Examining the early distribution of the artemisinin-resistant *Plasmodium falciparum kelch13* R561H mutation in Rwanda

Rebecca Kirby[1*], David Giesbrecht[1*], Corine Karema[2,3], Oliver Watson[4,5], Savannah Lewis[1], Tharcisse Munyaneza[6], Jean de Dieu Butera[6], Jonathan J Juliano[7*], Jeffrey Bailey[1*], Jean Baptiste Mazarati[8*]

[1]Department of Pathology and Laboratory Medicine, Brown University, Providence, Rhode Island 02912, USA
[2]Quality Equity Health Care, Kigali Po Box 6803, Rwanda
[3]Swiss Tropical and Public Health Institute, University of Basel, Basel 4123, Switzerland
[4]MRC Centre for Global Infectious Disease Analysis, Imperial College London, London W2 1PG, UK
[5]London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK
[6]National Reference Laboratory, Rwanda Biomedical Center, Kigali, Rwanda
[7]Department of Epidemiology, University of North Carolina, Chapel Hill, North Carolina 27599, USA
[8]Biomedical Laboratory Sciences, INES-Ruhengeri, Ruhengeri, Rwanda

*Authors contributed equally

1 **Abstract**

2 Artemisinin resistance mutations in *Plasmodium falciparum kelch13* (*Pfk13*) have begun to

3 emerge in Africa. *Pfk13*-R561H was the first reported African mutation found in Rwanda in

4 2014, but limited sampling left questions about its early distribution and origin. We detected 476

5 parasitemias among 1873 residual blood spots from a 2014-15 Rwanda Demographic Health

6 Survey. We sequenced 351 samples revealing 341/351 were wild type (97.03% weighted) and 4

7 samples (1.34% weighted) harbored R561H which were significantly spatially clustered. Our

8 study better defines the early distribution of R561H in Rwanda and suggests that the origin may

9 have involved higher-transmission regions.

10

11 **Introduction**

12   Artemisinin-based combination therapies (ACT) are currently the key antimalarials used

13 in Africa and emerging resistance threatens decades of public health gains. Concerningly,

14 mutations in the *Plasmodium falciparum kelch13* (*Pfk13)* gene causing in vitro artemisinin

15 resistance are emerging in Africa [1,2]. Additionally, recent reports have shown delayed parasite

16 clearance in Rwandan patients infected with R561H mutants after treatment with ACT [3]. While

17 high-grade clinical resistance to ACT remains low in Africa thanks to the continued efficacy of

18 partner drugs, decreased sensitivity to artemisinin is thought to have potentiated partner drug and

19 lead to ACT treatment failure in southeast Asia (SEA), making the emergence of validated *Pfk13*

20 resistance mutations a grave concern [4–6].

21   Early in the emergence of artemisinin resistance in SEA, multiple mutations in *Pfk13*

22 were found with the eventual emergence of C580Y as the dominant genotype. A similar pattern

23 may be occuring in Africa with multiple validated or candidate resistance mutations emerging

24   including R561H, C469F/Y, and A675V [1,2,7,8]. Understanding trends in resistance allele

25   frequencies can only be done through intensive genomic surveillance.

26   The first report of artemisinin resistant *Pfk13* mutants in Africa occurred in Rwanda in

27   2014, with 7.4% of samples collected in Masaka between 2014-15 harboring R561H [1]. By

28   2018, reported R561H prevalence increased to 19.6% in Masaka and 22% in Rukara [9]. While

29   ring-stage survival assays confirmed in vitro artemisinin resistance, these studies found no

30   association between clinical outcomes and *Pfk13* mutations. The expansion of a validated

31   resistance mutation is nevertheless concerning, and the lack of broader sampling in previous

32   studies makes it difficult to trace R561H's spread through the country. To our knowledge, only

33   two *Pfk13* genotyping studies were done in Rwanda between 2010-2015, the years preceding and

34   including the first report of R561H [1,10] (Figure 1A). These studies were limited to a few

35   sampling sites, leaving questions about the true distribution of R561H early in its emergence.

