.001</b>	<b>&lt;0.001</b>
<b>Branch site model (foreground: primates + Glires)</b>											
A1	InL	-2156.29	-3054.13	-3901.78	-5942.25	-7862.87	-8137.66	-13500.4	-13500.40	-6583.02	-6946.30
A	InL	-2155.03	-3054.13	-3901.78	-5942.25	-7862.66	-8137.59	-13500.4	-13500.40	-6576.2	-6943.39
	2ΔL	2.52	0	0	0	0.42	0.14	0	0	13.64	5.83
	P <sup>b</sup>	0.056	0.5	0.5	0.5	0.499	0.712	0.5	0.5	<b>&lt;0.001</b>	<b>0.008</b>
<b>Branch site model (foreground: primates)</b>											
A1	InL	-2158.87	-3051.84	-3906.69	-5942.6	-7861.9	-8138.56	-13516.91	-13607.14	-6587.23	-6956.98
A	InL	-2159.97	-3051.93	-3906.69	-5942.6	-7857.15	-8134.79	-13516.91	-13603.51	-6586.4	-6956.26
	2ΔL	-2.19	0.18	0	0	9.5	7.54	0	7.25	1.67	1.45
	P <sup>b</sup>	N.C.	0.854	0.5	0.5	<b>0.001</b>	<b>0.006</b>	0.5	<b>0.007</b>	0.133	0.228
<b>Branch site model (foreground: Glires)</b>											
A1	InL	-2153.3	-3056.64	-3906.4	-5943.54	-7865.11	-8142.63	-13500.37	-13500.37	-6585.23	-6950.58
A	InL	-2153.21	-3056.64	-3906.4	-5943.54	-7864.92	-8142.48	-13500.37	-13500.37	-6584.93	-6950.58
	2ΔL	0.18	0	0	0	0.38	0.32	0	0	0.6	0
	P <sup>b</sup>	0.336	0.5	0.5	0.5	0.536	0.287	0.5	0.5	0.382	0.5

Bold denotes  $P < 0.05$ .

N.C.: Not calculable.

<sup>a</sup>PAML analyses using the 'star' tree as the input tree. Note that the iNOS have inconsistent LRT results in branch site model test (foreground: primates), and the DARC have inconsistent LRT results in both free-ratio model test and site model test (M7 vs. M8).

<sup>b</sup>The P-Value is half of the P-Value using  $\chi^2$  test due to 50:50 mixture distribution of point mass 0 and  $\chi^2$  of null model.<sup>65</sup>



**Figure 2.** Summary of free-ratio model shows lineages that evolve with  $\omega > 1$  (bold branches). Tree topology is inferred by two reference genes, *FGG* and *RAG1*, by Bayesian inference. The  $\omega$  value of the positively selected branches is labeled beside the bold branches. The bold gray branches are the lineages with an estimated  $\omega > 100$ , representing relatively abundant non-synonymous variations or very rare synonymous variations.

*IL-10* at roughly 63.4 (65.0–61.7) Mya during the early Eocene (Figure 2).<sup>47</sup> In fact, selection signals were prevalently detected in *CCL2* and *IL-10* throughout evolutionary history of eutherians. The third phase describes the selection of hominidae genes, including human *IL-10* and *CD36*. This phase describes the selective pressure on primates, especially for the apes, during the late Miocene. The fourth phase focused on the selection of other eutherians (exclusive of primates and Glires) since the Eocene, including *CCL2* and *IL-10* (Figure 2). Frequent detection of positive selection signals on four phases of mammalian evolutionary history is consistent with the arms races evolutionary model, a hypothesis used to explain the high variability of immunogenes.

### Long-term selective pressures drove the evolution of *CCL2* and *IL-10*

*CCL2* and *IL-10* are both non-specific immune-related genes. The free-ratio model showed significant higher  $\omega$  in several branches in *CCL2* and *IL-10*, revealing prevalent genetic change in response to the pathogen infections since the common ancestors of eutherians in Palaeocene (Figure 2). Frequent episodic positive selection drives the genetic change of immunogenes in responses to pathogen infections.<sup>66,67</sup> The genetic diversity of *CCL2* is associated with its receptor *CCR2* regarding susceptibility to the immune-mediated disease.<sup>68</sup> *CCR2* has been evidenced to be positively selected in placental mammals.<sup>38</sup> However, the site



model does not reject the neutral or nearly neutral models (M1a, M7 and M8a) in *CCL2* (Table 4). Signals of positive selection were detected in the branch model test but were not seen in site model analysis. This suggests that the pathogen stress could merely drive gene divergence between organisms but matter less with the structural and functional change of *CCL2*.

