

# Exceptional Diversity, Maintenance of Polymorphism, and Recent Directional Selection on the *APL1* Malaria Resistance Genes of *Anopheles gambiae*

Susan M. Rottschaefer<sup>1</sup>, Michelle M. Riehle<sup>2</sup>, Boubacar Coulibaly<sup>3</sup>, Madjou Sacko<sup>3</sup>, Oumou Niaré<sup>3</sup>, Isabelle Morlais<sup>4</sup>, Sekou F. Traoré<sup>3</sup>, Kenneth D. Vernick<sup>5</sup>, Brian P. Lazzaro<sup>1\*</sup>

**1** Department of Entomology, Cornell University, Ithaca, New York, United States of America, **2** Department of Microbiology, University of Minnesota, Saint Paul, Minnesota, United States of America, **3** Malaria Research and Training Center, University of Bamako, Bamako, Mali, **4** Laboratoire de Recherche sur le Paludisme, Institut de Recherche pour le Développement IRD-OCEAC, Yaoundé, Cameroun, **5** Unit of Insect Vector Genetics and Genomics, Institut Pasteur, Paris, France

## Abstract

The three-gene *APL1* locus encodes essential components of the mosquito immune defense against malaria parasites. *APL1* was originally identified because it lies within a mapped QTL conferring the vector mosquito *Anopheles gambiae* natural resistance to the human malaria parasite, *Plasmodium falciparum*, and *APL1* genes have subsequently been shown to be involved in defense against several species of *Plasmodium*. Here, we examine molecular population genetic variation at the *APL1* gene cluster in spatially and temporally diverse West African collections of *A. gambiae*. The locus is extremely polymorphic, showing evidence of adaptive evolutionary maintenance of genetic variation. We hypothesize that this variability aids in defense against genetically diverse pathogens, including *Plasmodium*. Variation at *APL1* is highly structured across geographic and temporal subpopulations. In particular, diversity is exceptionally high during the rainy season, when malaria transmission rates are at their peak. Much less allelic diversity is observed during the dry season when mosquito population sizes and malaria transmission rates are low. *APL1* diversity is weakly stratified by the polymorphic 2La chromosomal inversion but is very strongly subdivided between the M and S “molecular forms.” We find evidence that a recent selective sweep has occurred at the *APL1* locus in M form mosquitoes only. The independently reported observation of a similar M-form restricted sweep at the *Tep1* locus, whose product physically interacts with APL1C, suggests that epistatic selection may act on these two loci causing them to sweep coordinately.

**Citation:** Rottschaefer SM, Riehle MM, Coulibaly B, Sacko M, Niaré O, et al. (2011) Exceptional Diversity, Maintenance of Polymorphism, and Recent Directional Selection on the *APL1* Malaria Resistance Genes of *Anopheles gambiae*. PLoS Biol 9(3): e1000600. doi:10.1371/journal.pbio.1000600

**Academic Editor:** David S. Schneider, Stanford University, United States of America

**Received:** August 23, 2010; **Accepted:** January 27, 2011; **Published:** March 8, 2011

**Copyright:** © 2011 Rottschaefer et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by National Institutes of Health grant R01 AI062995. The funding agency had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

**Abbreviations:** LRR, leucine-rich repeat; QTL, quantitative trait locus

\* E-mail: bplazzaro@cornell.edu

## Introduction

A *gambiae* 250 kb region on chromosome 2L encodes the *APL1* locus, a cluster of genes that confer natural resistance to malaria parasites [1]. The *APL1* locus is highly polymorphic, with over 100 alleles identified in natural populations [2]. High diversity is observed during the rainy season, when malaria transmission rates are at their peak [3,4]. The *APL1* locus is highly structured across geographic and temporal subpopulations. In particular, diversity is exceptionally high during the rainy season, when malaria transmission rates are at their peak. Much less allelic diversity is observed during the dry season when mosquito population sizes and malaria transmission rates are low. *APL1* diversity is weakly stratified by the polymorphic 2La chromosomal inversion but is very strongly subdivided between the M and S “molecular forms.” We find evidence that a recent selective sweep has occurred at the *APL1* locus in M form mosquitoes only. The independently reported observation of a similar M-form restricted sweep at the *Tep1* locus, whose product physically interacts with APL1C, suggests that epistatic selection may act on these two loci causing them to sweep coordinately.

(QTL) on chromosome 2L of *P. falciparum* [3–6]. The *APL1* locus is highly polymorphic, with over 100 alleles identified in natural populations [2]. High diversity is observed during the rainy season, when malaria transmission rates are at their peak [3,4]. The *APL1* locus is highly structured across geographic and temporal subpopulations. In particular, diversity is exceptionally high during the rainy season, when malaria transmission rates are at their peak. Much less allelic diversity is observed during the dry season when mosquito population sizes and malaria transmission rates are low. *APL1* diversity is weakly stratified by the polymorphic 2La chromosomal inversion but is very strongly subdivided between the M and S “molecular forms.” We find evidence that a recent selective sweep has occurred at the *APL1* locus in M form mosquitoes only. The independently reported observation of a similar M-form restricted sweep at the *Tep1* locus, whose product physically interacts with APL1C, suggests that epistatic selection may act on these two loci causing them to sweep coordinately.

**Author Summary**

Immune defense genes are sometimes highly variable in host populations, reflecting selective pressure to combat diverse pathogens. In other instances, where there are only a few dominant pathogens, natural selection may favor only one or a few defense alleles. Here, we show that both adaptive strategies can occur in the same genes under different circumstances. We examined diversity in the *APL1* genes of the human malaria vector mosquito *Anopheles gambiae*, which play a role in defense against malaria parasites. We found that the *APL1* genes are exceptionally polymorphic, being 10-fold more diverse than typical *A. gambiae* genes. The distribution of *APL1* allelic diversity, however, is strongly structured depending on whether the genes are carried by the M or S “molecular forms” of the vector, which are thought to constitute newly forming species. We show that despite the evolutionary maintenance of *APL1* diversity in the S form of *A. gambiae*, there is evidence of strong recent directional selection on *APL1* genes in the M form. Independent research has shown that *Tep1*, a gene which encodes a protein that physically interacts with the APL1C protein, also harbors high allelic diversity in the S form and shows evidence of recent directional selection in the M form, suggesting that the evolutionary trajectories of the *Tep1* and *APL1* defense loci may be correlated.

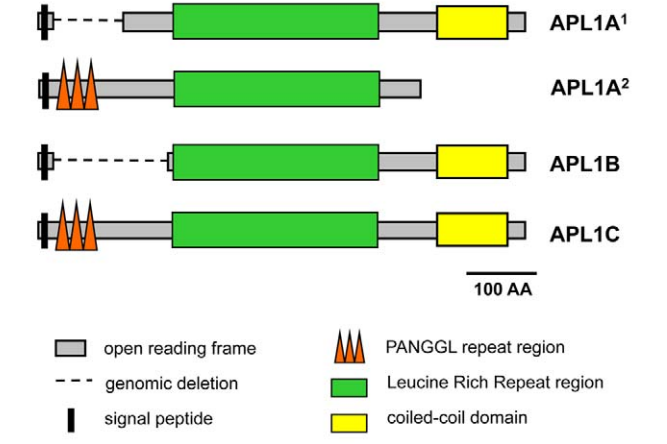
... d c ... a ... c a ... A. ca  
*A. g.* ... a l a ... a c d ... a *AP* ... a r d a c .  
 W ... c d ... *APA* , *APB* , a d *AEC* ... d A .  
*g.* ... c d ... b-Sa a a A. ca:  
 Ba ... a a, Ma ; T ... a -O ... a, Ma ; a d Ma ... c ...  
 Ca ... a . Ba ... a a ... a a d a a a a ...  
 ca ... a c , Ba ... T ... a -O ... a ... d  
 ... d ... a d Ma ... c ... a l d a ... a  
 ... a ... R ... b ... Sa ... d a ... a ...  
 ... ca ... d ... a ... a ... a a a a ...  
 ... a ... a d d ... a ... d ... d ... a ...  
 Ba ... a a ... a ... W ... d ... d ... c ... a ...  
 ... a ... c d ... a ... a ... a ... a ...  
 ... a ... a ... b ... d d ... a ... a ... W ... d *AB* ... c  
 ... a ... a ... b ... l ... c ... d ... a ... ca ... d a d b M/S  
 “ ... c a a ... ” ( ... d ... [25]) a d ... a ... d ... b  
 ... a ... c ... a ... a ... 2La. T ... l ... a  
 ... a ... *AP* ... l ... ca ... d ... a ... a ...  
 ... a ... *A. g.* ... a d ... a ...  
 ... a d a ... a ... a ... c ... a ... S ...  
 ... A ... a ... a ... c ... a d ... c ... a  
 ... d ... c d d ... a ... *AB* ... c ... M ...