36   Here we used an existing Demographic Health Surveys (DHS) sample set, originally

37   designed as a random representative sample of the Rwandan population for HIV prevalence,

38   conducted between 2014-15. Repurposing this nationally representative dataset provides an

39   opportunity for a nation-wide baseline study of R561H prevalence in Rwanda early in the

40   emergence timeline (Figure 1B). This study also employs a cost-conscious approach for

41   genotyping discarded samples that can be employed in resource-limited settings.

42   **Methods**

43   **Sample collection and preparation**

44   Sample collection for 2014-15 DHS in Rwanda included a subset of participants from whom

45   dried blood spots (DBS) were collected on filter paper for HIV testing. Residual DBSs were

46   provided by Rwanda Biomedical Center. The DHS included 12,699 GPS tagged households in

3

47   492 clusters in 30 regions who completed the initial questionnaire. Of these households, 50%

48   were selected for malaria and anemia testing in women and children, and the other 50% were

49   selected for HIV testing via DBS sample collection (Figure 1C). Complete survey methods are

50   described by the DHS [11]. The final anonymized DHS dataset was downloaded from

51   https://dhsprogram.com. All data cleaning and analysis was done in R version 4.1.2 using

52   packages detailed here: https://github.com/bailey-lab/Rwanda-DHS-2014-15. Analyses utilizing

53   parasite genomes from de-identified samples were deemed nonhuman subjects of research at the

54   University of North Carolina at Chapel Hill (NC, USA) and Brown University (RI, USA).

55   From the DHS survey database, malaria prevalence for each cluster was calculated based

56   on either positive rapid diagnostic test (RDT) or microscopy. Given we could not simply identify

57   positive DBS samples based on RDT testing as they were different groups, we focused on

58   clusters where we would expect to detect a reasonable number of positives to genotype, selecting

59   clusters with >15% malaria prevalence (n clusters = 67, n samples = 2255). Regional malaria

60   prevalence was calculated in R using survey-weighted means, using DHS household survey-

61   weights. We located 1873 of these samples and placed three 6 mm punches into wells of a 2 ml

62   plate. DNA was extracted using Chelex [13].

Amplification of short PCR amplicons including R561H site

64   A two-step PCR was optimized to amplify short amplicons with sample barcodes covering

65   amino acid positions 516-572 on the Pf*k13* gene (Supplementary Methods). PCR step 1 product

66   was resolved on an agarose gel, and the presence of a faintly visible or stronger band was used to

67   select samples for sequencing. First-step primers were based on previous designs [12] and

68   contained molecular inversion probe (MIP) compatible 5' linking sequences, enabling barcoding

69   [13]. Positive samples (n=476) were re-plated using an OT2 robot (Opentrons, Brooklyn, NY,

4

70   USA). PCR step 2 was used to add unique barcodes to each sample, and these products were

71   resolved on a gel.

72   **Library preparation**

73   To reduce over sequencing of high parasitemia samples, 5 uL of each PCR step 2 reaction were

74   sorted into high and low parasitemia pools based on the PCR step 2 band intensity, using an

75   Opentrons OT2. To prepare libraries for sequencing, a 1.2x SPRI bead cleanup was performed

76   on the two pools eluted in 30uL TE low EDTA pH=8 (Thermo Fisher Scientific, Waltham MA,

77   USA). The eluted pools were combined for a total volume of 20 uL, and amplicons were

78   resolved on a 1.0 % agarose gel. Fragments of the correct size were excised and DNA was

79   purified using the NEB DNA Gel Extraction kit T1020S (Ipswich, MA, USA). Library quality

80   was assessed using a fragment analyzer (Agilent, Santa Clara, CA), the library pooled with other

81   libraries and sequenced on an Illumina Nextseq 550 using a 300-cycle mid output kit at 5.3 pM.

82   **Analysis**

83   Samples were demultiplexed using MIPtools software available at https://github.com/bailey-

84   lab/MIPTools and analyzed using SeekDeep pipeline available at https://github.com/bailey-

85   lab/SeekDeep [14]. Samples that had < 50 reads on the first pass of sequencing (n=243) were

86   repooled, gel extracted, and resequenced. Samples with ≥ 20 total combined reads were included

87   in subsequent analysis (n=351). Mutation prevalence was calculated using HIV survey-weights.