*IL-10* also carried several genetic signals of positive selection throughout its evolutionary history based on the free-ratio model analysis. However, in contrast to the absence of selective signals on specific codons of *CCL2*, the *IL-10* was detected with positive selection at its hydrophobic N terminals (25G and 28S; Figure 1) in most of the primate lineages and in certain eutherians (Table 4). The protein structure in which these two putative positively-selected amino acids are located is beyond the published crystal structure of human IL-10 (PDB: 2H24). In the genetic survey by Neves et al.,<sup>29</sup> one codon (the 26th codon) at the N-terminal of IL-10 was also detected to be positively selected in mammals, which is consistent with our results of positive selection on N-terminal region. IL-10 plays an important function as a component of local defense (inflammation) against pathogen infections, with significantly higher  $\omega$  driving the change of protein composition; its genetic divergence was apparently a genetic response to the diverse pathogens, supporting the arms race model of immune evolution.

### Positive selection on *HO1*

Evidence for positive selection on primates and rodents *HO1* was only found in the site model analysis (M8 model) but not in the free-ratio and branch-site models (Table 4, Figure 2), suggesting that the selective pressure could be pervasive in mammals, rather specifically in primates, rodents or in other specific taxa. Four codons (35R, 115T, 191I and 226P) have been estimated at  $\omega > 1$  ( $\omega = 1.773$ ) in the M8 model but only 115T has a high posterior probability (0.943), and these putatively positively selected codons are distally located from any known ligand binding site.<sup>69</sup> *HO1* is ubiquitously expressed and catalytically active for stress responses in multiplicity and have broad functions but are conserved along mammalian evolution.<sup>70,71</sup> Therefore, the functionally important ligand binding sites were selectively constrained with low residual replacement rates among species, whereas non-binding regions and non-coding regions (e.g. promoters) could be more highly polymorphic, associated with the susceptibility to multiple diseases.<sup>72–74</sup> The bending region that links to binding domains could be more flexible and has higher amino acid replacement rates. For example, the 115th codon that is located between two  $\alpha$ -helices beside the binding domain of residues HEM\_A300, HEZ\_A401 and Q86\_A301 has

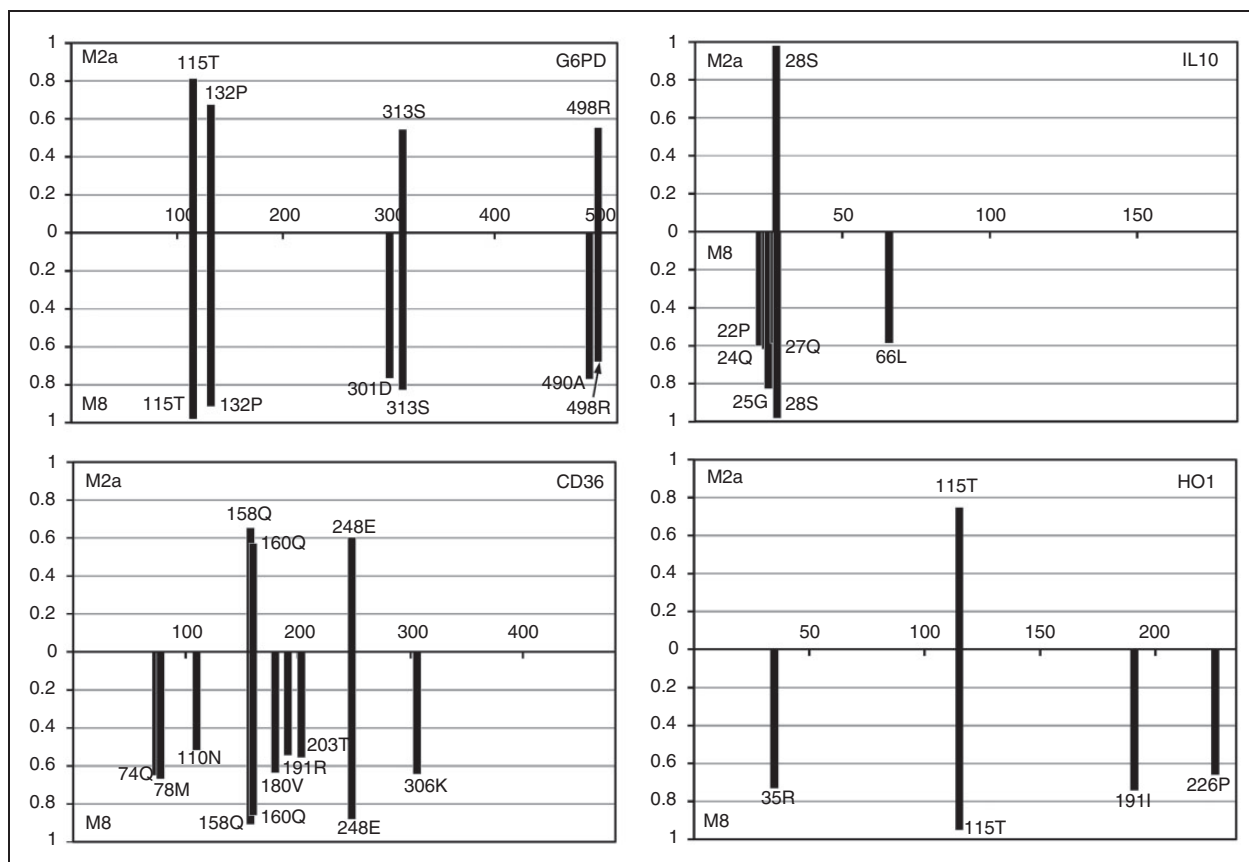
a higher replacement rate and is suggested to be positively selected in primates, rodents, cattle and horses (Figure 1).<sup>69</sup> In addition, a trend of identical or similar adaptive signals on *HO1* and *iNOS* lineages (Figure 1) suggests parallel evolution at these two synergistically expressed genes,<sup>75–77</sup> like that of the *RLR* genes *RIG-I* and *LGP2*, which suffer independent but parallel selective pressures at the same sites.<sup>28</sup>