**Results**

**Structure of the *APL1* Genes and Encoded Proteins**

*APA* , *APB* , a d *AEC* a ... a c ... d ... a ... a 5'  
 ... a d ... d ... d ... a a d b a ... [7].  
 S ... a c ... d ... d ... a ... [7] a d F ... l .  
 E a c ... c a a c ... d b a N ... a ... a ... d , a  
 ... c ... c ... a (LRR) ... a ... a ... a  
 300 a ... a c d ... d d ... d d a ... d ... d ...  
 d ... a ... a ... C ... *APA* <sup>2</sup> a ... d ... a ... a ...  
 ... d ... a ... a ... d ... a ... LRR  
 d ... a ... a ... C ... a ... d ... d ... c d c d  
 a ... d ... W ... b ... d 5 *APA* <sup>1</sup> a ... ( ... 38 ... a  
 a ... d ) ... c ... a ... a ... d ... a b ... l ... a d  
 ... ATA ; ... c ... a ... a ... l ... a ... a ... a ...  
 ATG ... a ... a ... a ... *AEC* a ... d ... a ... a ...  
 ... a d ... a ... a ... a c d P-A-N-G-G-L a d ... a d

... a ... a a a *Ap* ... a ... a c ...  
 LRIM1 [10] ... a ... a c ... a ... a d ... a b ...  
 ... TEPI , a d ... *P. g.* ... a ... a d ... [11,12].  
 T “G3” a b ... a ... *A. g.* ... d ...  
 a ... a l ... a ... a *APA* , *APB* , a d *AEC* (d ... d  
 ... c ... l a d 2 ; ... [7]). M ... G3 ...  
 ... a ... a ... *APA* <sup>2</sup> / *APB* <sup>2</sup> / *AEC* <sup>2</sup> ... a ...  
 ... a ... d ... a ... c ... *P. g.* ... c ... [7] , ...  
 ... a ... a ... a ... a *AE* ... b ... a ... a ... c ...  
 a a a a d .  
 T ... d a ... l ... a ... c ... d ... d ... d ... d ...  
 ... d ... d ... b ... d ... d ... *A. g.* ... l ... d ... d ... a ...  
 ... d ... d ... d ... c ... e - a d a ... - a ... l ... a ... a  
 d a c [13 21] , a ... l ... d ... d ... a ... a ... b ...  
 ... d ... d ... d ... d ... a ... a ... a ... c ... l ... c ...  
 a ... a d b ... l ... l ... a ... c ... l ... c ...  
*Ap* ... , ... a a ... *A. g.* ... a ... c ... a d  
 ... e ... a a ... a d ... d ... a ... l ... b ... l ... a ...  
 ... a ... a ... a d ... d ... a ... l ... b ... l ... a ...  
 ... a ... a ... a ... [14]. D ... a ... a ... c ...  
 ... l ... a ... l ... a d *BM* ... , a d ... c ... ca  
 ... a c ... a ... a ... APL1C , a b ... a ... d ... d ... a ...  
 ... l ... c ... a ... a ... a ... a ... a c d a d ... c ... d  
 ... , ... b ... d ... a ... c ... c a ... l ...  
 ... a ... l ... e ... [18]. D ... a ... a ... l ...  
 ... a ... b ... b ... e ... a ... a ... a c ... a d  
 ... l ... c ... b ... c ... b ... *P. g.* ... a d *P. g.* ... [22 24].  
 I ... e ... a ... , ... a ... a ... *BM* ... ca ... *A.*  
*g.* ... [14,16] , a ... l ... *BM* ... a ... d ... c ...  
 a d a ... d ... d ... a ... l ... *A. g.* ... c ... *A.*  
 ... l ... l ... a ... l ... a ... d a ... b ... l ...  
 ... TEPI-  
 LRIM1-APL1C ... l ... d ... d ... c ...  
 ... l ... c ... e ... , d ... c ... a ... a d a ... a ... e ... b ... a ...  
 ... C ...  
 Ma ... l ... c ... a ... a ... a ... *AP* ... l ... a ... b ...  
 ... d [7] , b ... l ... a ... l ... a ... c ... a ... a ... *AP* ...  
 ... d ... l ... a ... b ... d ... d ... l ... d ...



**Figure 1. Schematic representation of proteins encoded by *APL1A*, *APL1B*, and *APL1C* genes.** The two major structural variants of *APL1A* are shown separately. *APL1A*<sup>1</sup> alleles are characterized by the deletion of the PANGGL region. *APL1A*<sup>2</sup> alleles carry an early stop codon that eliminates the coiled-coil domain.  
 doi:10.1371/journal.pbio.1000600.g001

... (a, c, d, a) PANGGL... T...  
*APB*... a a PANGGL... I...  
 PANGGL... *APA*<sup>1</sup>... a... b... ab...  
*APA*<sup>1</sup> [7]. I... c... d... a...  
*APA* a... c... *A. g.* (*A. h.*, *A.*  
*#*, and *A. m.*) a a PANGGL... a d...  
 a *APA*<sup>1</sup> a a a *APA*<sup>2</sup> a a b...  
 c... a... *A. g.* T... d... a...  
 a PANGGL... *APB* a... *AEC* 207 b...  
 a a d a... b a... /  
 d a d... *APA*<sup>1</sup> *APA*<sup>2</sup> T...  
 a b... a... d... a...  
 a... PANGGL... *AP* T...  
 a... c b... PANGGL... a a d a...  
*AEC* a d *APA*<sup>2</sup> a... a... a...  
 ab... *APA*<sup>2</sup> a... *A. m.*, *A. h.*, and *A.*  
*#* (F... S1), a PANGGL... a a...  
 a b... d... *APA* a a a...  
 e... *AEC* *A. g.* E... a d... a...  
 b... *AEC* a d *APA*<sup>2</sup> a... a... c d...  
 b... Pa... e... a... a... b...  
 d... a... d... a...  
*A. g.* [18]. N... a b... d...  
 PANGGL... a... b... c... a PANGGL...  
 c... *APB* a d *APA*<sup>1</sup> a... a... PANGGL...  
*AEC* a d *APA*<sup>2</sup> a... a... /ab... c...  
 PANGGL d a... a... *AP* a... a d... a...  
 T... a... a... a... a...

**APL1 Genes Are Exceptionally Polymorphic**

S... (a... c d a...)  
 a d a... *APA*, *APB*, and *AEC*,  
 c... d... a... d... ( $\pi_{\text{tot}}$ )...  
 5.9%, 3.1%, and 2.4%, c... (Tab 1). T... a... a...  
 a... a... 10... d... a... a... ca... b... d...  
*A. g.* c... d... ([14, 17,  
 19–21], b... [18]). T... a...  
 ac... *AP*... a... e...  
 a... a... d... b... d... c... b... c...  
 c... b... c... R... a... e... a...  
 b... d... ac... *AP*...  
 ca... b... a... d... b... c... e...  
 G... a... [17], a... d... a...  
 109... d... b... d... a... d... *A. g.* c... d... 72  
 b... b... d... c... T... da... ca...  
 b... a... a... "d... b... c... *AP*...  
 c... ca... b... e... a... d... A... *AP*... b... a...  
 c... b... d... a... a... d... da... b... G...  
 a... [17]... c... a... a... a...  $\pi_{\text{tot}}$ ... 0.3% and a...  
 a... 2.1%. W... e... a... d... d...  
 da... a... a... *AP*... a... a...  
 ca... c... a... a... c... d... *A. g.* a... d... a...  
 d... c... (a... a... b... *A. g.* a... d... a...  
*#* ( $\chi^2_{(1)} = 5.79$ ;  $p = 0.016$ , a... a 2×2  
 e... c... ab... a... d... b... e...  
 and *A. h.* a... d... a... b... *A. g.*  
 and *A. h.* ac... *AP*... a... d...  
 a... T... *AP*... a... a... ca... c...  
 a... c... d... c... a...  
 ( $\chi^2_{(1)} = 7.54$ ;  $p = 0.006$ ). T... a... b... d... a... *AP*...  
 a... a... b... d... d... c... a... a...  
 d... c... acc... a... b... c... a... d... a... d... e... a...

**Table 1.** Population genetic parameter estimates at the *APL1* locus in four collections.

Collection	n <sup>a</sup>	bp <sup>b</sup>	$\pi_{\text{tot}}$ <sup>c</sup>	$\theta_{\text{tot}}$ <sup>d</sup>	TajD <sup>e</sup>	$\pi_{\text{syn}}$ <sup>f</sup>	$\pi_{\text{non}}$ <sup>g</sup>
<b>APL1A</b>							
Bancoumana dry	19	1,669	0.019	0.032	-1.734	0.033	0.016
Bancoumana rainy	9	1,665	0.048	0.054	-0.587	0.086	0.039
Toumani-Oulena	12	1,541	0.084	0.074	0.668	0.119	0.066
Makouchetoum	8	1,541	0.088	0.078	0.678	0.119	0.070
All pooled	48	1,537	0.075	0.065	0.541	0.114	0.059
<b>APL1B</b>							
Bancoumana dry	19	2,005	0.014	0.017	-0.685	0.017	0.015
Bancoumana rainy	12	2,077	0.030	0.032	-0.213	0.046	0.030
Toumani-Oulena	16	2,067	0.042	0.043	-0.079	0.074	0.036
Makouchetoum	12	1,968	0.039	0.047	-0.832	0.070	0.033
All pooled	59	1,902	0.034	0.046	-0.921	0.057	0.031
<b>APL1C</b>							
Bancoumana dry	15	2,569	0.006	0.009	-1.308	0.010	0.005
Bancoumana rainy	10	2,569	0.028	0.028	-0.055	0.059	0.021
Toumani-Oulena	16	2,410	0.027	0.027	-0.002	0.050	0.020
Makouchetoum	12	2,393	0.030	0.029	0.231	0.064	0.020
All pooled	53	2,393	0.031	0.025	0.556	0.061	0.024

The Bancoumana dry season collection is almost entirely M form mosquitoes, the Toumani-Oulena and Makouchetoum collections are almost entirely S form mosquitoes, and the Bancoumana rainy collection is a mixture of M and S form. These same parameter estimates are given separately for M form and S form mosquitoes in Table S1 and for *APL1A*<sup>1</sup> and *APL1A*<sup>2</sup> alleles in Table S2.