88   Spatial correlation of R561H was calculated in GeoDa (https://geodacenter.github.io)

89   using Moran's I with Empirical Bayes to take varying cluster sample sizes into account.

90   Significance was calculated in Geoda with 999 permutations. A 97.5% confidence interval of

91   inferred R561H prevalence was created using a spatial Binomial logistic model.

92   Map figures were created in RStudio using spatial data downloaded from the DHS

93   website.

Results

95   By genotyping parasites from discarded DHS samples, we were able to investigate the

96   distribution of diversity of *Pfk13* mutations in Rwanda early in the emergence of R561H.

97   Short amplicon PCR was an effective way to determine infections

98   Amplification identified 25.4% of *Pf* unknown samples positive (476/1873), similar to the

99   expected 25.7% positivity found by DHS by RDT. On a per cluster level, we found that

100   amplicon success had moderate correlation with the expected *P. falciparum* prevalence (Pearson

101   coefficient = 0.357). However, some clusters fell above or below the expected value based on

102   DHS data due to inherent variance from testing different subsets of individuals such as

103   differences in average age (RDT was 18 years compared to 22 years for PCR). While PCR is

104   normally more sensitive than RDT, sample degradation likely decreased our sensitivity and our

105   assays did not leverage multicopy genes. Additionally, RDTs can be false positives due to the

106   persistence of HRP2 protein despite clearance of the parasite.

107   Targeted deep sequencing of amplicon revealed R561H and other *Pfkelch13* mutations

108   Of the 476 PCR positive samples, 351 (73.7%) were successfully sequenced, revealing a

109   majority of wild type parasites (341/351, 97.03%) (Supplementary Table 1). One sample (0.27%)

110   had a synonymous mutation (G533G) and 9 samples (2.70%) had nonsynonymous K13

111   mutations, including R561H (4), V555A (3), C532W (1), and G533A (1) (Supplementary Table

112   1).

113   The four R561H isolates were spatially clustered across three DHS clusters in the

114   southeast part of the country, representing 5.47% of samples in Kirehe and 2.74% in Ngoma.

6

115   (Figure 2a, Supplementary Figure 1, Supplementary Table 2). The Moran's I value of spatial

116   correlation of clusters with R561H was 0.133 (p=0.046), indicating significant spatial

117   clustering. Additionally, a 97.5% confidence interval of inferred R561H prevalence based on the

118   number of mutant isolates, sampling scheme, and geographic proximity to other isolates

119   highlighted localization in the southeast, higher confidence of low or no prevalence in the

120   southwest, but lower of confidence in unsampled northern regions (Figure 2b).

121   **Discussion**

122   Previous *Pfk13* genotyping in Rwanda in the years leading up to the first observation of R561H

123   failed to capture the larger picture across the country. This study provides expanded geographic

124   sampling and an improved baseline of *Pfk13* genotypes in relatively high-prevalence regions

125   early in the R561H emergence timeline. Previous studies only observed the mutation in Masaka

126   as of 2014, but our study demonstrates that it was also present and significantly clustered in

127   high-transmission regions in the southeast of the country. In keeping with prior reports, we found

128   a small number of clinically unvalidated polymorphisms, three previously not observed in

129   Rwanda as of 2014-15 (Supplementary Table 3).

130   The presence of R561H in three spatially clustered high-transmission clusters outside of

131   previously reported areas could point towards a more southern origin or spread. While additional

132   data is needed to better understand the origin and spread, it is concerning that the resistant

133   parasites were already present in high-transmission areas as of 2014. This was unexpected, as it

134   is hypothesized that low-transmission areas drive selection of resistant parasites, due to low

135   levels of host immunity and reduced within-host competition [15]. A possible explanation of the

136   current findings could be periods of low- and high-transmission due to intermittent indoor

137 residual spraying in these regions. Regardless of cause, the presence of mutant parasites in high-

138 transmission zones is of significant concern for the control of malaria.

