Malaria Policy Advisory Committee Meeting
2–4 October 2019, Geneva, Switzerland
Background document for Session 7

Technical consultation on the role of parasite and anopheline genetics in malaria surveillance

5–7 June 2019, Geneva, Switzerland

WHO/CDS/GMP/MPAC/2019.17

Background

The WHO Global Malaria Programme (GMP) convened a three-day Technical Consultation from 5 to 7 June 2019. The aim of the Technical Consultation was to review the existing evidence from genetic epidemiological research studies and use cases and assess the role of such research in the development of future policies and the potential for malaria programmes to make practical use of genomics. The Technical Consultation also aimed at establishing a list of global research priorities for the future strategic use of genetic epidemiology, in the hopes of accelerating progress towards achieving the goals of the Global Technical Strategy for malaria.

Objectives

- Review existing evidence across the use cases of genetic epidemiology in malaria surveillance
- Identify priority research questions relevant to policy and operational national programme activities for each use case
- Discuss appropriate study protocols and issues related to ethics, data sharing and coordination mechanisms

Outcomes

- A meeting report summarizing the content of the presentations, discussions and outcomes of the meeting
- A list of key research questions relevant to policy and operational activities of national programmes for each use case
- A work plan to implement the key action points of the meeting (post MPAC)

Proposed policy relevant priority research areas/questions

Please refer to accompanying table

Strategic next steps

1. The table of research priority areas (Table A1) identified during this meeting should be made available online and updated on an annual basis by WHO with help from research networks and individuals.
2. A database of researchers and institutions involved in policy-relevant malaria genetic epidemiology studies should be developed by WHO, and this database should be updated annually.
3. Use cases share several overlapping themes across the spectrum of transmission in terms of understanding gene flow in insecticide and drug resistance. Studies should maximize these linkages so that common data generation platforms and samples can be used, wherever possible.
4. In addition to research studies, there are opportunities to explore drug and insecticide resistance monitoring sites: collecting genetic samples during case detection and investigations in elimination settings, and in burden reduction settings, passive case detection systems and household surveys could become the mainstay for genomic surveillance. A structured approach that will not add unnecessary burden on health system is needed.

5. Stakeholders should work with researchers to ensure that study protocols are designed to generate evidence in formats relevant to policy and programmes. For example, studies exploring the relevance of genomic surveillance metrics must include a comparison to metrics currently recommended by WHO and used by countries in terms of their relevance, reliability, accuracy, precision, cost and sustainability. WHO to work with network of research during study design stage.
6. Established global databases should be harnessed to develop information products relevant for policy and country operations. WHO to work with groups such as Sanger Institute and BROAD on appropriate information products once policy relevance is established.
7. Investment in regional and national capacities for genetic epidemiology should be sought. WHO to work with researchers and funders such as BMGF on pathways to increased national capacity.

Request to MPAC

- Provide guidance on and approve key research areas/questions
- Provide guidance on and approve the strategic next steps

Table of contents

List of acronyms

AFP	acute flaccid paralysis
COI	complexity of infection
DRC	Democratic Republic of the Congo
EEMS	estimated effective migration surfaces
GMP	Global Malaria Programme
GPEI	Global Polio Eradication Initiative
GPLN	Global Polio Laboratory Network
IBD	identity/identical by descent
IBS	identity/identical by state
IRS	indoor residual spraying
ITN	insecticide-treated net
kdr	knock-down resistance
LLIN	long-lasting insecticidal net
MDR-TB	multidrug-resistant tuberculosis
MFO	mixed-function oxidase
MPAC	Malaria Policy Advisory Committee
NMCP	national malaria control programme
PAHO	Pan American Health Organization
PBO	piperonyl butoxide
PCR	polymerase chain reaction
PLA	plasmepsin amplification
QA	quality assurance
QC	quality control
RDT	rapid diagnostic test
SNP	single nucleotide polymorphism
TB	tuberculosis
TES	therapeutic efficacy study
WGS	whole genome sequencing
WHO	World Health Organization

1. Introduction

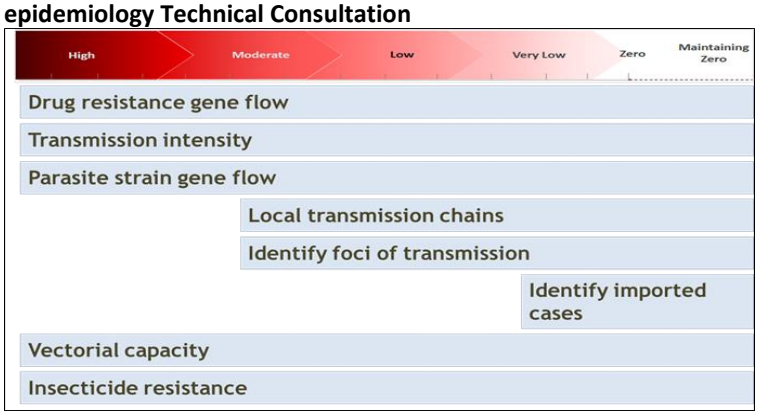
Advances in genetic epidemiology are creating new opportunities for the surveillance, prevention and control of infectious diseases. Emerging evidence shows that mosquito genotyping can improve the understanding of mechanisms of speciation and the processes that influence the mosquitoes' ability to transmit malaria parasites to humans. Such knowledge can foster a better understanding of vectorial capacity and consequently how to better target interventions. Research on parasite genotyping also indicates potential applications in the understanding of parasite gene flow, including drug-resistance genes, *Pfhrp2/3* deletions, quantification of malaria importation risk, as well as characterization of changing transmission intensity. Most of the work in malaria genetic epidemiology, however, has remained within the realm of research and has not been guided by clearly defined policy-relevant questions. There have been few examples of how such work could improve the operational decisions made by national malaria programmes.

For these reasons, the WHO Global Malaria Programme (GMP) convened a three-day Technical Consultation from 5 to 7 June 2019. The aim of the Technical Consultation was to review the existing evidence from genetic epidemiological research studies and use cases, and assess the role of such research in the development of future policies and the potential for malaria programmes to make practical use of genomics. The Technical Consultation also aimed at establishing a list of global research priorities for the future strategic use of genetic epidemiology, in the hopes of accelerating progress towards achieving the goals of the Global Technical Strategy for malaria *(1)*. The Technical Consultation was approved by the Malaria Policy Advisory Committee (MPAC) during its October 2018 meeting and was jointly convened by the GMP Units responsible for Surveillance, Monitoring and Evaluation (SUR); Drug Efficacy and Resistance (DER); Entomology & Vector Control (EVC); and Elimination (ELI). The meeting was chaired by Professor Dyann Wirth.

2. Objectives and expected outcomes

The main objectives of the consultation were to understand the role of genetic epidemiology (specifically parasite and anopheline genetic signals and gene flow) in malaria surveillance and control, and to define priority research questions that are relevant to policy and operational activities of national programmes (see Fig. 1).

Fig. 1. Topics across the transmission continuum recommended by MPAC for discussion during the genetics epidemiology Technical Consultation



Specifically, the consultation served to:

- review existing evidence across the use cases of genetic epidemiology in malaria surveillance;
- identify key research questions relevant to policy and operational activities of national programmes for each use case;

- Discuss appropriate study protocols and issues related to ethics, data sharing and coordination mechanisms.

Expected outcomes were:

- a meeting report summarizing the content of the presentations, discussions and outcomes of the meeting;
- a list of key research questions relevant to policy and operational activities of national programmes for each use case;
- a work plan to implement the key action points of the meeting.

This report summarizes:

- presentations given by meeting participants
- major discussion points
- the list of key research questions relevant to policy and operational activities of national programmes for each use case related to the topics of transmission and resistance
- next steps.

3. Summary of presentations and associated discussions

All presentations can be found at the following link:
https://www.dropbox.com/sh/zsuu5p3ls7m2l6w/AAAqO-Jfa0f73wXTBRpp-Puga?dl=0

3.1. The current and potential future role of genetic epidemiology in malaria surveillance

Presenter: Abdisalan Noor, WHO-GMP

Within WHO-GMP, work is ongoing in the use of molecular epidemiology for monitoring drug and insecticide resistance and the *pfhrp2/3* parasite gene deletions that evade detection by rapid diagnostic tests (RDTs). Data from monitoring sites and research studies around the world are displayed on the Malaria Threats Map[1], which provides a global spatial and temporal overview of vector insecticide resistance, parasite drug resistance and parasite *pfhrp2/3* gene deletions.

With the growing acceptance that genomics could play an integral role in policy and programmatic decisions, there have been increased investments, demonstration studies and refinement of sampling and analytical methods that could prove optimal in expanding the use of genomics as a tool in malaria control. There are, however, still significant unresolved issues related to priority research questions, programmatic applications, ethics of use of genetic material, data sharing and data use. A review of the range and complexity of genomic methods is also required to assess whether such methods are comparable and representative in different geographical contexts, and to determine the feasibility of implementation in countries with limited resources and limited capacity for data generation and analysis. Additionally, the fragmentation of genomics research has resulted in very few joint genetic and epidemiological analyses, which could provide practical applications for operational use and translation into policy. Without clear guidance on priority policy-relevant research questions, most of the studies may not have immediate policy relevance. It is our hope that this Technical Consultation will discuss these issues and identify evidence that could immediately contribute to policy

[1] The Malaria Threats Map is an online mapping platform that collates information on the biological challenges to malaria control and elimination, including insecticide and drug resistance, and gene deletions. The Malaria Threats Map App is available at: https://apps.who.int/malaria/maps/threats/.

recommendations, along with evidence that may be relevant but is only likely to be available within medium (five years) and long (10 years) timeframes.

3.2. Session 1: Experiences from other diseases: polio, Ebola and tuberculosis

Opening remarks by facilitator, Dyann Wirth

Genomic data have been applied to understanding the epidemiology of various infectious diseases, ranging from support for outbreak investigations to providing the foundation for elimination programmes. Understanding the practical application of genomics by other disease control programmes can offer insight into the potential uses and challenges in implementation at the national and subnational levels. The lessons learned from polio, Ebola and tuberculosis (TB) were presented in order to stimulate further discussion on potential use cases for malaria genomic epidemiology.