### Positive selection on the primate *CD36* gene

CD36 is a major receptor for the *Plasmodium*-infected erythrocytes for sequestering the infected erythrocytes.<sup>78,79</sup> Recent positive selection on variations of CD36 was suggested to be responsible for differential susceptibility to *P. falciparum* malaria in Africa.<sup>30,32</sup> In the present study, although we did not comprehensively survey the human population, signals of selective pressures were also detected in human lineage and the common ancestor of primates and Glires based on free-ratio model analysis (Figure 2). No other lineages showed significantly high average  $\omega$  under the free-ratio model test, but signals of positive selection in primates were found by site model and branch-site model tests. This indicates that adaptive divergence of *CD36* between mammals occurred only on specific codon positions.

The positive selection of *CD36* is evident in primate lineages by M8 site model analysis, but the nearly neutral (M1a) model and the M2a positive-selection model failed to detect positive selection signals (Table 4). Under the M8 model analysis, three codons, 158Q, 160Q and 248E, evolving with  $\omega > 1$  (posterior probability  $> 0.80$ ; Figure 3) were mostly found in primates, supporting the existence of selective pressures that impose rapid protein composition changes in primates. *Plasmodium*-infected erythrocytes can adhere to the CD36 protein and prevent parasites from adhering to neurovascular adhesion molecules (e.g. ICAM-1) in order to protect from cerebral malaria.<sup>80,81</sup> The rapid non-synonymous evolution of *CD36* genes driven by positive selection in human or in primates could not only be related to the multiple-function divergence of CD36, but also to multiple malaria parasitism with host shift through *Plasmodium* evolution.<sup>8–11</sup>

Three foreground branches were designated for the branch-site model test: primates, Glires and primates + Glires. Among all comparisons of all genes between model A1 ( $\omega_s = 1$  fixed) and model A, *CD36* was a better fit in model A only when the primates were designated as the foreground branch ( $P = 0.001$ , Table 4). Three codons were inferred to be  $\omega > 1$  with a posterior probability  $> 0.9$  in BEB analysis: 18V (probability = 0.972, for human, chimpanzee and macaque), 208L (probability = 0.953, for all primates except macaque) and 412N [probability = 0.830, for hominidae (i.e. human, chimpanzee and orang utan)]. The inter-specific