139 Though these samples allowed for retrospective genotyping, working with old, discarded

140 samples poses challenges. Some bags were not fully sealed, some had water damage, and the

141 desiccant was exhausted in most sample bags. The DNA was thus variably degraded, forcing us

142 to focus on just a short amplicon containing the R561H mutation site. While this did not capture

143 all mutations of interest in Rwanda (i.e. C469F, P574L, and A675V), these could be pursued

144 with additional amplicons [1,7,10].

145 Relatively low sequencing success can likely be attributed to selection of PCR step 1

146 positive amplicons, as any sign of a band was classified as positive. The presence of non-specific

147 bands could have led to misidentification of positives at this step.

148 The sampling scheme should also be taken into consideration. This method led to

149 samples from older patients, with sequenced samples from patients averaging 22 years. As

150 factors such as pre-existing immunity are thought to have an impact on the evolution of parasites,

151 the relative proportion of mutant parasites may be different from studies with a lower median age

152 such as Uwimana et al 2020, who only took samples from children 1-14 years [1].

153 This approach requires significant manual labor for large studies. Sorting through

154 thousands of samples, punching dried blood spots, and running gels to identify *Plasmodium*

155 positive samples was not a small task. This in practice required us to sample from high malaria

156 prevalence areas to not waste resources and time on unknown samples in low-prevalence areas

157 for little output. This is a key limitation as low-prevalence are historically the emergence sites of

158 resistant parasites, so it is possible that R561H existed in 2014-15 in areas we did not sample.

159 Finally, we did not sample directly from Masaka, where R561H was first observed in 2014.

8

160   However, our sampling scheme included multiple clusters nearby Masaka and did not find

161   R561H. When possible, future national sampling should coordinate HIV and RDT sampling for

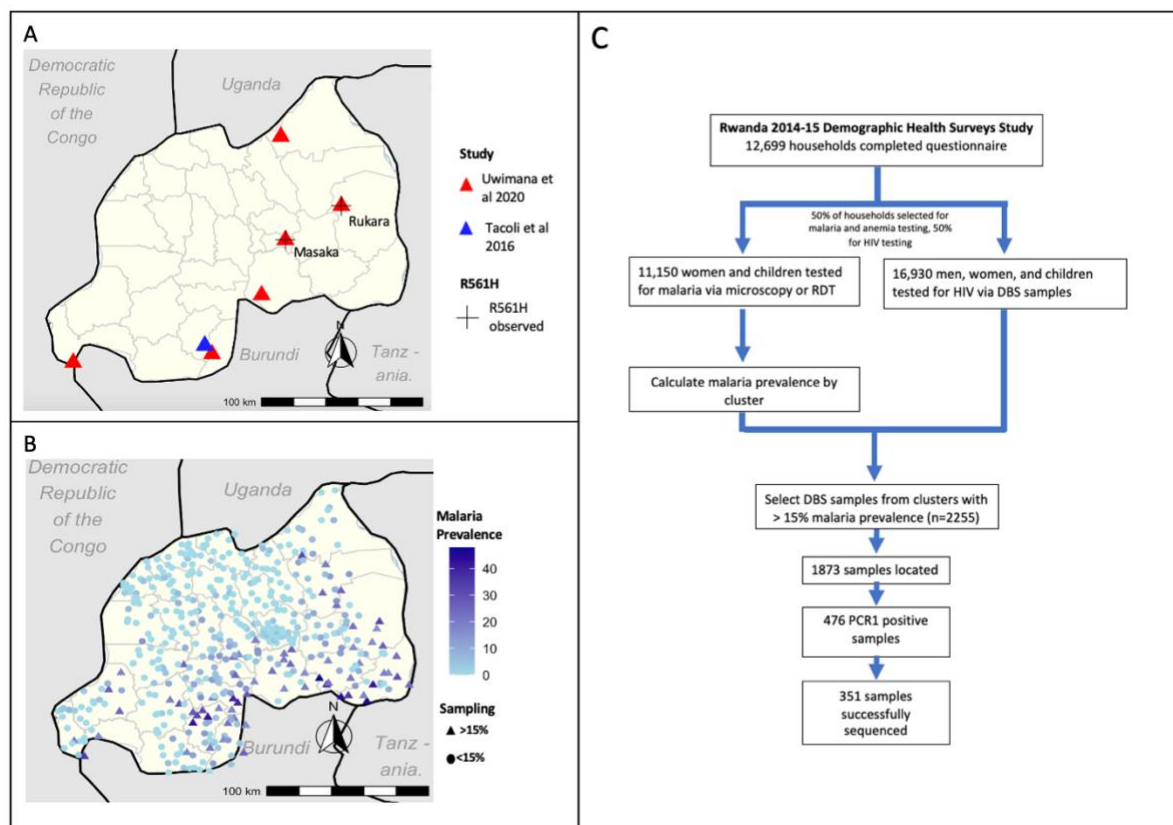162   ease of genotyping residual samples.