<sup>a</sup>Number of alleles sequenced.  
<sup>b</sup>Locus size, in base pairs, excluding insertions and deletions.  
<sup>c</sup>Average number of differences per pair of alleles, per nucleotide.  
<sup>d</sup>Watterson's estimator of the population genetic parameter 4N<sub>e</sub>μ.  
<sup>e</sup>Tajima's D test statistic.  
<sup>f</sup>Average number of difference per pair of alleles, per nucleotide, synonymous sites only.  
<sup>g</sup>Average number of difference per pair of alleles, per nucleotide, nonsynonymous sites only.  
 doi:10.1371/journal.pbio.1000600.t001

a... c... a... b... *A. g.*  
 and *A. h.*... c... a... b...  
 d... c... ada... d... c... b... c...  
 [17]. T... c... d... a... d...  
 b... d... a... b... a... d... a... d... *AP*...  
 a... ada... a... ac... a... a...  
 ada... [26], *AP* a... b... c... [27].  
 A... ad... a... McD... a... d... K... a... [28]. e... a...  
 a... a... ad... a... d... a...  
*A. g.* a... d... a... d... c... b...  
*A. g.* and *A. h.*... ca... d... a... [17]  
 (*AP* : P... = 342, P... = 478, F... = 4, F... = 11,  
 G = 1.45,  $p = 0.23$ ; a... : P... = 1967, P... = 731, F... = 86,  
 F... = 38, G = 0.74,  $p = 0.73$ ). T... McD... a... d...  
 K... a... d... b... a...  
 a... d... c... b... *A. g.* and *A. h.* T...  
 ac... a... "d... " *A. h.* a... *AP*... a...  
 a... ca... d... *A. g.* a... ad... a... a...  
 McD... a... d... K... a... [28] and a... a... da... T...  
 d... b... d... d... ca...

*An* . . . d . a . d . d . *A.g* . . . a . *A* . . .  
 . . . ca . . . b . . . 4% 11% d . . . c . b . . .  
 . . . c . a . . . d . . . ([14]). W . a . . . d . a . . .  
 HKA . . . a . a . . . d . a . . . [29] . . .  
 . . . a . *AP* . . . a . a . d . . . a . a . c . . .  
 . . . a . a . . . 50 . . . a . d . a . d . . . d . . .  
 . . . c . b . . . d . a . d . c . b . . . *A.g* . . .  
 a . d . d . . . c . b . . . *A.g* . . . a . d . *A.m* . . . a . a . a . b . . .  
 [14,16,19,21,30,31]. *A* . . . l . i . a . . . d . a . . . d . . .  
 . . . *AP* . . . b . . . b . . . a . d . a . . . a . . . a . r . c . . .  
 . . . . . . . . . ca . da . a . . . . . . . . . ca . . . b . . . . . . . . .  
 . . . d . . . a . a . . . d . . . a . . . a . . . a . . .  
 ( $\chi^2_{(3)} = 32.8, p = 3.63 \times 10^{-7}$ ), *AP* . . . l . i . a . . . d . . .  
 . . . b . 12 . . . d . 35 . . . d . a . . . d . . . a . . . l . d . b . . . c . d . . .  
 . . . d . . . a . . . T . . . a . . . a . b . . . . . . . . . a . d . b . . .  
 . . . d . . . d . . . c . . . . . . . . . d . . . c . d . b . . .  
 . . . a . a . . . l . . . . . . . . . *AP* . . . b . . . . . . . . .  
 . . . c . a . b . . . a . . . c . . . a . . . a . . . c . d . . . a . d . . .  
 . . . c . c . d . . . c . . . *AP* . . . b . . . a . d . s . . . . . . . . .  
 a . d . a . . . a . . . a . r . c . . . . . . . . .

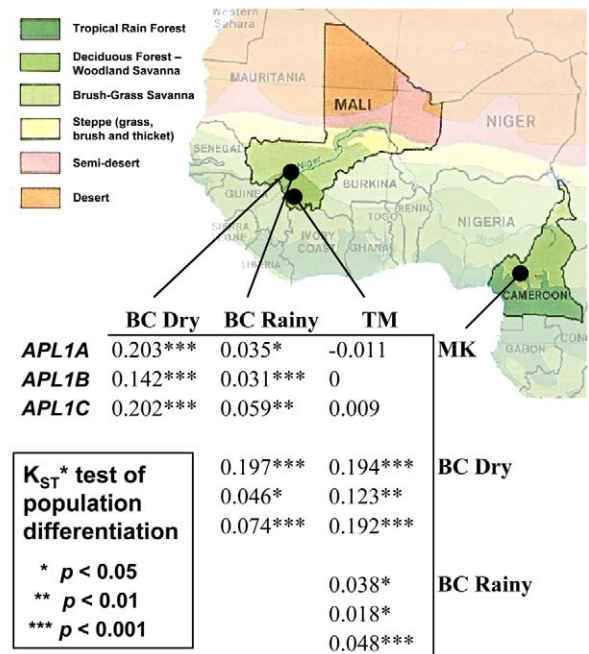
**APL1 Diversity Is Not Due to Degradation or Pseudogenization**

T . . . d . . . b . . . d . . . *AP* . . . a . . . a . . .  
 . . . e . . . a . . . c . . . a . . . l . . . a . . . a . . . d . . . a . . . c . . .  
 . . . l . . . a . . . T . . . l . . . a . . . d . a . . . . . . . . . d . . . . . . . . .  
 . . . l . . . *AP* . . . . . . . . . a . . . c . . . l . . . a . . . l . . . a . . .  
 . . . b . . . l . . . a . . . a . . . a . . . a . . . a . . . l . . . d . . . b . . .  
 . . . c . . . d . . . a . . . a . . . c . . . c . . . d . . . c . . . a . . . . . . . . .  
 . . . a . . . b . . . a . . . d . . . . . . . . . a . . . *AP* . . . . . . . . .  
 . . . a . . . b . . . a . . . d . . . *A.g* . . . a . . . a . . . (F . . . S2), e . . .  
 . . . e . . . l . . . d . . . a . . . a . . . a . . . a . . . d . . . a . . . c . . . . . . . . .  
 . . . a . . . a . . . a . . . a . . . a . . . a . . . a . . . a . . . a . . . *AP* . . .  
 . . . d . . . e . . . a . . . d . . . l . . . . . . . . . a . . . a . . . l . . . a . . .  
 . . . c . . . e . . . a . . . b . . . . . . . . . l . . . . . . . . .  
 . . . c . . . a . . . b . . . . . . . . . a . . . b . . . a . . . d . . . a . . . c . . . l . . .  
 . . . l . . . a . . . a . . . b . . . c . . . a . . . a . . . a . . . a . . . a . . . 38  
 . . . l . . . a . . . a . . . b . . . . . . . . . (d . . . ) . . . a . . . l . . .  
*AP* . . . . . . . . . a . . . l . . . a . . . e . . . a . . . a . . . a . . . 11 . . . d . . .  
 . . . PANGGL . . . (F . . . S1). O . . . 3 . . . . . . . . . 38 . . . d . . .  
 . . . d . . . l . . . a . . . d . . . a . . . . . . . . . b . . . b . . . 13 . . . c . . . d . . . b . . . c . . . a . . . c . . . e . . . d . . .  
 . . . a . . . c . . . a . . . a . . . b . . . d . . . d . . . a . . . a . . . d . . . d . . . d . . . a . . .  
 . . . a . . . S . . . a . . . . . . . . . b . . . d . . . 341 . . . c . . . d . . . . . . . . .  
 . . . i . . . a . . . l . . . *AP* . . . . . . . . . b . . . . . . . . .  
 . . . d . . . a . . . a . . . *APA* . . . 2 . . . *APA* . . . a . . . . . . . . . a . . . l . . .  
 . . . l . . . a . . . d . . . c . . . c . . . a . . . *APL1A* . . . . . . . . .). O . . .  
 . . . . . . . . . a . . . l . . . c . . . a . . . a . . . a . . . c . . . d . . . b . . . a . . . C . . .  
 . . . *APL1B*, a . . . d . . . a . . . a . . . a . . . l . . . a . . . . . . . . .  
 . . . a . . . O . . . a . . . c . . . . . . . . . a . . . a . . . a . . . c . . . a . . . l . . .  
 . . . e . . . d . . . a . . . . . . . . . l . . . b . . . . . . . . . l . . . c . . .  
 . . . c . . . a . . . a . . . b . . . a . . . l . . . a . . . a . . . c . . . b . . . a . . . c . . .  
 . . . a . . . a . . . c . . . l . . . a . . . b . . . d . . . a . . . e . . . . . . . . . c . . .  
 . . . c . . . d . . . . . . . . . d . . . d . . . c . . . d . . . . . . . . . ( . . .  
 [32–34]). *A* . . . . . . . . . a . . . 30% . . . e . . . d . . . *AP* . . . . . . . . .  
 . . . l . . . a . . . a . . . a . . . a . . . b . . . e . . . l . . . a . . . s . . . e . . . d . . .  
 . . . a . . . a . . . 1/9 . . . . . . . . . l . . . a . . . a . . . e . . . d . . . . . . . . .  
 . . . a . . . l . . . a . . . l . . . a . . . a . . . c . . . *AP* . . . a . . .  
 . . . a . . . c . . . a . . . d . . . d . . . d . . . c . . . . . . . . . a . . . l . . . a . . .  
 . . . l . . . c . . . . . . . . . a . . . a . . . 3.3% . . . . . . . . .  
 . . . b . . . d . . . *AP* . . . . . . . . . l . . . d . . . b . . . a . . . l . . . T . . .

