medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20159624; this version posted October 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Genetic surveillance in the Greater Mekong Subregion and South Asia to support malaria control and elimination

Christopher G Jacob PhD[1], Nguyen Thuy-Nhien PhD[2], Mayfong Mayxay MD[3,4,5], Richard J Maude FRCP[6,5,7], Huynh Hong Quang PhD[8], Bouasy Hongvanthong MD[9], Viengxay Vanisaveth MD[9], Thang Ngo Duc PhD[10], Huy Rekol MD[11], Rob W van der Pluijm MD[6,5], Lorenz Von Seidlein MD[6,5], Rick M Fairhurst PhD[12,#], Francois H Nosten MD[13,5], Md Amir Hossain MD[14], Naomi Park PhD[1], Scott Goodwin MSc[1], Pascal Ringwald MD[15], Keobouphaphone Chindavongsa MD[9], Paul N Newton FRCP[3,6,5], Elizabeth A Ashley MD[3,5], Sonexay Phalivong[3], Rapeephan R Maude MD[6,16], Rithea Leang PhD[11], Cheah Huch MD[11], Le Thanh Dong PhD[17], Kim-Tuyen Nguyen MSc[2], Tran Minh Nhat MD[2], Tran Tinh Hien MD[2], Hoa Nguyen MPM[18], Nicole Zdrojewski MA[18], Sara E Canavati PhD[18], Abdullah A Sayeed MD[14], Didar Uddin MSS[6], Caroline Buckee PhD[7], Caterina I Fanello PhD[6,5], Marie Onyamboko PhD[19], Thomas Peto PhD[6,5], Rupam Tripura MD[6,5], Chanaki Amaratunga PhD[12,^], Aung Myint Thu MSc[13,5], Gilles Delmas MD[13,5], Jordi Landier PhD[13,20], Nguyen Hoang Chau MD[2], Dysoley Lek MD[11], Seila Suon MD[11], James J Callery MBCHB[6,5], Podjanee Jittamala MD[21], Borimas Hanboonkunupakarn MD[21], Sasithon Pukrittayakamee DPhil[21,22], Aung Pyae Phyo MD[23,5], Frank Smithuis MD[23,5], Khin Lin MD[24], Myo Thant MSc[25], Tin Maung Hlaing MD FRCP[25], Parthasarathi Satpathi MD[26], Sanghamitra Satpathi MD[27], Prativa K Behera MD[27], Amar Tripura MD[28], Subrata Baidya MD[28], Neena Valecha MD[29], Anupkumar R Anvikar MD[29], Akhter ul Islam MD[30], Abul Faiz PhD[31], Chanon Kunasol MSc[6], Eleanor Drury BSc[1], Mihir Kekre MSc[1], Mozam Ali[1], Katie Love BSc[1], Shavanthi Rajatileka PhD[1], Anna E Jeffreys MSc[32], Kate Rowlands MSc[32], Christina S Hubbart MSc[32], Mehul Dhorda PhD[33,6,5], Ranitha Vongpromek MSc[33,6], Namfon Kotanan BSc[21], Phrutsamon Wongnak MSc[6], Jacob Almagro Garcia PhD[34], Richard D Pearson PhD[34,1], Cristina V Ariani PhD[1], Thanat Chookajorn PhD[21], Cinzia Malangone MSc[1], T Nguyen MSc[1], Jim Stalker MA[1], Ben Jeffery DPhil[34], Jonathan Keatley BSc[1], Kimberly Johnson BSc[34,1], Dawn Muddyman PhD[1], Xin Hui S Chan DPhil[6,5], John Sillitoe BEng[1], Roberto Amato PhD[1], Victoria Simpson PhD[34,1], Sonia Gonçalves PhD[1], Kirk Rockett PhD[1,32], Nicholas P Day FRCP[6,5], Arjen M Dondorp MD[6,5], Dominic P Kwiatkowski FRS[1,34], Olivo Miotto PhD[1,6,34,*]

[1] Wellcome Sanger Institute, Hinxton, UK
[2] Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam
[3] Lao-Oxford-Mahosot Hospital-Wellcome Research Unit (LOMWRU), Microbiology Laboratory, Mahosot Hospital, Vientiane, Lao PDR
[4] Institute of Research and Education Development (IRED), University of Health Sciences, Ministry of Health, Vientiane, Lao PDR
[5] Centre for Tropical Medicine and Global Health, University of Oxford, Oxford, UK
[6] Mahidol-Oxford Tropical Medicine Research Unit, Mahidol University, Bangkok, Thailand

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20159624; this version posted October 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

[7] Harvard TH Chan School of Public Health, Harvard University, Boston, USA
[8] Institute of Malariology, Parasitology and Entomology (IMPEQN), Quy Nhon, Vietnam
[9] Centre of Malariology, Parasitology, and Entomology, Vientiane, Lao PDR
[10] National Institute of Malariology, Parasitology and Entomology (NIMPE), Hanoi, Vietnam
[11] National Center for Parasitology, Entomology, and Malaria Control, Phnom Penh, Cambodia
[12] National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville MD, USA
[13] Shoklo Malaria Research Unit, Mae Sot, Thailand
[14] Chittagong Medical College Hospital, Chittagong, Bangladesh
[15] World Health Organization, Geneva, Switzerland
[16] Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand
[17] Institute of Malariology, Parasitology and Entomology (IMPEHCM), Ho Chi Minh City, Vietnam
[18] Vysnova Partners, Inc., Hanoi, Vietnam
[19] Kinshasa School of Public Health, University of Kinshasa, Kinshasa, Democratic Republic of Congo
[20] Aix-Marseille Université, INSERM, IRD, SESSTIM, Marseille, France
[21] Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand
[22] The Royal Society of Thailand, Bangkok, Thailand
[23] Myanmar-Oxford Clinical Research Unit, Yangon, Myanmar
[24] Department of Medical Research, Pyin Oo Lwin, Myanmar
[25] Defence Services Medical Research Centre, Yangon, Myanmar
[26] Midnapore Medical College, Midnapur, India
[27] Ispat General Hospital, Rourkela, India
[28] Agartala Medical College, Agartala, India
[29] National Institute of Malaria Research, Indian Council of Medical Research, New Delhi, India
[30] Ramu Upazila Health Complex, Cox's Bazar, Bangladesh
[31] Malaria Research Group and Dev Care Foundation, Dhaka, Bangladesh
[32] Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK
[33] Worldwide Antimalarial Resistance Network (WWARN), Asia Regional Centre, Bangkok, Thailand
[34] MRC Centre for Genomics and Global Health, Big Data Institute, Oxford University, Oxford, UK

[#]Current affiliation: AstraZeneca, Gaithersburg, MD 20878, USA

[^]Current affiliation: Mahidol-Oxford Tropical Medicine Research Unit, Mahidol University, Bangkok, Thailand; Centre for Tropical Medicine and Global Health, University of Oxford, Oxford, UK

[*]Address correspondence to Olivo Miotto (olivo@tropmedres.ac)

Running title: Genetic surveillance of malaria in the Greater Mekong Subregion

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20159624; this version posted October 3, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

ABSTRACT

The use of parasite genetic data by National Malaria Control Programmes (NMCPs) is currently limited, and typically focused on specific genetic features or a small number of study sites. We have developed GenRe-Mekong, a platform for genetic surveillance of malaria in the Greater Mekong Subregion (GMS). By integrating simple sample collection procedures in the routine operations of public health facilities, GenRe-Mekong enables NMCPs to conduct large-scale surveillance project in endemic regions. Samples are processed by the SpotMalaria platform, which uses high-throughput technologies to produce a broad set of genotypes, including most known drug resistance markers, species markers and a genomic barcode. Through the application of heuristics based on published evidence, GenRe-Mekong delivers Genetic Report Cards, a compendium of genotypes and phenotype predictions that are used to map prevalence of resistance to multiple drugs. To date, GenRe-Mekong has worked with NMCPs in five countries, and with several large-scale research projects, processing 9,623 samples from clinical cases. The monitoring of resistance markers has been especially valuable for NMCPs tracking the recent rapid spread of DHA-piperaquine resistant parasites across the region. In Vietnam and Laos, data from GenRe-Mekong have provided novel knowledge about the spread of these resistant strains in provinces previously thought to be unaffected. GenRe-Mekong facilitates data sharing by aggregating at regional level results from different countries, providing cross-border views of the spread of resistant strains.

Funding

Bill & Melinda Gates Foundation, Wellcome Trust, UK Medical Research Council, UK Department for International Development, NIAID

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20159624; this version posted October 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

INTRODUCTION

In low-income countries, particularly in sub-Saharan Africa, malaria continues to be a major cause of mortality, and intense efforts are underway to eliminate *Plasmodium falciparum* parasites, which cause the most severe form of the disease. However, *P. falciparum* has shown a remarkable ability to develop resistance to antimalarials, rendering therapies ineffective and frustrating control and elimination efforts. This problem is most acutely felt in the Greater Mekong Subregion (GMS), a region of relatively low malaria prevalence and mortality that has repeatedly been the origin of drug resistant strains,[1-5] and in neighbouring countries including Bangladesh and India, where resistance could be imported. In the past, drug resistance alleles spread multiple times from the GMS to Africa, rolling back progress against the disease at the cost of many lives.[6,7] In view of the recent regional emergence of parasite strains resistant to artemisinin[1,8,9] and its partner drug piperaquine,[10-13] the elimination of *P. falciparum* from the GMS has become a global health priority. Elimination from this region presents significant challenges and, to ensure the most effective outcomes, NCMPs have to evaluate multiple changing factors: efficacy of frontline treatments, available alternatives, routes of spread, location of transmission hubs, importation of cases, and so on. In these assessments, NMCPs make extensive use of clinical and epidemiological data, such as those from routine clinical reporting and therapy efficacy studies. Parasite genetic data is less frequently available, and typically restricted to single genetic variants,[14] or small numbers of sites where quality sample collection protocols could be executed.[15] However, the increased affordability of high-throughput sequencing technologies now offers new opportunities for public health to leverage on big data to optimize interventions where resources are limited.[16] Cost-effective implementation of genomic technologies, aimed at supporting public health decision-making, can make important contributions to malaria elimination.[17]

Here, we describe GenRe-Mekong, a genetic surveillance project conceived to provide public health experts in the GMS with timely and actionable knowledge, to support their decision-making in malaria elimination efforts. GenRe-Mekong analyzes small dried blood spots samples, which are easy to collect at public health facilities from patients with symptomatic malaria, and uses high-throughput technologies to extract large amounts of parasite genetic information from each sample. The results are captured in Genetic Report Cards (GRCs), datasets regularly delivered to NMCPs to keep them abreast of rapid epidemiological changes in the parasite population. The underlying technological platform is designed for low sample processing costs, promoting large-scale genetic epidemiology surveys with dense geographical coverage and large sample sizes.