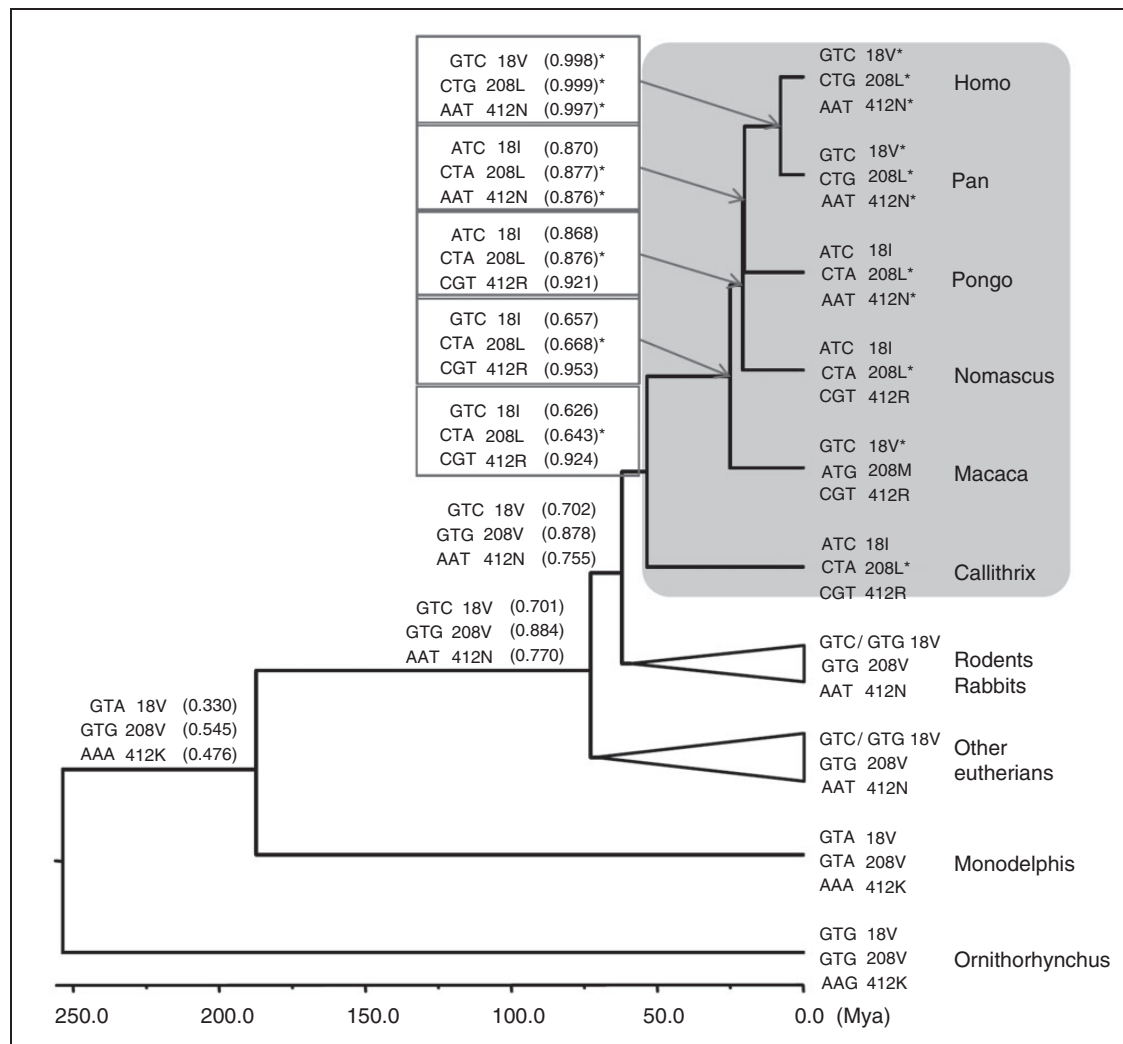
**Figure 3.** Posterior probability of the inferred positively selected sites estimated from site models of M2a and M8. The x-axis indicates the codon position and the y-axis is the posterior probability of the codon with  $\omega > 1$ .

test of selective pressures on specific codons indicated that the extracellular domain of CD36 might be the earlier target of positive selection since the Eocene (the MRCA of primates, 208L) and since the early Miocene (the MRCA of hominidae, 412N).<sup>33</sup> Selected pressure on 18V, which is located on the *N*-terminal transmembrane domain of CD36, happened for humans and chimpanzees during the late Miocene (c. 7.77 Mya). It should be noted that codon 18 and codon 412 of all or most background taxa are also Val and Asn, respectively, suggesting the reversible amino acid change for keeping ancestral states to be positively selected.

We further examined whether the positively selected amino acids in foreground lineages that are identical to backgrounds are a consequence of reversible evolution (i.e. identical codons) or only a coincidence (i.e. synonymous codons) based on ancestral sequence inferences. The ancestral states of the 18th codon of the common ancestor of primates, common ancestor of primates + Glires and common ancestor of all eutherians are all GTC (Val), identical to human, chimpanzee and macaque; the identical codon of the 412th Asn (AAT) was also observed between the hominidae and the

common ancestors of primates + Glires and of all eutherians (Figure 4).

A higher proportion of radical amino acid changes were found between primates and Glires (24.556 radical changes/467 amino acids) clades and between primates and other eutherians (28.185 radical changes/467 amino acids) than between Glires and other eutherians (19.852 radical changes/467 amino acids) (Table 5). In addition, the proportion of fixed radical change (F00,R) between primates and Glires (0.012) and between primates and other eutherians (0.011) is higher than between Glires and other eutherians (0.003) (Table 5). A high proportion of the radical changes in primates that are distinguishable from other eutherians suggest early episodes of positive selection driving functional divergence of CD36 in primates, especially on its extracellular domain, which is involved in recognition or binding activity to the malaria-infected erythrocytes. In fact, functions of CD36 are likely to shift from nutritional properties to extracellular matrix-binding capacities,<sup>82</sup> which incriminate many other disease-response functions.<sup>83–90</sup> Both branch site model analysis on positive selection and radical change estimates for functional divergence analysis indicated that the positive selection



**Figure 4.** Ancestral states of positively selected codons of the primate *CD36* gene. The codons, codon positions, amino acids and posterior probabilities of the ancestral state reconstruction (parentheses) are listed. The positively selected codons are inferred by the branch site model analysis using the primates (gray area) as the foreground branch and are indicated by asterisks (\*).

