

# The changing landscape of *Plasmodium falciparum* drug resistance in the Democratic Republic of Congo

Molly N/A Deutsch-Feldman (✉ [mollydf1@live.unc.edu](mailto:mollydf1@live.unc.edu))

University of North Carolina at Chapel Hill <https://orcid.org/0000-0001-6191-930X>

**Ozkan Aydemir**

Brown University

**Margaret Carrel**

University of Iowa

**Nicholas F. Brazeau**

University of North Carolina at Chapel Hill

**Samir Bhatt**

Imperial College London

**Jeffrey A. Bailey**

Brown University

**Melchior Kashamuka**

Universite de Kinshasa

**Antoinette K. Tshetu**

Universite de Kinshasa

**Steve M Taylor**

Duke University School of Medicine

**Jonathan J. Juliano**

University of North Carolina at Chapel Hill

**Steven R. Meshnick**

University of North Carolina at Chapel Hill

**Robert Verity**

Imperial College London

---

## Research article

**Keywords:** malaria, drug resistance, spatial-temporal modeling

**Posted Date:** October 2nd, 2019

**DOI:** <https://doi.org/10.21203/rs.2.11285/v2>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published on October 22nd, 2019. See the published version at <https://doi.org/10.1186/s12879-019-4523-0>.

# Abstract

**Background:** Drug resistant malaria is a growing concern in the Democratic Republic of the Congo (DRC), where previous studies indicate that parasites resistant to sulfadoxine/pyrimethamine or chloroquine are spatially clustered. This study explores longitudinal changes in spatial patterns to understand how resistant malaria may be spreading within the DRC, using samples from nation-wide population-representative surveys. **Methods:** We selected 552 children with PCR-detectable *Plasmodium falciparum* infection and identified known variants in the *pfdhps* and *pfcr*t genes associated with resistance. We compared the proportion of mutant parasites in 2013 to those previously reported from adults in 2007, and identified risk factors for carrying a resistant allele using multivariate mixed-effects modeling. Finally, we fit a spatial-temporal model to the observed data, providing smooth allele frequency estimates over space and time. **Results:** The proportion of co-occurring *pfdhps* K540E/A581G mutations increased by 16% between 2007 and 2013. The spatial-temporal model suggests that the spatial range of the *pfdhps* double mutants expanded over time, while the prevalence and range of *pfcr*t mutations remained steady. **Conclusions:** This study uses population-representative samples to describe the changing landscape of SP resistance within the DRC, and the persistence of chloroquine resistance. Vigilant molecular surveillance is critical for controlling the spread of resistance.

## Background

With 17 million confirmed cases in 2016, the burden of malaria in the Democratic Republic of the Congo is one of the highest in the world (1). Understanding malaria dynamics in DRC is critical in eliminating malaria from sub-Saharan Africa. Previous studies of *P. falciparum* genetic diversity within the DRC have shown a mixture of both West and East African strains, indicating that the DRC serves as a nexus of regional transmission, incorporating parasites from both sides of the continent (2–4). Insights from genetic *P. falciparum* studies within the DRC therefore have important implications for reducing disease burden within the country and for Central and sub-Saharan Africa.

Efforts to halt transmission in the DRC, and across Africa, are being threatened by growing resistance to commonly used antimalarial drugs (5,6). Molecular markers can be used to identify resistant infections to monitor the spread of resistance (7,8). These markers include mutations in the dihydropteroate synthase (*pfdhps*) gene, which, along with mutations of the dihydrofolate reductase (*pfdhfr*) gene, confer resistance to sulfadoxine (7–9). Specifically, the *pfdhps* A437G, K540E, and A581G mutations are associated with sulfadoxine/pyrimethamine (SP) treatment failure (8,9). Co-occurrence of the K540E and A581G mutations has been associated with failure of SP for intermittent preventive therapy during pregnancy (IPTp), a program recommended for all pregnant women in the DRC (10,11). Additionally, mutations of the chloroquine resistance transporter (*pfcr*t) gene, specifically mutations in amino acids 72–76 (wild type CVMNK), confer chloroquine resistance (12,13). Past work has demonstrated that the K76T mutation alone increases the risk of chloroquine treatment failure, though a K76T-containing CVIET triple mutant haplotype has emerged in many African countries, including the DRC (5,7,14). Studies have demonstrated that this haplotype is also associated with amodiaquine treatment failure; amodiaquine is

used as part of the first line therapy for malaria in the DRC (10,15–17). Additionally, another *pfcr* haplotype, SVMNT, is also associated with resistance to amodiaquine (12,17,18). Monitoring these molecular markers is critical for halting the spread of resistance.

Studies conducted in the DRC have demonstrated spatial structure of parasites resistant to SP and chloroquine (5,6,19). Amongst adult respondents to the 2007 Demographic and Health Survey (DHS), those infected with parasites with a single *pfdhps* A437G mutation were spread throughout the country, though mostly located in the western part of the DRC (6). Co-occurring K540E and A581G mutants were rarer, but showed geographic clustering in the northeast region of the country (6). Past studies from the DRC have demonstrated that the *pfdhfr* mutations are nearly fixed within the population (20). There was no apparent clustering of the *pfcr* CVIET haplotype (21). More recent work demonstrated similar patterns for *pfdhps* mutations amongst children sampled in 2013 (19). However, the *pfcr* CVIET haplotype displayed a pattern of concentrated cases on the eastern and western borders, with fewer mutations observed in the center of the DRC (19). These findings highlight hotspots of resistance within the DRC.

Previous geospatial studies of drug resistance in the DRC often do not use population based samples and are therefore not necessarily nationally representative (22). Better estimates of the burden and distribution of these mutations can be obtained using nationally representative surveys (19,23). This is the first study, to our knowledge, that uses data from a nationally representative database to evaluate risk factors for carrying a resistant infection and to study longitudinal changes in resistance.

Here, we describe the changing spatial patterns of SP and chloroquine resistance over time in the DRC by comparing samples from the 2013-2014 DHS to previously published data drawn from the 2007 DHS. We begin by describing the proportion of single and co-occurring mutations in both 2007 and in 2013. Next, we present an epidemiologic risk factor analysis to identify covariates associated with increased prevalence of resistant infections. The findings from this analysis will help identify individuals and communities that may be of higher risk for resistant infections. Finally, we use Markov chain Monte Carlo to fit a spatial-temporal model to the observed data to explore whether the geographic range of drug resistant mutations has shifted between 2007 and 2013. This model allows us to directly compare allele frequencies across space and time.