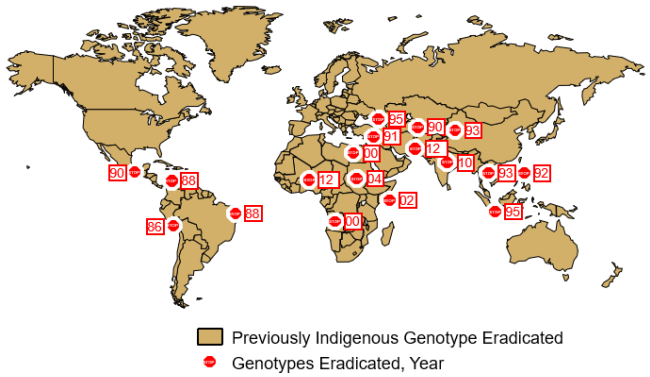
3.2a. Genetic epidemiology and disease surveillance for elimination: polio

Presenter: Ousmane Diop

Applications of genetic epidemiological methods have been a critical component in the success of the Global Polio Eradication Initiative (GPEI). The combination of an effective vaccine and strong collaboration between field disease surveillance and laboratory virologic surveillance teams has been crucial in achieving progress in the eradication of polio. Acute flaccid paralysis (AFP) surveillance, environmental surveillance, and special targeted studies within the control programme framework have allowed for strategic use of genetic epidemiology to determine the source of infection and inform whether transmission has occurred. In clinical surveillance, a suspected case has a sample sent for culture. Polymerase chain reaction (PCR) is used for intratypic differentiation, and any non-Sabin-like or indeterminate virus is sequenced. Genotypic indicators can then offer insight into transmission, including characterization of new virus introductions, epidemiological linkages between cases and surveillance quality. Environmental surveillance, including in areas where wild-type virus transmission has been interrupted but vaccine-derived virus transmission persists, can shed light on regional migration and flow patterns of the virus.

Inclusion of genetic epidemiology in the eradication initiative has provided a mechanism for accurately measuring key programme indicators: reduction in number of cases, geographic extent and genetic diversity. Identifying reductions in genetic diversity speaks to the overall progress towards eradication, but requires knowledge of both the natural existing reservoir of viruses (so that new virus types can be identified) and the origin of introduced viruses. In 1988, there were over 30 identified genotypes and three serotypes circulating globally. Two serotypes were eradicated in 1999 and 2012: wild-type poliovirus (wt-PV) type 2 and type 3, respectively (Fig. 2). Only two genotypes remain in circulation: the SOAS genotype circulating in Afghanistan and Pakistan (most recent case in May 2019) and the WEAF-B1 genotype, which was last detected in Nigeria in 2016.

Fig. 2. Eradication of WPV3 genotypes, 1986–2012



Similarly, reduction and disappearance of genetic clusters represent progress regionally and locally for control programmes. Clusters include isolates sharing ≥5% of VP1 identity. For example, in 2018, six distinct genetic clusters were detected from AFP cases and environmental samples from the SOAS genotype. Expansion and reduction of genetic clusters are linked to transmission reservoirs, indicator communities and cross-border transmission, and vary with seasonality and peak transmission seasons. As surveillance quality indicators, genomic data have been used to identify orphan viruses (>1.5% different from the closest matching VP1 capsid sequence), which are indicative of possible missing cases in the transmission chain. Such data can also inform local population targets for improved vaccination campaigns. Furthermore, genomic data have been used as a mechanism for quality assurance and quality control (QA/QC) by identifying contaminants and providing evidence on mismanagement of samples in order to facilitate improvements in surveillance protocols and data management. The successful integration of genomic data into the polio eradication initiative has been in part due to the comprehensive understanding of the poliovirus molecular clock, including the rate of natural evolution, which allows for accurate classification of nucleotide divergence in isolates to discern genetic lineages and chains of transmission.

Despite the successful use of genomic data at the local, regional and international levels of collaboration, the programme is not without operational challenges. Key operational challenges and areas with opportunity for improvement include capacity, utility, meaningful collection and use of data, quality of information, coordination and data sharing. The quality and use of sequence information are dependent on the quality of all aspects of surveillance. Within the Global Polio Laboratory Network (GPLN), there is still a need for increased capacity in sequencing capabilities, including standardization of methodologies, training, and QA/QC. Coordination and data sharing occur between WHO (three levels), national ministries of health, and other organizational partners. However, acceptance of genetic epidemiology data and use as part of routine surveillance and decision-making requires that data sharing be comprehensive and decision-making consensual. The GPEI is structured around the WHO regions and a laboratory network (GPLN) that has the necessary capabilities for conducting the molecular work that supports the programme. Information exchange is critical and considered a significant success of the programme. This example highlights the need for timeliness in communication and in the management and movement of samples to appropriate laboratories within the network in order to allow for genomic data generation and analysis. Timeliness is highly dependent on local capacity. Generally, sequencing can be completed within one month of sample arrival, although in countries such as Pakistan where there is local capacity for molecular work, sequencing can be completed within a few weeks. Despite the time needed for sequencing, information exchange occurs in “real-time”. Current gaps in data sharing and the availability of whole genome data to support interpretation of locally generated data remain a concern, as such gaps can delay the use of information in decision-making. Ideally, a full database of all virus isolations that is shared and managed collectively would ensure the availability of data for accurate interpretation in a timely manner. Maintenance of such a comprehensive surveillance strategy and eradication programme approach would require the partnership of multiple organizations. Most importantly, commitment and buy-in at the country level would be required to make the strategy possible. While external funders could provide additional support, without consensus among partners for a robust programme, the approach would not be sustainable.

3.2b. Genetic epidemiology and disease surveillance for outbreaks: Ebola

Presenter: Mark Perkins

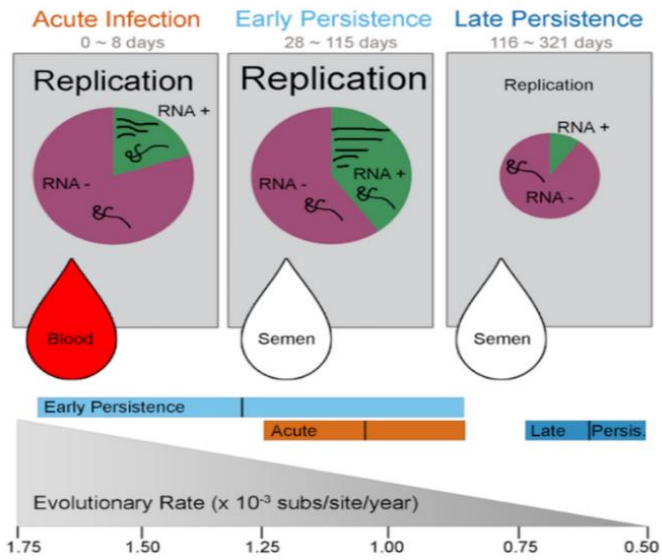
The accessibility of molecular tools to support outbreak investigations and emergency situations has increased dramatically due to declining costs and increased ease of use, including the recent dramatic simplification of sequencing platforms. In this context, genetic epidemiology can help to supplement and fill gaps where conventional epidemiological methods have failed. Particularly in the current outbreak of Ebola in the Democratic Republic of the Congo (DRC), the utility and depth of information obtained through conventional epidemiology have been hampered by social issues within the

community. Inadequate data yield an unclear or incomplete picture of the chains of transmission, which subsequently impacts outbreak management. Other key areas where flaws can exist in traditional methods include the over-/under-reporting of clinical cases, imperfect sensitivity or specificity of diagnostic assays, avoidance of health care facilities by at-risk groups or infected patients, misjudged or no information on close contacts, and inaccurate assumptions regarding the transmission source.

Incorporating sequencing routinely into Ebola outbreak management presents opportunities to improve accuracy in understanding the transmission and spread of the pathogen within the population. For zoonotic pathogens such as Ebola, it is important to distinguish between single introductions versus multiple transmission chains from animal reservoirs linked to human cases. Genetic markers can also identify transmission between individuals within the community, identify nosocomial spread and aid in the detection of infection control failures in health facilities, for example, transmission due to reused needles in pharmacies. The data can then be used to support the implementation of specific control measures to prevent human–human transmission or introductions from animal reservoirs by minimizing the risks of exposure related to clinical or community practices or the environment.

Another key utility of genetic data is to identify novel mechanisms of transmission, for example, through sexual contact or breastfeeding (Fig. 3). Additionally, using specific mathematical algorithms to analyse metagenomic and sequencing data can identify how many transmission chains have been missed, and help to estimate the true size and scope of an outbreak. This information can then be applied to decision-making for improving the surveillance system and expanding or targeting control measures in a given area. Genomic data can also reveal threats to the current medical countermeasures, such as mutations in PCR primer/probe sites, which could prevent the detection of the disease.

Fig. 3. Evolutionary rate of virus by stage of infection



While there are many opportunities and applications for genomic data to support outbreak management, as has been evident in the case of Ebola, significant challenges remain. Communication and information sharing, particularly in low-resource settings, can be critical. A platform or database would be needed to improve monitoring of traditional epidemiologic data and to ensure appropriate integration of the supportive information that genetic epidemiology data could provide. Furthermore, despite clear use cases for genomic data in decision-making, there has been no strategic change in the overall outbreak response to Ebola to make data use more systematic and streamlined. It is important

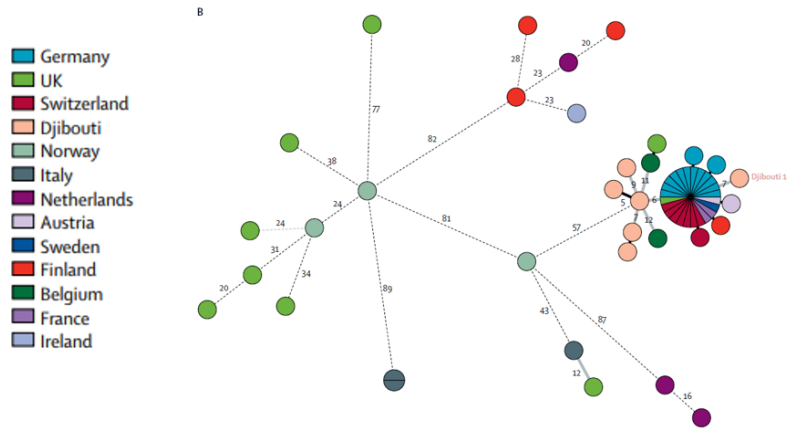
to note that security issues in outbreak and crisis situations play a significant role in the ability to roll out molecular-based surveillance approaches. Even when capabilities exist in an area, the response can be thwarted by security and access concerns, minimizing the ability to implement a robust lab-based component to outbreak control. For example, in the current outbreak in the DRC, compared to the West African Ebola epidemic or previous epidemics in the DRC, serious security issues have prevented access to the geographical area of the outbreak. This situation has created major difficulties in managing the outbreak, despite past experience, collaboration and training in effective Ebola response using genomic epidemiology. As a result, minimal sequencing has been done, which has minimized the availability of statistical predictions and response algorithms to support decision-making. While some sequencing data are available, accurate assumptions cannot be made for Bayesian analyses because meta-data are lacking, rendering the sequencing data useless in that context.