**Table 5.** Summary of pairwise comparison of functional divergence between primates, glires and other eutherians of *CD36*.

	Primates/glires	Primates/others	Glires/others
N	413.222	405.037	426.926
C	29.222	33.778	20.222
R	24.556	28.185	19.852
F00,N	0.730	0.648	0.645
F00,R	0.012	0.011	0.003
F00,C	0.012	0.011	0.002

N: number of sites with no change between two clusters; C: number of sites with conserved change between two clusters; R: number of sites with radical change between two clusters; F00,N: proportion of sites with no change within and between gene clusters; F00,R: proportion of sites with no change within gene clusters, but conserved change between clusters; F00,C: proportion of sites with no change within gene clusters, but radical change between clusters.

might drive functional shift of *CD36* from other mammals to primate.

### Positive selection of non-primate-mammalian *G6PD*

We found evidence of positive selection in five of the eight immunogenes analyzed; *G6PD* is the only one gene that showed no positive selection signal in primates or the earliest ancestors. However, selective signatures *G6PD* were detected in ancestor lineages of other Eutheria excluding primates and Glires clade, based on free-ratio model analysis (Figure 2), and positive selection in codon sites were found in all Eutheria excluding primates, according to the site model analysis (Figure 1). Although *G6PD* diversity has been reported to show its positive selection regarding malaria,<sup>30</sup> the estimated origin of malaria-resistance alleles of *G6PD* is

roughly from around 10,000 years ago.<sup>31</sup> The presumptive positive selection on *G6PD* in non-primate ancestors of Eutheria was approximately at 50 Mya (Figure 2), far before the origination of malaria-resistance alleles. Therefore, the positive selection at *G6PD* detected in this study is probably unrelated to malaria but it can be driven by other factors.

## Conclusions

*Plasmodium*-induced malaria parasitism found in most mammals implies the long term inter-relationship between the *Plasmodium* parasite and mammalian immune system. In this study, primates showed the highest rate of evolution in *CD36* than that of other mammals. A higher ratio of radical amino acid changes suggests that primate *CD36* might have undergone functional divergence from other eutherians in response to the blood-borne infectious diseases, such as malaria. In addition to *CD36*, other immunogenes were also detected to be under positive selection in different lineages and on several codons. For example, multiple lineages of *CCL2* and *IL-10* were evolved with a higher proportion of non-synonymous mutations, revealing a prevalent evolutionary response against infections. Ancient recombination detected in the pro-inflammatory cytokine gene *TNF- $\alpha$*  among primates is likely an alternative mechanism in maintaining gene polymorphism. Genetic signatures of positive selection during evolutionary trajectories of mammalian immune-related genes indicate the prevalence of a particular type of adaptive mutations offering protection against malaria, which clues the arms race co-evolution of infectious *Plasmodium* and eutherians.

## Funding

This research was financially supported by the National Science Council in Taiwan (NSC 102-2621-B-003-005-MY3).

## Conflicts of interest

The authors do not have any potential conflicts of interest to declare.

## Acknowledgements

Our gratitude goes to the Academic Paper Editing Clinic, NTNU.

## References

1. Oliveira TYK, Harris EE, Meyer D, et al. Molecular evolution of a malaria resistance gene (DARC) in primates. *Immunogenetics* 2012; 64: 497–505.
2. Tung J, Primus A, Bouley AJ, et al. Evolution of a malaria resistance gene in wild primates. *Nature* 2009; 460: 388–391.
3. Silvie O, Mota MM, Matuschewski K, et al. Interactions of the malaria parasite and its mammalian host. *Curr Opin Microbiol* 2008; 11: 352–359.
4. Garnham PC. Recent research on malaria in mammals excluding man. *Adv Parasitol* 1973; 11: 603–630.