To date, GenRe-Mekong has worked with NMCPs in Cambodia, Vietnam, Lao PDR (Laos), Thailand and Bangladesh, and has supported large-scale multisite research and elimination projects across the region.[11,18-20] The project has processed 9,623 samples from eight countries, delivering data to the 12 studies that submitted samples. In its initial phase, GenRe-Mekong has focused on applications relevant to the urgent problem of drug resistance. To facilitate integration into NMCP decision-making workflows, our analysis pipelines translate genotypes into predictions of drug resistance phenotypes, and present these as maps which are easily interpreted by public health officials with no prior training in genetics. In Laos and Vietnam, where GenRe-Mekong is implemented in dozens of public health facilities in endemic provinces, results from GenRe-Mekong

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20159624; this version posted October 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

have been used by NMCPs in assessments of frontline therapy options and resource allocation to combat drug resistance.

GenRe-Mekong protects individual patient privacy, while encouraging aggregation and sharing of standardized data across national borders to answer regional questions about epidemiology, gene flow, and parasite evolution.[21] Aggregated data from multiple studies within GenRe-Mekong have powered large-scale genetic and clinical studies of resistance to dihydroartemisinin-piperaquine (DHA-PPQ), revealing a regional cross-border spread of specific strains.[11,21] To power such high-resolution genetic epidemiology analyses of population structure and gene flow, GenRe-Mekong conducts whole-genome sequencing of selected high-quality samples, contributing to the open-access MalariaGEN Parasite Observatory (www.malariagen.net/resource/26).[22] In this article, we summarize some key results from GenRe-Mekong, highlighting how they are used by public health to improve interventions. The data used in this paper are openly available, together with detailed methods documentation and details of partner studies, at www.malariagen.net/resource/29.

RESULTS

Collaborations, Site selection and Sample Collections

As of August 2019, GenRe-Mekong has partnered with NMCPs in five countries to conduct large-scale genetic surveillance (Vietnam, Laos), smaller-scale pilot projects (Cambodia, Thailand) and epidemiological surveys (Bangladesh). GenRe-Mekong also worked with large-scale research projects investigating drug efficacy and malaria risk, or piloting elimination interventions. A total of 9,623 samples from eight countries have been processed in this period (Supplementary Tables 1 and 2). The majority of samples (n=6,905, 72%) were collected in GMS countries (Vietnam, Laos, Cambodia, Thailand, Myanmar), but GenRe-Mekong also supported projects submitting samples from Bangladesh, India and DR Congo. The vast majority of processed samples were collected prospectively, under partnership agreements with GenRe-Mekong (n=9,002, 93.5%, Supplementary Figure 1); two research projects submitted retrospective samples collected in the period 2012-2015 (n=621, 6.5%). Approximately 59% of samples (n=5,716) were submitted by NMCP partnerships, whose contribution increased over time as surveillance projects ramped up (43.4% in 2016, vs. 94.6% in 2018, Supplementary Figure 2). Details of the partnerships, the nature of the studies conducted and the number of processed samples are given in Table 1.

Partnerships with NMCPs are often supported through collaborations with local malaria research groups, which provide support in implementing sample collections, and assist in the interpretation of results. To facilitate implementation in public health infrastructures, GenRe-Mekong provides template study protocols and associated documents; standardized kits of collection materials and documentation; and training for field and health centre staff. Study protocols are adapted to harmonize with local practices, and then approved by both a local ethical review board and the Oxford Tropical Research Ethics Committee (OxTREC). Informed consent forms and participant information sheets are translated to the local language(s), and public health facility staff are trained to execute sample collection procedures. Collection sites are mostly district-level or subdistrict-level health facilities, selected by NMCPs to cover the most informative endemic areas, often based on reported prevalence (Figure 1). Research studies and elimination projects included in their study

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20159624; this version posted October 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

protocol a sample collection procedure compatible with the standard GenRe-Mekong procedure, and sites were selected based on the study's requirements.

Sample processing and Genotyping

GenRe-Mekong samples consist of dried blood spots (DBSs) on filter paper. DNA extracted from the samples was selectively amplified[23] to increase the proportion of parasite DNA and reduce human DNA contamination before genotyping (see Methods). The production of genetic report cards involves genotyping different types of variants: single nucleotide polymorphisms (SNPs), haplotypes, copy number variations and gene domain sequences. These operations were performed by SpotMalaria, the genotyping platform underpinning GenRe-Mekong, whose implementation evolved during the course of the project; details of the methods used in different versions are provided in the Supplementary Materials. In the initial phase, SpotMalaria used a mixture of technologies: capillary sequencing of the *kelch13* gene to detect SNPs associated with artemisinin resistance;[8,24] and high-throughput mass spectrometry to genotype SNP variants. This was later replaced with an amplicon sequencing process, based on short-read deep sequencing of specific portions of the parasite genome, supporting a high degree of multiplexing (see Methods). A total of 3,473 samples (36%) were processed by the amplicon sequencing platform, which delivered a higher genotyping success rate than the earlier process (94% vs 82% mean success rate for genetic barcode positions).

The vast majority of samples were taken from malaria patients upon admission (92%, n=8,866). The remainder were from recurrent clinical episodes, or collected as part of post-admission time series to study infection dynamics (n=757, 7.9%), and were excluded from epidemiological analyses in order to minimize biases and avoid duplicates. Genotypes at mitochondrial positions provided confirmation of the infecting parasite species: *P. falciparum* (*Pf*), *P. vivax* (*Pv*), *P. knowlesi* (*Pk*), *P. malariae* (*Pm*) and *P. ovale* (*Po*). All five species were detected in our dataset: non-Pf parasites were found in 8.8% of samples (n=745 out of 8,486 samples for which species could be determined). A proportion of samples (n=414, 4.9%) only tested positive for non-*Pf* species, possibly due to misdiagnosis or extremely low *Pf* parasitaemia, and were excluded from epidemiological analyses. *Pv* was the most commonly detected non-*Pf* species (317 *Pf*/*Pv* mixed infections, and 405 *Pv*-only infections), followed by *Pk* (11 *Pf*/*Pk* and 6 *Pk*-only infections), while *Pm* and *Po* were detected in three and two samples respectively.

Genetic Barcodes

GenRe-Mekong produces a *genetic barcode* for each sample to enable analyses of relatedness, diversity, multiplicity of infection and population structure. Genetic barcodes are constructed by concatenating the alleles at 101 SNPs distributed across all nuclear chromosomes (see Methods), chosen on the basis of their geographically widespread variability and their power to recapitulate genetic distance. Genetic barcodes can be used to detect loss of diversity due to demographic effects,[25] or to compare parasites from the same patient to distinguish recrudescences from reinfections.[26] They can also produce estimates of genetic distance, which can discriminate between populations on the regional scale. For example, a neighbour-joining tree derived from genetic distance estimates from barcodes (Figure 2) clearly separates parasites from the Thai-Myanmar border region from those circulating along the Thai-Cambodian border, consistent with findings from WGS analyses.[27] Hence, while genetic barcodes produce lower resolution results than WGS data,

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20159624; this version posted October 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

they can be used for rapid low-cost detection of candidate imported parasites, to be further analysed using higher-definition approaches. We used genetic barcode results to discard 827 samples (9.8%) that failed to produce barcodes due to low *Pf* DNA content, yielding a final set of 7,626 *Pf* samples for epidemiological analyses.

Survey of drug resistance mutations

GenRe-Mekong produces genotypes covering a broad range of known variants associated to drug resistance (Table 2) to support assessment of the spread and risk of drug resistance. These include single nucleotide polymorphisms (SNPs) in genes *kelch13*, *arps10*, *ferredoxin*, *mdr2* (resistance to artemisinin),[24,27] *crt* (chloroquine), *mdr1* (multiple drugs), *dhfr* (pyrimethamine), *dhps* (sulfadoxine), as well as a marker amplification breakpoint sequence in *plasmepsin 2/3* (resistance to piperaquine).[28] The interpretation of these genetic markers in phenotypic terms requires extensive knowledge of relevant literature, which is sometimes outside the domain of expertise of public health officers. To bridge this gap, we used genotypes to derive *predicted phenotypes* based on a set of rules (see Methods) derived from peer-reviewed publications. These rules classify samples as *resistant* or *sensitive* to a particular drug or treatment, or *undetermined*. Several samples had missing genotype calls which were required for phenotype prediction; therefore, we also devised a number of rules for *imputation* of missing genotypes based on information from linked alleles. These imputation rules (see Methods) are based on an analysis of allele associations using data from over 7,000 samples in the MalariaGEN *Pf* Community Project,[22] and are applied prior to phenotype prediction rules. Phenotypic predictions allow simple estimations of the proportions of resistant parasites at the population level, which can be readily tabulated and mapped for use in public health decision-making. By aggregating sample data at various geographic levels (site, district, province, region, country), GenRe-Mekong delivers to NMCPs maps that capture the current drug resistance landscape, and can be compared to detect changes over time. Most GenRe-Mekong maps use intuitive “traffic light” colour schemes, in which red signifies presence of resistance, and green its absence. Below, we illustrate some results at regional level for the GMS and nearby countries, which are also summarized in Table 3.

The spread of artemisinin resistance (ART-R) is an urgent concern in the GMS. We estimated ART-R frequencies based on the presence of nonsynonymous mutations in the *kelch13* gene, as listed by the World Health Organization.[29] The resulting map indicates that ART-R has reached very high levels in the lower Mekong region (Cambodia, northeastern Thailand, southern Laos and Vietnam), nearing fixation in Cambodia and around its borders, with the exception of very few provinces of Laos and the Vietnam coast (Figure 3A). ART-R frequency declines to the west of this region, with no resistant samples in India and Bangladesh detected in this study, providing no evidence of spread beyond the GMA. An analysis of the distribution of *kelch13* ART-R alleles (Supplementary Figure 3, Supplementary Table 3) reveals a marked difference between the lower Mekong region, where the *kelch13* C580Y mutation is the dominant allele, and the region comprising Myanmar and western Thailand, where a wide variety of non-synonymous *kelch13* variants are found, and C580Y is not dominant. This reflects a recent increase of C580Y mutant prevalence in Cambodia and neighbouring regions, resulting from the rapid spread of the KEL1/PLA1 strain of multidrug-resistant parasites.[21,30] This hard selection sweep has replaced a variety of ART-R alleles previously present in that region, resulting from multiple soft sweeps;[27,31] this process has not occurred along the Thai-Myanmar

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20159624; this version posted October 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

border, where allele diversity is still very pronounced. The spread of the KEL1/PLA1 strain in the lower Mekong region is further observed when we map the frequency of *plasmepsin 2/3* amplifications conferring piperaquine resistance (PPQ-R, Supplementary Figure 4), showing they occur where C580Y are most prevalent. Mapping the combined presence of C580Y and *plasmepsin 2/3* amplification shows that KEL1/PLA1 spread is confined to a well-defined area of the lower Mekong region, and the strain has not made its way into provinces of Laos and Vietnam where ART-R and PPQ-R alleles are circulating independently (Figure 3B). Over time, GenRe-Mekong will continue track the spread of KEL1/PLA1 across the region.