. . . c . a . . . . . . . . . a . . . b . . . b . . . a . d . . . e . . . . . . . . .  
 . . . b . . . d . . . d . . . a . . . (1/110 . . . *APA* . . . , 2/114 . . . *APB* . . . , 0/117 . . .  
*APC* . . .). T . . . a . . . c . . . a . . . a . . . a . . . a . . . d . . . a . . . l . . .  
 . . . a . . . b . . . d . . . a . . . a . . . d . . . a . . . . . . . . . c . . . . . . . . .  
 2% . . . . . . . . . d . . . c . . . a . . . a . . . . . . . . . c . . . . . . . . .  
 . . . l . . . c . . . a . . . d . . . a . . . c . . . . . . . . . F . . . a . . . . . . . . . b . . .  
 . . . a . . . a . . . c . . . . . . . . . d . . . *AP* . . . M . . . . . . . . .  
 (d . . . c . . . d . . . b . . . ) . . . d . . . c . . . a . . . *AE* . . . . . . . . . l . . . b . . . c . . .  
 . . . . . . . . . a . . . a . . . d . . . a . . . . . . . . .

**Population Substructure at APL1**

G . . . c . . . l . . . . . . . . . c . . . d . . . . . . . . . a . . . *AP* . . . . . . . . . d . . . b . . . d . . .  
 . . . . . . . . . a . . . c . . . l . . . i . . . a . . . a . . . . . . . . . b . . . . . . . . . a . . . d . . . l . . . b . . . l . . . c . . . l . . . d . . .  
 . . . a . . . a . . . c . . . e . . . d . . . . . . . . . c . . . e . . . . . . . . . c . . . a . . . c . . . e . . . l . . . c . . . a . . . l . . . d . . .  
 . . . . . . . . . c . . . e . . . c . . . e . . . a . . . d . . . . . . . . . a . . . T . . . T . . . l . . . a . . . -O . . . u . . . l . . . e . . . n . . . a . . . a . . . d . . . M . . . a . . .  
 . . . c . . . l . . . i . . . a . . . a . . . d . . . M . . . a . . . l . . . i . . . a . . . 2005 . . . a . . . . . . . . . a . . . a . . . l . . . d . . . . . . . . . a . . . d . . . . . . . . .  
 . . . a . . . a . . . (p > 0.15; F . . . l . . . 2), b . . . l . . . b . . . a . . . a . . . d . . . d . . . . . . . . . a . . . d . . .  
 . . . 2005 . . . a . . . a . . . a . . . e . . . c . . . e . . . c . . . d . . . a . . . a . . . d . . . a . . . d . . . B . . . a . . . e . . . l . . . a . . . a . . .  
 . . . (p < 0.05 . . . a . . . . . . . . . ; F . . . l . . . 2). T . . . B . . . a . . . e . . . l . . . a . . . a . . . e . . . c . . .  
 . . . 2003 . . . d . . . a . . . . . . . . . l . . . a . . . . . . . . . c . . . a . . . d . . . . . . . . . a . . . d . . .  
 . . . a . . . a . . . a . . . e . . . c . . . e . . . c . . . a . . . a . . . . . . . . . (p ≤ 10<sup>-4</sup> . . . a . . .  
 . . . e . . . a . . . a . . . e . . . a . . . a . . . T . . . l . . . a . . . -O . . . u . . . l . . . e . . . n . . . a . . . a . . . d . . . M . . . a . . .  
 . . . c . . . l . . . i . . . a . . . a . . . d . . . M . . . a . . . l . . . i . . . a . . . p ≤ 1.9 × 10<sup>-2</sup> . . . e . . . a . . . d . . . B . . . a . . . e . . . l . . . a . . . a . . . a . . .  
 . . . e . . . c . . . a . . . a . . . (p > 0.15; F . . . l . . . 2). B . . . . . . . . . 2L . . . a . . . c . . . a . . . a . . . a . . . a . . . d . . .  
 . . . “M” . . . a . . . d . . . “S” . . . . . . . . . c . . . a . . . a . . . a . . . a . . . l . . . i . . . a . . . c . . . a . . .  
 . . . a . . . d . . . e . . . l . . . c . . . a . . . . . . . . . e . . . d . . . d . . . d . . . . . . . . . c . . . l . . .



**Figure 2. Population differentiation among *A. gambiae* collections at the *APL1* locus as estimated by  $K_{ST}^*$ .** Statistical significance determined by permutation of alleles among subpopulation pairs [48]. Mosquitoes sampled during the 2005 rainy season from the humid Toumani-Oulena (TM) and Makouchetoum (MK) regions are undifferentiated. These populations are mildly differentiated from a collection drawn from Bancoumana in the 2005 rainy season (BC rainy). All collections are highly differentiated from a collection drawn from Bancoumana in the 2003 dry season (BC dry). The analysis presented in this figure pools all mosquitoes by site and date of collection and does not take into account 2La karyotype or M/S molecular form. doi:10.1371/journal.pbio.1000600.g002

a... a... d... a... a... *AP*... b... a... b... ab...  
 d... c... c... c... 2La<sup>a</sup> M/S...  
 T... *AP*... c... a... a... 1 Mb... d... d... a...  
 b... a... c... c... c... a... 2La<sup>a</sup>... c...  
 a... a... b... b... b... a... c... ad... c...  
 c... ca... a... a... c... T... "d" (2La<sup>a</sup>)...  
 a... d... "a... da... d" (2La<sup>a</sup>)... d... a...  
 ca... [35,36]. W... d... a... a... *AP*...  
 a... e... d... b... a... ca... d... d... c... 2La<sup>a</sup>... a... ad...  
 a... d... c... c... c... a... a... 2La<sup>a</sup>... a...  
 ac... c... c... T... 2La<sup>a</sup>... a... a... d...  
 Ba... e... l... a... e... c... b... b... a... a... a...  
 T... a... -O... a... ad... Ma... t... c... (F... 3).  
 T... a... d... c... b... 2La<sup>a</sup> and 2La<sup>+</sup>  
 c... a... d... d... b... b... d... b... c... a... *AP*...  
 a... d... d... a... a... *AP*... a... a...  
 T... a... b... 2La<sup>a</sup>... c... c... c...  
 S... c... b... d... c... c... d... d... a...  
 c... d... da... a... *AP*... c... d... d... a... d... d...  
 a... ca... b... e... d... c... d... a... a...  
 c... d... da... T... 2La<sup>a</sup>... d... a... a...  
 d... da... a... e... d... Ba... e... a... ad... a...  
 e... c... (a... 2La<sup>a</sup>/2La<sup>a</sup>... a...),  
 e... a... c... d... a... a... a... c... a...  
 e... c... T... a... d... d... a... b... 2La<sup>a</sup>/2La<sup>a</sup>  
 and 2La<sup>+</sup>/2La<sup>+</sup>... a... a... *AP*... a... S...  
 (*APA*:  $K_s^* = 0.059$ ,  $p = 0.016$ ; *APB*:  $K_s^* = 0.014$ ,  
 $p = 0.094$ ; *APC*:  $K_s^* = 0.050$ ,  $p = 0.004$ ). I... c... a...  
 c... d... 2La<sup>a</sup>/2La<sup>a</sup>... d... a... a...  
 Ba... e... l... a... a... a... d... a... a...  
 a... *AP*... a... a... c... c... a...  
 e... a... c... 2La and "M" and "S" ... c... a...  
 (d... d... b...). N... a... d... c... a... a...  
*AP*... c... d... b... d... b... 2La<sup>a</sup>/2La<sup>a</sup> and  
 2La<sup>+</sup>/2La<sup>+</sup>... T... d... a... a... b...  
 a... a... c... a... d... a... d... a... d... b... ca...  
 a... a... c... d... da... b... e... b... a... a...  
 d... d... a... b... d... a... ca... d... b...  
 M/S... c... a...  
 T... DNA... d... "M" and "S" ... c... a...  
 a... a... b... a... ca... d... a... d... a... c...  
 ( [25]), and... a... c... M and S...  
 a... a... a... ab... *A. g.*... l... a...  
 [36,37]. T... a... a... M/S... a...