## Methods

### *Study population*

Samples were drawn from the DHS survey, conducted in the DRC in 2013-2014 (24,25). The DHS Program conducts cross-sectional, nationally representative population health surveys in over 90 countries. In the DRC, the DHS survey uses a randomized cluster sampling method (26). For the 2013-2014 survey, 536 geographic clusters across the DRC were randomly selected. Next, households were randomly selected from these clusters for inclusion in the DHS. The 2013-2014 survey included adolescents and adults ages 15-49 and children under age 5. DHS survey conductors visited selected households and obtained informed consent from each individual age 18 or older, or from a parent or legal guardian for children and

adolescents under age 18. Survey conductors administered an extensive questionnaire covering topics such as socioeconomic status, education, and health history. Each individual was administered a malaria rapid diagnostic test and blood samples were collected on filter paper and shipped to the University of North Carolina for molecular diagnostic testing. All DHS questionnaires and procedures have been approved by the ICF Institutional Review Board and comply with the United States Department of Health and Human Services regulations for the protection of human subjects. This study was approved by the Internal Review Board at The University of North Carolina, Chapel Hill and at the Kinshasa School of Public Health.

A previous DHS survey was conducted in the DRC in 2007 (24). Similar to 2013-2014, a two-stage random cluster sampling scheme was used to select households for inclusion, though only adults were asked to participate. The 2007 survey used 300 sampling clusters (fewer than in 2013-2014) and the clusters were not the same between years.

### *DNA amplification and genotyping*

Findings from the 2007 DHS have been previously published (3–6,27). As described, 220 samples were previously genotyped at the *pf dhps* and *pf crt* loci (6,28).

Molecular diagnostic testing for malaria parasites was completed for all individuals included in the 2013-2014 DHS (23,29). Unlike children included in previous studies, these children participated in the DHS and thus have extensive individual demographic data (19). DNA was extracted from filter paper using a Chelex-100 Kit (Bio-Rad, Richmond, CA). Samples were tested in duplicate using a real-time PCR assay to target the *P.f. lactate dehydrogenase* gene; human *beta-tubulin* was used as a positive control. Primer sequences for both genes have been previously published (18,30). Samples in which both replicates amplified parasite DNA were considered positive. If one replicate failed to amplify but the other amplified with a PCR cycle threshold ( $C_T$ ) value below 38 that sample was also considered positive (23).

A total of 552 children with PCR-confirmed *P.falciparum* infections from 536 clusters were selected from the 2013-2014 DHS for inclusion in this study. Overall, 7137 children included in the DHS had complete data; malaria prevalence by PCR was 38.6% (23). Children with  $C_T$  values under 30 were chosen for this analysis to ensure sufficient sequencing coverage. Children were selected from throughout the DRC, providing ample geographic representation. Samples from the selected children were amplified using molecular inversion probes (MIPs), a multistep protocol that allows for highly multiplexed deep sequencing (19). MIPs were designed to flank the *pf dhps* and *pf crt* targets. Each sample was individually barcoded in order to de-multiplex sequences and yield individual level data. Sequencing data was processed using the MIPWrangler software, as previously described (19). Paired-end reads were stitched and filtered by base quality scores, expected length, a minimum unique molecular index (UMI) count of 3 and minimum relative abundance of 0.5% within sample. SNP calls were further filtered to have a minimum Phred quality score of 20. Mixed infections were identified as those with heterozygous SNP calls at any of the genotyped loci.

### *Comparison of allele frequencies*

Proportions of each SNP were calculated for each year and compared using the UpSet package in the R statistical language (31,32). Mixed infections with both referent and mutant genotype calls were considered mutant. The UpSet package does not accommodate missing data; therefore, this analysis only included observations with SNP calls at all sites. Frequencies were statistically compared between years using chi-squared tests. For these tests, individuals missing a genotype call at any given site were not included in the analysis for that site only. COI estimates were determined using THE REAL McCOIL software (33).

### *Epidemiologic risk factor analysis*

Risk factor data was drawn from all surveyed individuals within a cluster. Potential cluster and individual level demographic risk factors were based on biological plausibility and by consulting relevant literature (27,34). Both cluster-level and individual-level risk factors were evaluated as several studies of malaria conducted in the DRC have demonstrated the role of community level factors on individual infection risk (27,34,35). Selected cluster-level covariates included: malaria prevalence by PCR, percentage of individuals in the lowest wealth category, percentage of individuals without education, percentage of pregnant women who took SP, percentage of children who took chloroquine for a fever or cough, cluster size, and urban vs rural status. Individual covariates included wealth index and biological sex. Age could not be included as the DHS does not collect data from individuals aged 5-15, thus there is a gap in the age distribution.

Multivariate mixed-effect Poisson models were used to identify associations between selected covariates and the probability of having a drug resistant infection. Using a Poisson distribution and estimating a robust variance is an alternative to fitting log-risk models (that use a binomial distribution), which often do not converge (36,37). Full specification of the model is available in **Supplemental Text 1**. We used backwards selection to identify significant associations, initially fitting a full model with all previously mentioned potential risk factors. Covariates were subsequently removed one at a time based on the highest p-value until only covariates with p-values less than or equal to 0.05 were left. Secondary analyses were conducted using univariate models for each potential risk factor to determine if the marginal associations from the univariate models matched those of the multivariate model. To account for dependency between individuals living in the same province, all models fit random effects for DHS province; DHS cluster could not be used as there were too few observations per cluster to estimate random effects. All analyses were conducted in the R statistical language using the lme4 package (38).

### *Spatial Prediction Models*

Spatial prediction maps were generated by fitting a spatial-temporal model to the data. This model assumed a smooth surface based on the logistic Gaussian process(39) to describe the underlying frequency of resistant alleles as a function of space, time, and a number of covariates. Observed counts of resistant alleles were modelled as binomial draws from the underlying frequency distribution. The

complete model specification can be found in **Supplemental Text 1**. Covariates used in the model included accessibility, night time lights (a measure of population density), and proportion urban/rural (40–42). Every covariate was given a weighting parameter allowing it to have a greater or lesser effect on the data, and these parameters were given suitable priors. To facilitate model fitting and to ensure our method scaled well with the number of dimensions and DHS clusters, the full model was approximated using 250 random Fourier features (**Supplemental Text 1**) (43). Model fitting was conducted via Hamiltonian Monte Carlo (HMC) using the GRETA package (44) in the R statistical language. In each analysis the HMC was run for 10,000 burn-in iterations and 1 million sampling iterations, thinning to every 100th sample to remove autocorrelation. Posterior parameter values were sampled at random to generate 1000 maps for each of the *pf dhps* mutations A437G, K540E, A581G, and the *pf crt* CVIET haplotype, and these 1000 maps were summarized in the form of a mean prediction map and standard deviation (error) map.

## Results

### *MIP analyses of 2013-2014 samples*

Following MIPWrangler processing, a 250 bp paired end MiSeq run following a single MIP capture yielded 9 million paired end reads and 4 million UMIs. Sequencing was successful for 514/552 children. The geolocation data indicates that these 514 children live throughout the DRC (**Figure 1**). Complete *pf crt* SNP data was available for 513 children, and 307 had data available across all *pf crt* and *pf dhps* loci of interest.