Resistant populations can revert to sensitive haplotypes after drugs are discontinued, as was the case for chloroquine-resistant parasites in East Africa.[32,33] To help detect similar trends in the GMS, GenRe-Mekong reports on markers of resistance to previous frontline antimalarials that have been discontinued because of reduced efficacy. The resulting data show that, decades after the replacement of chloroquine as frontline therapy, the frequency of resistant parasites (CQ-R) remains exceptionally high across the GMS (Supplementary Figure 5). The reasons for such sustained levels of resistance are unclear; the continued use of chloroquine as frontline treatment for *P. vivax* malaria could be a major contributing factor. Similarly, we found high levels of the *dhfr* and *dhps* markers associated with resistance to sulfadoxine-pyrimethamine (SP, Supplementary Figures 6 and 7). It is unclear why resistance to SP is so widespread, several years after discontinuing this therapy in the GMS, although similar results have been seen in Malawi[34]. It is interesting that resistance is lowest in India, where SP is still used with artesunate as the frontline ACT.[35]

Case Study: Vietnam

In Vietnam, sample collections were carried out by two NMCP institutes (IMPEQN and NIMPE), covering approximately 70 sites in seven provinces. Genetic report cards were delivered to public health officials over two malaria seasons, communicating new findings for malaria control. Prior to the surveillance activity, evidence of artemisinin resistance had been found in the provinces of Binh Phuoc, Gia Lai, Dak Nong, Khanh Hoa and Ninh Thuan province.[36] GenRe-Mekong data confirmed the presence of resistant parasites in these provinces, and showed that the province of Dak Lak also has extremely high levels of ART-R (Supplementary Figure 8A). Furthermore, our data showed that nearly all resistant parasites collected near the border with Cambodia belonged to the KEL1/PLA1 strain, carrying the *kelch13* C580Y mutation (Supplementary Figure 9) and *plasmepsin 2/3* amplification (Supplementary Figure 8B-C) .[21,30] C580Y parasites were also found in the coastal provinces of Ninh Thuan, Khanh Hoa and Quang Tri, but they did not carry the *plasmepsin 2/3* amplification; it is therefore likely they were introduced by an earlier sweep of ART-R parasites. Several parasites in Khang Hoa carried the *kelch13* P553L mutation, previously associated with an ART-R founder population in Binh Phuoc province,[27,37] supporting the hypothesis they belong to an earlier sweep.

Data from consecutive seasons offers a view of the dynamics of drug resistance spread. In the 2018/2019 season, there was a marked increase in the number of cases in the Krong Pa district of Gia Lai province (Figure 4). In 2017/2018, this district accounted for 15% of cases in the three central provinces that border with Cambodia (n=96 of 656); the following season, this increased to 64% (n=341 of 529, $p<10^{-15}$). In the same timeframe, KEL1/PLA1 parasites in Krong Pa rose from 65%

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20159624; this version posted October 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

(n=40 of 62) to 98% (n=298 of 305, $p<10^{-14}$). These results suggest that an outbreak occurred in this district in 2018/2019, underpinned by strong selection of KEL1/PLA1 genetic background, because of its ability to survive the frontline ACT DHA-PPQ.

Case Study: Laos

The Lao NMCP implemented genetic surveillance in five provinces of southern Laos, at over 50 public health facilities. Artemisinin resistant parasites were found in all five provinces, at frequencies higher in districts bordering Thailand and Cambodia (Figure 5A). The *kelch13* C580Y mutation was found in four of the five provinces, and was the most common ART-R allele (Supplementary Figure 10). However, parasites carrying both C580Y and the *plasmepsin 2-3* amplification were restricted to the two southernmost provinces (Champasak and Attapeu, referred to as "Lower Zone", Figure 5B), and completely absent from Savannakhet and Salavan provinces ("Upper Zone") where C580Y parasites lack the PPQ-R amplification. In other words, it appears that KEL1/PLA1 parasites, possibly imported from Cambodia or Thailand, have migrated into the Lower Zone but not the Upper Zone, where a different population of ART-R parasites circulates.

Given the very recent aggressive spread of KEL1/PLA1, it is likely that Upper Zone parasites are remnants of a previous ART-R sweep which may also have spread from the south, as suggested by the higher frequency in Salavan province than in Savannakhet. To confirm the presence of distinct ART-R populations, we used genetic barcodes to construct a tree that recapitulates population structure in Laos (Figure 6), which clearly separates Upper Zone and Lower Zone parasites. In this tree, KEL1/PLA1 parasites form a large, tight cluster clearly separated from the *kelch13* wild-type samples from the Upper Zone. The Upper Zone C580Y mutants, cluster separately from both these groups, and appear more similar to some C580Y mutants from the Lower Zone which are not PPQ-R, corroborating the hypothesis that Upper Zone mutants migrated from the South. It is likely that the northward spread of KEL1/PLA1 has been contained by the use of artemether-lumefantrine in Laos, which diminishes the survival advantage of PPQ-R parasites. However, the spread of KEL1/PLA1 across the Lower Zone, probably displacing previous ART-R strains, suggests that it is a well-adapted artemisinin-resistant strain, highly competitive even in the absence of piperaquine pressure.

DISCUSSION

GenRe-Mekong provides a genetic surveillance platform suitable for endemic regions of low- and middle-income countries, which delivers to NMCPs detailed knowledge about the genetic epidemiology of malaria parasites, to support decision-making. Pilot studies have been conducted in all GMS countries, with the Vietnam and Laos NMCPs having implemented GenRe-Mekong on a long-term basis. GenRe-Mekong has multiple features that facilitate NMCP engagement: a sample collection procedure that easily integrates with standard medical facility workflows; standardized protocols and training to support implementation; clear presentation of results, including translation to phenotype predictions, to provide intuitive understanding and rapid communication; and support by our regional analysis team and local partners to deliver and discuss findings. GenRe-Mekong has also worked closely with research projects, contributing to their analyses of the genotyping data and supporting publication of key findings. The genetic data produced were valuable for a wide range of

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20159624; this version posted October 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

research applications, such as clinical studies of drug efficacy,[11] evaluation of elimination interventions,[20] and epidemiological investigation of malaria importation.[19]

Collaborations with public health have rapidly translated into real impact for malaria control, especially where GenRe-Mekong has been implemented over multiple seasons. Genetic surveillance results were used by the Vietnam NMCP and Ministry of Health in reviews of national drug policy, which have led to revisions of frontline therapy for three provinces including Dak Lak, where an early report by GenRe-Mekong in 2018 was the first evidence of ART-R, confirmed by treatment failure data from *in vivo* therapy efficacy studies (TES) in 2019. In addition, our report of a KEL1/PLA1 outbreak in Gia Lai province has alerted authorities to the need to review the use of DHA-PPQ in that province. In Laos, authorities have been equally responsive, using GenRe-Mekong reports in their review of frontline therapy choices: the Ministry of Health opted against adopting DHA-PPQ based on our evidence of the expansion of KEL1/PLA1 in the Lower Zone of southern Laos. The impact has not been limited to the national level: data shared by surveillance and research projects participating in GenRe-Mekong has powered regional large-scale epidemiological analyses in the GMS and beyond, revealing patterns of spread and evolution of multidrug-resistant malaria.[11] By combining results from areas populated by multidrug resistant strains with those from countries where these strains could potentially spread, such as Bangladesh and India, GenRe-Mekong maps support risk assessment and preparedness. GenRe-Mekong will continue to encourage public sharing to increase the value of genetic data generates, while respecting patient anonymity and giving recognition to those who contributed to the project.

A major advantage of genetic surveillance, compared to more costly clinical studies, is the potential for dense coverage across all endemic areas, which can identify important spatial heterogeneities across the territory. Our data suggests that a single efficacy study in Savannakhet province could have convinced authorities that DHA-PPQ could be used in Laos, with disastrous effects in the southernmost provinces. Similarly, the adoption of DHA-PPQ in Thailand was based on the drug's efficacy in the western provinces; unfortunately, this facilitated the spread of KEL1/PLA1 into the northeast of the country. The detailed knowledge generated by GenRe-Mekong may avert such chains of events in the future.

The SpotMalaria genotyping platform is designed for extensibility, and has been expanded twice in the course of the project: to test for the newly discovered marker for the *plasmepsin 2/3* amplification,[28] and to add new mutations in *crt* which are associated to higher levels of piperaquine resistance in KEL1/PLA1 parasite.[21,38,39] Such improvements will continue as new markers are identified, and new techniques developed. However, there are newer drugs such as pyronaridine, and established drugs such as lumefantrine and amodiaquine, for which clinical drug resistance markers are yet to be identified. GenRe-Mekong will support the identification of new markers in practical ways, by performing WGS on selected surveillance samples, and contributing these data to public repositories to study epidemiological effects, such as reductions in diversity, increases in cases and founder populations,[31] and identify genomic regions under selection that may lead to discovering new markers. As the project develops, Genetic Report Cards will be expanded, to address new public health use cases, including those not directly related to drug resistance. For example, genetic barcodes and WGS data can be used to detect imported cases; to distinguish

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20159624; this version posted October 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

recrudescences from reinfections; and to measure connectedness between sites, and routes of spread.[19]

In the future, the integration of genetic surveillance data in public health decision-making processes will be a major focus for GenRe-Mekong, which will be addressed in several ways. First, we will make available online platforms for selecting, visualizing and retrieving genetic epidemiology data, which will provide customized views of the data. Second, we will integrate with public health information systems, such as NMCPs' dashboards, at both national and international level. This includes sharing GenRe-Mekong data through the World Health Organization's data visualization platform, Malaria Threats Map (http://apps.who.int/malaria/maps/threats/). Third, we will provide training and support to expand in-country expertise, developing local capacity to evaluate drug resistance data and other outputs that GenRe-Mekong will deliver in the future. Finally, we will promote in-country implementations of the SpotMalaria amplicon sequencing platform that underpins the system, to enable faster turnaround times and long-term self-sufficiency. As the adoption cycle continues, we envisage that a growing global network of public health experts will leverage on genetic surveillance to maximize the impact of their interventions, and accelerate progress towards malaria elimination.