3.2c. Genetic epidemiology and disease surveillance for ongoing transmission: TB

Presenter: Anna Dean

For pathogens with ongoing transmission and varying burden around the globe, genetic epidemiology can support surveillance efforts to understand disease trends, changes in transmission over time, and threats to countermeasures such as emergence of drug resistance. The Global Project on Anti-TB Drug Resistance Surveillance, hosted by WHO, was initiated in 1994 and has since become the oldest and largest antimicrobial resistance project. The project estimates prevalence of drug resistance among people with TB, captures trends, and guides resource allocation, planning and policy development. Through its network of supranational TB reference laboratories, the project integrates whole genome sequencing (WGS) to conduct global surveillance and monitor trends in drug resistance. In high-income settings, WGS is being increasingly incorporated into investigations of cases of multidrug-resistant (MDR-) TB, such as in a recent cross-border outbreak of MDR-TB in Europe among migrants from the Horn of Africa (Fig. 4). This investigation was possible because of the level of capacity and collaboration present throughout the region. However, the capacity for routine surveillance varies. Paradoxically, the countries with the highest disease burden are often the ones with the lowest capacity; they must rely on nationally representative surveys conducted periodically to estimate disease burden. Low capacity for continuous surveillance impacts the timeliness and availability of data for decision-making on drug resistance patterns or transmission, and for understanding the true prevalence of disease.

Fig. 4. Cross-border outbreak of MDR-TB *(2)*



58 patients with related MDR-TB strains

At country level, data generated from the project are used to identify local and regional needs for medicines and resources and can inform programmatic decisions based on the resistance patterns identified. Countries can use the data to set national targets for rifampicin-resistant TB through case finding, calculate second-line drug needs, assess the feasibility of new treatment regimens and guide national diagnostic algorithms. Conventional diagnostic methods followed by phenotypic drug susceptibility testing on culture isolates takes weeks to months. Rapid molecular testing using the GeneXpert platform or other tools presents a significant advantage in terms of the timeliness of results at the peripheral level, although this is currently limited to rifampicin only. In addition to the decreased length of time required for accurate results, sequencing also provides information on transmission chains and case clusters.

There are significant advantages associated with the incorporation of sequencing data into national drug resistance surveys. In addition to improved accuracy and reliability of testing for drug resistance patterns, these data allow for the assessment of the feasibility of new drug regimens and decision-making to guide programme efforts. Unfortunately, in most places, sequencing is not a local capability and must be outsourced to laboratories within the network. Capacity-building is required to make sequencing sustainable in regions with the highest density of transmission. Other considerations for scaling up sequencing in TB surveillance include the technological requirements (e.g., sample types: culture needed for WGS compared to either culture or sputum for targeted gene sequencing), logistics for sample transport systems, biosafety, and expertise required. There also needs to be increased local capacity for data storage, standardization of data, data reporting and interpretation, and cost–benefit analysis for implementation of an ongoing surveillance programme.

Incremental next steps that can improve access to next-generation sequencing technologies will be implemented with WHO support. WHO released policy guidance in 2018 on standardizing the approaches to conducting sequencing and interpreting results through standardized pipelines *(3)*. WHO has also developed and houses a multi-country database for countries to directly send sequencing data where the information can be safeguarded. The database currently contains population-representative isolates from 13 countries, with approximately 12 000 isolates in total. It will serve as a repository that can be used to support WHO analyses and data aggregation by region and to improve understanding of disease trends globally. To minimize concerns over data ownership and management by participating countries, the database will be closed and not available for manipulation and use externally. Data sharing depends on the country: While the majority of countries are willing to share data and send samples to partners for analysis, other countries prefer to rely on local data analysis and choose not to outsource molecular work to regional partners or share data that are generated locally.

WHO is also playing a role in promoting regional and country capacity for local sequencing efforts; for example, in Africa, there are efforts underway to increase capacity beyond the laboratories in South Africa, Benin and Uganda. WHO has also produced policy guidance and translation for action. At the country level, interest in incorporating sequencing efforts into national programmes may be shaped by different priorities. For example, some countries may be more interested in support for individual clinical case management, although interpreting the results is not straightforward. In the context of national surveys, sequencing data cannot be used for clinical decision-making, since data generation is too slow to support case management and such surveys are only conducted periodically (approximately every five years). However, data are being used to support revisions and adaptations of national diagnostic algorithms. For policy guidance, collaboration with the WHO Global TB Programme has supported guideline development and provided lessons learned to other diseases such as HIV that are further behind in their capacity for global drug resistance surveillance. Overall, the WHO Global TB Programme will continue to support national TB programmes and help meet their needs to establish quality surveillance programmes.

3.2d. Discussion: Key considerations in application of genetic epidemiology

There are several key considerations in the application of genetic epidemiology to the surveillance of polio, Ebola and TB. These include the type of pathogen, mode of transmission, and current situation for management and control within the population. Lessons learned from the use cases in ongoing surveillance programmes, eradication efforts and outbreak responses can be used to inform next steps in malaria genetic epidemiology efforts. Use cases that have been effective in supporting surveillance include 1) understanding transmission links and 2) identifying missed transmission chains. These use cases have been significant in gauging the intensity of the disease event and prevalence of disease, and helping to target prevention and control strategies based on the populations affected and at risk.

Despite the clear potential for use of genomic data in polio, TB and Ebola, key challenges were also highlighted across other diseases in terms of i) cost, ii) capacity and iii) data generation, sharing and use of information. Countries have varying capabilities and capacity for local genetic epidemiology methods and interpretation of data. In addition, depending on health system structures, facility set-up and maintenance can be difficult. Particularly in crisis settings, security is a major concern, presenting additional needs for maintaining adequate biosecurity around facilities, equipment and samples. For example, in the current Ebola epidemic, despite the presence of local facilities and technical capabilities, armed groups have targeted health facilities and laboratories, endangering the safety and security measures needed to keep the facilities open. In cases where network availability can support implementation of a genetic epidemiology programme, there are still local and national network needs for managing quality assurance and ensuring consistency in the logistics for sample movement, storage and testing results. In addition to data reliability, data ownership and reluctance to share data present further data challenges. In outbreak settings such as Ebola, the WHO R&D blueprint on pathogen genetic sequencing data and code of conduct for open and timely sharing of data have proved useful. However, this has not been translated for ongoing surveillance and elimination programmes in order to provide guidance on consensual data sharing and use. Lessons learned from previous applications of genomic data and the key considerations for further use cases will prove fruitful in informing future scale-up efforts and the incorporation of such data into other disease control and elimination programmes.

3.3. Session 2: Overview: malaria parasite, anopheline gene flow, modelling

Opening remarks by facilitator, Dyann Wirth

Gene flow is a generic term that describes the spread of genetic material between populations and/or locations. For example, gene flow between two locations implies that there is migration between these two locations, whereas gene flow between two genetic subpopulations implies that there is interbreeding between the subpopulations. Understanding gene flow in malaria parasite populations has the potential to drive the implementation of surveillance strategies to control spread, monitor resistance and evaluate the effectiveness of interventions. Gene flow can be measured across the whole genome or at specific loci. Differences in the rate of gene flow are related to the mutation rate, which varies at different loci through recombination or evolutionary selective pressure. Measuring genome-wide gene flow can provide estimated rates of dispersal, migration and interbreeding, whereas locus-specific gene flow can estimate rates of spread of drug resistance, insecticide resistance, and gene drive.

3.3a. Tracking gene flow in malaria parasite populations

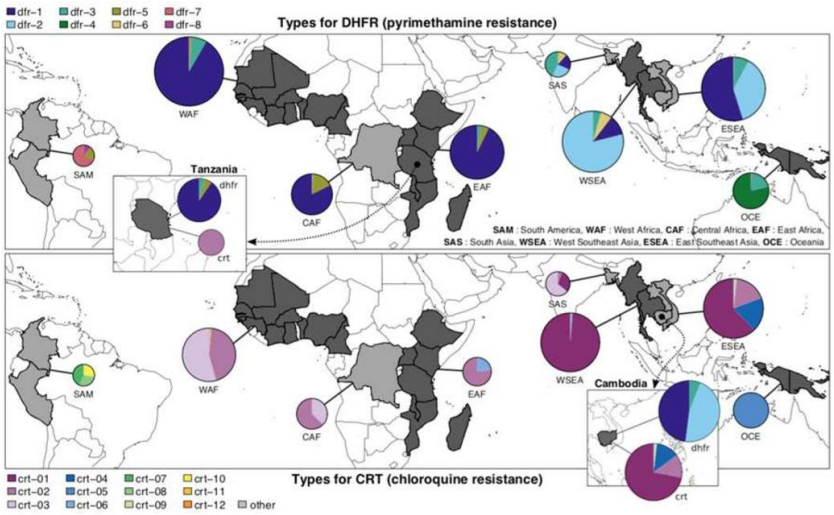
Presenter: Dominic Kwiatkowski

For malaria control, it is essential to distinguish between analytical use cases and operational use cases for understanding gene flow. The former involves understanding changes in epidemiology, whereas the latter is about applying genetic information to the decision-making process – a subtle but significant difference. Analytical use cases include efforts to understand changes in transmission

intensity, identify hotspots, and determine rates and routes of transmission. By contrast, an operational use case of malaria genomic data would inform plans for elimination zones, containment strategies for multidrug resistance, or approaches to tackle the resurgence of malaria in an area. When establishing a genetic epidemiology surveillance system, it is important to consider the type of use cases that are anticipated in order to ensure that the appropriate methodology and approach are being used to generate the type of data that can support analytical and/or operational use. For *Plasmodium*, this can be particularly challenging. It is necessary to maintain centralized, open genome sequencing data in order to understand lineages and recent common ancestry. For example, if two parasites have the same sequence at a large haplotype locus (e.g., >30kb or ~2cM), this implies that they must have a recent common ancestor at that locus and are of the same lineage. It is also necessary to understand the type of genotyping technology needed to generate these data. In the context of the emergence of resistance in malaria vectors, surveillance programmes can use specific markers (e.g., SNP barcodes or known markers of drug resistance), amplicon sequencing (e.g., haplotypes and new mutations at known resistance loci), or genome sequencing (e.g., signals of recent selection due to new forms of resistance). Chromosomes in a eukaryotic parasite like *Plasmodium* undergo meiotic recombination with every sexual generation. Consequently, there is high variability in the genome such that two randomly sampled parasites are unlikely to have the same chromosomal haplotypes. Therefore, sequencing technology in a surveillance programme must meet certain requirements so that it can provide useful data for understanding malaria epidemiology and key elimination concerns, such as imported cases in elimination zones or the emergence of resistance markers in a region.

Keeping in mind the complexities of gene flow in malaria parasites at specific loci, in the use case of understanding resistance, it is important to note that most forms of drug and insecticide resistance have multiple lineages with different patterns of spread and that some lineages can spread more aggressively than others. For national malaria control programmes (NMCPs), epidemiological interests lie in what resistant lineages are present in the region and which ones are newly emerging. For example, to understand the spread of artemisinin resistance caused by *kelch13* mutations (KEL1) *(4–7)*, a tiered phase approach was implemented. Phase 1 investigated the emergence of KEL1 in different parts of South-East Asia with notable localized geographic distribution. Phase 2 then investigated the rapid expansion of a related group of parasites that shared a specific lineage of KEL1 and a specific lineage of plasmepsin amplification (PLA1) that caused DHA-PPQ treatment failure in western Cambodia. The current phase 3 is investigating the KEL1/PLA1 co-lineage that has spread across the region and differentiated into sub-lineages that vary in geographical distribution and phenotype. While there are a number of resistant lineages, only certain lineages are sustained and continue to spread. MalariaGEN is producing global estimates of *Plasmodium falciparum* multidrug resistance based on genome sequencing of 7000+ samples to identify the most successful lineages of pyrimethamine resistance and chloroquine resistance. The project also promotes longitudinal genomic surveillance to support further analytical and eventual operational use cases for sequencing data (Fig. 5).