5. Garnham PC. Malaria in mammals excluding man. *Adv Parasitol* 1967; 5: 139–204.
6. Schaer J, Perkins SL, Decher J, et al. High diversity of West African bat malaria parasites and a tight link with rodent *Plasmodium* taxa. *Proc Natl Acad Sci U S A* 2013; 110: 17415–17419.
7. Weatherall D and Clegg JB. Genetic variability in response to infection: malaria and after. *Genes Immun* 2002; 3: 331–337.
8. Liu WM, Li YY, Learn GH, et al. Origin of the human malaria parasite *Plasmodium falciparum* in gorillas. *Nature* 2010; 467: 420–425.
9. Ollomo B, Durand P, Prugnolle F, et al. A new malaria agent in African hominids. *Plos Pathog* 2009; 5: e1000446.
10. Prugnolle F, Durand P, Ollomo B, et al. A fresh look at the origin of *Plasmodium falciparum*, the most malignant malaria agent. *Plos Pathog* 2011; 7: e1001283.
11. Rich SM, Leendertz FH, Xu G, et al. The origin of malignant malaria. *Proc Natl Acad Sci U S A* 2009; 106: 14902–14907.
12. Bonneaud C, Balenger SL, Zhang JW, et al. Innate immunity and the evolution of resistance to an emerging infectious disease in a wild bird. *Mol Ecol* 2012; 21: 2628–2639.
13. Eizaguirre C, Lenz TL, Kalbe M, et al. Rapid and adaptive evolution of MHC genes under parasite selection in experimental vertebrate populations. *Nat Commun* 2012; 3: 621.
14. van Oosterhout C. A new theory of MHC evolution: beyond selection on the immune genes. *Proc Biol Sci* 2009; 276: 657–665.
15. Oliver MK and Piertney SB. Selection maintains MHC diversity through a natural population bottleneck. *Mol Biol Evol* 2012; 29: 1713–1720.
16. Martinsohn JT, Sousa AB, Guethlein LA, et al. The gene conversion hypothesis of MHC evolution: a review. *Immunogenetics* 1999; 50: 168–200.
17. Piertney SB and Oliver MK. The evolutionary ecology of the major histocompatibility complex. *Heredity* 2006; 96: 7–21.
18. Garrigan D and Hedrick PW. Perspective: Detecting adaptive molecular polymorphism: Lessons from the MHC. *Evolution* 2003; 57: 1707–1722.
19. Edwards SV and Hedrick PW. Evolution and ecology of MHC molecules: from genomics to sexual selection. *Trends Ecol Evol* 1998; 13: 305–311.
20. Meyerson NR and Sawyer SL. Two-stepping through time: mammals and viruses. *Trends Microbiol* 2011; 19: 286–294.
21. Barreiro LB and Quintana-Murci L. From evolutionary genetics to human immunology: how selection shapes host defence genes. *Nat Rev Genet* 2010; 11: 17–30.
22. Akira S and Takeda K. Toll-like receptor signalling. *Nat Rev Immunol* 2004; 4: 499–511.
23. Takeuchi O, Hemmi H and Akira S. Interferon response induced by Toll-like receptor signaling. *J Endotoxin Res* 2004; 10: 252–256.
24. Wlasiuk G and Nachman MW. Adaptation and constraint at toll-like receptors in primates. *Mol Biol Evol* 2010; 27: 2172–2186.
25. Areal H, Abrantes J and Esteves PJ. Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. *BMC Evol Biol* 2011; 11: 368.
26. de Matos AL, McFadden G and Esteves PJ. Evolution of viral sensing RIG-I-like receptor genes in Leporidae genera *Oryctolagus*, *Sylvilagus*, and *Lepus*. *Immunogenetics* 2014; 66: 43–52.
27. de Matos AL, McFadden G and Esteves PJ. Positive evolutionary selection on the RIG-I-like receptor genes in mammals. *Plos One* 2013; 8: e81864.
28. Cagliani R, Forni D, Tresoldi C, et al. RIG-I-like receptors evolved adaptively in mammals, with parallel evolution at *LGP2* and *RIG-I*. *J Mol Biol* 2014; 426: 1351–1365.
29. Neves F, Abrantes J, Steinke JW, et al. Maximum-likelihood approaches reveal signatures of positive selection in IL genes in mammals. *Innate Immun* 2014; 20: 184–191.