METHODS

Additional detailed documentation on the methods used in this study is available from the article's Resource Page, at www.malariagen.net/resource/29.

Sample Collection

GenRe-Mekong samples were collected and contributed by independent studies with different goals, geographical coverage, and sampling strategies. Studies were managed by a local partner, such as a NMCP or a research organization, and often supported by a local technical partner. Most sampling sites were district or subdistrict health centres or provincial hospitals, selected by the local partner according to their public health or research needs. Each site was assigned a code, and its geographical coordinates recorded to support result mapping. GenRe-Mekong uses a common genetic surveillance study protocol covering the entire GMS, which can be locally adapted; this protocol was used for NMCP surveillance projects, after obtaining approval by a relevant local ethics review board and by the Oxford University Tropical Research Ethics Committee (OxTREC). Research studies included in their own protocol provisions for sample collection procedures, informed consent, patient privacy protection, and data sharing compatible with those in the GenRe-Mekong protocol, and obtained ethical approval from both a relevant local ethics review board, and their relevant institutional research ethics committee.

Samples were collected from patients of all ages diagnosed with *P. falciparum* malaria (including patients co-infected by other *Plasmodium* species) confirmed by positive rapid diagnostic test (RDT) or blood smear microscopy. Participation in the study required written informed consent by patient, parent/guardian, or legally authorised representative (plus patient assent wherever required by national regulations), with the exception of Laos, where the Ministry of Health classified GenRe-Mekong as a surveillance activity for national benefit, requiring no additional informed consent. After obtaining consent, and before administering treatment, three 20µL dried blood spots (DBS) on

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20159624; this version posted October 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

filter paper were obtained from each patient through finger-prick. GenRe-Mekong supplied study sites with kits containing all necessary materials, including strips of Whatman 31ET CHR filter paper, disposable lancet, 20 µl micropipette, cotton swab, alcohol pad, and plastic bag with silica gel for DBS storage. Scannable barcode stickers with unique identifiers were applied on the filter paper, the sample manifest where the collection date was recorded, and the site records. Samples were identified by means of these anonymous barcodes, and no patient identifying information or clinical data were collected by GenRe-Mekong.

A number of participating studies also collected an optional anonymous questionnaire, to capture location of abode and work, occupation and travel history of the previous two months. These data are intended for in-depth epidemiological studies, such as analyses of the contribution of travel to gene flow.[19] Data from these questionnaires were stored in a separate system, and linked to genetic data by means of the tracking barcodes. They were not used in the present work.

Sample Preparation and Genotyping

DBS samples were received and stored either at the Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam, or at the MORU/WWARN molecular laboratory, Bangkok, Thailand. Samples were registered and tracked in a secure bespoke online database, where location and date of collection were recorded. DNA was extracted from samples using high-throughput robotic equipment (Qiagen QIAsymphony) according to manufacturer's instructions. Extracted DNA was plated and shipped to the MalariaGEN Laboratory at the Wellcome Sanger Institute (WSI), Hinxton, UK, for genotyping and whole genome sequencing. Parasite DNA was amplified by applying selective whole genome amplification (sWGA) as previously described.[23]

Genotyping was performed by the SpotMalaria platform, described in the separate document "SpotMalaria platform - Technical Notes and Methods" available from the Resource Page, which includes the complete list of genotyped variants and the details of the genotyping procedures for these variants. Briefly, the first version of SpotMalaria used multiplexed mass spectrometry arrays on the Agena MassArray system for typing most SNPs, and capillary sequencing for the artemisinin resistance domains of the *kelch13* gene. This was eventually replaced by an amplicon sequencing method, using Illumina sequencing of specific genome segments amplified by PCR reaction. The two implementations genotype a common set of variants, each iteration extending or improving on previous versions. Amplicon sequencing also offers greater portability, since it can be deployed on smaller sequencers in country-based laboratories.

Genetic Report Cards Generation

For each sample, genotypes were called for each variant analysed by SpotMalaria, and further processed to determine commonly recognized haplotypes associated with drug resistance (e.g. in genes *crt*, *dhfr*, *dhps*). Genetic barcodes were constructed by concatenating 101 SNP alleles. The generated genotypes, combined with sample metadata, were returned in tabular form to those partners who had submitted the samples along with explanatory documentation for the interpretation of the reports.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20159624; this version posted October 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

The genotypes generated were used to classify samples by their predicted resistance to different drugs. The prediction rules were based on the available data and current knowledge of resistance markers, and are detailed in the separate document "Mapping genetic markers to resistance status classification" available from the Resource Page. For each drug, samples were classified as "sensitive", "resistant", "undetermined" or "missing"- the latter identifying samples that failed to produce a valid genotype for the classification. Heterozygous samples, i.e. those containing genomes carrying both sensitive and resistant alleles, were classified as undetermined, due to lack of evidence for the drug resistance phenotype of such mixed infections.

In order to minimized the impact of call missingness, we also applied a set of *imputation rules* that predict missing alleles in the *crt*, *dhfr* and *dhps* genes, based on statistically significant association with alleles at other positions. Associations were tested (using the threshold $p < 0.05$ by Fisher's exact test) using over 7,000 samples in the MalariaGEN *Pf* Community Project Version 6.[22] The rules for imputations were applied before phenotype prediction rules. They are detailed in the separate document "Imputation of genotypes for markers of drug resistance" available from the Resource Page.

Data aggregation and Mapping of Drug Resistance

To determine the frequency of resistant parasite for a drug, we selected samples at the desired level of geographical aggregation (e.g. province/state or district), based on sampling location. After removing samples with missing and undetermined phenotype predictions for the desired drug, we counted the individuals predicted to be resistant (n_r) and sensitive (n_s), giving a total aggregation sample size $N=n_r+n_s$. Resistant parasite frequency was then computed as $f_r=n_r/N$. Maps of resistance frequency were produced using Tableau Desktop 2020.1.8 (www.tableau.com). To indicate levels of resistance, markers were coloured with a custom green-orange-red palette. Pie chart markers, used to represent allele proportions, were also derived from the same set of N aggregated samples.

Population Structure Analysis

Pairwise genetic distances between parasites were estimated by comparing genetic barcodes. To reduce error due to missingness, we first eliminated samples with more than 50% missing barcode genotypes; then we removed SNPs with missing calls in >20% of the remaining samples; and finally discarded samples with >25% missingness in the remaining SNPs. This produced a dataset of 87-SNP barcodes for 7,490 samples from which genetic distances were estimated. For each sample s, we assigned a within-sample non-reference frequency g_s at each position carrying a valid genotype, as follows: $g_s=0$ if the sample carried the reference allele, $g_s=1$ if it carried the alternative allele, $g_s=0.5$ if both alleles were present. The distance between two samples at that position was then estimated by: $d = g_1(1- g_2) + g_2(1- g_1)$ where g_1 and g_2 are the g_s values for the two samples. The pairwise distance was estimated as the mean of d across all positions where d could be computed (i.e. where neither of the two samples had a missing call). Neighbour-joining trees (NJTs) were then produced using the `nj` implementation in the R package `ape` on R v4.0.2 from square distance matrices.

Role of the funding source

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20159624; this version posted October 3, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

The funders had no role in study design, data collection, data analysis, data interpretation, or report writing. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

ACKNOWLEDGMENTS

The GenRe-Mekong project is funded by the Bill & Melinda Gates Foundation (OPP11188166, OPP1204268). Genotyping and sequencing at Wellcome Sanger Institute and University of Oxford were funded by the Wellcome Trust (098051, 206194, 203141, 090770, 204911, 106698/B/14/Z) and Medical Research Council (G0600718). A proportion of samples were collected with the support of the UK Department for International Development (201900, M006212), and Intramural Research Program of the National Institute of Allergy and Infectious Diseases. We are grateful all patients and health workers who participated in samples collections. This study used data from the MalariaGEN Pf3k Project and *Plasmodium falciparum* Community Project. We thank the staff of Wellcome Sanger Institute Sample Logistics, Sequencing, and Informatics facilities for their contribution; in particular, we are grateful to the Wellcome Sanger Institute DNA Pipelines Informatics for supporting the development of the methods used in this work. We thank the many collaborators who contributed to the GenRe-Mekong Project, and especially: Pannapat Masingboon, Narisa Thongmee, Zoë Doran, Salwaluk Panapipat, Ipsita Sinha, Rapeephan Maude, Vilasinee Yuwaree, Tran Minh Nhat, Hoang Hai Phuc, Ro Mah Huan, Nguyen Minh Nhat, Tran Van Don. PR is a staff member of the World Health Organization; PR alone is responsible for the views expressed in this publication and they do not necessarily represent the decisions, policy or views of the World Health Organization.