The results of THE REAL McCOIL analysis estimated an average complexity of infection (COI) of 2 (range = 1-17). Of children with complete genotyping data, one-hundred and eight (35% of the total) had polyclonal infections, compared with 20% of infections that were polyclonal in 2007 ( $X^2 = 7.28$ ,  $df = 1$ ,  $p < 0.01$ ). However, this is likely an underestimate of the true number of polyclonal infections as we are only looking at three loci.

### *Frequency of pf dhps and pf crt variants over time*

The overall proportion of *pf dhps* mutations remained relatively steady from 2007 to 2013, (80% [95% CI = 72-86%] vs 86% [95% CI = 83-89%], **Figure 2**). However, the proportions of K540E mutations increased significantly from 17% in 2007 (95% CI = 11-24%) in 2007 to 41% (95% CI = 36-47%) in 2013 ( $X^2 = 25.57$ ,  $df = 1$ ,  $p < 0.01$ ). A581G mutations also increased significantly between years, from 3% (95% CI = 1-8%) in 2007 to 18% (95% CI = 14-23%) in 2013 ( $X^2 = 15.27$ ,  $df = 1$ ,  $p < 0.01$ ). Only one individual in 2007 had a single A581G mutation, in all other cases, in both years, A581G was only found in the presence of a K540E mutation. Thus, the proportion of double K540E/A581G mutants also increased significantly across years, from 2% (95% CI = 1-7%) in 2007 to 18% (95% CI = 14-23%) in 2013 ( $X^2 = 19.27$ ,  $df = 1$ ,  $p < 0.001$ ).

Amongst monoclonal infections, there were similar patterns of allele frequencies over time. The proportions of infections carrying any of the three *pf dhps* SNPs increased slightly; 62% (95% CI = 51-73%) in 2007 versus 73% (95% CI = 66-79%) in 2013 ( $X^2 = 2.71$ ,  $df = 1$ ,  $p = 0.10$ ). However, the proportion of double K540E and A581G mutant parasites increased from 4% (95% CI = 1-8%) in 2007 to 12% (95% CI = 7-17%) in 2013 ( $X^2 = 3.03$ ,  $df = 1$ ,  $p = 0.08$ ).

The proportion of *pf crt* CVIET haplotypes did not change significantly from 2007 (58% [95% CI = 50-65%] to 2013 (54% [95% CI = 49-58%];  $X^2 = 0.80$ ,  $df = 1$ ,  $p = 0.37$ ). No parasites harbored the SVMNT haplotype. Among monoclonal infections, the proportion of *pf crt* CVIET haplotypes also remained steady; 55% (95% CI = 46-63%) in 2007 and 56% (95% CI = 51-61%) in 2013 ( $X^2 = 0.012$ ,  $df = 1$ ,  $p = 0.91$ ).

### *Risk factor analysis*

Complete *pf dhps* and DHS covariate data were available for 492 individuals from both the 2007 and 2013-2014 studies; complete *pf crt* and DHS covariate data was available for 675 individuals. Reported antimalarial use was low, with a cluster average of only 12% of pregnant women receiving SP in 2007 and 24% in 2013. In 2007, an average of only 4% of children per cluster reporting a cough or fever received amodiaquine, and only about 1% in 2013. A summary of the cluster and individual level characteristics by *pf dhps* and *pf crt* genotype is available in **Table 1**.

### **Table 1: Individual and cluster level characteristics of all study participants, stratified by *Pfdhps* and *Pfcrt* genotype**

	<i>Pfdhps</i>			<i>Pfcr</i>		
	Wildtype (N = 81)	Any <i>pfdhps</i> mutation (N = 434)	P- value <sup>***</sup>	Wildtype (N = 306)	CVIET haplotype (N = 369)	P- value <sup>***</sup>
Malaria prevalence (SD)	59.3 (20.4)	58.9 (21.8)	0.872	60.04 (21.74)	57.42 (22.24)	0.125
Anti-malarial use during pregnancy* (SD)	16.7 (16.6)	22.2 (18.1)	0.011	2.0 (6.2)	1.9 (6.0)	0.955
Anti-malarial use amongst children (SD)**	3.0 (6.4)	2.0 (5.6)	0.126	1.7 (4.4)	1.6 (4.0)	0.745
Mean DHS Cluster size (SD)	17.9 (18.0)	19.3 (18.8)	0.652	17.8 (19.07)	20.0 (22.14)	0.266
% without education (SD)	32.7 (23.48)	23.1 (21.5)	<0.001	28.8 (24.41)	22.0 (20.16)	<0.001
% in lowest wealth category (SD)	30.2 (23.0)	21.4 (22.5)	0.001	27.2 (22.5)	20.5 (22.5)	<0.001
Number living in urban area (%)	28 (34.6)	154 (35.4)	0.975	90 (29.4)	135 (36.6)	0.059
<b>Individual covariates:</b>						
Number female (%)	41 (50.6)	228 (52.5)	0.845	153 (50.0)	192 (52.0)	0.654
Median Individual Wealth Index (IQR)	2 (1-3)	3 (1-4)	0.015	2 (1-3)	3 (2-4)	<0.001

\*Percentage of pregnant women reporting drug use; SP use is described by *pfdhps* status and chloroquine use by *pfcr* status

\*\* Percentage of children with a cough or fever that received SP or chloroquine; SP use is described by *pfdhps* status and chloroquine use by *pfcr* status

\*\*\* p-values for tests conducted for comparisons between wildtype and mutant groups (Chi-squared tests for categorical data and t-tests for continuous data)

The mixed-effects model identified several risk factors for *pfdhps* mutations and the *pfcr* CVIET haplotype (**Table 2**). Increasing cluster-level use of SP was a risk factor for carrying a K540E mutation (PR = 1.14, 95% CI = 1.09 – 1.20, p <0.01) as was increasing cluster prevalence of *P. falciparum* infections (PR = 1.11, 95% CI = 1.06 – 1.17, p = 0.02). The results from the *pfcr* model indicated an inverse

relationship between the prevalence of mutations and the proportion of uneducated individuals (PR = 0.92, 95% CI = 0.90 – 0.95,  $p < 0.01$ ). Education may be a proxy for access to medications.

**Table 2: Risk factors identified from final backwards selection multivariate risk factor model.**

<i>Covariate</i>	<i>Prevalence Ratio (95% CI)</i>	<i>p-value</i>
	<i>Pfdhps K540E</i>	
10% increase in malaria prevalence	1.11 (1.06 – 1.17)	0.024
10% increase in cluster SP use*	1.14 (1.09 – 1.20)	<0.01
	<i>Pfcr1 CVIET</i>	
10% increase in lowest education category	0.92 (0.90 – 0.95)	< 0.01

\* reported SP use amongst pregnant women

Increasing cluster level SP use amongst pregnant women and malaria prevalence were both identified as risk factors for carrying the K540E mutation (including those with the A581G mutation also), while education was the only risk factor identified for carrying the CVIET haplotype.