a... c... ca... *A. g.*... l... b... l... a...  
 a... d... b... a... d... c... ba... ( [25]).  
 W... d... a... M/S... e... b...  
 a... a... c... b... c... a... *AP*... I... d... d... 95%...  
 a... e... c... d... Ba... e... l... a... d... 2003 d... a...  
 a... M... a... M... a... ab... T... a...  
 O... a... ad... Ma... t... c... (F... 3),  
 a... l... a...  
 a... a... d... d... a... a... *AP*... b...  
 ac... a... d... b... d... c... a... b... M and S...  
 B... M and S... a... a... a... d... a...  
 c... c... Ba... e... l... a... d... a... a... (F... 3),  
 c... c... d... d... a... a... d... ac...  
 e... c... d... a... d... c... d... a...  
 l... b... l... a...  
 S... c... b... M and S... a... d... d...  
 2005 a... a... Ba... e... l... a... e... d... d... c...  
 a... d... d... a... b... M and S... c... a...  
 e... b... l... b... d... a... *AP*... A... c... d... d...  
 a... d... a... M...  
 Ba... e... l... a... 2005 a... a... e... c... a... l... d... a... d...  
 Ba... e... l... a... 2003 d... a... M... a... b...  
 a... d... d... a... d... 2005 a... a... S...  
 e... c... d... T... a... -O... a... ad... Ma... t... c...  
 (Tab... S3). R... c... ca... S... 2005  
 Ba... e... l... a... a... a... l... d... a... d... S...  
 T... a... -O... a... ad... Ma... t... c... b... a...  
 c... a... d... a... d... 2003 d... a... e... c...  
 c... a... M... (Tab... S3). T... a... a...  
 a... b... M and S... c... a...  
 c... c... c... b... a... *AP*... c... d... *APA*...  
*APB*... and *APC*... add... a... e... c... d... *A. g.*...  
 F... b... a... d... e... d... a... e... c... *A. g.*...  
 Ba... e... l... a... e... c... d... 2007. L... 2003 d... a...  
 e... c... 2007 d... a... a... a... M... a... da...  
 d... c... 2005 a... a...  
 e... c... T... 2007 d... a... a... a... ca...  
 d... l... ab... 2003 d... a...  
 a... d... a... a... ba... l... a... ( $p > 0.05$  a...  
 ; Tab... S3), b... a... c... d... a... d... a... d...  
 T... a... -O... a... ad... Ma... t... c... l... a...  
 ( $p \leq 0.005$  a... ; Tab... S3). I... a... e... d... e... a...  
 a... a... d... d... c... M... e... c... d... a...  
 Ba... e... l... a... d... a... a... 1997 and 1999. T...  
*AP*... a... a... a... M... a... a...  
 c... a... d... l... ab... M... 2003 and  
 2007 d... a... ( $p > 0.05$  a... ; Tab... S3) b...  
 a... a... d... a... d... S... T... a... -O... a... ad...  
 Ma... t... c... l... a... ( $p < 10^{-3}$  a... ; Tab... S3).

2La	Bancoumana dry		Bancoumana rainy		Toumani-Oulena		Makouchetoum		2La		
a/a	20	1	100%	5	24%	1	4	25%	3	21%	a/a
a/+				1	10%		9	45%	6	43%	a/+
+/+				1	67%		6	30%	5	36%	+/+
unk.							2				unk.
	95% 5%		33% 67%		5% 95%		100%				
	M	S	M	S	M	S	M	S			

**Figure 3. Number of observations of mosquitoes with each 2La inversion karyotype in each molecular form (M and S) over each sample collection.** Population frequencies (in percentages) are given in the margins of each table. "Unk." indicates that 2La karyotype was not determined.  
 doi:10.1371/journal.pbio.1000600.g003

W... a... d... a... c... a... d... a... d...  
 ... a... d... S... b... a... a...  
 ... ca... d... a... d... M... b... a... a...  
 ... ( $p < 10^{-4}$  a... ac...). W... c... d... a...  
 ... a... b... c... b... *AP*... a... d...  
 ... d... a... b... c... M... a... d... S... *A*... a... d...  
 ... a... c... a... d... a... a... b... c... d... c... b...  
 ... c... M... a... d... S... a... c...

**A Recent Selective Sweep in M Form *A. gambiae* at *APL1***

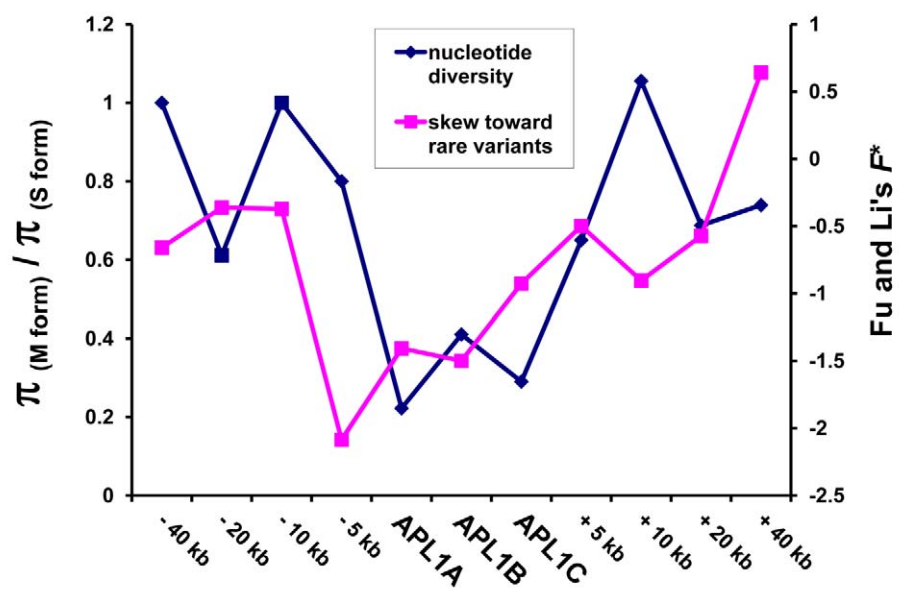
T... M... b... d... a... d... c... d...  
 a... *AP*... a... d... d... S... (Tab 1, Tab S1), a...  
 ... b... a... a... c... c... a... a... a... d...  
*AP*... c... a... a... M... a... a... T... ca... ca...  
 ... d... ca... a... c... c... c... c... d... a... d... c...  
 ... [38], a... d... a... c... c... c...  
 ... a... d... a... c... a... [39] a... c... a... b... a... d... a... a...  
 ... a... a... c... c... a... T... a... *D* (Tab 1; [40]), F... a... d...  
 L...  $F^*$  (F... 4; [41]) a... d... a... d... c... a... d... [42]. T...  
 M... a... a... b... a... a... c... a... c... c... a...  
*AP*... (Tab 1, F... 4, Tab 2, Tab S1).

L... *AP*... a... b... a... a... a... c... c... c...  
 M... c... a... a... a... d... c... a... a... a...  
 ... c... a... a... c... a... a... a... a... *AP*...  
 ... a... d... a... a... a... d... a... a... c... a...  
 ... (c... b... a... a... d... a... c... a... a... c... T...  
 ... c... d... c... c... a... a... a... 5, 10, 20, a... d... 40... b... a...  
 ... d... *AP*... c... b... M... a... d... S...  
 B... c... a... a... T... M... d... a... a...  
 ... d... d... d... a... d... d... S... a...  
*AP*... c... a... a... a... a... b... 5... 10... b...  
 ... d... *AP*... (F... 4). T... M... a... a... a...  
 ... a... a... c... d... a... d... a... a... a... d... a...  
 ... a... d... d... a... *AP*... a... a... b... d... S...  
 ... d... a... c... (F... 4, Tab 2). W... a... S...

... l... a... d... d... a... ca... l... c... a... *AP*...  
 ... d... a... d... a... a... a... a... b... c...  
 ... d... a... a... c... *AP*... M... a... d... S... d... a...  
 S2). T... c... b... c... b... M... a... d... S... d... a...  
 ... d... a... c... d... d... *AP*... c... a... M...  
 ... a... d... a... d... d... a... c... a... l... c... a... d... a... c...  
 ... S... (F... S3). B... c... a... a...  
 ... a... a... c... d... a... c... d... a... *AP*... c... a... b... d...  
 ... d... d... c... c... d... a... c... b... M... a... d... S...  
 ... G... c... d... a... d... d... d... c... a... *AP*...  
 ... a... c... b... a... a... d... c... a... c... c... c... a...  
 ... M... a... a...  
 I... a... a... a... d... c... a... a... a... d...  
 ... a... a... l... a... a... M... (a... d...  
 ... a... a... c... d... a... a... c... c... a... a... a... d... a...  
 ... a... b... a... a... T... c... c... d... a... a... a... c... a...  
 ... S... a... a... a... a... a... a... a...  
 ... d... a... M... (F... S2) a... d... a... M...  
 ... a... a... a... d... a... a... a... a... a... S...  
 ... (F... S2). O... a... a... a... d... a... a... d... b... a...  
 ... c... d... d... a... a... c... a... *AP*... a... b...  
 ... M... a... d... S... G... d... d... d... c... a... a...  
 ... a... a... d... c... d... c... b... d... b... M... a... d... S...  
 ... d... [25], ... a... a... a... a...  
 ... b... a... d... a... d... c... a... a... a... a...  
 ... a... a... a... l... d... c... M... a... a... a... *AP*...  
 I... c... a... a... a... c... a... b... d... a...  
 ... [43] (c... a... [18]). T... a... a... TEPI a... d... APLIC... c... a...  
 ... a... c... [11-12], a... a... a... b... a... a... c...  
 ... a... b... d... d... a... d... a... a... c...

**Discussion**

I... a... a... a... a... a... a...  
 ... a... E... a... d... d... l... a... d... c... a... a...