Fig. 5. Winning lineages of pyrimethamine resistance and chloroquine resistance based on genome sequencing of 7000 samples



A key application for understanding the gene flow of resistance in *P. falciparum* includes tracking outbreaks to determine the development and movement of resistant lineages as opposed to simply identifying whether resistance is present or not. Considering the diversity in the global population compared to local parasite populations, comparative analysis of point mutations at specific loci are not informative about parasite migration. Rather, understanding migration scenarios requires the ability to discern between external introduction, parasite movement between areas, and mixing or interbreeding in a given area. This analytical use case then gives way to operational applications in decision-making. Maintenance of shared resources, open data sharing, and capacity for data generation and sample testing are necessary to advance the field towards effective use cases for decision-making, programmatic and intervention impact, and guidance for resource allocation. Moreover, a framework for understanding the connection between genomic data and their application to interventions or policy-making needs to be clearly defined in order to facilitate the use of data in decision-making. For example, genotypic signals can be informative for understanding or expecting a phenotype in a region. These data can then inform survey strategies and further data collection to confirm genetic implications, thus providing stronger evidence to support decision-making based on genetic epidemiological information. However, this flow and approach to programmatic work would imply a significant change in the general framework for interpreting specific data to inform decision-making for malaria.

3.3b. Tracking gene flow in anopheline populations

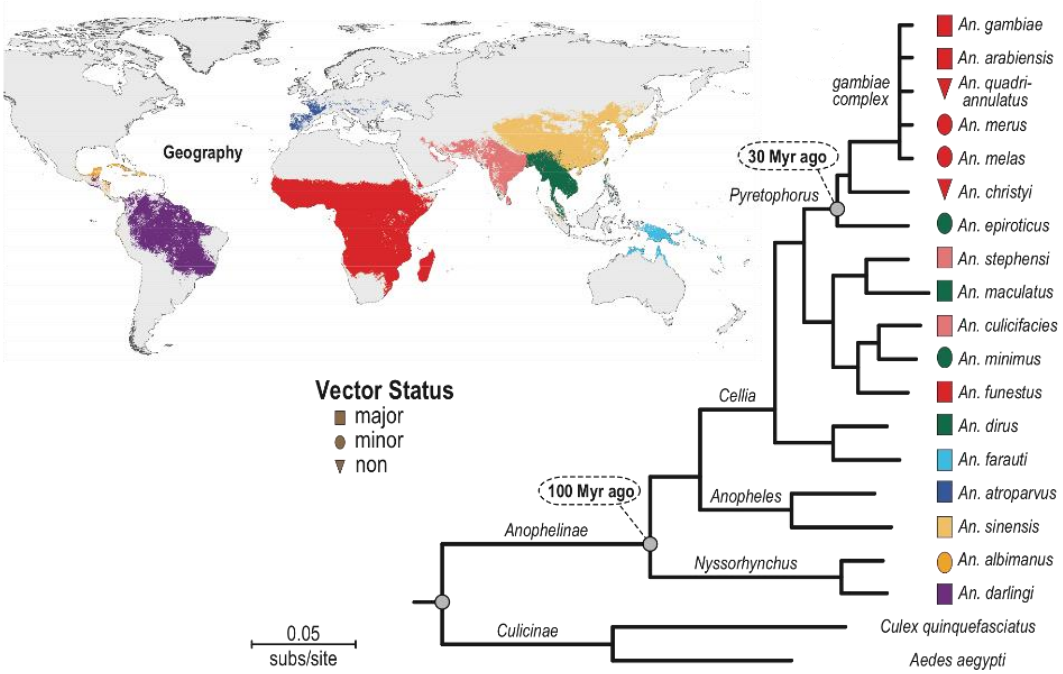
Presenter: Daniel Neafsey

Understanding and tracking gene flow in anopheline mosquitoes is complex given the amount of genetic diversity that exists within mosquito populations. There has been a long-standing need to improve integration between the field of molecular and medical entomology and the field of public health, along with a long history of understanding the underlying genotypes that are linked to phenotypes observed. Chemosensation, which governs many phenotypes including host feeding preference, and mosquito immunity to infectious microbes (including human *Plasmodium* parasites) are among the most rapidly evolving traits in mosquitoes. These traits contribute to the heterogeneity of mosquito populations and changes in complex traits such as vectorial capacity over short evolutionary periods. Vectorial capacity is dependent on multiple factors that vary among species, such as chemoreception, circadian rhythm, immunity to the parasite, insecticide resistance, reproduction, larval development habitat, and aridity tolerance, among others. Local vectorial capacity is therefore a function of species composition, and changes in this composition can impact

malaria transmission. Comparative genomics can be applied to understand differences in vectorial capacity and their impact on malaria transmission (Fig. 6).

A multi-locus approach is needed to understand anopheline gene flow because inversions, divergence and introgressions can occur, making a single-marker approach less than informative. Patterns in gene flow and species divergence are not uniform and can occur at different locations and rates along chromosomes. Such patterns can be missed if using a single-marker approach. Specific mutations can either lead to gene flow in populations or suppress it. Inversions are linked to niche specialization and lead to suppressed recombination and subsequent suppression of gene flow. Introgressions or interspecific genetic exchange, on the other hand, can lead to rapid changes in vectorial capacity. It is also important to note that there are still undiscovered species, cryptic species and aspects of vector ecology that are not fully understood. Consequently, attributes of unknown vector parasite interactions and vectorial capacity that have yet to be measured can have potential impacts on key concerns for malaria control, such as resistance patterns and intervention impact.

Fig. 6. Mosquito comparative genomics to understand differences in vectorial capacity *(8)*



Among other considerations in control programmes, new technological approaches such as gene drive and the use of genetically modified mosquitoes leave unknowns related to impact on the population dynamics of natural mosquito populations. For example, forced selection and gene drive may allow for introgression that may not otherwise occur. Sporadic hybridization events are also possible, but not sufficiently accounted for in current uses and studies related to gene drive. A whole genome strategy may be useful, namely in discerning cryptic barriers to gene flow and generating evidence on the emergence of hybrid mosquito populations. In addition, identification of common lineages and ancestors can also foster better understanding of migration patterns of mosquito populations and connectivity in a region.

In general, there is a need to improve surveillance of shifts in populations, understand rates of migration and insecticide resistance patterns, and consider the necessary studies and scales to determine the effects and impacts of gene drive. Opportunities for applying genomic tools to further understand mosquito populations and their movements or introductions to new areas, with

consideration for environmental changes and emerging issues such as climate change, are also important. A WGS strategy for local vector and non-vector species, identification of key local markers for taxonomy and insecticide resistance, and large sample collections for genotyping can help further strategies for malaria control and elimination.

3.3c. Discussion

In identifying use cases for malaria genetic epidemiology, the key is in discerning where the evidence is strong enough to consider policy development, and where additional information is needed to strategically develop research guidance that could later inform policy and operational use. In reviewing the use cases based on gene flow, a major concern is the exclusion of genomics in diagnostics from the Technical Consultation. Clinical applications of genomics and the impact of *pfhrp2/3* deletions on diagnosis need to be addressed. There are concerns that, in the future, rapid diagnostic tests (RDTs) will no longer be viable in Africa and therefore genomics and proteomics could be used to identify new markers for the development of new diagnostic tests, in particular rapid diagnostics that can be used in the field. For some, the advancements in genomics and the available technology for molecular diagnostics and speciation of parasites make identifying use cases for clinical applications ideal. However, development of new tools such as rapid diagnostics and identification of new markers are outside the scope of the current consultation. Rather, it is important to distinguish between operational and analytical use cases in surveillance or elimination certification contexts and identify use cases that are more actionable in a research space, such as the use of genomic data to drive identification of new resistance markers.

With expectations of operational use cases in genetic epidemiology, issues surrounding the generation of genomic data and subsequent data storage and data sharing need to be considered. Data guidelines and agreements are needed in a normative context in which WHO can offer support for longitudinal data generation in order to strengthen use cases with conceptual evidence that lack data for comparison and evidence in the field. In this respect, a repository or data storage platform that can support metadata aggregation during data collection could enhance data interpretation and integration for the decision-making process. This would also require capacity for management of big data and the possibility of data sharing within a region. Assessing what health system structures exist and what NMCP capacity is available will inform the development of or recommendations for any data storage platform. It is important to ensure that such a platform is not only functional now, but can also be adapted to future needs. Knowing what decisions the data will support can inform system development so that relevant data are generated and stored, with the awareness that NMCPs and policy-makers may have different data needs to support alternative decision-making processes.

Other use cases considered integral to understanding malaria transmission dynamics and informing elimination programmes include application to questions on importation. Currently, determining importation is often reliant on travel history and conventional epidemiology methods. Genomics can offer improvements in the process and more accurate data. When conventional metrics are used, despite their utility, there are often gaps; such gaps have been evident in Ebola outbreak management for example, and are even more complex in low transmission density settings for malaria. Applying genetic epidemiology methods using geospatial frameworks alongside genomic data on transmission chains can provide further inference into population-level transmission that may otherwise be incomplete.

From a funding perspective, understanding transmission dynamics on a more refined level and the application of genomic data in decision-making could help to elicit financial support for resources required to implement the use of genomic data in malaria control in both high-burden and elimination settings. Understanding current needs and projecting future needs will help to inform areas for capacity-building, appropriate settings for genetic epidemiology use, and generation of rich data that can support future policy recommendations and programmatic decisions.