Results from the secondary univariate models matched those from the multivariate models (**Supplemental Table 2**). Like the multivariate model, the univariate models did not identify any risk factors for carrying any *pfdhps* mutation. The univariate models of K540E identified both increasing SP use and increasing cluster *Pf.* prevalence as risk factors, though the p-value for prevalence was not significant at the 5% level. Like the multivariate model, the univariate models of *pfcr1* identified only increasing cluster level education as a risk factor for the CVIET haplotype. Similarly, increasing cluster level proportion of poor individuals showed a protective effect against the CVIET haplotype, though the association had a p-value that was not significant at the 5% level. Full results for the univariate models are available in **Supplemental Table 1**.

#### *Spatial-temporal prediction maps:*

The prediction maps generated from the logistic Gaussian model indicate that the allele frequency distribution of the A437G mutation shifted range slightly between 2007 and 2013, decreasing in the east and west of the country but increasing in the south (**Figure 3**). The results also demonstrate the geographic spread of both the K540E and A581G mutations from east to west, showing both an increase in the frequency of each mutation and a geographic expansion, indicated by the shift in the 10% contour lines (marked in black). *Pfcr1* results demonstrate that there has been no significant change in the spatial distribution of the CVIET haplotype between 2007 and 2013; the prevalence of the haplotype is highest across the central part of DRC. The wide 95% credible intervals on posterior parameter weights indicate that there is large uncertainty as to which components are driving the signal (**Supplementary Figure 1**).

Similarly, the posterior error maps show that there is large uncertainty in the predicted allele frequency at most points in space (**Supplementary Figure 2**). Hence, it is important to recognize that the maps in **Figure 3** show only the average prediction, and there are alternative maps that are plausible under the posterior distribution. However the general patterns described above, such as the east-west expansion of K540E and A581G mutations, remain consistent over the majority of posterior draws, and therefore are well-supported in spite of uncertainty in any specific prediction.

## Discussion

Monitoring the changing spatial distribution of drug resistance markers is necessary for developing efficient interventions to halt the spread of resistance and eliminate malaria. Here, we leverage geolocated samples from the DHS to measure resistance mutations across the DRC and map changes that occurred between 2007 and 2013 (6,19,21). Studies using nationally representative samples like the DHS are less susceptible to selection bias; however, longitudinal comparisons of DHS surveys have been impeded by the fact that the individual survey clusters change between surveys. Here, we use a spatial prediction model that overcomes this by assuming a continuous surface of underlying allele frequencies, allowing us to integrate information at different points in space and time.

This study found that the 540 and 581 *pfdhps* mutations have increased in the DRC since 2007, both in numbers and in geographic spread. This agrees with recent findings of an increase in *pfdhps* mutations between 2014 and 2015 amongst individuals living in southwest DRC (45). Evidence of geographic expansion from the eastern part of DRC is also supported by previous research that demonstrated higher prevalence of both mutations in East Africa compared to West Africa (9,46). This expansion is particularly concerning as these mutations are associated with SP failure during IPTp (11,19,46). The risk factor analysis indicates that these increases may be in part driven by SP use, which was associated with increased prevalence of *pfdhps* mutations. Further, this study indicates that increasing community level drug use, not necessarily individual use, is associated with increases in resistance. This is consistent with previous work that demonstrated associations between community level interventions and malaria risk (34,35).

Chloroquine resistance has remained relatively steady since 2007; the proportion of CVIET parasites is unchanged and the spatial distribution remains similar. These findings are troubling as the DRC halted chloroquine use as a first line treatment in 2001 due to concerns about growing resistance (47,48). This sustained resistance may be in part driven by demographic factors; the risk factor models results indicate that cluster-level education and wealth are associated with chloroquine resistance. There may also be unregulated chloroquine use, as has been reported in other sub-Saharan African countries (49).

Additionally, there is evidence that the CVIET haplotype is associated with amodiaquine resistance (15,16,50). Since amodiaquine is used as part of the first line treatment ASAQ in the DRC, this association may explain why the prevalence of CVIET has remained steady over time (16,48,50). Reported ASAQ use was too low in this study for us to evaluate this relationship statistically. However, we did not detect the SVMNT haplotype, also found to be associated with ASAQ resistance, in this population (17,18).

The findings from this study have direct implications for malaria control programs in the DRC. As mentioned, SP is still used in the DRC as the primary drug for IPTp (47,48). Increasing SP resistance may threaten these preventive efforts. Additionally, though chloroquine is no longer a recommended treatment for malaria, reports from other sub-Saharan African countries show a steep drop in the proportion of resistant parasites after ending chloroquine use (48,51–53). The sustained prevalence of chloroquine resistance seen in this study is alarming and warrants further investigation.

Effective monitoring of drug resistance requires sensitive molecular tools that can accommodate a large number of samples. Using MIPs to amplify resistance loci allows for highly multiplexed and efficient deep sequencing of *Plasmodia*. This study demonstrates the utility of MIPs for drug resistance surveillance, and the ability to answer critical epidemiological questions. This novel method can also be used to investigate questions of parasite population structure, gene flow, and selective sweeps, amongst others. The spatial-temporal approach used here also represents a step forward compared with previous mapping efforts (19). The random Fourier features (RFF) method allows us to explore complex models in a computationally efficient way, thereby reducing the time and resources required to perform this kind of advanced spatial analysis and opening the door to much larger datasets in the future.

There are several limitations to this study. First, we only have access to a relatively small number of samples distributed over a wide geographic area, and this is reflected in the large credible intervals around our spatial-temporal predictions. We can therefore only draw large-scale conclusions about changes that have occurred over the study time period, based on patterns that are consistent over the majority of posterior draws. Second, this study compared genotype data generated using different approaches: data from 2013-2014 was obtained using MIPs and Illumina sequencing, while data from 2007 was obtained with standard PCR amplification and alternate sequencing methods. However, the sequencing coverage is approximately the same across studies, providing assurance that the methods are comparable. Additionally, the MIPs did not amplify across all of *pf dhps* in a single sequence but rather used multiple MIP probes to target the regions of interest. Therefore, we could not create true haplotypes across *pf dhps*.

## Conclusion

The findings from this study indicate that the prevalence and geographic spread of SP resistance increased between 2007 and 2013. In contrast, the proportion and pattern of chloroquine resistance stayed the same, potentially a result of ASAQ use or informal chloroquine use. These findings indicate a need to continue monitoring these resistant mutations to prevent additional spread, and to further investigate the factors driving these patterns.