AUTHOR CONTRIBUTIONS

HHQ, BHo, VV, TND, HR, NTN, MM, RM, RvdP, LVS, RF, FN, MAH, KC, PNN, EAA, SPh, RRM, RL, CHuc, LTD, KTN, TMN, TTH, HN, NZ, AAS, MO, RT, AMT, JL, DL, PJ, SPu, FS, MT, PS, PKB, SB, ARA, AF, OM organized or carried out sample collections. ED, MA, SR, MD conducted laboratory analyses. CGJ, KaR, KiR produced genomic data. CGJ, CM, JSt, JK developed informatics systems. RM, PR, CB, DM, JSi, VS, KiR, AMD, OM designed and coordinated the project. CGJ, NP, ScG, ED, KaR, KiR developed laboratory systems. CGJ, OM drafted the manuscript.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20159624; this version posted October 3, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

REFERENCES

1. Dondorp AM, Nosten F, Yi P, et al. Artemisinin resistance in Plasmodium falciparum malaria. *N Engl J Med* 2009; **361**(5): 455-67.
2. Noedl H, Se Y, Schaecher K, et al. Evidence of artemisinin-resistant malaria in western Cambodia. *N Engl J Med* 2008; **359**(24): 2619-20.
3. Plowe CV. The evolution of drug-resistant malaria. *Trans R Soc Trop Med Hyg* 2009; **103 Suppl 1**: S11-4.
4. Roper C, Pearce R, Nair S, Sharp B, Nosten F, Anderson T. Intercontinental spread of pyrimethamine-resistant malaria. *Science* 2004; **305**(5687): 1124.
5. Mita T, Venkatesan M, Ohashi J, et al. Limited geographical origin and global spread of sulfadoxine-resistant dhps alleles in Plasmodium falciparum populations. *J Infect Dis* 2011; **204**(12): 1980-8.
6. Mita T, Tanabe K, Kita K. Spread and evolution of Plasmodium falciparum drug resistance. *Parasitology International* 2009; **58**(3): 201-9.
7. Trape JF, Pison G, Preziosi MP, et al. Impact of chloroquine resistance on malaria mortality. *C R Acad Sci III* 1998; **321**(8): 689-97.
8. Ashley EA, Dhorda M, Fairhurst RM, et al. Spread of artemisinin resistance in Plasmodium falciparum malaria. *The New England Journal of Medicine* 2014; **371**(5): 411-23.
9. MalariaGEN Plasmodium falciparum Community Project. Genomic epidemiology of artemisinin resistant malaria. *Elife* 2016; **5**.
10. Amaratunga C, Lim P, Suon S, et al. Dihydroartemisinin-piperaquine resistance in Plasmodium falciparum malaria in Cambodia: a multisite prospective cohort study. *The Lancet Infectious diseases* 2016; **16**(3): 357-65.
11. van der Pluijm RW, Imwong M, Chau NH, et al. Determinants of dihydroartemisinin-piperaquine treatment failure in Plasmodium falciparum malaria in Cambodia, Thailand, and Vietnam: a prospective clinical, pharmacological, and genetic study. *Lancet Infect Dis* 2019.
12. Leang R, Taylor WR, Bouth DM, et al. Evidence of Plasmodium falciparum Malaria Multidrug Resistance to Artemisinin and Piperaquine in Western Cambodia: Dihydroartemisinin-Piperaquine Open-Label Multicenter Clinical Assessment. *Antimicrob Agents Chemother* 2015; **59**(8): 4719-26.
13. Spring MD, Lin JT, Manning JE, et al. Dihydroartemisinin-piperaquine failure associated with a triple mutant including kelch13 C580Y in Cambodia: an observational cohort study. *Lancet Infect Dis* 2015; **15**(6): 683-91.
14. Menard D, Khim N, Beghain J, et al. A Worldwide Map of Plasmodium falciparum K13-Propeller Polymorphisms. *N Engl J Med* 2016; **374**(25): 2453-64.
15. Lim P, Dek D, Try V, et al. Ex vivo susceptibility of Plasmodium falciparum to antimalarial drugs in western, northern, and eastern Cambodia, 2011-2012: association with molecular markers. *Antimicrobial agents and chemotherapy* 2013; **57**(11): 5277-83.
16. Nagar SD, Moreno AM, Norris ET, et al. Population Pharmacogenomics for Precision Public Health in Colombia. *Front Genet* 2019; **10**: 241.
17. Desmond-Hellmann S. Progress lies in precision. *Science* 2016; **353**(6301): 731.
18. von Seidlein L, Peto TJ, Landier J, et al. The impact of targeted malaria elimination with mass drug administrations on falciparum malaria in Southeast Asia: A cluster randomised trial. *PLoS Med* 2019; **16**(2): e1002745.
19. Chang HH, Wesolowski A, Sinha I, et al. Mapping imported malaria in Bangladesh using parasite genetic and human mobility data. *Elife* 2019; **8**.
20. Landier J, Parker DM, Thu AM, et al. Effect of generalised access to early diagnosis and treatment and targeted mass drug administration on Plasmodium falciparum malaria in Eastern Myanmar: an observational study of a regional elimination programme. *Lancet* 2018; **391**(10133): 1916-26.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.23.20159624; this version posted October 3, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

21. Hamilton WL, Amato R, van der Pluijm RW, et al. Evolution and expansion of multidrug-resistant malaria in southeast Asia: a genomic epidemiology study. *Lancet Infect Dis* 2019.
22. Pearson RD, Amato R, Kwiatkowski DP, MalariaGEN *Plasmodium falciparum* Community Project. An open dataset of Plasmodium falciparum genome variation in 7,000 worldwide samples. *bioRxiv* 2019; **824730**.
23. Oyola SO, Ariani CV, Hamilton WL, et al. Whole genome sequencing of Plasmodium falciparum from dried blood spots using selective whole genome amplification. *Malar J* 2016; **15**(1): 597.
24. Ariey F, Witkowski B, Amaratunga C, et al. A molecular marker of artemisinin-resistant Plasmodium falciparum malaria. *Nature* 2014; **505**(7481): 50-5.
25. Daniels RF, Schaffner SF, Wenger EA, et al. Modeling malaria genomics reveals transmission decline and rebound in Senegal. *Proceedings of the National Academy of Sciences of the United States of America* 2015; **112**(22): 7067-72.
26. Felger I, Snounou G, Hastings I, Moehrle JJ, Beck H-P. PCR correction strategies for malaria drug trials: updates and clarifications. *The Lancet Infectious Diseases* 2019.
27. Miotto O, Amato R, Ashley EA, et al. Genetic architecture of artemisinin-resistant Plasmodium falciparum. *Nature Genetics* 2015; **47**(3): 226-34.
28. Amato R, Lim P, Miotto O, et al. Genetic markers associated with dihydroartemisinin-piperaquine failure in Plasmodium falciparum malaria in Cambodia: a genotype-phenotype association study. *The Lancet Infectious diseases* 2016.
29. World Health Organization. Status report on artemisinin resistance and ACT efficacy (August 2018). Geneva: World Health Organization, 2018.
30. Amato R, Pearson RD, Almagro-Garcia J, et al. Origins of the current outbreak of multidrug-resistant malaria in southeast Asia: a retrospective genetic study. *Lancet Infect Dis* 2018.
31. Miotto O, Almagro-Garcia J, Manske M, et al. Multiple populations of artemisinin-resistant Plasmodium falciparum in Cambodia. *Nature Genetics* 2013; **45**(6): 648-55.
32. Laufer MK, Thesing PC, Eddington ND, et al. Return of chloroquine antimalarial efficacy in Malawi. *N Engl J Med* 2006; **355**.
33. Frosch AE, Laufer MK, Mathanga DP, et al. Return of Widespread Chloroquine-Sensitive Plasmodium falciparum to Malawi. *The Journal of Infectious Diseases* 2014; **210**(7): 1110-4.
34. Artimovich E, Schneider K, Taylor TE, et al. Persistence of Sulfadoxine-Pyrimethamine Resistance Despite Reduction of Drug Pressure in Malawi. *J Infect Dis* 2015; **212**(5): 694-701.
35. Directorate of National Vector Borne Disease Control Programme DGoHS, Ministry of Health and Family Welfare, Government of India. National Drug Policy on Malaria, 2013.
36. World Health Organization. Update on artemisinin resistance - April 2017. Geneva: World Health Organization, 2017.
37. 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): 670-9.
38. Ross LS, Dhingra SK, Mok S, et al. Emerging Southeast Asian PfCRT mutations confer Plasmodium falciparum resistance to the first-line antimalarial piperaquine. *Nat Commun* 2018; **9**(1): 3314.
39. Agrawal S, Moser KA, Morton L, et al. Association of a Novel Mutation in the Plasmodium falciparum Chloroquine Resistance Transporter With Decreased Piperaquine Sensitivity. *J Infect Dis* 2017; **216**(4): 468-76.

Table 1 – Participating Studies in GenRe-Mekong

For each study we list the NMCP and Research partners involved, the type of study, the geographical region covered and the number of collection sites. In the last two columns, we show the total number of samples submitted, and the number included in the final set of quality-filtered sample used in epidemiology analyses.

NMCP Partner	Research / Technical Partner	Study Type	Regions Surveyed	Sites	Submitted Samples	Filtered Samples
Center for Malaria Parasitology and Entomology of Lao PDR (CMPE)	Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit (LOMWRU), Vientiane	Genetic Surveillance	South Laos (5 provinces)	51	1555	1387
Institute of Malariology, Parasitology, and Entomology Quy Nhon (IMPE-QN), Vietnam	Oxford University Clinical Research Unit (OUCRU), Ho Chi Minh City	Genetic Surveillance	Central Vietnam (7 provinces)	51	1632	1492
National Institute of Malariology, Parasitology, and Entomology (NIMPE), Vietnam	Vysnova Partners, Mahidol-Oxford Research Unit (MORU)	Epidemiological Study	South Vietnam (3 provinces)	19	292	265
National Center for Parasitology, Entomology, and Malaria Control (CNM), Cambodia		Genetic Surveillance	Northeast Cambodia (2 provinces)	19	182	174
Bangladesh National Malaria Control Programme	Mahidol-Oxford Research Unit (MORU)	Epidemiological Study	Bangladesh (Chittagong Division)	55	2055	1575
-	Mahidol-Oxford Research Unit (MORU)	Clinical Efficacy Study	Cambodia, Vietnam, Thailand, Lao PDR, Myanmar, Bangladesh, India, DR Congo	17	1875	1123
-	National Institutes of Health (NIH)	Clinical Efficacy Study	Cambodia	3	592	502
-	Oxford University Clinical Research Unit (OUCRU)	Epidemiological Study	South Vietnam	4	184	175
-	Mahidol-Oxford Research Unit (MORU)	Elimination Study	West Cambodia	1	69	32
-	Mahidol-Oxford Research Unit (MORU)	Epidemiological Study	Northeast Thailand	7	87	60
-	Shoklo Malaria Research Unit (SMRU)	Clinical Efficacy Study	Thailand (Tak province)	4	29	28
-	Shoklo Malaria Research Unit (SMRU)	Elimination Study	Myanmar (Kayin State)	51	1071	813
				Total	**9623**	**7626**

Table 2 – Drug resistance-related SNPs genotyped by GenRe-Mekong (excludes *kelch13*)