3.4. Session 3: Parasite gene flow and the spread of drug resistance – setting the scene

Opening remarks by facilitator, Pascal Ringwald

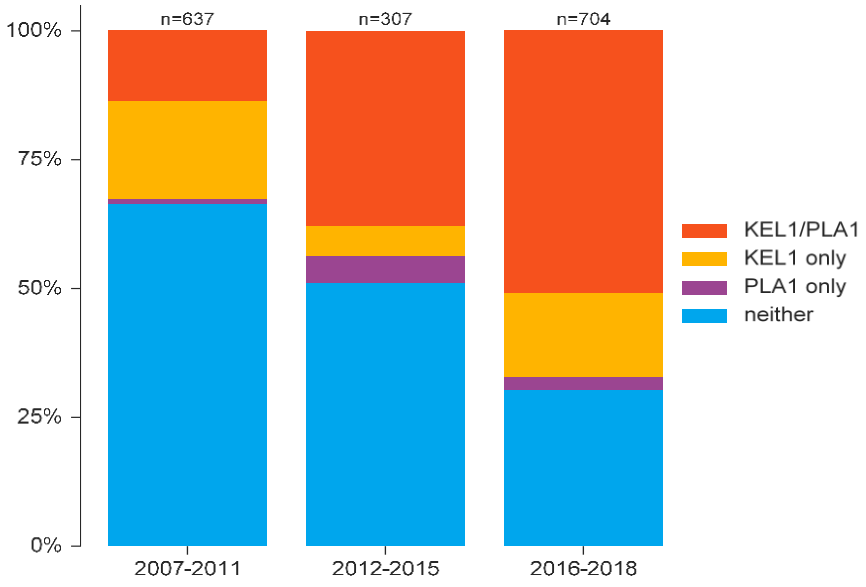
An important aspect of understanding the spread of drug resistance in the context of parasite gene flow is in determining the geographic origin of resistant parasites. It is necessary to discern if resistance emerged locally, or if a resistance gene was imported and subsequently disseminated within the local parasite population, thereby becoming established in a given area. There are several approaches to determining natural emergence or introduction when tracking resistance spread. However, depending on the genomic technology applied, results may be reliable and accurate, or leave questions and uncertainties. In terms of tracking drug resistance as a use case for malaria genetic epidemiology, the key questions are related to the minimum and optimal information requirements needed, and how to ensure precision of methods when determining the geographical origin of resistance genes.

3.4a. Tracking the spread of drug resistance using gene flow data

Presenter: Olivo Mioto

In South-East Asia, the GenRe-Mekong project has demonstrated the use of gene flow to track the spread of drug resistance. By integrating genotyping of dried blood spot samples and reporting on marker genotypes into routine NMCP activities, it has been possible to determine which use cases have some utility and identify gaps in how data have been translated. Understanding the gene flow of resistance in the region is crucial, as there has likely been a combination of a spillover event and selective pressure, but this is poorly understood. The project has shown clear geographic differences in resistance patterns for artemisinin and piperaquine, and pressure in certain areas that has aggressively contributed to the parasite population being replaced by introduced parasites (Fig. 7). The spread of the introduced parasites has been linked to gene acquisition, which has then facilitated spread to surrounding countries. Low resolution data from 101 single nucleotide polymorphism (SNP) barcodes have been used to discover the population structure and the movement of the resistance genes. However, even in areas where the KEL1/PLA1 co-lineage has excelled in terms of spread, there are still nuances in the population structure and variation in distribution that need to be understood *(9)*. It is also important to understand regional differences within a country because, due to differences in local resistance patterns, genomic data from one region may not appropriately inform decision-making in another area.

Fig. 7. DHA-PPQ resistance in the Greater Mekong Subregion *(9)*



In determining appropriate methods for understanding the emergence of resistant parasite populations, it is essential to have baseline whole genome data available for comparison. For example, in Papua New Guinea (PNG), where there was limited access to regional parasite whole genome data for comparison, a reference database of sequences from other South-East Asian countries was a critical aspect in accurately identifying the origin of the resistance patterns observed. Initial investigations compared microsatellite/SNP genotyping from parasites isolated from PNG data with resistant reference parasites and local parasites. In investigating CY580 mutant parasites to determine natural emergence versus importation from South-East Asia, identity by descent (IBD) analyses could not confirm where the parasites were from. There was confirmation of similar haplotypes corresponding to parasites from the South-East Asia region, and confirmation that there was no recent introduction. At the same time, there was some suggestion that there may have been early spread from South-East Asia, which allowed for similarities in the patterns of resistance emergence. Multiple lines of evidence using various genomic methods were needed to identify these different attributes and understand the parasite population, especially since current tools for IBD are not entirely clear when reconstructing the origin of resistance alleles. Most importantly, a large database of parasites from multiple regions needs to be available for comparison. This will allow for improved efficiency in comparisons across regions. While this approach is useful for understanding parasite origin, there are several limitations and results are not definitive. Extensive whole genome surveillance would be key to more reliable analyses, especially as emergence and importation may be more complex than expected.

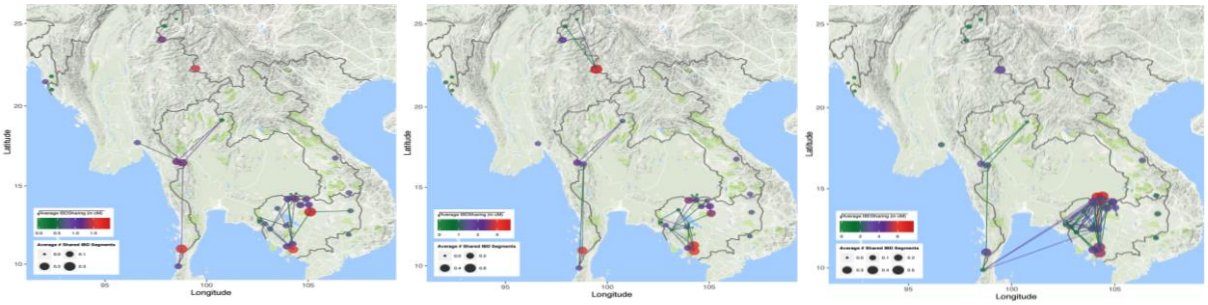
3.4b. Genomic structure, diversity and migration of *P. falciparum* in South-East Asia

Presenter: Shannon Takala-Harrison

In areas such as South-East Asia, where there have been ongoing efforts to eliminate malaria in the Greater Mekong Subregion, it is important to further understand the factors driving malaria risk in order to prioritize resources and optimize elimination strategies. Estimates of parasite migration are important in stratifying malaria risk, providing information about where parasites are moving or where there are barriers to parasite movement. Parasite migration has often been inferred based on human movement (regardless of infection status) from areas of high malaria prevalence. These studies are informative, but do not directly measure parasite migration. Thus, use of malaria parasite genomic data to understand parasite population demography can augment studies of human movement to understand parasite migration. In efforts to understand genomic structure, diversity and migration of *P. falciparum* in South-East Asia, different approaches have been applied.

A recent study aimed at mapping patterns of parasite population structure and inferring migration patterns using IBD analyses to determine the age of shared ancestry among sample isolates and estimate regional relatedness *(10)*. The study identified parasite genetic population substructure at a district level, based on shared IBD genomic segments. Parasites sampled from sites along the China–Myanmar border and in Bangladesh were relatively isolated from parasite populations in other regions of the Greater Mekong Subregion (Fig. 8), showing low genetic relatedness with parasites from other study sites based on IBD sharing. In addition, IBD estimates indicated connectivity between parasite populations along the Thailand–Myanmar border and within northern, western and southern Cambodia, but very little connectivity between parasite populations on the Thailand–Myanmar border and Thailand–Cambodia border, consistent with other studies that have indicated genetic differences between parasites in the western and eastern Greater Mekong Subregion. This study was also able to explore directional parasite migration based on admixture estimates, as well as likely drivers of increased IBD sharing in recent timeframes among parasites sampled in Cambodia, such as the selection and spread of multidrug resistance. While the IBD analysis in this context proved useful for understanding the parasite population structure in this geographic region, this approach did not explicitly model the spatial structure of the genomic data.

Fig. 8. Regional relatedness between parasites in South-East Asia *(10)*



A tool called estimated effective migration surfaces (EEMS) can be used to visualize spatial patterns in the data, allowing for visualization of geographic regions of more or less effective parasite migration for a given geographic distance between different sampling locations. EEMS does not currently allow for inference of directionality of migration. Additionally, genetic distance metrics need to be modified in order to better reflect more recent migration patterns and inform decision-making for operational use. The tool and approach can aid in understanding parasite population structure and migration and could potentially identify geographic units for interventions. However, the tool will require optimization to make it more spatially explicit to estimate local versus long-range migration patterns, and account for the impact of sample size and grid density to ensure accuracy in analyses.

3.4c. Discussion

In discussing approaches to understanding gene flow of resistance by geographic origin, there was a consensus that IBD analysis is a useful research tool but is not practical for NMCPs in its current state. The methodological approach would need to be distilled down to something that could be applied as a use case for control programmes in the future. For this to be feasible, it is clear that more rapid techniques and standardized markers would be needed at a minimum. In addition, data generation using a WGS approach across varying parasite populations from different countries and from geographical areas within the same country would be required for comparison. Considering the time needed to collect these data, a database would also be needed to store information over time. Some concluded that the information would be useful to uncover drug resistance patterns across the genome, instead of just at specific resistance foci, but that proof of concept is still needed in areas of mixed infection. Moreover, at a global level, although these data could be used to answer questions on whether certain drug-resistant parasite lineages are spreading between countries or regions, the approach is not timely and cannot be used in its current state to make policy decisions on treatment. Confirming population structures and understanding migration are currently research priorities; yet, these approaches do not provide conclusive evidence that can be used for programmatic action at this time. More research is therefore needed before IBD analysis can be used in operational use case scenarios.

3.5. Session 4: Parasite and mosquito genetics to understand transmission intensity – setting the scene

Opening remarks by facilitator, Jan Kolaczinski

When it comes to understanding gene flow in parasites and mosquitoes as it relates to transmission intensity, there are priority questions in malaria surveillance, vector surveillance, insecticide resistance management and evaluation of new vector control tools. For surveillance, it is necessary to better understand importation risk and receptivity, and changes in transmission intensity over space and time. Within the vector population, drivers of population change, spatial and temporal variations in the patterns of resistance, and adequate methods and sampling strategies for measuring such changes are important considerations. In addition, in terms of novel control tools, such as gene drive

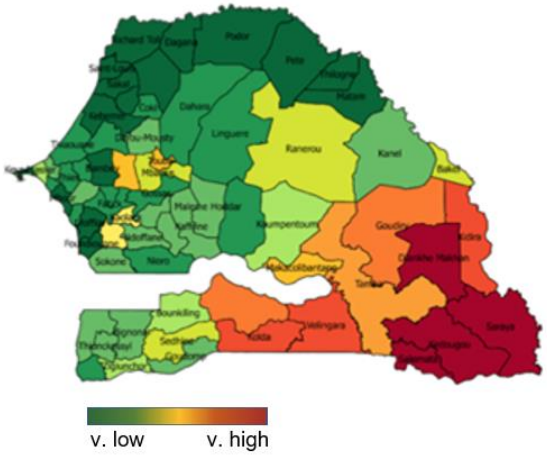
and genetically modified mosquitoes, genomics has the potential to contribute to field evaluations designed to assess the effects on local populations and selection for resistance to these new tools. WHO requires high-quality evidence to support the development of guidelines and practical manuals to support implementation of surveillance activities and deployment of interventions by NMCPs.

3.5a. Parasites: tracking gene flow and its relevance to transmission intensity

Presenter: Daouda Ndiaye

The elimination programme in Senegal has successfully applied genetic epidemiology in the control of malaria as a result of good local capacity for implementation, operational research partnerships and collaboration to monitor and evaluate elimination progress. Community engagement has aided in the acceptance of genetic epidemiology as a means to monitor progress towards elimination and support outbreak investigations. In Senegal, interventions are stratified according to transmission intensity, and use cases are targeted towards answering key questions for malaria control and elimination in order to measure intervention progress and impact (Fig. 9).