## Declarations

**Ethics Approval and Consent to Participate:** This study was approved by the Internal Review Board at The University of North Carolina, Chapel Hill and at the Kinshasa School of Public Health. Informed

consent was obtained from each individual age 18 or older, or from a parent or legal guardian for children and adolescents under age 18.

**Consent for Publication:** Not Applicable

**Availability of Data and Materials:** Sequencing data used in this study are available in the NCBI SRA (Accession number PRJNA545347). Data used in this study from the Demographic Health Surveys Program are available upon request at <https://dhsprogram.com/> (25). Previously published data included in this study are available from the corresponding author upon reasonable request.

**Competing Interests:** The authors declare that they have no competing interests

**Funding:** This work was supported by the National Institutes of Health (R01AI107949, K24AI134990, R21AI121465, R01AI139520). RV is jointly funded by the UK Medical Research Council (MRC) and the UK Department for International Development (DFID) under the MRC/DFID Concordat agreement and is also part of the EDCTP2 programme supported by the European Union.

**Author contributions:** MDF conducted the epidemiological analyses, generated figure 2 and tables 1 and 2, and prepared the manuscript. OA performed MIP analyses. MC aided with data collection and prepared Figure 1. NB helped conduct epidemiological analyses. SB conducted geospatial analyses. JB designed and oversaw MIP analyses and initial study design. MK and AT aided in initial sample collection as well as study design. ST conducted initial genetic analyses and data analysis. JJ, SM, and RV designed the study; JJ and SM oversaw study set up and data collection. RV conducted geospatial analyses and generated figure 3. All authors edited and approved the final manuscript.

**Acknowledgements:** The authors would like to thank everyone who participated in the study and all members of the study team.

## References

1. The World Health Organization. Democratic Republic of the Congo [Internet]. 2016 [cited 2018 Jun 21]. Available from: [http://www.who.int/malaria/publications/country-profiles/profile\\_cod\\_en.pdf?ua=1](http://www.who.int/malaria/publications/country-profiles/profile_cod_en.pdf?ua=1)
2. Verity R, Hathaway NJ, Waltmann A, Doctor SM, Watson OJ, Patel JC, et al. Plasmodium falciparum genetic variation of var2csa in the Democratic Republic of the Congo. *Malar J*. 2018;17(1):1–8.
3. Taylor SM, Messina JP, Hand CC, Juliano JJ, Muwonga J, Tshefu AK, et al. Molecular malaria epidemiology: Mapping and burden estimates for the democratic Republic of the Congo, 2007. *PLoS One* [Internet]. 2011;6(1). Available from: <http://dx.doi.org/10.1371/journal.pone.0016420>
4. Carrel M, Patel JC, Taylor SM, Janko M, Mwandagilirwa MK, Tshefu AK, et al. The geography of malaria genetics in the Democratic Republic of Congo: A complex and fragmented landscape. *Social Science and Medicine*. 2015;133:233–41.

5. Taylor SM, Antonia AL, Harrington WE, Goheen MM, Mwapasa V, Chaluluka E, et al. Independent lineages of highly sulfadoxine-resistant *Plasmodium falciparum* haplotypes, Eastern Africa. *Emerg Infect Dis*. 2014;20(7):1140–8.
6. Taylor SM, Antonia AL, Parobek CM, Juliano JJ, Janko M, Emch M, et al. *Plasmodium falciparum* sulfadoxine resistance is geographically and genetically clustered within the DR Congo. *Sci Rep*. 2013;3:9–11.
7. Picot S, Olliaro P, Monbrison F, Bievenu AL, Price R, Ringwald P. A systematic review and meta-analysis of evidence for correlation between molecular markers of parasite resistance and treatment outcome in *falciparum* malaria. *Malar J*. 2009;8(89).
8. Hyde JE. Drug-resistant malaria - an insight. *FEBS J*. 2007;274(18):4688–898.
9. Pearce RJ, Pota H, Evehe M-SB, Ba E-H, Mombo-ngoma G, Malisa AL, et al. Multiple Origins and Regional Dispersal of Resistant dhps in African *Plasmodium falciparum* Malaria. *PLoS Med* [Internet]. 2009;6(4). Available from: <http://dx.doi.org/10.1371/journal.pmed.1000055>
10. Presidents Malaria Initiative. Democratic Republic of the Congo: Malaria Operational Plan FY 2018. President's Malaria Initiative; 2017.
11. Gesase S, Gosling RD, Hashim R, Ord R, Naidoo I, Madebe R, et al. High Resistance of *Plasmodium falciparum* to Sulphadoxine/Pyrimethamine in Northern Tanzania and the Emergence of dhps Resistance Mutation at Codon 581. *PLoS One*. 2009;4(2).
12. Alifrangis M, Dalgaard MB, Lusingu JP, Vestergaard LS, Staalsoe T, Jensen AT, et al. Occurrence of the Southeast Asian/South American SVMNT haplotype of the chloroquine-resistance transporter gene in *Plasmodium falciparum* in Tanzania. *J Infect Dis*. 2006;193:1738–41.
13. Gama B, Pereira-Carvalho G, Kosi F, Almeida de Oliveria N, Fortes F, Rosenthal P, et al. *Plasmodium falciparum* isolates from Angola show the SVMNT haplotype in the *pfcr*t gene. *Malar J*. 2010;9(174).
14. Sutherland CJ, Fifer H, Pearce RJ, bin Reza F, Nicholas M, Haustein T, et al. Novel *pf*dhps Haplotypes among Imported Cases of *Plasmodium falciparum* Malaria in the United Kingdom. *Antimicrob Agents Chemother*. 2009;53(8):3405–10.
15. Beshir KB, Sutherland CJ, Merinopoulos I, Durrani N, Leslie T, Rowland M, et al. Amodiaquine Resistance in *Plasmodium falciparum* Malaria in Afghanistan Is Associated with the *pfcr*t SVMNT Allele at Codons 72 to 76. *Antimicrob Agents Chemother*. 2010;54(9):3714–6.
16. Holmgren G, Hamrin J, Svard J, Martensson A, Gil JP, Bjorkman A. Selection of *pfmdr*1 mutations after amodiaquine monotherapy and amodiaquine plus artemisinin combination therapy in East Africa. *Infect Genet Evol*. 2007;7(5):562–9.
17. Sa J, Twu O. Protecting the malaria drug arsenal: halting the rise and spread of amodiaquine resistance by monitoring the *PfCRT* SVMNT type. *Malar J*. 2010;9(374).
18. Beshir KB, Hallett RL, Eziefula AC, Bailey R, Watson J, Wright SG, et al. Measuring the efficacy of anti-malarial drugs in vivo: quantitative PCR measurement of parasite clearance. *Malar J*. 2010;9(312).
19. Aydemir O, Janko M, Hathaway N, Verity R, Mwandagalirwa MK, Tshetu AK, et al. Drug Resistance and Population Structure of *Plasmodium falciparum* Across the Democratic Republic of Congo using