30. Sabeti PC, Reich DE, Higgins JM, et al. Detecting recent positive selection in the human genome from haplotype structure. *Nature* 2002; 419: 832–837.
31. Tishkoff SA, Varkonyi R, Cahinhinan N, et al. Haplotype diversity and linkage disequilibrium at human G6PD: Recent origin of alleles that confer malarial resistance. *Science* 2001; 293: 455–462.
32. Aitman TJ, Cooper LD, Norsworthy PJ, et al. Malaria susceptibility and CD36 mutation. *Nature* 2000; 405: 1015–1016.
33. Greenwalt DE, Lipsky RH, Ockenhouse CF, et al. Membrane glycoprotein CD36: a review of its roles in adherence, signal transduction, and transfusion medicine. *Blood* 1992; 80: 1105–1115.
34. Oliveira TYK, Harris EE, Meyer D, et al. Molecular evolution of a malaria resistance gene (DARC) in primates. *Immunogenetics* 2012; 64: 497–505.
35. Demogines A, Truong KA and Sawyer SL. Species-specific features of DARC, the primate receptor for *Plasmodium vivax* and *Plasmodium knowlesi*. *Mol Biol Evol* 2012; 29: 445–449.
36. Migot-Nabias F, Mombo LE, Luty AJF, et al. Human genetic factors related to susceptibility to mild malaria in Gabon. *Genes Immun* 2000; 1: 435–441.
37. Sotgiu S, Sannella AR, Conti B, et al. Multiple sclerosis and anti-*Plasmodium falciparum* innate immune response. *J Neuroimmunol* 2007; 185: 201–207.
38. Metzger KJ and Thomas MA. Evidence of positive selection at codon sites localized in extracellular domains of mammalian CC motif chemokine receptor proteins. *BMC Evol Biol* 2010; 10: 139.
39. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 1999; 41: 95–98.
40. Hersch-Green EI, Myburg H and Johnson MTJ. Adaptive molecular evolution of a defence gene in sexual but not functionally asexual evening primroses. *J Evol Biol* 2012; 25: 1576–1586.
41. Heled J and Drummond AJ. Bayesian inference of species trees from multilocus data. *Mol Biol Evol* 2010; 27: 570–580.
42. Gernhard T. The conditioned reconstructed process. *J Theor Biol* 2008; 253: 769–778.
43. Drummond AJ and Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 2007; 7: 214.
44. Tamura K, Peterson D, Peterson N, et al. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011; 28: 2731–2739.
45. Rambaut A and Drummond AJ. Tracer ver. 1.5, <http://beast.bio.ed.ac.uk/Tracer> (2009, accessed 18 August 2014).
46. Rambaut A. FigTree ver. 1.3.1, <http://tree.bio.ed.ac.uk/software/FigTree/> (2008, accessed 12 August 2014)
47. O'Leary MA, Bloch JI, Flynn JJ, et al. The placental mammal ancestor and the post-K-Pg radiation of placentals. *Science* 2013; 339: 662–667.
48. Schierup MH and Hein J. Consequences of recombination on traditional phylogenetic analysis. *Genetics* 2000; 156: 879–891.
49. Anisimova M, Nielsen R and Yang ZH. Effect of recombination on the accuracy of the likelihood method for detecting positive selection at amino acid sites. *Genetics* 2003; 164: 1229–1236.
50. Sawyer S. Statistical tests for detecting gene conversion. *Mol Biol Evol* 1989; 6: 526–538.
51. Yang ZH. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 2007; 24: 1586–1591.
52. Nielsen R and Yang ZH. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* 1998; 148: 929–936.
53. Anisimova M, Bielawski JP and Yang ZH. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. *Mol Biol Evol* 2001; 18: 1585–1592.
54. Yang ZH, Wong WSW and Nielsen R. Bayes empirical Bayes inference of amino acid sites under positive selection. *Mol Biol Evol* 2005; 22: 1107–1118.
55. Zhang JZ, Nielsen R and Yang ZH. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol* 2005; 22: 2472–2479.
56. Gu X, Zou YY, Su ZX, et al. An update of DIVERGE software for functional divergence analysis of protein family. *Mol Biol Evol* 2013; 30: 1713–1719.
57. Ashkenazy H, Penn O, Doron-Faigenboim A, et al. FastML: a web server for probabilistic reconstruction of ancestral sequences. *Nucleic Acids Res* 2012; 40: W580–W584.
58. Escalante AA and Ayala FJ. Phylogeny of the malarial genus *Plasmodium*, derived from ribosomal RNA gene sequences. *Proc Natl Acad Sci U S A* 1994; 91: 11373–11377.
59. Prugnolle F, Durand P, Neel C, et al. African great apes are natural hosts of multiple related malaria species, including *Plasmodium falciparum*. *Proc Natl Acad Sci U S A* 2010; 107: 1458–1463.
60. Urwin R, Holmes EC, Fox AJ, et al. Phylogenetic evidence for frequent positive selection and recombination in the meningococcal surface antigen PorB. *Mol Biol Evol* 2002; 19: 1686–1694.
61. Degnan JH and Rosenberg NA. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol Evol* 2009; 24: 332–340.
62. Sugawara T, Terai Y and Okada N. Natural selection of the rhodopsin gene during the adaptive radiation of East African Great Lakes cichlid fishes. *Mol Biol Evol* 2002; 19: 1807–1811.
63. Spady TC, Seehausen O, Loew ER, et al. Adaptive molecular evolution in the opsin genes of rapidly speciating cichlid species. *Mol Biol Evol* 2005; 22: 1412–1422.
64. Yang Z, Nielsen R, Goldman N, et al. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 2000; 155: 431–449.
65. Yang Z. *User Guide: PAML: Phylogenetic analysis by maximum likelihood version 4.6*. London: University College London, 2012.
66. Dheilly NM, Adema C, Raftos DA, et al. No more non-model species: The promise of next generation sequencing for comparative immunology. *Dev Comp Immunol* 2014; 45(1): 56–66.
67. Trowsdale J and Parham P. Defense strategies and immunity-related genes. *Eur J Immunol* 2004; 34: 7–17.
68. Navratilova Z. Polymorphisms in CCL2 and CCL5 chemokines/chemokine receptors genes and their association with diseases. *Biomed Pap Med Univ Palacky Olomouc Czech Repub* 2006; 150: 191–204.
69. Rahman MN, Vlahakis JZ, Szarek WA, et al. X-ray crystal structure of human heme oxygenase-1 in complex with 1-(adamantan-1-yl)-2-(1H-imidazol-1-yl)ethanone: A common binding mode for imidazole-based heme oxygenase-1 inhibitors. *J Med Chem* 2008; 51: 5943–5952.
70. Poss KD and Toneyawa S. Heme oxygenase 1 is required for mammalian iron reutilization. *Proc Natl Acad Sci U S A* 1997; 94: 10919–10924.
71. Wilks A. Heme oxygenase: evolution, structure, and mechanism. *Antioxid Redox Sign* 2002; 4: 603–614.
72. Arredondo M, Jorquera D, Carrasco E, et al. Microsatellite polymorphism in the heme oxygenase-1 gene promoter is associated with iron status in persons with type 2 diabetes mellitus. *Am J Clin Nutr* 2007; 86: 1347–1353.
73. Okamoto I, Krogler J, Endler G, et al. A microsatellite polymorphism in the heme oxygenase-1 gene promoter is associated with risk for melanoma. *Int J Cancer* 2006; 119: 1312–1315.
74. Yamada N, Yamaya M, Okinaga S, et al. Microsatellite polymorphism in the heme oxygenase-1 gene promoter is associated with susceptibility to emphysema. *Am J Hum Genet* 2000; 66: 187–195.
75. Datta PK, Dhupar S and Lianos EA. Regulatory effects of inducible nitric oxide synthase on cyclooxygenase-2 and heme oxygenase-1 expression in experimental glomerulonephritis. *Nephrol Dial Transpl* 2006; 21: 51–57.