Chromosome	Position	Gene Id	Gene Description	Mutation	Reference	Alternate
Pf3D7_04_v3	748239	PF3D7_0417200	*dhfr* (bifunctional dihydrofolate reductase-thymidylate synthase)	N51I	A	T
Pf3D7_04_v3	748262			C59R/Y	T	C
Pf3D7_04_v3	748263			C59R/Y	G	A
Pf3D7_04_v3	748410			S108N/T	G	AC
Pf3D7_04_v3	748577			I164L	A	T
Pf3D7_05_v3	958145	PF3D7_052300	*mdr1* (multidrug resistance protein)	N86Y	A	T
Pf3D7_05_v3	958440			Y184F	A	T
Pf3D7_05_v3	961625			D1246Y	G	T
Pf3D7_07_v3	403623	PF3D7_0709000	*crt* (chloroquine resistance transporter)	N75D/E	T	A
Pf3D7_07_v3	403625			K76T	A	C
Pf3D7_07_v3	405362			N326S	A	G
Pf3D7_07_v3	405600			I356T	T	C
Pf3D7_08_v3	549681	PF3D7_0810800	*dhps* (dihydropteroate synthetase)	S436A/Y/F/G	T	GC
Pf3D7_08_v3	549682			S436A/Y/F/G	C	TAG
Pf3D7_08_v3	549685			A437G	G	C
Pf3D7_08_v3	549993			K540E/N	A	GT
Pf3D7_08_v3	549995			K540E/N	A	TG
Pf3D7_08_v3	550117			A581G	C	G
Pf3D7_08_v3	550212			A613S/T	G	TA
Pf3D7_13_v3	748395	PF3D7_1318100	*fd* (ferredoxin)	D193Y	C	A
Pf3D7_13_v3	2504560	PF3D7_1362500	*exo* (exonuclease)	E415G	A	G
Pf3D7_14_v3	-	PF3D7_1408000& PF3D7_1408100	*pm23* (plasmepsin 2/3)	Breakpoint	-	-
Pf3D7_14_v3	1956225	PF3D7_1447900	*mdr2* (multidrug resistance protein 2)	T484I	G	A
Pf3D7_14_v3	2481070	PF3D7_1460900	*arps10* (apicoplast ribosomal protein S10)	V127M	G	A
Pf3D7_14_v3	2481073			D128Y/H	G	TC

Table 3 – Frequencies of resistant parasites in provinces/states/divisions surveyed, for different antimalarials.

Country	Province, State or Division	ART-R	PPQ-R	DHA-PPQ-R	CQ-R	PYR-R	SD-R	SP-R	SP-R (IPTp)
India	Odisha	0%	0%	0%	18%	57%	6%	1%	0%
	West Bengal	0%	0%	0%	47%	71%	14%	5%	0%
	Tripura	0%	0%	0%	85%	100%	99%	55%	0%
Bangladesh	Chittagong	0%	0%	0%	97%	100%	87%	46%	16%
Myanmar	Rakhine	0%	0%	0%	71%	100%	100%	51%	26%
	Bago	1%	0%	0%	88%	100%	100%	91%	74%
	Mandalay	29%	0%	0%	96%	98%	98%	29%	24%
	Kayin	54%	2%	0%	100%	100%	56%	73%	27%
Thailand	Tak	61%	-	0%	100%	100%	96%	100%	88%
	Sisakhet	100%	90%	90%	100%	100%	100%	100%	100%
	Ubon Ratchathani	80%	75%	56%	100%	100%	85%	100%	17%
Cambodia	Pailin	93%	97%	90%	100%	100%	100%	100%	56%
	Battambang	100%	100%	100%	100%	100%	88%	100%	29%
	Pursat	88%	98%	67%	100%	100%	92%	98%	44%
	Preah Vihear	61%	100%	11%	100%	100%	94%	98%	21%
	Steung Treng	93%	75%	70%	100%	100%	97%	100%	0%
	Ratanakiri	49%	79%	42%	99%	100%	76%	90%	5%
Laos	Champasak	66%	75%	56%	100%	100%	88%	94%	12%
	Attapeu	46%	43%	31%	100%	100%	82%	100%	18%
	Sekong	26%	6%	0%	100%	100%	91%	74%	5%
	Salavan	17%	2%	1%	89%	97%	28%	38%	1%
	Savannakhet	10%	1%	0%	87%	96%	21%	41%	2%
Vietnam	Binh Phuoc	92%	93%	83%	100%	100%	100%	100%	14%
	Dak Nong	94%	92%	88%	100%	100%	97%	96%	22%
	Dak Lak	96%	90%	86%	100%	100%	100%	99%	15%
	Gia Lai	84%	83%	76%	99%	100%	98%	95%	4%
	Khanh Hoa	22%	5%	2%	95%	100%	97%	74%	2%
	Ninh Thuan	13%	18%	0%	28%	100%	98%	75%	0%
	Quang Tri	16%	9%	0%	75%	76%	59%	26%	5%
Congo PDR	Kinshasa	0%	0%	0%	58%	98%	72%	88%	0%

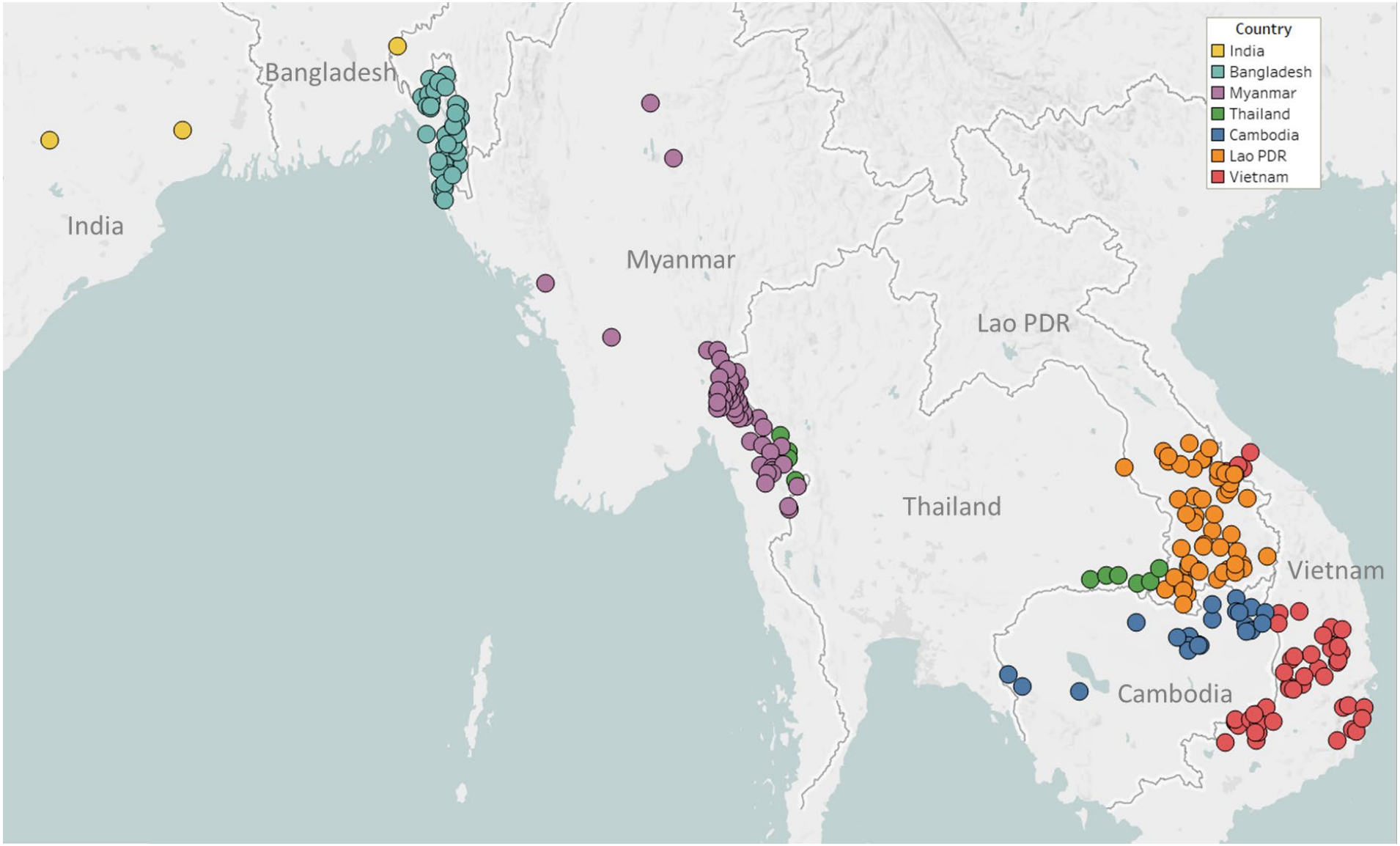


Figure 1 – Map of GenRe-Mekong sample collection sites in Asia.
Sites markers are coloured by country. One site in Kinshasa (DR Congo) not shown.