- high-throughput Molecular Inversion Probes. *J Infect Dis.* 2018;
20. {Alker AP, Kazadi WM, Kutelemani AK, Bloland PB, Tshefu AK, Meshnick SR}. dhfr and dhps genotype and sulfadoxine-pyrimethamine treatment failure in children with falciparum malaria in the Democratic Republic of the Congo. *Trop Med Int Health.* 2008;13(11):1384–91.
  21. Antonia AL, Taylor SM, Janko M, Emch M, Tshefu AK, Meshnick SR. Short report: A cross-sectional survey of Plasmodium falciparum pfcr1 mutant haplotypes in the Democratic Republic of Congo. *Am J Trop Med Hyg.* 2014;90(6):1094–7.
  22. Mvumbi DM, Kayembe J-M, Situakibanza H, Bobanga TL, Nsibu CN, Mvumbi GL, et al. Falciparum malaria molecular drug resistance in the Democratic Republic of Congo: a systematic review. *Malar J.* 2015;14(354).
  23. Doctor S, Liu Y, Whitesell A, Thwai K, Taylor SM, Janko M, et al. Malaria Surveillance in the Democratic Republic of the Congo: Comparison of Microscopy, PCR, and Rapid Diagnostic Test. *Diagn Microbiol Infect Dis.* 2017;85(1):16–8.
  24. Ministère du Plan and Macro International. Enquête Démographique et de Santé, République Démocratique du Congo 2007. Calverton, MD, USA: Ministère du Plan and Macro International.; 2008.
  25. Ministère du Plan et Suivi de la Mise en œuvre de la Révolution de la Modernité Ministère de la Santé Publique and ICF International [Distributors]. Democratic Republic of the Congo Demographic and Health Survey 2013-14 [Dataset]. Data Extract from CDPR61.SAV [Internet]. Available from: <http://idhsdata.org>
  26. Ministère du Plan et Suivi de la Mise en œuvre de la Révolution de la Modernité (MPSMRM) Ministère de la Santé Publique (MSP) and ICF International. Democratic Republic of Congo Demographic and Health Survey 2013-14: Key Findings. Rockville, Maryland: MPSMRM, MSP and ICF International; 2014.
  27. Messina JP, Taylor SM, Meshnick SR, Linke AM, Tshefu AK, Atua B, et al. Population, behavioural and environmental drivers of malaria prevalence in the Democratic Republic of Congo. *Malar J.* 2011;10(1):161.
  28. Vinakay S, Alam M, Mixson-Hayden T, McCollum A, Sem R, Shah N, et al. Origin and Evolution of Sulfadoxine Resistant Plasmodium falciparum. *PLoS Pathog.* 2010;6(3).
  29. Plowe C, Djimdé AA, Bouare M, Doumbo O, Wellems TE. Pyrimethamine and Proguanil resistance-conferring mutations in Plasmodium Falciparum Dihydrofolate Reductase : polymerase chain reaction methods for surveillance in Africa. *Am J Trop Med Hyg.* 1995;52(6):565–8.
  30. Taylor SM, Juliano JJ, Trottman PA, Griffin JB, Landis SH, Kitsa P, et al. High-Throughput Pooling and Real-Time PCR-Based Strategy for Malaria Detection. *J Clin Microbiol.* 2010;48(2):512–9.
  31. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria [Internet]. 2017. Available from: <https://www.r-project.org/>
  32. Lex A, Gehlenborg N, Strobel H, Vuillemot R, Pfister H. UpSet: Visualization of Intersecting Sets. *IEEE Trans Vis Comput Graph.* 2014;20(12):1983–92.

33. Chang H-H, Worby C, Yeka A, Nankabirwa J, Kanya M, Staedke SG, et al. THE REAL McCOIL: A method for the concurrent estimation of the complexity of infection and SNP allele frequency for malaria parasites. *PLoS Comput Biol.* 2017;13(1):e1005348.
34. Janko M, Irish SR, Reich BJ, Peterson M, Doctor SM, Mwandagilirwa MK, et al. The links between agriculture, Anopheles mosquitoes, and malaria risk in children younger than 5 years in the Democratic Republic of the Congo: a population-based, cross-sectional, spatial study. *The Lancet Planetary Health.* 2018;2(2):PE74–E82.
35. Levitz L, Janko M, Mwandagilirwa K, Thwai KI, Likwela JL, Tshetu AK, et al. Effect of individual and community-level bed net usage on malaria prevalence among under-fives in the Democratic Republic of Congo. *Malar J.* 2018;17(1).
36. Chen W, Qian L, Shi J, Franklin M. Comparing performance between log-binomial and robust Poisson regression models for estimating risk ratios under model misspecification. *BMC Med Res Methodol [Internet].* 2018;18(63). Available from: <http://dx.doi.org/10.1186/s12874-018-0519-5>
37. Barros A, Hirakata V. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. *BMC Med Res Methodol [Internet].* 2003;3(21). Available from: <http://dx.doi.org/10.1186/1471-2288-3-21>
38. Bates, Douglas Machler, Martin Bolker, Ben Walker, Steve. Fitting Linear Mixed-Effects Models Using lme4. *J Stat Softw.* 2015;61(1):1–48.
39. Leonard T. Density Estimation, Stochastic Processes and Prior Information. *J R Stat Soc Series B Stat Methodol.* 1978;40(2):113–46.
40. Weiss DJ, Nelson A, Gibson HS, Temperley W, Peedell S, Lieber A, et al. A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature.* 2018 Jan 10;553:333.
41. Noor AM, Alegana VA, Gething PW, Tatem AJ, Snow RW. Using remotely sensed night-time light as a proxy for poverty in Africa. *Popul Health Metr.* 2008 Oct 21;6(1):5.
42. Esch T, Heldens W, Hirner A, Keil M, Marconcini M, Roth A, et al. Breaking new ground in mapping human settlements from space – The Global Urban Footprint. *ISPRS J Photogramm Remote Sens.* 2017 Dec 1;134:30–42.
43. Rahimi A, Recht B. Random features for large-scale kernel machines. In: *Advances in neural information processing systems.* 2008. p. 1177–84.
44. Golding N. greta: Simple and Scalable Statistical Modelling in R [Internet]. 2018. Available from: <https://cran.r-project.org/web/packages/greta/index.html>
45. Nkoli Mandoko P, Rouvier F, Matendo Kakina L, Moke Mbongi D, Latour C, Losimba Likwela J, et al. Prevalence of Plasmodium falciparum parasites resistant to sulfadoxine/pyrimethamine in the Democratic Republic of the Congo: emergence of highly resistant pfdhfr/pfdhps alleles. *J Antimicrob Chemother [Internet].* 2018;(August). Available from: <https://academic.oup.com/jac/advance-article/doi/10.1093/jac/dky258/5057668>
46. Okell LC, Griffin JT, Roper C. Mapping sulphadoxine-pyrimethamine-resistant Plasmodium falciparum malaria in infected humans and in parasite populations in Africa. *Sci Rep.* 2017 Aug 7;7(1):7389.