76. Sumi D and Ignarro LJ. Regulation of inducible nitric oxide synthase expression in advanced glycation end product-stimulated RAW 264.7 cells – The role of heme oxygenase-1 and endogenous nitric oxide. *Diabetes* 2004; 53: 1841–1850.
77. Cheng PY, Chen JJ and Yen MH. The expression of heme oxygenase-1 and inducible nitric oxide synthase in aorta during the development of hypertension in spontaneously hypertensive rats. *Am J Hypertens* 2004; 17: 1127–1134.
78. Ockenhouse CF, Tandon NN, Magowan C, et al. Identification of a platelet membrane glycoprotein as a falciparum malaria sequestration receptor. *Science* 1989; 243: 1469–1471.
79. Oquendo P, Hundt E, Lawler J, et al. CD36 directly mediates cytoadherence of *Plasmodium falciparum* parasitized erythrocytes. *Cell* 1989; 58: 95–101.
80. Cortes A, Mellombo M, Mgone CS, et al. Adhesion of *Plasmodium falciparum*-infected red blood cells to CD36 under flow is enhanced by the cerebral malaria-protective trait South-East Asian ovalocytosis. *Mol Biochem Parasit* 2005; 142: 252–257.
81. Gomez F, Ko W-Y, Davis A, et al. Impact of natural selection due to malarial disease on human genetic variation. In: Brinkworth JF and Pechenkina K (eds) *Primates, pathogens, and evolution*. New York: Springer, 2013, pp.117–160.
82. Ozbek S, Balasubramanian PG, Chiquet-Ehrismann R, et al. The evolution of extracellular matrix. *Mol Biol Cell* 2010; 21: 4300–4305.
83. Garcia-Monzon C, Lo Iacono O, Crespo J, et al. Increased soluble CD36 is linked to advanced steatosis in nonalcoholic fatty liver disease. *Eur J Clin Invest* 2014; 44: 65–73.
84. Daviet L and McGregor JL. Vascular biology of CD36: Roles of this new adhesion molecule family in different disease states. *Thromb Haemostasis* 1997; 78: 65–69.
85. Coraci I, Husemann J, Berman J, et al. CD36, a class B scavenger receptor, is expressed on microglia in Alzheimer's disease brains and mediates production of reactive oxygen species in response to beta-amyloid fibrils. *Mol Biol Cell* 2000; 11: 414a–414a.
86. Aitman TJ. CD36, insulin resistance, and coronary heart disease. *Lancet* 2001; 357: 651–652.
87. Susztak K, Ciccone E, Wu D, et al. CD36 is involved in tubular epithelial apoptosis during the progression of diabetic kidney disease. *J Am Soc Nephrol* 2003; 14: 94a–94a.
88. Febbraio M and Silverstein RL. CD36: Implications in cardiovascular disease. *Int J Biochem Cell B* 2007; 39: 2012–2030.
89. de las, Fuentes L, Gu CC, Waggoner AD, et al. The effect of interactions among variants of CD36 on hypertensive heart disease phenotypes. *Circulation* 2008; 118: S1100–S1101.
90. He JH, Lee JH, Febbraio M, et al. The emerging roles of fatty acid translocase/CD36 and the aryl hydrocarbon receptor in fatty liver disease. *Exp Biol Med* 2011; 236: 1116–1121.