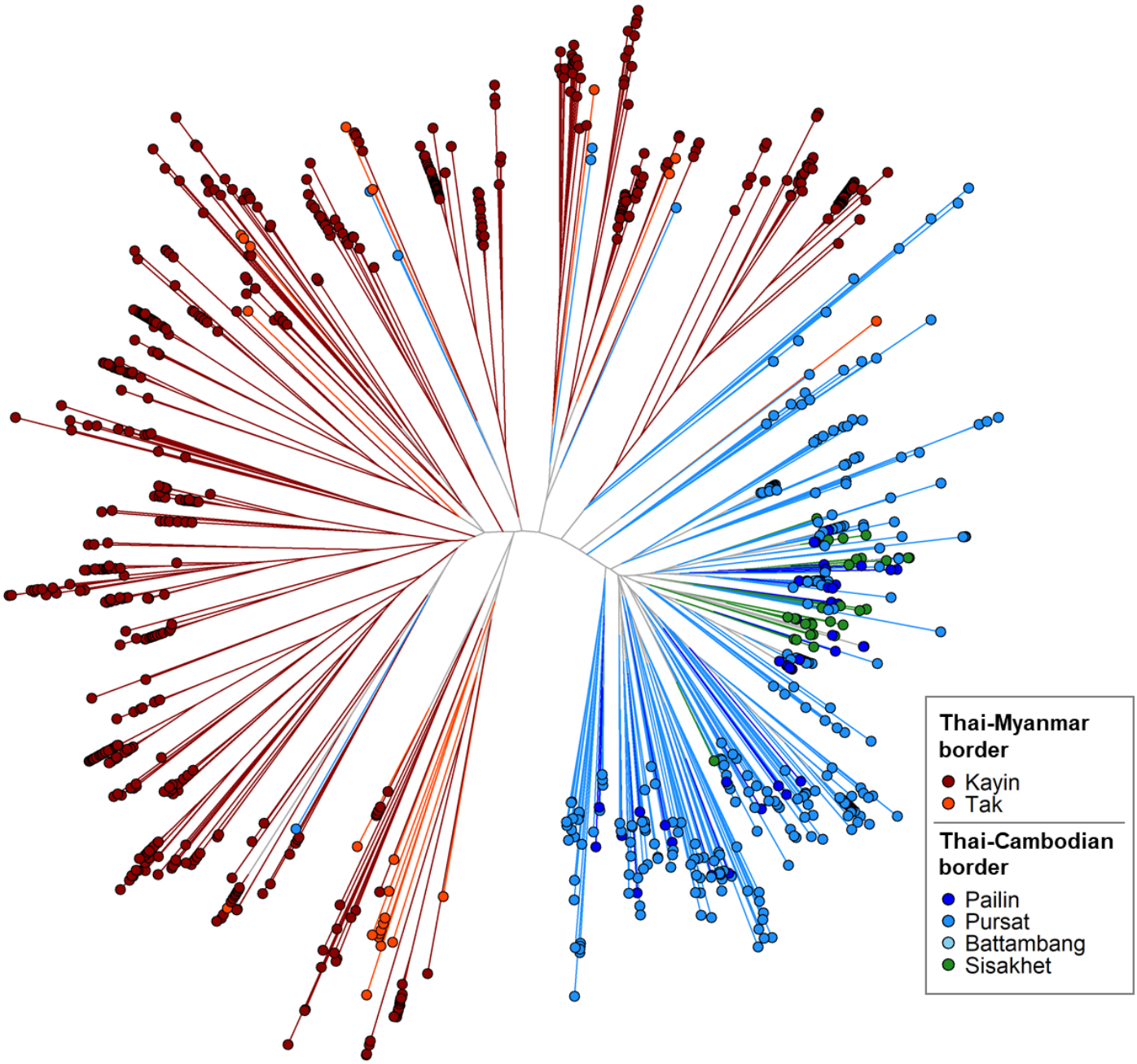


Figure 2 - Neighbour-joining tree using barcode data to show genetic differentiation between parasites in the Thai-Myanmar and Thai-Cambodian border regions.

The tree was derived from a matrix distance matrix, computed by comparing the genetic barcodes of samples. The branch length separating each pair of parasites represents the amount of genetic differentiation between them: individuals separated by shorter branches are more similar to each other. Samples from provinces/states of Myanmar, Thailand and Cambodia near to the borders were included. Each circular marker represents a sample, coloured by the province/state of origin.

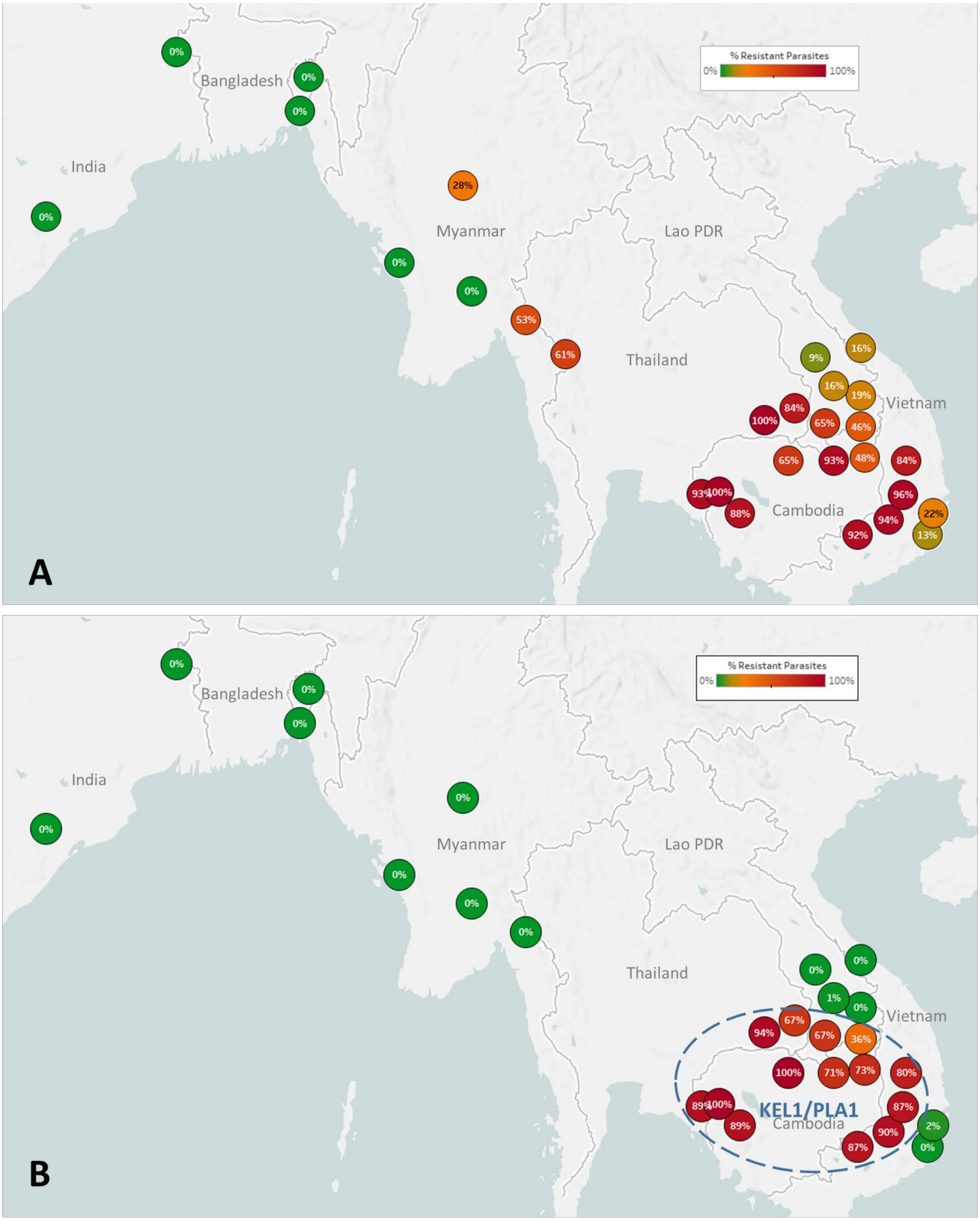
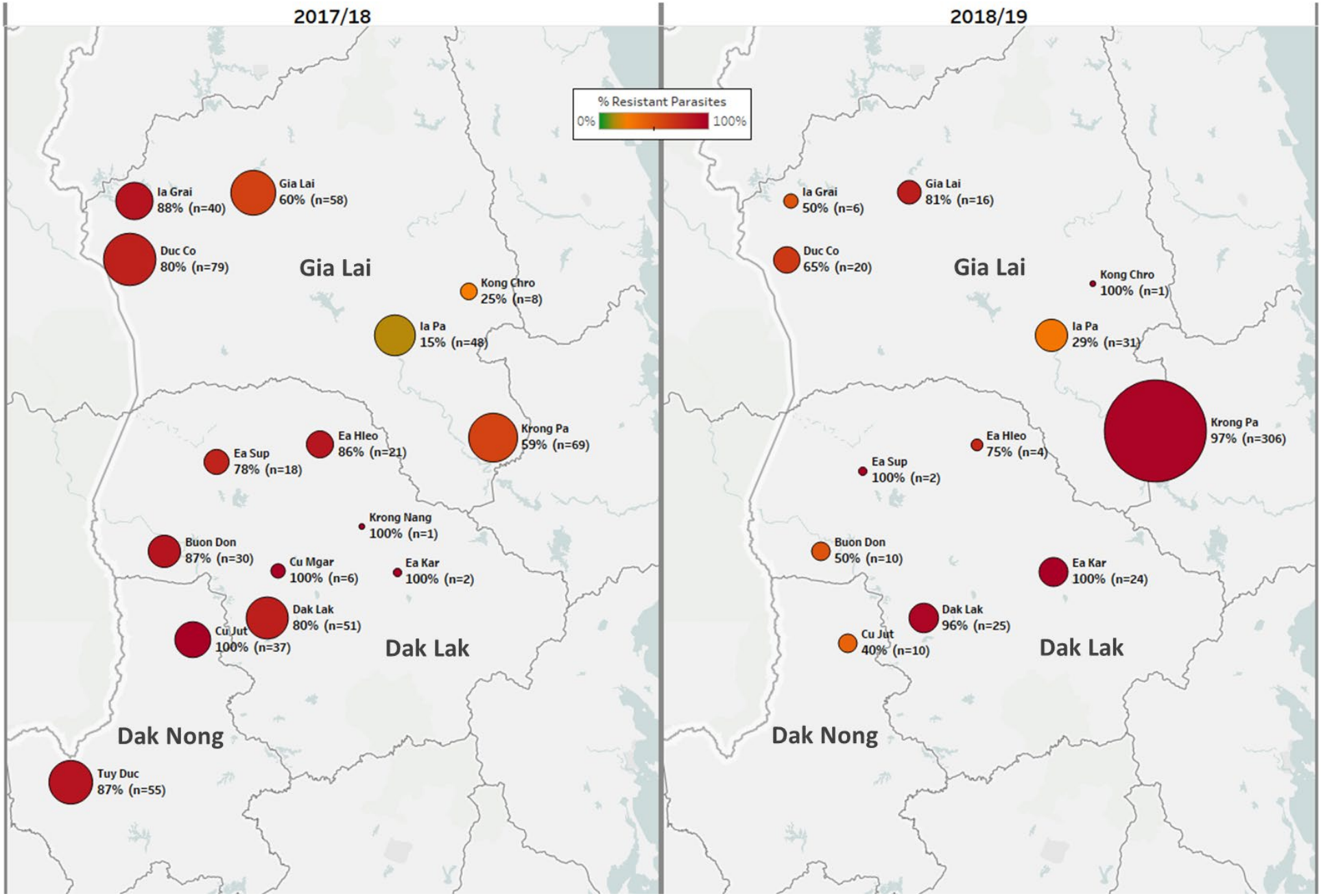


Figure 3 – Map of the spread of (A) artemisinin resistance (ART-R) and (B) dihydroartemisinin-piperaquine resistance (DHA-PPQ-R) in Asian countries.

Marker text and colour indicate the proportion of sample classified as *resistant* in each province/state/division surveyed. A total of 6,762 samples were included in (A) and 3,395 samples in (B), after excluding samples with *undetermined* phenotypes. The results are summarized in Table 3.