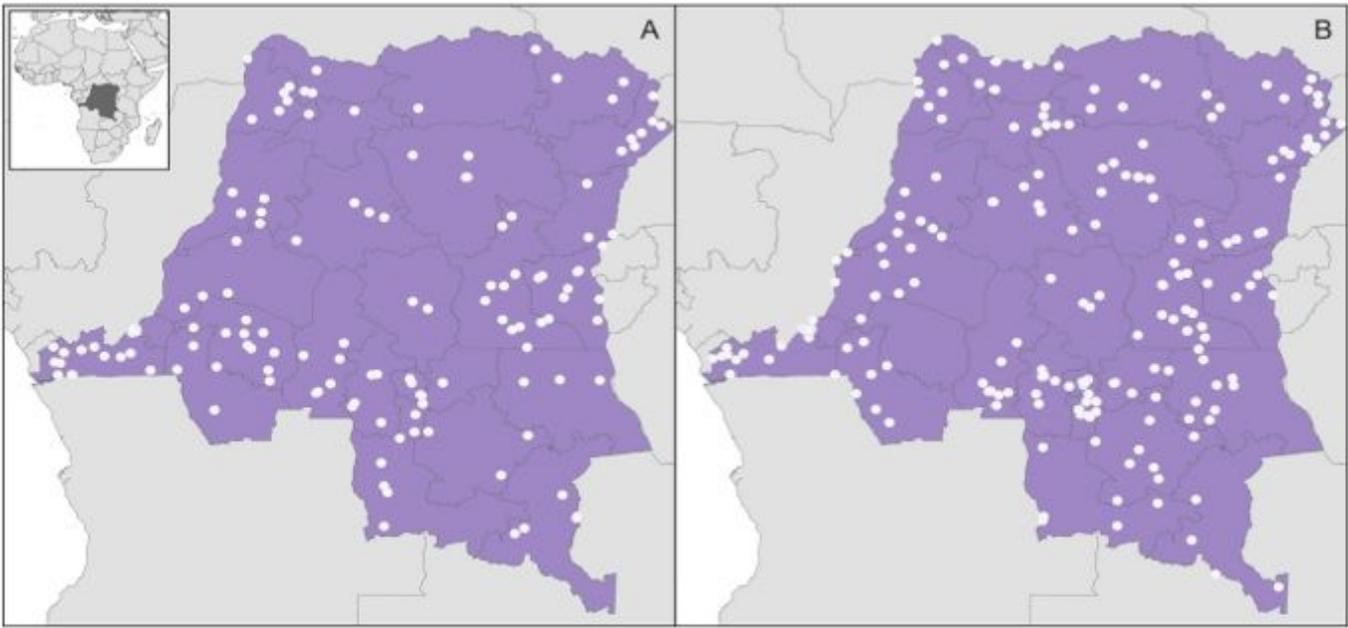
47. Ruh E, Bateko JP, Imir T, Taylan-Ozkan A. Molecular identification of sulfadoxine-pyrimethamine resistance in malaria infected women who received intermittent preventive treatment in the Democratic Republic of Congo. *Malar J.* 2017;17(17).
48. ACTWatch. Epidemiological Profile of the DRC [Internet]. 2015. Available from: <http://www.actwatch.info/countries/democratic-republic-of-the-congo/profile>
49. Wang L, Bwambale R, Keeler C, Reyes R, Muhindo R, Matte M, et al. Private sector drug shops frequently dispense parenteral anti-malarials in a rural region of Western Uganda. *Malar J.* 2018;17(305).
50. Humphreys GS, Merinopoulos I, Ahmed J, Whitty CJ, Mutabingwa TK, Sutherland CJ, et al. Amodiaquine and artemether-lumefantrine select distinct alleles of the *Plasmodium falciparum* *mdr1* gene in Tanzanian children treated for uncomplicated malaria. *Antimicrob Agents Chemother.* 2007;51(3):991–7.
51. Fansidar replaces chloroquine in DR Congo's malaria treatment. *PanaPress.* 2001 Dec 28;
52. Takala-Harrison S, Laufer M. Antimalarial drug resistance in Africa: key lessons for the future. *Ann N Y Acad Sci.* 2013;1342:62–7.
53. Laufer M, Takala-Harrison S, Dzinjalama F, Stine OC, Taylor TE, Plowe C. Return of Chloroquine-Susceptible *Falciparum* Malaria in Malawi Was a Reexpansion of Diverse Susceptible Parasites. *J Infect Dis.* 2012;202(5):801–8.

## Additional Files

File name: BMC\_Additional\_file.doc

Description: Additional description of the modeling methods used in the study as well as additional tables and figures to support our findings.

## Figures



**Figure 1**

DHS cluster locations of the children included in the analysis. Clusters are from 2007 (A) and 2013 (B). The 26 DRC municipal province borders are outlined in black.