163   Our study is the first country-wide examination of early distribution of artemisinin

164   resistant *Pfk13* R561H parasites in Rwanda using samples derived from a representative sample

165   of circulating parasites. This data will allow for better modeling of the initial rate of spread when

166   combined with other retrospective analyses and current surveys.

References

1. Uwimana A, Legrand E, Stokes BH, et al. Emergence and clonal expansion of in vitro artemisinin-resistant Plasmodium falciparum kelch13 R561H mutant parasites in Rwanda. Nat Med. **2020**; 26(10):1602–1608.

2. Asua V, Conrad MD, Aydemir O, et al. Changing Prevalence of Potential Mediators of Aminoquinoline, Antifolate, and Artemisinin Resistance Across Uganda. J Infect Dis. **2021**; 223(6):985–994.

3. Straimer J, Gandhi P, Renner KC, Schmitt EK. High Prevalence of *Plasmodium falciparum* K13 Mutations in Rwanda Is Associated With Slow Parasite Clearance After Treatment With Artemether-Lumefantrine. J Infect Dis. **2022**; 225(8):1411–1414.

4. Ashley EA, Dhorda M, Fairhurst RM, et al. Spread of Artemisinin Resistance in *Plasmodium falciparum* Malaria. N Engl J Med. **2014**; 371(5):411–423.

5. Ariey F, Witkowski B, Amaratunga C, et al. A molecular marker of artemisinin-resistant Plasmodium falciparum malaria. Nature. **2014**; 505(7481):50–55.

6. Takala-Harrison S, Jacob CG, Arze C, et al. Independent Emergence of Artemisinin Resistance Mutations Among Plasmodium falciparum in Southeast Asia. J Infect Dis . **2015**; 211(5):670–679.

7. Bergmann C, Loon W van, Habarugira F, et al. Increase in *Kelch 13* Polymorphisms in *Plasmodium falciparum* , Southern Rwanda. Emerg Infect Dis. **2021**; 27(1):294–296.

8. Loon W van, Oliveira R, Bergmann C, et al. In Vitro Confirmation of Artemisinin Resistance in *Plasmodium falciparum* from Patient Isolates, Southern Rwanda, 2019. Emerg Infect Dis. **2022**; 28(4):852–855.

9.   Uwimana A, Umulisa N, Venkatesan M, et al. Association of Plasmodium falciparum kelch13 R561H genotypes with delayed parasite clearance in Rwanda: an open-label, single-arm, multicentre, therapeutic efficacy study. Lancet Infect Dis. **2021**; 21(8):1120–1128.

10.   Tacoli C, Gai PP, Bayingana C, et al. Artemisinin Resistance–Associated K13 Polymorphisms of Plasmodium falciparum in Southern Rwanda, 2010–2015. Am J Trop Med Hyg. **2016**; 95(5):1090–1093.

11.   National Institute of Statistics of Rwanda, Rwanda, DHS Program, editors. Rwanda demographic and health survey, 2014-15: final report. Kigali, Rwanda : Rockville, Maryland, USA: National Institute of Statistics of Rwanda, Ministry of Finance and Economic Planning : Ministry of Health ; The DHS Program, ICF International; 2016.

12.   Makunin A, Korlević P, Park N, et al. A targeted amplicon sequencing panel to simultaneously identify mosquito species and *Plasmodium* presence across the entire *Anopheles* genus. Mol Ecol Resour. **2022**; 22(1):28–44.