2017/18
2018/19
% Resistant Parasites
0%
100%
Gia Lai
Dak Lak
Dak Nong
Ia Grai
88% (n=40)
Gia Lai
60% (n=58)
Duc Co
80% (n=79)
Kong Chro
25% (n=8)
Ia Pa
15% (n=48)
Krong Pa
59% (n=69)
Ea Hleo
86% (n=21)
Ea Sup
78% (n=18)
Krong Nang
100% (n=1)
Buon Don
87% (n=30)
Cu Mgar
100% (n=6)
Ea Kar
100% (n=2)
Dak Lak
80% (n=51)
Cu Jut
100% (n=37)
Tuy Duc
87% (n=55)
Ia Grai
50% (n=6)
Gia Lai
81% (n=16)
Duc Co
65% (n=20)
Kong Chro
100% (n=1)
Ia Pa
29% (n=31)
Krong Pa
97% (n=306)
Ea Hleo
75% (n=4)
Ea Sup
100% (n=2)
Buon Don
50% (n=10)
Ea Kar
100% (n=24)
Dak Lak
96% (n=25)
Cu Jut
40% (n=10)

(previous page)

Figure 4 – Longitudinal sample counts and proportions of KEL1/PLA1 parasites in three provinces of Central Vietnam.
The same geographical area (Gia Lai, Dak Lak and Dak Nong provinces) is shown for two malaria seasons: 2017/18 (12 months from May 2017, n=523) and 2018/2019 (the following 12 months, n=455). Districts are represented by markers whose size is proportional to the number of samples, and whose colour indicates the frequency of samples carrying both the *kelch13* C580Y mutation and the *plasmepsin2/3* amplification, and thus assigned to the KEL1/PLA1 strain. Marker labels show district name, resistant parasite frequency and sample count.

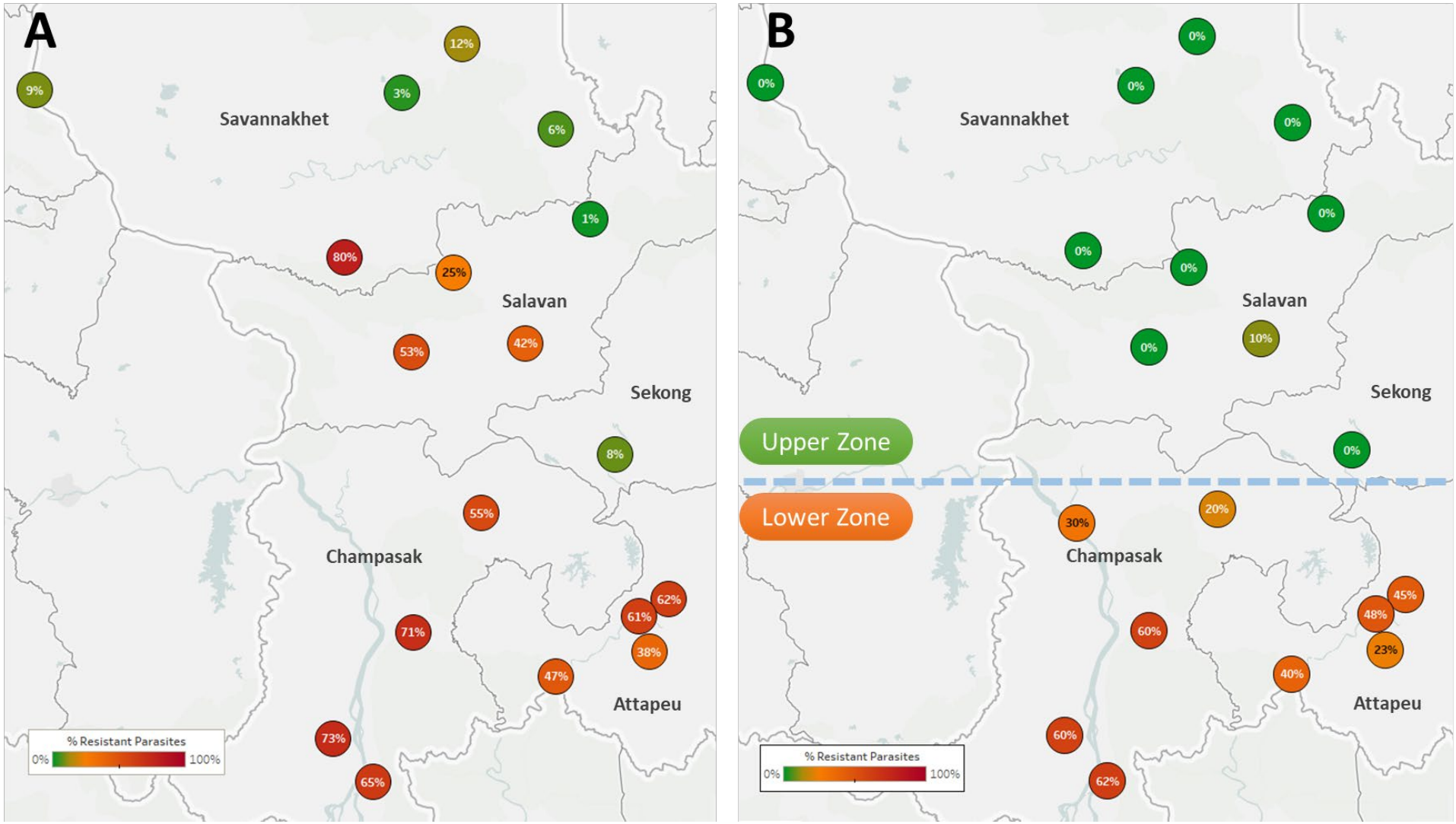


Figure 5 – Proportions of ART-R and KEL1/PLA1 parasites in southern Laos districts.
Districts in five provinces of southern Laos are represented by markers whose colour and label indicates the frequency of samples classified as artemisinin-resistant (A) and as belonging to the KEL1/PLA1 strain, resistant to both artemisinin and piperaquine. Only districts with more than 10 samples with valid genotypes are shown. In panel (B), a dashed line denotes a hypothetical demarcation line between a Lower Zone, where KEL1/PLA1 has spread, and an Upper Zone, where it is absent and ART-R parasites have a different origin.

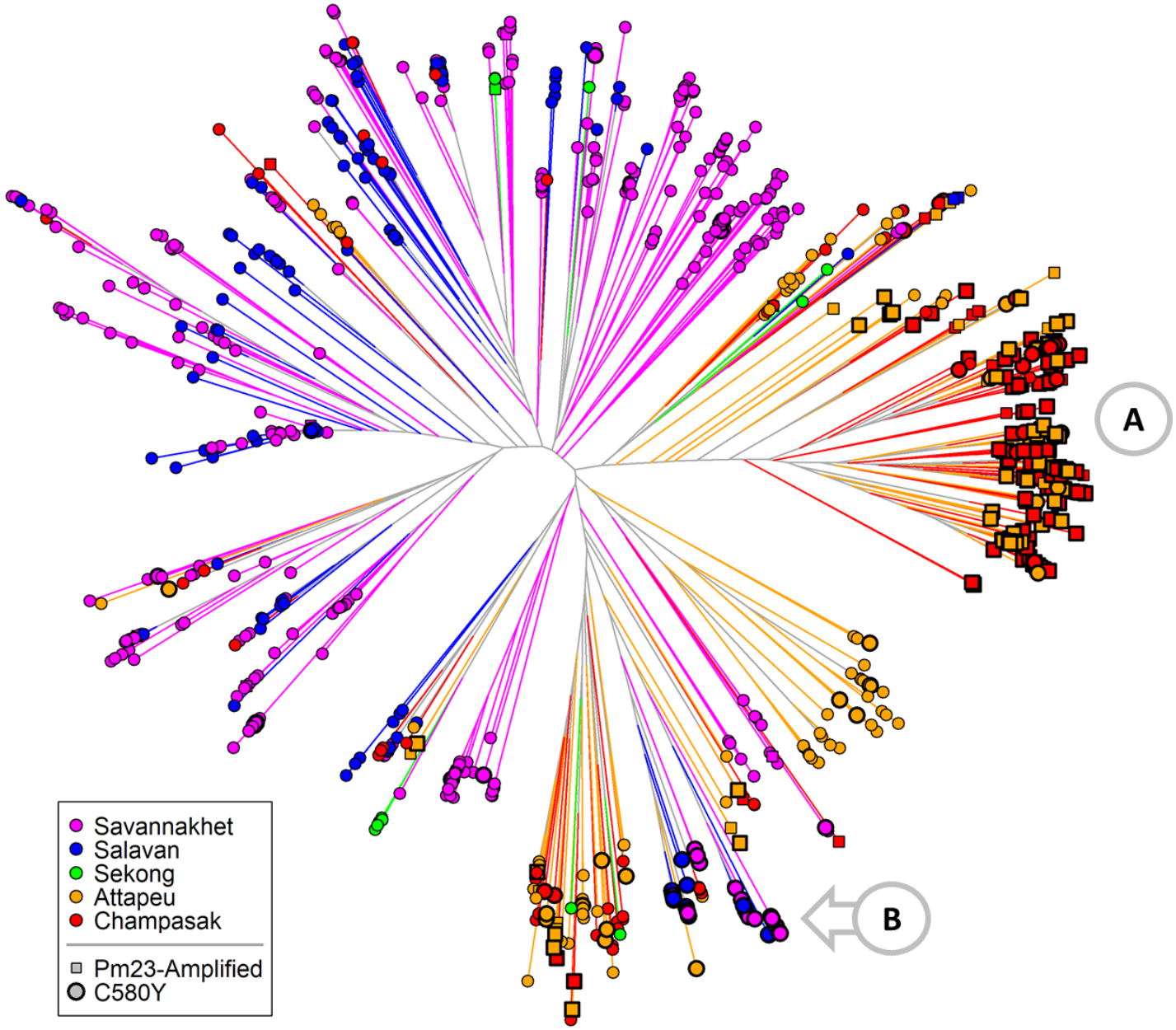


Figure 6 - Neighbour-joining tree using barcode data to show genetic differentiation between groups of parasites collected in Southern Laos.

The tree was derived from a matrix distance matrix, computed by comparing the genetic barcodes of samples collected in the Lao PDR (n=1,332 with <25% barcode genotypes missing). Each marker represents a parasite sample, coloured by province. The branch length separating each pair of parasites represents the amount of genetic differentiation between them: individuals separated by shorter branches are more similar to each other. Thicker marker borders indicate parasites carrying the *kelch13* C580Y mutation, while square markers indicate samples with *plasmepsin 2/3* amplification. Orange circular callouts show notable features of this tree. (A) shows a large cluster of parasites from the Lower Zone (Attapeu and Champasak provinces) carrying both C580Y and *plasmepsin 2/3* amplification (KEL1/PLA1 strain). (B) indicates that C580Y mutants from the Upper Zone (Savannakhet and Salavan provinces) are genetically distinct from KEL1/PLA1, but also from Upper Zone wild-type parasites.