**Figure 4. Plot of nucleotide diversity and skew in the site frequency spectrum as a function of physical distance from the *APL1* locus.** M form mosquitoes exhibit a sharp drop in polymorphism at the *APL1* locus relative to S form mosquitoes, plotted as the ratio of nucleotide diversity ( $\pi$ ) in the M form to diversity in the S form. The M form mosquitoes also exhibit an enhanced skew toward rare variants, indicated by negative values of Fu and Li's  $F^*$  [41]. The data are consistent with a recent selective sweep at *APL1* in the M form only. doi:10.1371/journal.pbio.1000600.g004

**Table 2.** Genetic diversity and haplotype homozygosity in the M and S molecular forms at the *APL1* locus and flanking regions are indicative of selective maintenance of diversity at *APL1* in the S molecular form and a recent selective sweep at *APL1* within the M form.

Position	M Form					EW <sup>f</sup>	S Form					
	bp <sup>a</sup>	$\pi$ <sup>b</sup>	F <sup>*</sup> <sup>c</sup>	n <sup>d</sup>	# haps <sup>e</sup>		bp <sup>a</sup>	$\pi$ <sup>b</sup>	F <sup>*</sup> <sup>c</sup>	n <sup>d</sup>	# haps <sup>e</sup>	EW <sup>f</sup>
-30 kb	1,181	0.014	-0.660	10	10	0.100	1,181	0.014	-1.667	10	10	0.100
-20 kb	981	0.011	-0.362	10	10	0.100	981	0.018	-0.260	10	10	0.100
-10 kb	1,144	0.014	-0.372	10	10	0.100	1,144	0.014	-0.523	10	10	0.100
-5 kb	1,162	0.008	-2.086	10	10	0.350 ( $p=0.001$ )	1,162	0.010	0.584	10	10	0.120
<i>APL1A</i>	<b>1,669</b>	<b>0.018</b>	<b>-1.407</b>	<b>20</b>	<b>14</b>	<b>0.155 (<math>p&lt;10^{-3}</math>)</b>	<b>1,537</b>	<b>0.081</b>	<b>0.568</b>	<b>27</b>	<b>27</b>	<b>0.037</b>
<i>APL1B</i>	<b>2,005</b>	<b>0.016</b>	<b>-1.499</b>	<b>26</b>	<b>17</b>	<b>0.172 (<math>p=0.002</math>)</b>	<b>1,966</b>	<b>0.039</b>	<b>-1.081</b>	<b>34</b>	<b>33</b>	<b>0.033</b>
<i>APL1C</i>	<b>2,587</b>	<b>0.009</b>	<b>-0.927</b>	<b>20</b>	<b>12</b>	<b>0.145 (<math>p=0.014</math>)</b>	<b>2,393</b>	<b>0.031</b>	<b>0.404</b>	<b>33</b>	<b>30</b>	<b>0.036</b>
+5 kb	1,198	0.013	-0.500	10	9	0.120	1,198	0.020	-0.021	9	9	0.111
+10 kb	1,135	0.019	-0.905	10	9	0.120	1,135	0.018	-0.616	10	10	0.100
+20 kb	1,119	0.011	-0.573	10	10	0.100	1,119	0.016	-0.602	10	10	0.100
+40 kb	1,313	0.034	0.642	10	9	0.120	1,313	0.046	-0.178	10	9	0.120

Nucleotide diversity ( $\pi$ ) at *APL1* is greatly reduced in the M form relative to the S form with a strong skew toward rare variants (indicated by negative values of  $F^*$ ) that is not observed in the S form (see also Figure 4). Nucleotide diversity at *APL1* is slightly reduced relative to flanking loci within the M form but is greatly elevated relative to flanking loci in the S form. Haplotype diversity is prominently depressed at the *APL1* locus, as indicated by high values of the EW statistic [42] that are significantly incompatible with neutral evolution. No such decrease in haplotype diversity is observed in flanking regions progressively distant from the *APL1* locus in the M form or at any of the S form loci. Most of the haplotypes in M form mosquitoes differ by only one or a few nucleotide substitutions at the *APL1* locus, whereas S form mosquitoes show deeper genealogical structure (see Figure S2).

<sup>a</sup>Locus size, in base pairs, excluding insertions and deletions.  
<sup>b</sup>Average number of differences per pair of alleles, per nucleotide.  
<sup>c</sup>Skew in the site frequency spectrum, with negative values indicating an excess of rare variants.  
<sup>d</sup>Number of alleles sequenced.  
<sup>e</sup>Number of distinct haplotypes observed.

<sup>f</sup>Haplotype homozygosity, calculated as the sum of squared observed haplotype frequencies. The use of the EW statistic to contrast the *APL1* genes to the flanking regions is very conservative for the detection of selective sweeps in our experimental framework, as many more alleles were sampled at the *APL1* locus and the physical region surveyed is larger in the *APL1* genes than in flanking regions, both of which allow greater opportunity for recombination to generate distinct haplotypes.  $p$  values are given only for loci that depart significantly from the neutral expectation.

doi:10.1371/journal.pbio.1000600.t002

a d b . a . a . a . d . . . . . l . i . a . . . . . a . a . e . . . . . l . i . c . . . .  
 . . . . . a . l . a . . . . . c . . . . . d . . . . . e . . . . . b . a . . . . . d . . . . . a . . . . . b . . . . . a . . . . . d . . . . .  
 . . . . . c . . . . . a . . . . . d . . . . . c . . . . . a . . . . . l . . . . . c . . . . . a . . . . . a . . . . . d . . . . . d . . . . . c . . . . . a . . . . .  
 . . . . . c . . . . . c . . . . . d . . . . . d . . . . . T . . . . . A P L 1 . . . . . A . . . . . h . . . . .  
 . . . . . e . . . . . . . . . . . l . . . . . i . . . . . d . . . . . d . . . . . c . . . . . a . . . . . d . . . . . a . . . . . d . . . . . a . . . . .  
 . . . . . a . . . . . a . . . . . a . . . . . c . . . . . . . . . . . l . . . . . b . . . . . l . . . . . a . . . . . a . . . . . r . . . . . d . . . . .  
 . . . . . d . . . . . c . . . . . a . . . . . a . . . . . a . . . . . I . . . . . S . . . . . A P L 1 . . . . .  
 . . . . . c . . . . . d . . . . . A . . . . . h . . . . . . . . . . . a . . . . . a . . . . . a . . . . . d . . . . . b . . . . . 10 . . . . . d . . . . . d . . . . .  
 . . . . . d . . . . . a . . . . . a . . . . . d . . . . . a . . . . . a . . . . . b . . . . . b . . . . . d . . . . . d . . . . . A . . . . . h . . . . .  
 . . . . . d . . . . . c . . . . . a . . . . . a . . . . . a . . . . . b . . . . . b . . . . . d . . . . . d . . . . . l . . . . . d . . . . .  
 . . . . . c . . . . . c . . . . . a . . . . . d . . . . . c . . . . . l . . . . . d . . . . . l . . . . . d . . . . .  
 . . . . . a . . . . . d . . . . . a . . . . . c . . . . . [ 1 3 1 7 , 1 9 2 1 ] . T . . . . . a . . . . . . . . . . . a . . . . . d . . . . . d . . . . .  
 . . . . . b . . . . . d . . . . . A P L 1 . . . . . . . . . . . e . . . . . l . . . . . d . . . . . a . . . . . c . . . . . a . . . . .  
 . . . . . c . . . . . c . . . . . d . . . . . l . . . . . . . . . . . a . . . . . d . . . . . a . . . . . c . . . . . a . . . . .  
 . . . . . d . . . . . d . . . . . a . . . . . a . . . . . l . . . . . a . . . . . l . . . . . a . . . . . a . . . . . c . . . . .  
 . . . . . d . . . . . a . . . . . c . . . . . a . . . . . a . . . . . T . . . . . e . . . . . a . . . . . . . . . . . c . . . . . c . . . . .  
 . . . . . d . . . . . l . . . . . c . . . . . a . . . . . A P L 1 . . . . . a . . . . . c . . . . . a . . . . . A . . . . . h . . . . .  
 . . . . . d . . . . . a . . . . . d . . . . . a . . . . . c . . . . . c . . . . . d . . . . . l . . . . . T . . . . . b . . . . . d . . . . . a . . . . .  
 . . . . . a . . . . . d . . . . . a . . . . . a . . . . . c . . . . . c . . . . . d . . . . . l . . . . . [ 2 6 ] . A . . . . . a . . . . .  
 . . . . . a . . . . . d . . . . . e . . . . . . . . . . . d . . . . . c . . . . . a . . . . . a . . . . . c . . . . .  
 . . . . . a . . . . . a . . . . . c . . . . . d . . . . . A P L 1 . . . . . d . . . . . a . . . . . a . . . . . a . . . . . a . . . . . a . . . . .  
 . . . . . a . . . . . a . . . . . c . . . . . a . . . . . M . . . . . c . . . . . l . . . . . b . . . . . l . . . . . a . . . . .  
 . . . . . a . . . . . a . . . . . c . . . . . a . . . . . d . . . . . a . . . . . d . . . . . a . . . . . a . . . . .  
 . . . . . l . . . . . c . . . . . c . . . . . a . . . . . d . . . . . a . . . . . a . . . . . T . . . . . a . . . . . a . . . . .  
 . . . . . b . . . . . e . . . . . d . . . . . a . . . . . a . . . . . d . . . . . d . . . . . d . . . . . a . . . . .  
 . . . . . l . . . . . [ 4 3 ] . . . . . a . . . . . a . . . . . a . . . . . a . . . . . c . . . . . a . . . . . c . . . . .  
 . . . . . c . . . . .