Fig. 9. Stratification of malaria incidence in Senegal



One key area where genetic epidemiology has been applied is in detecting persistent local transmission in low transmission areas and determining changes in parasite populations in high transmission areas. In areas with very low incidence (<1/1000) and presumed no local transmission, the identification of identical parasites persisting across multiple transmission seasons suggests that local transmission is ongoing despite low rates. Genomics can confirm that there has been persistence of the same parasite population over time rather than importation or new emergence of another parasite population from a different area. This information can then contribute to progress in malaria control in the area and intervention success. This use case of understanding parasite composition nationally and regionally can then inform elimination progress. Similarly, in detecting changes in high transmission areas, increasing parasite relatedness can be considered a possible early indicator of intervention impact. Alternatively, in an outbreak scenario in which parasites are confirmed in a location where transmission should not be occurring, the use case of understanding transmission foci can allow for case cluster investigations and confirmation of elimination success by characterizing an introduced parasite.

In Senegal, where programmatic application of genetic epidemiology methods has been established, there is an opportunity to scale up surveillance approaches. The current way forward is to scale up the capacity to apply tools in use case scenarios with past success and continue to integrate data from routine surveillance activities for both parasite and vector populations, monitor resistance, and conduct spatial risk mapping. Continued use of traditional epidemiologic data alongside supplementary genomic data to provide additional evidence and precision in complex transmission

scenarios will support programme decisions and exemplify the application of genetic epidemiology at the NMCP level.

3.5b. Tracking gene flow and its relevance to insecticide resistance – the example of *Anopheles gambiae*

Presenter: Alistair Miles

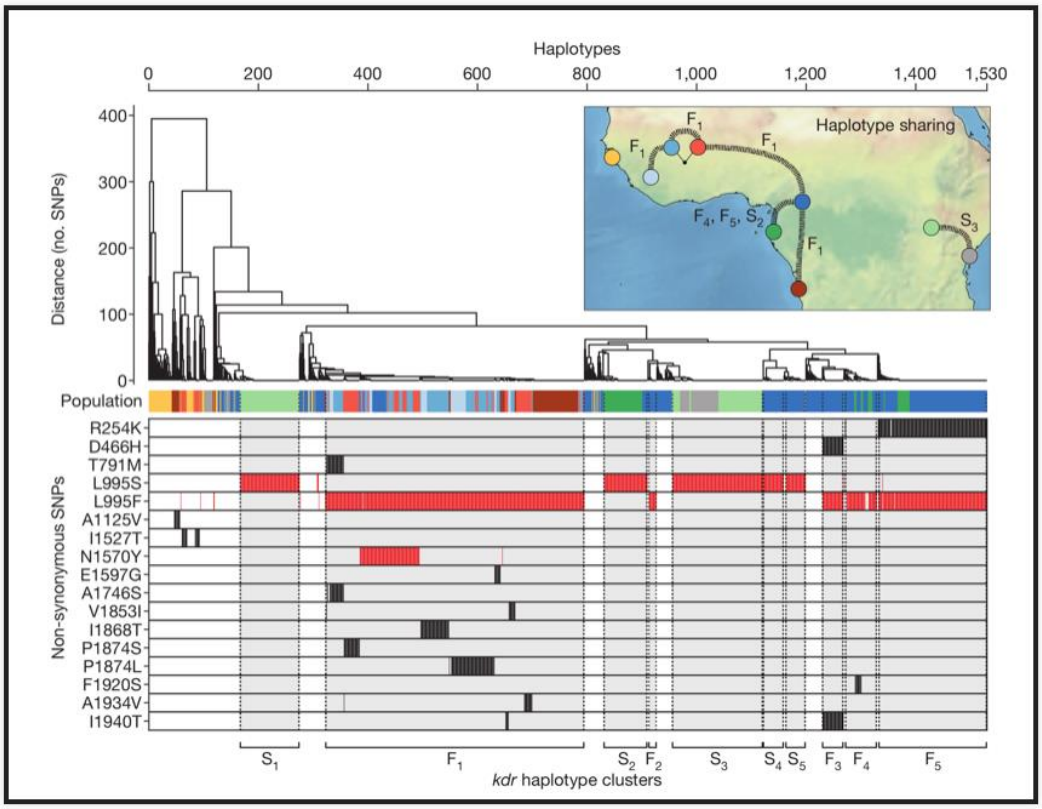
In determining use cases relevant to tracking gene flow in vector populations for insecticide resistance, the irony is that decision-making consistently excludes molecular data that could support decisions that are inherently molecular. Practical use cases for incorporating genetic epidemiology data include assisting decision-making on whether to procure next-generation long-lasting insecticidal nets (LLINs) or whether to deploy indoor residual spraying (IRS) and facilitating the coordination of cross-border resistance issues. Many of NMCPs' key questions on where and what resources are needed often require some genomic data, especially if the decisions made are to be truly informed by the transmission situation.

As previously mentioned, access to WGS data representative of a region can facilitate research and surveillance for improved understanding of the gene flow that is occurring. For mosquito populations, efforts are underway through the *Anopheles gambiae* 1000 Genomes Project (Ag1000G) to create an open database of anopheline genomes in order to aid investigation of genetic variation and evolution in natural mosquito populations. The project employs a three-phased approach to collect data on *An. coluzzi*, *An. gambiae* and *An. arabiensis* across a broad geographic range of up to 18 countries. These efforts have provided insight into the genetic variation that exists in areas of malaria transmission.

When investigating gene flow, in addition to understanding flow between species and locations, it is also important to understand changes across population generations. Certain genes may be under stronger selection, for example, and increase in frequency within the population. Understanding these changes can aid in making predictions for future generations of mosquitoes and identifying genetic indicators or markers of resistance patterns that may emerge. Specifically, in understanding selection of resistance genes, evidence of selective emergence related to a particular driver or isolated instances across populations can also provide the basis to infer influences on gene flow.

Two examples of selection and spread of resistance genes include pyrethroid target-site resistance and pyrethroid metabolic resistance. With pyrethroid target-site resistance, the spread of "knock-down resistance" (*kdr*) mutations in the voltage-gated sodium channel gene (*Vgsc*) was of concern. Two known *kdr* mutations were in circulation and there were questions as to whether the mutations were spreading and where gene flow was occurring. Gene flow could be inferred by identifying the same *kdr* haplotype in two different locations using over 1700 biallelic SNPs from all mutations within the *Vgsc* gene (Fig. 10).

Fig. 10. *kdr* haplotype clusters



With pyrethroid metabolic resistance, it was important to determine the spread of copy number variations in cytochrome P450 genes, because increased gene copy number implies increased expression and increased expression implies pyrethroid resistance. There are multiple P450 genes in the genome; for example, *Cyp6p/aa* and *Cyp9k1* are two loci where gene amplifications are common. Gene flow could then be inferred by confirming duplicate resistance loci across different populations. Interestingly, it is clear that gene flow is occurring and that different patterns of spread exist. While multiple independent events drive resistance, some spread, whereas others remain localized. The key lies in using this information to address strategic programmatic questions such as where best to deploy piperonyl butoxide (PBO) LLINs based on the evidence of resistance gene flow occurring within a region, while also considering factors related to cost and logistics.

While there have been many advances in understanding particular vector species, there are still gaps in the broader understanding of mosquito populations. There are still unknowns over what potential vector species may be present that may be contributing to transmission. These species may even be contributing to selection pressure for genes and influencing gene flow without any notable data or information to reveal the full scope of what is happening. Efforts such as Ag1000G can further support by scaling up genome sequencing of vector populations, increasing geographical coverage, conducting regular seasonal sampling in different areas and including other vector species. In addition, there is a need to bridge the gap between research and national programmes to begin to investigate and apply gene flow information in more analytical and operational use cases.

3.5c. Discussion

Despite clear examples of genomic epidemiology use cases in understanding intervention impact, surveillance of resistance, and progress in control programmes, there are still some questions as to what types of data are truly needed to inform decisions and support NMCPs. Rather than isolated use cases where genomic data have been deemed useful in verifying conventional epidemiology information or in supporting a single procurement decision, there is a need to determine the utility of

robust genomic data generation for supporting strategic decision-making over time and across various transmission scenarios. For example, in discussing the use case of procuring PBO LLINs, a combination of factors such as cost and logistics would need to be considered in decision-making, not solely the understanding or awareness that there may be gene flow occurring in the locale that is contributing to the presence of resistance patterns. There are concerns related to the implementation of new tools in terms of further driving gene flow and selection of future resistant mosquito populations. Could rolling out new chemicals to address resistance lead to similar issues as seen with the poor use of new drugs and antibiotics that has contributed to the emergence of antimicrobial resistance? With regard to using genetic epidemiology to provide information on decreased parasite population diversity and reduced intermixing by region, what programmatic changes or best practices could be identified to support use of this information for implementation in other geographical areas? There is still a need to simplify tools and address limitations in the timeliness of data collection so that the data generated remain relevant while supporting decision-making. Additionally, the identification of use cases in understanding parasite and vector population genetics for malaria surveillance is being carried out disparately. If new resources and tools are to be introduced, there needs to be increased coordination between parasite and vector control applications in order to ensure strategic implementation of tools by NMCPs.

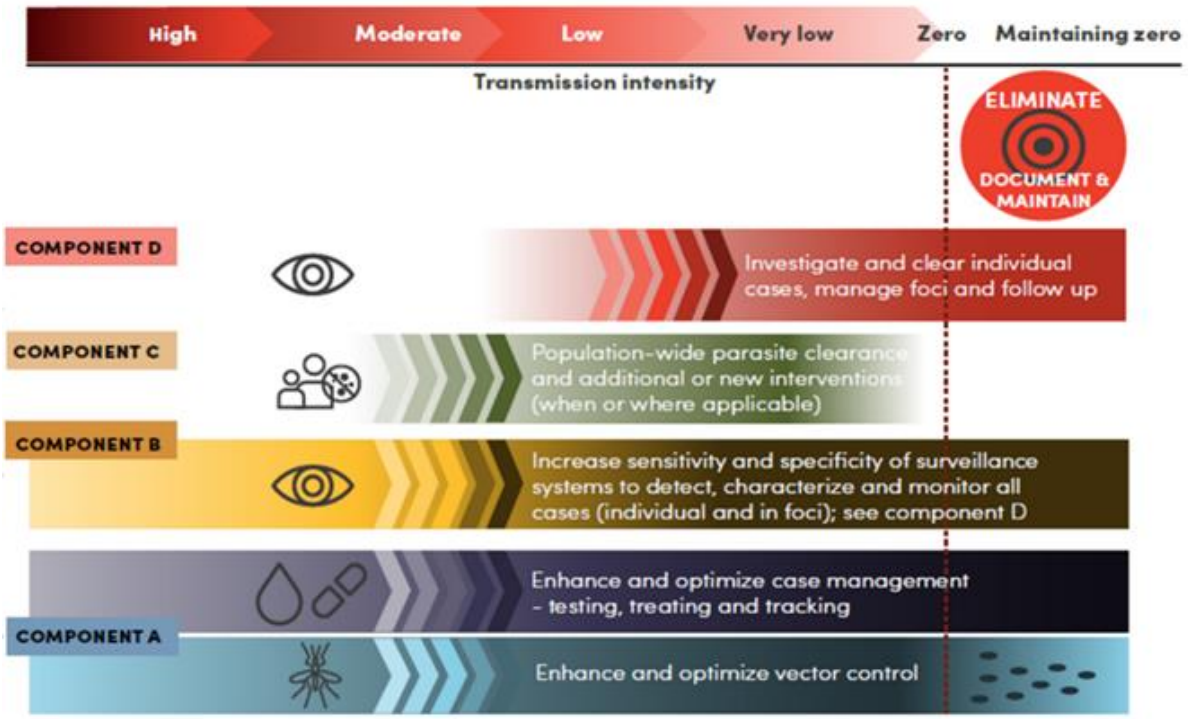
3.6. Session 5: Parasite and anopheline gene flows to understand importation and identify foci of transmission