13.   Aydemir O, Janko M, Hathaway NJ, 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**; 218(6):946–955.

14.   Hathaway NJ, Parobek CM, Juliano JJ, Bailey JA. SeekDeep: single-base resolution de novo clustering for amplicon deep sequencing. Nucleic Acids Res. **2018**; 46(4):e21–e21.

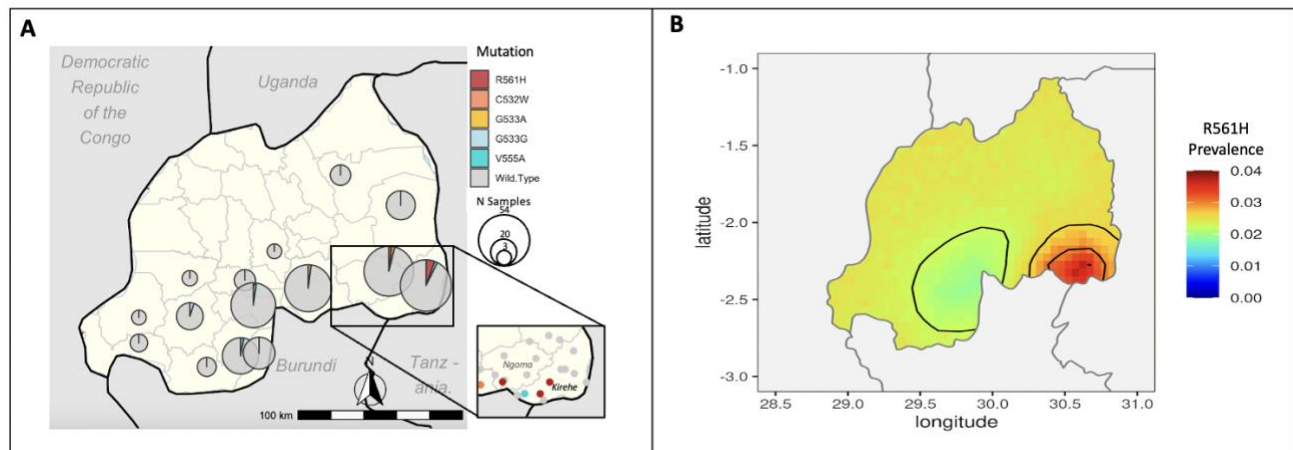15.   White NJ. Antimalarial drug resistance. J Clin Invest. **2004**; 113(8):1084–1092.

Figures

Figure 1: Previous Studies and Our Study Design

A: Sampling sites of P. falciparum genotyping studies done in Rwanda 2010-2015. Sites included Masaka, Ruhuha, Bugarama, Kibirizi, Nyarurema and Rukara (Uwimana et al 2020) and the Huye district 2010-15 (Tacoli et al 2016).

B: DHS 2014 sampling clusters. We sequenced samples from clusters with >15% malaria prevalence.

C: Study design. We used DBS samples from the HIV arm of a 2014-15 DHS study, selecting 2255 samples taken from 67 clusters with >15% malaria, as calculated in the malaria/anemia arm.

12

Figure 2: Geographic distribution of *Pfk13* mutations in high malaria prevalence demographic health survey clusters in Rwanda.

A: Mutations confirmed by sequencing amplicons corresponding to *Pfk13* amino acids 516-572 are presented as pie charts, the center of each corresponding to the center of the DHS survey district. Mutation presence at the DHS cluster level shown for Kirehe and Ngoma districts.

B: The 97.5% confidence interval of inferred R561H prevalence using a spatial Binomial logistic model.

Footnotes

The authors declare no conflicting interests.

This laboratory work and sequencing using residual dried blood spot samples from the Rwanda Demographic and Health Survey 2014-15 was supported by NIH R01AI156267 to JAB.

Sequence files are accessible under the accession number PRJNA885045.

Correspondence should be made to Jeffrey Bailey (jeffrey_bailey@brown.edu) or Jean Baptiste Mazarati (jmazarati@gmail.com).