W . . . . . l . . . . . A P L 1 . . . . . d . . . . . a . . . . . a . . . . . A . . . . . h . . . . .  
 . . . . . d . . . . . . . . . . . b . . . . . a . . . . . . . . . . . a . . . . . a . . . . . l . . . . . T . . . . .  
 . . . . . A P L 1 C , T E P 1 , a . . . . . d . . . . . L R I M 1 . . . . . . . . . . . a . . . . . c . . . . . a . . . . .  
 . . . . . a . . . . . c . . . . . a . . . . . a . . . . . d . . . . . a . . . . . b . . . . . T E P 1 . . . . . a . . . . . c . . . . . a . . . . . a . . . . . [ 1 1 , 1 2 ] ,  
 . . . . . a . . . . . a . . . . . b . . . . . a . . . . . e . . . . . a . . . . . a . . . . . e . . . . . d . . . . . a . . . . .  
 . . . . . L . . . . . A P L 1 . . . . . l . . . . . a . . . . . . . . . . . d . . . . . a . . . . . a . . . . . a . . . . . d . . . . . l . . . . . a . . . . .  
 . . . . . . . . . . . l . . . . . d . . . . . a . . . . . a . . . . . a . . . . . a . . . . . a . . . . . A P L 1 . . . . . [ 1 8 ] ,  
 . . . . . a . . . . . l . . . . . [ 1 4 , 1 6 ] . A . . . . . h . . . . . . . . . . . a . . . . . a . . . . . S . . . . .  
 . . . . . A . . . . . h . . . . . a . . . . . b . . . . . a . . . . . a . . . . . c . . . . . c . . . . . c . . . . . A . . . . . h . . . . .  
 . . . . . A . . . . . h . . . . . a . . . . . d . . . . . A . . . . . m . . . . . a . . . . . A P L 1 . . . . . l . . . . . , a . . . . . d . . . . . B M . . . . .  
 . . . . . ( F . . . . . S 2 , [ 1 6 , 1 8 ] ) . W . . . . . . . . . . . l . . . . . e . . . . . . . . . . . d . . . . .  
 . . . . . l . . . . . a . . . . . c . . . . . a . . . . . d . . . . . a . . . . . c . . . . . a . . . . . a . . . . . d . . . . . ( a . . . . .  
 . . . . . l . . . . . a . . . . . c . . . . . a . . . . . . . . . . . b . . . . . d . . . . . ) . . . . . c . . . . . A . . . . . h . . . . . a . . . . . d . . . . . A . . . . .  
 . . . . . h . . . . . [ 3 0 ] , . . . . . d . . . . . a . . . . . a . . . . . d . . . . . A . . . . . m . . . . . c . . . . . a . . . . . a . . . . .  
 . . . . . a . . . . . l . . . . . A . . . . . h . . . . . . . . . . . A . . . . . . . . . . . A . . . . .  
 . . . . . . . . . . . b . . . . . d . . . . . a . . . . . b . . . . . A . . . . . h . . . . . a . . . . . d . . . . . A . . . . . h . . . . .  
 . . . . . A . . . . . m . . . . . , . . . . . a . . . . . a . . . . . e . . . . . . . . . . . a . . . . . a . . . . . A P L 1 . . . . . , a . . . . .  
 . . . . . a . . . . . a . . . . . l . . . . . a . . . . . d . . . . . B M . . . . . , . . . . . e . . . . . . . . . . . d . . . . . l . . . . . a . . . . .  
 . . . . . a . . . . . a . . . . . - d . . . . . a . . . . . a . . . . . c . . . . .  
 . . . . . W . . . . . b . . . . . 2 L a . . . . . a . . . . . d . . . . . a . . . . . c . . . . . / e . . . . . c . . . . .  
 . . . . . a . . . . . . . . . . . c . . . . . a . . . . . d . . . . . d . . . . . b . . . . . c . . . . . l . . . . . A P L 1 . . . . . , b . . . . . a . . . . .  
 . . . . . b . . . . . l . . . . . c . . . . . S . . . . . l . . . . . c . . . . . l . . . . . c . . . . . a . . . . . A P L 1 . . . . . d . . . . . c . . . . .  
 . . . . . b . . . . . M a . . . . . d . . . . . S . . . . . d . . . . . a . . . . . T . . . . . M a . . . . . d . . . . . S . . . . . d . . . . . a . . . . .  
 . . . . . a . . . . . a . . . . . a . . . . . d . . . . . c . . . . . a . . . . . a . . . . . d . . . . . d . . . . . d . . . . .  
 . . . . . c . . . . . c . . . . . a . . . . . c . . . . . a . . . . . [ 2 5 ] a . . . . . d . . . . . a . . . . . B a . . . . . e . . . . . a . . . . . M a . . . . .  
 . . . . . a . . . . . l . . . . . A . . . . . l . . . . . A P L 1 . . . . . d . . . . . a . . . . . [ 4 4 , 4 5 ] , . . . . . d . . . . . M . . . . .  
 . . . . . d . . . . . c . . . . . b . . . . . d . . . . . a . . . . . d . . . . . c . . . . . a . . . . .

a d S b ca d a  
*AP* d c a a c  
a c d M S a a  
d a a a c d M ([43];  
a [18]). I d b a b a a d a  
a c a a d APLIC-LRIM1-TEPI  
M Obba d a. [14],  
d c a a *EM* M  
c d Ca  
I c a a c *AP* a d *F*  
d b c d M a a  
bab ca d c b  
[25]. B a a a b c a d a  
c a a a b d c a a ab a  
a a c a d a a a  
a c a A *AP*, *F*, a d *EM*  
a b c a a d a a a d a c [4,7,8  
10,22], bab a a c d  
c K a c b a M a. [8] a  
*AEC* TD66([14dd]-246.5- a a *F*



Shi, C. K. 3.1(ABI). T...  
S... (G... G... G...). *AP*...  
d... G... ba... d... acc... b... HQ702785-  
HQ702849 and HQ860124-HQ860265.

I... d... a... a... *AP*...  
c... M... a... l... ba...  
da... a... ba... d... 10 M... and 10 S...  
e... d... Ba... a... a... d... c... a... 5' b...  
10' b... 20' b... and 40' b... d... *AP*... ba... d...  
e... d... a... "A... P3" a... b... *A. g.*  
I... d... O... 9 S... a... e... d... a...  
5' b... 5'... *AP*... c... b... ca... 10' DNA... a...  
e... a... d... PCR a... N... c... a... ca... d...  
d... d... c... b... d... a... d... a... b... M... and S...  
a... d... A... c... a... d... d... d...  
a... c... ba... d... PEST... [49], and  
d... c... d... a... d... c... b... d... ab... T... a...  
I... c... a... b... d... d... G... ba... d... acc...  
b... HQ859966-HQ860123.

I... d... e... d... b... a... ca... a... d... d... c...  
PCR... d... c... add... a... d... d... a... d...  
and e... d... d... c... G... c... DNA... a... d... a...  
a... a... d... a... ca... a... d... d...  
G... P... (GE... H... a... ca... )... a... da...  
W... d... a... d... d... c... d... d... l... 1:100, and  
l... d... d... a... d... DNA... a... l... da... a... a... 20'  
PCR... l... ca... d... d... e... d... T...  
a... a... e... a... l... d... a... a... a... e... da...  
PCR... c... a... d... l... d... b...  
a... a... e... a... e... a... b... a... da... d...  
U... e... a... d... and dNTP... ac... a... d...  
e... da... a... ca... d... c... b... c... ba... 60' a...  
37°C... E... I... and SAP (b... a... ac... d... b... USB),  
a... b... l... ac... a... d... b... 10' c... ba... a... 65°C...  
A... ca... d... c... d... c... d... c... d...  
B... D... T... a... C... c... S... c... K... 3.1 (A... d... B...  
). Ac... a... *AP*... 470... 581...  
a... a... da... d... (80.9% a... da...). T... a...  
e... c... l... a... e... and l... c... l... ad... a...  
a... a... 3' 10,000' c... d...

PCR a... ca... *APA*... d... da...  
cca... a... d... d... d... c... l... c... d... a... DNA...  
l... c... a... d... a... ba... d... a... a... l... c...  
*APA*... l... a... b... ca... da... a... c... c... d... a...  
a... a... a... e... d... a... d... PANGGL...  
l... l... c... d... d... d... e... d... a...  
a... a... a... d... d... a... a... e... d... a...  
LRR... W... d... c... d... d... da... a... ca... d...  
l... c... d... *APA*... l... c... add...  
e... a... *APA*... l... and *APA*... 2' a... l... l... a...  
e... *APA*... l... c... a... b... a... c... d... ca... N... l... c...  
*APA*... d... ca... ca... b... l... d... e... l... c... d... A...  
l... [49], and *APA*... PCR... d... d... da...  
a... d... d... d... ba... d... d... da...  
d... d... d... d... c... N... d... da... l... c... a... d...  
d... da... a... ab... c... a... a... *APA*... a...  
U... l... a... a... ca... *APA*... l... d... ca... a...  
a... a... l... ab... ac... ca... a... ca...  
a... DNA... a... a... l... ab... d... c... a... a...  
l... a... l... c... d... *APA*... d... ca... N... a...  
l... ab... e... a... l... a... l... ca... a... d...  
e... c... a... l... a... ca... a... a... a...  
and l... b... a... a... l... d... ca... a... T...

d... ca... a... a... d... d... l... c... a... a... c... a...  
l... c... d... a... a... a... a... b... d... a... l... a... da... d... *APA*...  
a... l... l... a... d... ca... a... b... a... d... and  
l... d... d... *APA*...