Opening remarks by facilitator, Kimberly Lindblade

Genetic epidemiology can play an important but varying role in the control of malaria across the continuum from high to zero transmission (Fig. 11) *(11).*[2] In elimination settings, countries experience many years of low-level transmission before reaching and maintaining zero cases. In low transmission settings, strong passive surveillance, comprehensive case investigations that elucidate the likely location of infection, active case detection, and targeted interventions in foci of active transmission are important. In areas with high malariogenic potential, preventing re-establishment is essential. An identified need and opportunity for genetic epidemiology is in providing evidence to support the correct classification of cases as imported, introduced or indigenous, particularly in the absence of epidemiologic data. In "getting to zero", it is necessary to understand whether resurgences are due to poor case detection and surveillance, or whether persistent transmission is a result of repeated importations or indigenous transmission. Adequate spatial resolution could improve our understanding of cross-border transmission and case origin, thus also supporting efforts to prevent re-establishment. Genomic data are likely to be useful in improving our understanding of the transmission dynamics in eliminating countries. However, their operational utility will depend on the quality of surveillance and epidemiologic data collection, the availability of recent genomic data, data accuracy and the time it takes to generate the data.

[2] WHO has guidance on the tools, activities and strategies required to achieve malaria elimination and prevent re-establishment of transmission in countries, regardless of where they lie across the spectrum of transmission intensity. The framework informs national malaria elimination strategic plans and should be adapted to local contexts. Download the framework at: http://apps.who.int/iris/bitstream/10665/254761/1/9789241511988-eng.pdf

Fig. 11. WHO Framework for malaria elimination 2017



3.6a. Use of genetic evidence in the Pan American Health Organization (PAHO)

Presenter: Kumar V. Udhayakumar

The application of molecular tools to support programmes in post-elimination settings in the PAHO region requires a database of parasite genomes to help identify the origin of the parasites and subsequently inform public health responses. Genomic data on drug resistance markers and genotypes could support identification of imported cases, transmission foci and parasite migration, including in post-elimination settings in the region.

In an example of outbreak investigation from Peru, determining the parasite origin was only possible because of the rich data that existed – more than a decade of longitudinal data relevant to understanding the malaria parasites in the region. In the Peruvian Amazon, there was emergence of a drug resistance profile, Bv1 clonal lineage, that was distinctly different from the previous genotype found in the region. The Bv1 lineage profile posed a significant problem because the strain is multidrug-resistant and escapes detection by *Pfhrp2*-based RDTs secondary to *pfhrp2* and *pfhrp3* deletions. The hypothesis was that the Bv1 strain had emerged as a highly successful parasite lineage for transmission by different vectors and had contributed to the increased malaria burden recently observed in some Amazonian regions. Genetic connectivity was found between *P. falciparum* populations in Colombia and Ecuador, where there were also outbreaks. This is a critical use case scenario for understanding gene flow within a region, considering the significant impact resistance emergence could have on case detection and management.

The presentation highlighted an example of genetic epidemiology being used to identify imported cases to Guatemala in UN Peacekeepers who had spent nine months in the DRC. Genotyping showed that the infections in the returning peacekeepers were caused by parasites that were genetically related to DRC parasites and distinct from local Guatemalan parasites. This finding supported the case classification as imported (Fig. 12) and led to the implementation of a screening policy, treatment protocols and prophylaxis in peacekeepers. Similarly, confirming the parasite origin of imported cases in non-endemic countries and understanding onward transmission can inform prevention guidelines

for travellers. These analyses, among others, were successful due to regional partnerships, a longitudinal database, and the availability of data from multiple countries in the PAHO region.

Fig. 12. Neighbour-joining tree of three P. falciparum populations showing genetic connectivity between peacekeepers' parasites and DRC parasites *(12)*

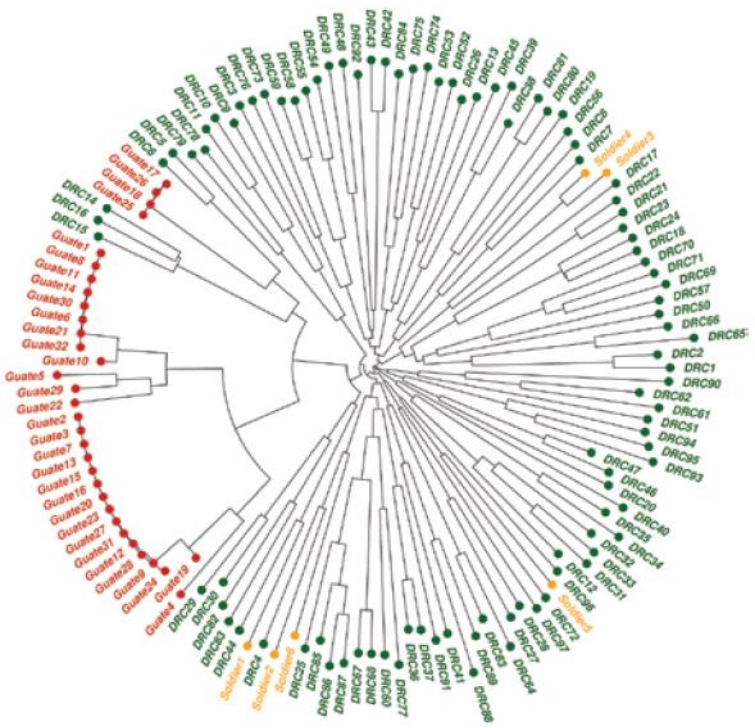


Figure 3. Neighbor-joining tree of 3 *Plasmodiium falciparum* populations. Prefixes of genomes indicate parasite origins: Green text indicates parasite populations from the Democratic Republic of the Congo (DRC); orange indicates parasite populations detected in soldiers who were returning from the DRC to Guatemala; red indicates parasite populations from Guatemala.

3.6b. Use of genetic evidence in Greece

Presenter: Danai Pervanidou

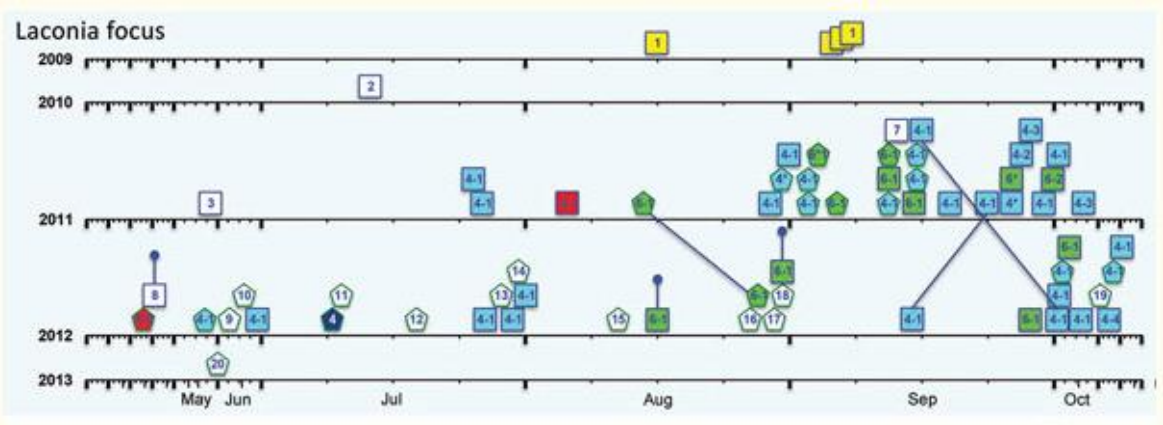
In Greece, there were several examples of genomic data supporting risk assessments, decision-making and case classification in the NMCP. Greece – a malaria-free country since 1974 – reports between 20 and 110 imported cases per year. A high receptivity risk combined with influxes of migrants from the Indian subcontinent has led to sporadic introductions and local acquisitions of *P. vivax* cases and one event of indigenous transmission in a particular area in 2011–2012. Given that this combination of continuous recording of imported cases and high receptivity risk increases the country's risk for malaria reintroduction, it became essential to establish an action plan for the management of malaria. The action plan is supervised by a multisectoral national committee and includes a series of prevention and response activities. In this context, the use of both genetic and epidemiologic data has supported risk assessment, surveillance and response.[3]

A key issue for Greece is that the WHO classification and definition of an introduced case requires documentation of the index imported case, which is often difficult to detect. Vulnerable populations, including refugees and undocumented migrants from malaria-endemic countries carrying out seasonal agricultural work, pose a local risk for malaria reintroduction (especially when the parasite is imported into a vulnerable area). However, due to barriers in vulnerable individuals accessing the health care system (due to e.g. fear of arrest/deportation, suspicion of the government and health services, language barriers, transport difficulties, etc.), their high turnover from one area to another,

[3] For more information on the historical context and current epidemiological surveillance of malaria by the National Public Health Organization in Greece, see: https://eody.gov.gr/en/disease/malaria/.

and mild manifestations of *P. vivax* relapses, it is often difficult to detect and record all imported cases that may have led to or will lead to introduced cases. As a result, the epidemiological criteria for classification of an imported case were adapted to account for these facts. The adapted criteria considered all *P. vivax* cases in migrants from endemic countries with symptom onset within three years post-arrival to be imported cases. It was necessary to apply genomic epidemiology to improve case classification and understanding of local malaria transmission routes. One case study in an area with a cluster of *P. vivax* malaria cases (2011–2012) among both migrants from endemic countries and Greek residents sought to distinguish between imported and locally acquired cases in order to understand transmission routes and risk in the community, and interrupt local transmission. Following genotyping, some of the cases were reclassified from imported to locally acquired because evidence suggested that there were clusters of linked cases based on similar haplotypes (Fig. 13). Similarly, in another case study that genotyped *P. vivax* cases occurring in two neighbouring households, genomic data allowed for reclassification of the cases (migrants from endemic countries) as locally acquired even though they were initially classified as imported; all cases, however, had the same haplotypes. While some of these analyses were conducted retrospectively, the approach could still be used to assess the transmission risk and inform timely decision-making and response within the programme.