2007

2013

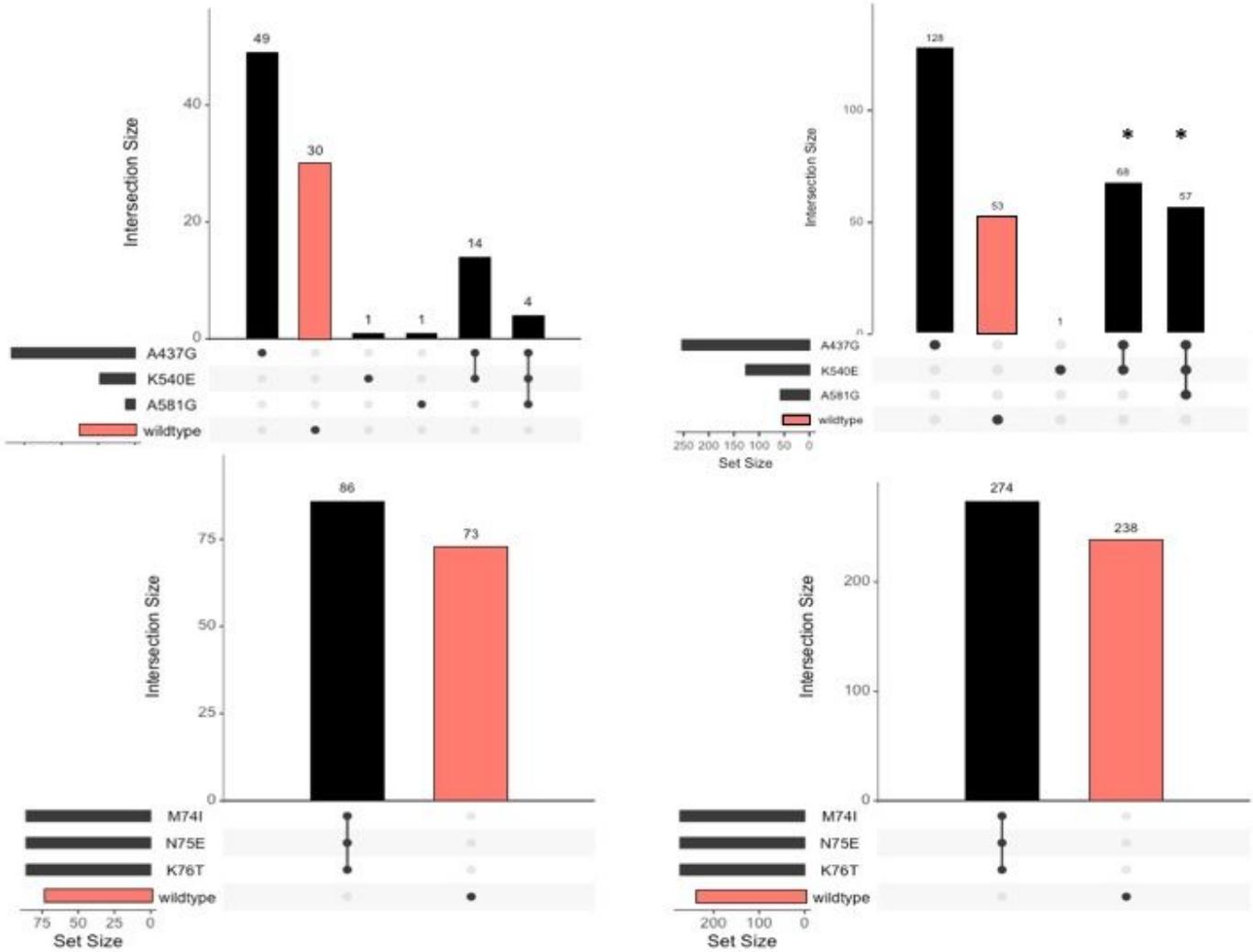
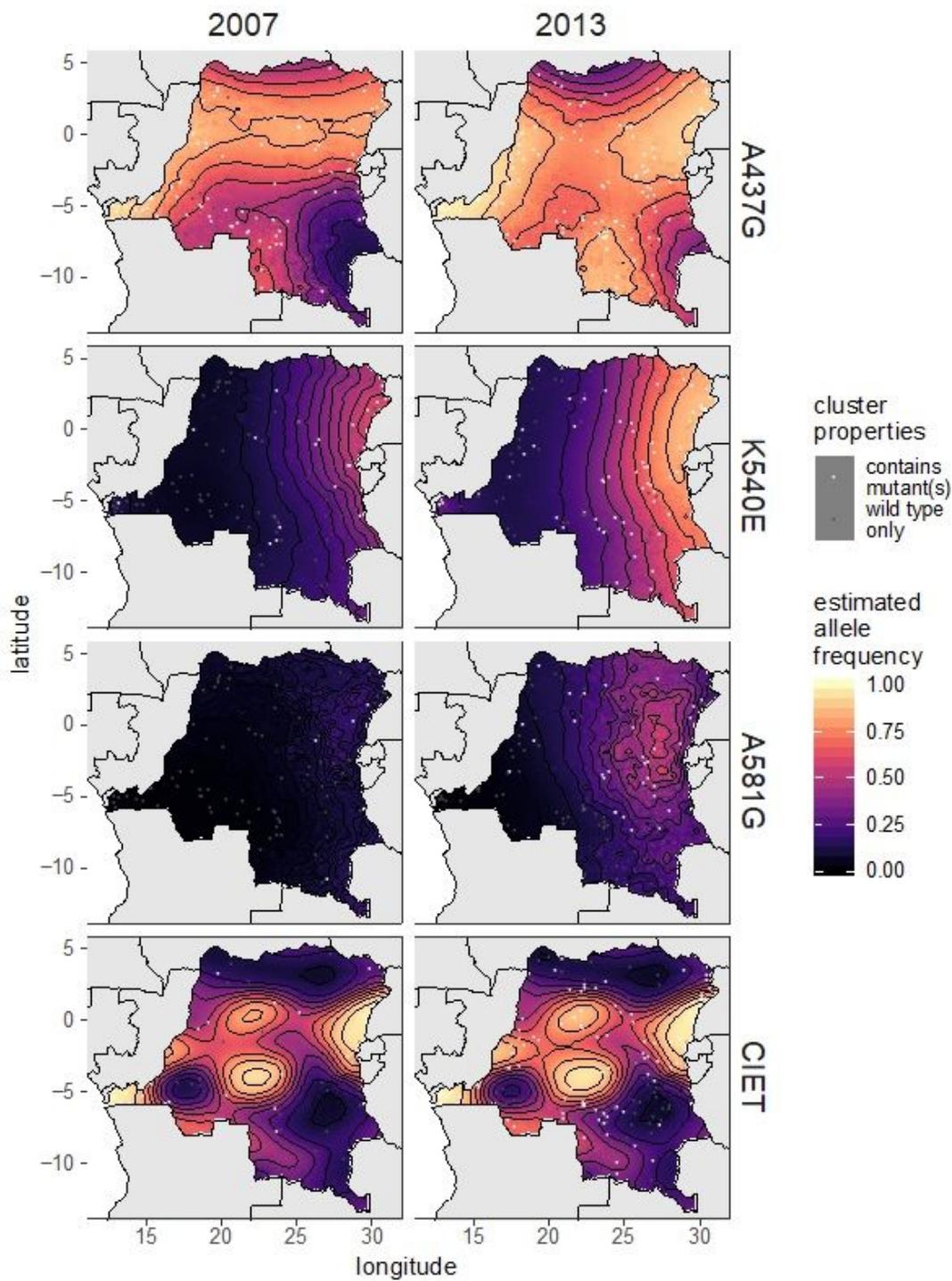


Figure 2

Frequencies of pfdhps and pfCRT mutations in 2007 and 2013. Wild-type genotypes are highlighted in red. Chi-squared tests were performed to statistically compare proportions; asterisks indicate a statistically significant difference in proportion between years.



**Figure 3**

Spatial prediction maps comparing prevalence and spatial distribution of pfdhps and pfcr1 mutations. Predictions were generated for 2007 (left panels) and 2013 (right panels). Clusters with at least one mutation are marked with a white "x", clusters with no mutations are marked in grey circles. Horizontal black lines represent a 5% increase in prevalence.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [BMCAAdditionalFile.pdf](#)