### Molecular Form and 2La Inversion State

T... M/S... c... a... ac... d... da... l... a...  
d... d... l... PCR... da... c... d... d... b... Fa... a...  
[50]. S... *AP*... ca... d... 2La... a... a...  
2La... a... a... d... ac... d... da... l... a... PCR...  
da... c... d... d... b... W... a... [51]. M/S... and 2La...  
l... d... a... d... a... l... DNA... a...

### Population Genetic Analyses

E... a... l... a... d... ba... d... l... b...  
c... c... ( $\theta_w$ ) and a... a... l... b... a...  
d... c... a... a... ( $\pi$ )... ca... a... d... a...  
l... ad... a... a... a... a...  
e... b... a... l... D... aSP 5.1 [52]. T... a... d... c...  
b... a... a... Ta... a' *D* [40], a... a... Fl...  
and L' *F\** [41]... a... ca... a... d... D... aSP... Ha...  
e... (EW) a... d... da... l... a... d...  
ac... d... c... a... b... da... d... c... b... d... Z... a... [42]  
ad... a... ca... a... d... l... a... c... c... C... T...  
d... b... EW... a... c... l... d... a... a...  
d... d... 1,000... a... d... l... a... a... a...  
a... a... ad... l... b... l... a... l... a... ac... ca... da...  
N... a... a... l... a... d... l... a... [53]  
e... a... a... l... e... b... a... T... d... l... c...  
l... b... d... a... a... e... c... a... a... d... l...  
K<sub>ST</sub>\*... a... c... [48] a... d... D... aSP... K<sub>ST</sub>\*... a...  
l... b... l... a... a... a... a... l... c... a... a...  
ac... a... l... b... l... a... S... ca... a... l... a... c... d... ca...  
a... d... da... a... a... l... b... l... a... d... b...  
l... ca... a... a... ac... a... a... a... [ (d... d... a... c... )],  
l... b... d... K...

$K_{ST}^*$ ... a... a... d... b... e... a... a... l... d... b...  
...  $K_{ST}^*$ ... e... l... c... d... ac... a... l... a... a... ac... d...



17. Gilchrist A, Kaur A, Saha S, and F. M. A. I. K. A. A. (2008) SNP diversity and selection at the *Anopheles APL1* locus. *BMC Genomics* 9: 227.

18. Obbard DJ, Caputo DM, J. L. F. M. S. A. DC, Yang G. A. (2008) The *Anopheles APL1* locus is under positive selection. *BMC Evol Biol* 8: 274.

19. Pagnano A, Saha MA, Malaria JC, Anand A, B. PH, Anand N. C. A. (2008) The *Anopheles APL1* locus is under positive selection. *BMC Evol Biol* 8: 79.

20. L. A. T. H. J. C. L. C. M. B. CS, W. B. K. A. (2009) Molecular evolution of the *Anopheles APL1* locus. *PLoS One* 4: 4549. doi:10.1371/journal.pone.0004549.

21. Obbard DJ, W. C. J. L. T. J. (2009) The *Anopheles APL1* locus is under positive selection. *Malaria J* 8: 117.

22. Bard S, Saha SH, Malaria LF, J. A. C. J. W. A. P. A. (2004) Genetic diversity and selection at the *Anopheles APL1* locus. *Genetics* 166: 661-670.

23. D. Y. A. R. X. Z. W. E. M. E. A. (2006) *Anopheles APL1* locus is under positive selection. *PLoS* 2: 52. doi:10.1371/journal.pone.0020052.

24. Bard SA, W. S. R. L. A. acc. a M. G. J. L. C. G. A. (2009) Diversity and selection at the *Anopheles APL1* locus. *Science* 326: 147-150.

25. L. A. T. D. A. A. (2008) The *Anopheles APL1* locus is under positive selection. *Genetics* 178: 737-746.

26. H. d. RR, K. a. NL (1988) The *Anopheles APL1* locus is under positive selection. *Genetics* 120: 831-840.

27. C. V. B. C. J. S. C. M. H. V. A. X (2008) The *Anopheles APL1* locus is under positive selection. *PLoS Genetics* 4: 1000168. doi:10.1371/journal.pgen.1000168.

28. McDavid JH, K. A. M (1991) The *Anopheles APL1* locus is under positive selection. *Nature* 351: 652-654.

29. W. S. L. C. A. B (2004) The *Anopheles APL1* locus is under positive selection. *Genetics* 168: 1071-1076.

30. B. A. NJ, K. A. J. L. A. T. S. A. F. K. M. A. (2003) The *Anopheles APL1* locus is under positive selection. *Proc Natl Acad Sci U S A* 100: 10818-10823.

31. Pagnano A, Malaria M, R. A. N. L. C. S. A. MA. A. (2010) The *Anopheles APL1* locus is under positive selection. *PLoS One* 5: 8885. doi:10.1371/journal.pone.0008885.

32. K. AD, J. CD, B. DJ (2004) The *Anopheles APL1* locus is under positive selection. *Genetics* 167: 725-735.

33. L. A. BP (2005) The *Anopheles APL1* locus is under positive selection. *Genetics* 169: 2023-2034.

34. Y. A. C. K. A. Y. S. A. M. K. H. A. A. Y. M. A. S. C. A. R. A. (2008) The *Anopheles APL1* locus is under positive selection. *PLoS One* 3: 3393. doi:10.1371/journal.pone.0003393.

35. Gilchrist M, Saha A, P. A. V. D. D. S. MA (1979) The *Anopheles APL1* locus is under positive selection. *Malaria J* 73: 483-497.

36. R. J. R. P. A. V. d. A. T. A. C. A. G. M (1999) The *Anopheles APL1* locus is under positive selection. *Genetics* 151: 101-113.

37. T. Y. T. P. A. V. T. A. S. F. G. A. A. M. A. H. M. A. (1998) The *Anopheles APL1* locus is under positive selection. *Malaria J* 40: 477-511.

38. M. A. d. S. J. H. A. J (1974) The *Anopheles APL1* locus is under positive selection. *Genetics* 23: 23-35.

39. B. A. J. M. H. d. RR, K. a. NL, L. A. C. H. S. A. W (1995) The *Anopheles APL1* locus is under positive selection. *Genetics* 140: 783-796.

40. T. A. F (1989) The *Anopheles APL1* locus is under positive selection. *Genetics* 123: 585-595.

41. F. YX, L. WH (1993) The *Anopheles APL1* locus is under positive selection. *Genetics* 133: 693-709.

42. Z. K. M. S. S. S. W. C. I (2007) The *Anopheles APL1* locus is under positive selection. *Malaria J* 24: 1562-1574.

43. W. B. J. L. A. C. A. M. K. L. C. C. G. A. M. B. W. M. D. A. (2011) The *Anopheles APL1* locus is under positive selection. *Proc Natl Acad Sci* 108: 244-249.

44. T. A. TL, H. A. MW, N. d. SV (2005) The *Anopheles APL1* locus is under positive selection. *PLoS* 3: 285. doi:10.1371/journal.pone.0030285.

45. T. A. TL, H. A. MW (2007) The *Anopheles APL1* locus is under positive selection. *Malaria J* 24: 2132-2138.

46. L. A. EA, M. A. LF, B. A. S. V. d. G. L. A. M. A. (2001) The *Anopheles APL1* locus is under positive selection. *Genetics* 104: 709-718.

47. E. W. J. (1972) The *Anopheles APL1* locus is under positive selection. *Trends* 3: 87-112.

48. H. d. RR, B. D. D. K. a. NL (1992) The *Anopheles APL1* locus is under positive selection. *Malaria J* 9: 138-151.

49. H. RA, S. A. A. GM, H. A. A. S. G. G. C. A. B. R. A. (2002) The *Anopheles APL1* locus is under positive selection. *Science* 298: 129-149.

50. F. A. G. L. A. A. S. A. S. L. S. d. K. A. I. L. C (2001) The *Anopheles APL1* locus is under positive selection. *Genetics* 157: 19-23.

51. W. B. J. S. A. A. A. F. K. A. L. R. B. M. G. O. A. (2007) The *Anopheles APL1* locus is under positive selection. *Malaria J* 24: 334-339.

52. L. A. P. R. A. J (2009) The *Anopheles APL1* locus is under positive selection. *Genetics* 151: 1451-1452.

53. H. d. RR (2002) The *Anopheles APL1* locus is under positive selection. *Genetics* 161: 337-338.

54. H. d. RR, S. A. M. M. A. d. W. P (1992) The *Anopheles APL1* locus is under positive selection. *Genetics* 132: 583-589.

55. K. A. S. T. A. K. N. M (2004) The *Anopheles APL1* locus is under positive selection. *Genetics* 166: 150-163.