Fig. 13. Clusters revealed through genotyping, Evrotas, 2011–2012



It is important to note that, in Greece, genotyping enabled a more comprehensive understanding of the malaria transmission risk in the country. Multiple importations and distinct introductions were confirmed, inferred from the detection of significant haplotype diversity and identification of small clusters. Insight into the transmission chains led to confirmation that there was no ongoing local transmission. Understanding the epidemiological situation facilitated decision-making that enhanced efforts to prevent the re-establishment of *P. vivax* in specific areas. Guided responses included mass drug administration programmes among migrants in high-risk areas (i.e., the area with the indigenous transmission event in 2011–2012, and the area with the cluster of introduced cases in two neighbouring households); reclassification of local risk levels to enhance the surveillance activities of the national malaria prevention programme; and intensification of response including proactive case finding and vector control in areas of high risk. In some cases, however, there were still limitations because the link between cases could not be confirmed and there was no supporting evidence to understand transmission routes. Questions that emerged included the possibility that the cases were linked by the same haplotype that was common in a particular area of the world, but were not linked epidemiologically; in other words, the cases were not in the same transmission chain. In this case, it was not possible to properly interpret the genetic analyses due to i) the need for more information on *P. vivax* genetics, including the geographical distribution and frequency of certain haplotypes in malaria-endemic countries, and ii) the lack of standardization of the loci used for genotyping, which is required for cross-border comparisons and to provide a better interpretation of results. Moreover,

knowledge is limited on the impact and potential role of the local vector population, and the species vectorial capacity of various imported *P. vivax* strains.

3.6c. Use of genetic evidence in China

Presenter: Junhu Chen

The National Institute of Parasitic Diseases, China CDC also presented significant use cases for genetic epidemiology in documenting progress towards elimination. One key focus was in understanding the parasite population in order to discern regional gene flow from introduced parasites and support case investigations. A use case that worked involved determining parasite relatedness in situations where cases were suspected to be linked by a local vector through nosocomial transmission, but genomic data were required for confirmation. Similarly, genomic data were able to provide evidence of transmission where conventional epidemiology could not in the investigation of a local case with no travel history in an area considered *P. falciparum*-free for two decades. In general, population and comparative analyses are effective when supporting genomic data can aid in case classification, determine parasite relatedness and trace geographic origin using IBD analyses. Limitations still exist depending on the parasite density, complexity of infection (COI) and sample size. In addition, access to a database of parasite genomes for comparison is key to identifying potential links, migration patterns and the geographic origin of parasites in supporting the case classification of imported and local transmission.

3.6d. Discussion

The PAHO region, Greece and China have different malaria transmission settings, but there are similarities in how genetic epidemiology has been applied to support NMCPs to identify transmission foci, determine risk of transmission and improve case classification. The discovery of *pfhrp2/3* deletions among clinical isolates in Peru in 2008 prompted retrospective investigations and prospective surveys in multiple neighbouring countries, including Brazil, Bolivia, Guyana, Honduras, Suriname and Colombia *(14)*. These findings directly informed policy around the use of RDTs in the PAHO region. While it is clear that the application of genetic epidemiology is useful in local settings, the need for a data repository of available local and global genomic data, as well as a network to facilitate data sharing is evident. Although genetic epidemiology allows for increased precision in case classification, there are concerns over the lack of standardization of methods for genotyping and subsequent data analysis across different geographical settings. It is also necessary to understand the local environment and susceptibility for transmission risk, so that this information can be used in conjunction with conventional and genetic epidemiology data to inform programmatic or policy decisions.

3.6e. Imported versus local transmission

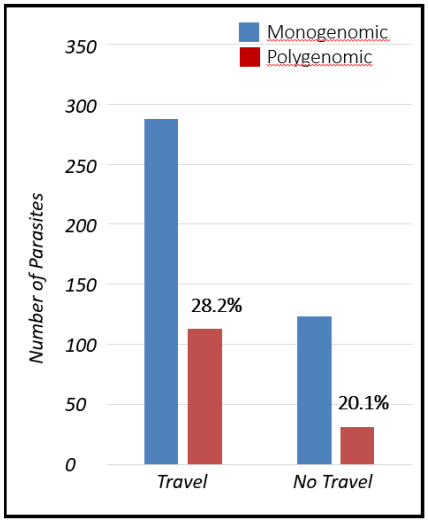
Presenter: Sarah Volkman

For NMCPs, the issues of importation and identification of transmission foci are important for tracking elimination progress and maintaining status as a malaria-free country. To further understand how genomic epidemiology can support case classification, it is necessary to first understand how parasite populations change with changing transmission intensity. As transmission intensity decreases from high to low, there is decreased COI, increased proportion of monogenomic infections (COI=1), appearance of clonal parasites and persistence of clonal lineages. This means that measures of parasite relatedness and connectivity can be used to understand transmission in an area and signify programme progress.

Genetic relatedness is measured using metrics of identity by state (IBS) and identical by descent (IBD): alleles that are genetically the same, and alleles that come from a common ancestor, respectively. Methods can be used to estimate IBD from IBS. This requires a number of informative markers (molecular barcode genotyping) that vary depending on the level of transmission in the geographic

area under examination. A barcode is considered informative for relatedness by IBS at >0.95 relatedness. When using this measure in a low transmission setting, relatedness can serve as a key indicator for distinguishing imported and local transmission and understanding the persistence of transmission in the area. For example, in Senegal, it was determined that there was an increased likelihood of polygenomic infections in travellers versus infections in households with no travel history, which were more likely to be genetically similar (Fig. 14). Additionally, there were identical parasites persisting across multiple years, which suggested maintenance of local transmission, albeit at low levels.

Fig. 14. Increased likelihood of polygenomic infections from travellers in a low transmission region: Richard Toll; Senegal



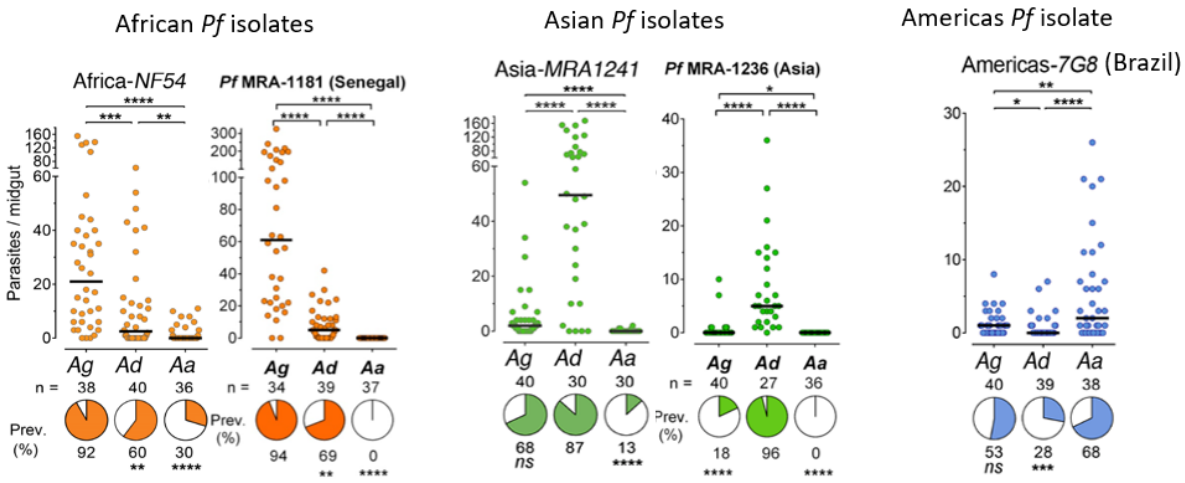
The ability to distinguish between imported and local infections can be critical to elimination and provides evidence for decreasing COI in parasite populations. Sequencing/amplicon data can reveal genetic connectivity to resolve questions in complex transmission settings; however, there is still a need for more genetic markers to expand the use of this methodology. In the case of Senegal, a 24 SNP barcode was used to characterize parasites. While these data were consistent with conventional epidemiologic data in terms of importation, directionality could not be confirmed. In this case, good traditional surveillance was necessary to ensure accurate interpretation of the genomic data.

3.6f. Pfs47 SNPs as a risk indicator for transmission of imported malaria

Presenter: Alvaro Molina-Cruz

An alternative approach to understanding receptivity risk for imported and onward indigenous transmission of malaria is to investigate parasite markers in parasite–vector interactions that determine whether the parasite can successfully infect the mosquito. A target of interest is *Pfs47*, which allows the ookinete to evade the immune response of the mosquito midgut and successfully develop into an oocyst. The allele is polymorphic with signatures of natural selection relevant to the geographic origin of the parasite. *P. falciparum* isolates are more compatible with *Anopheles* species from their region of origin (Fig. 15). This is related to parasite evolution as a result of both natural and forced human migration through which the parasite, but not the vector, was moved to different regions of the world. The allele is then linked to parasite development in the mosquito with adaptations in local mosquito populations within their respective regions globally. *Pfs47* SNPs can therefore be used to predict the transmission risk of imported *P. falciparum* and help establish its geographic origin. More data are needed to discern the boundaries for changes in which haplotypes are more or less prevalent. Additional research is needed to develop the evidence base for this phenomenon in *P. vivax.*

Fig. 15. *Pf* isolates are more compatible with *Anopheles* species from their region of origin *(13)*



3.6g. Discussion

The approach used for importation classification in low transmission settings in Senegal was based on the detection of decreased genetic diversity in local parasite populations. It is evident that more research is needed to build the evidence base on the underlying diversity of the local parasite population. In areas where knowledge of the local parasite population is poor, it will not be possible to identify imported cases, despite observations of decreasing polygenomic infections. In the case of *Pfs47*, this marker can offer information on the geographical origin of the parasite and whether the local mosquito populations are receptive to the genotype – a factor that could lead to ongoing transmission from imported cases.

3.7. Session 6: Data standardization, modelling and use

Facilitated by Abdisalan Noor

It is critical that genomics data be accessible so that important policy questions can be explored and platforms developed in order to eventually provide information products that are relevant to national malaria programmes. While the collection of these data is likely to remain within the realm of research in the near future, more routine processes for data assembly will increasingly become the main source of such data. This poses logistical as well as governance and ethical issues. Existing platforms for drug and insecticide resistance surveillance monitoring could also function as effective mechanisms for collecting genetic samples. As sample collection moves into the realm of routine surveillance systems, the burden on the health system and the ownership ethics involved in collecting samples in non-study settings will become important considerations. This will require the development of protocols and potentially normative guidance to advise countries on the way forward. Appropriate, statistical, geospatial and mathematical analysis methods should be explored so that results can be packaged in a way that is relevant to policy.

3.7a. Combining modelling and genomic surveillance data: insights for malaria elimination campaigns

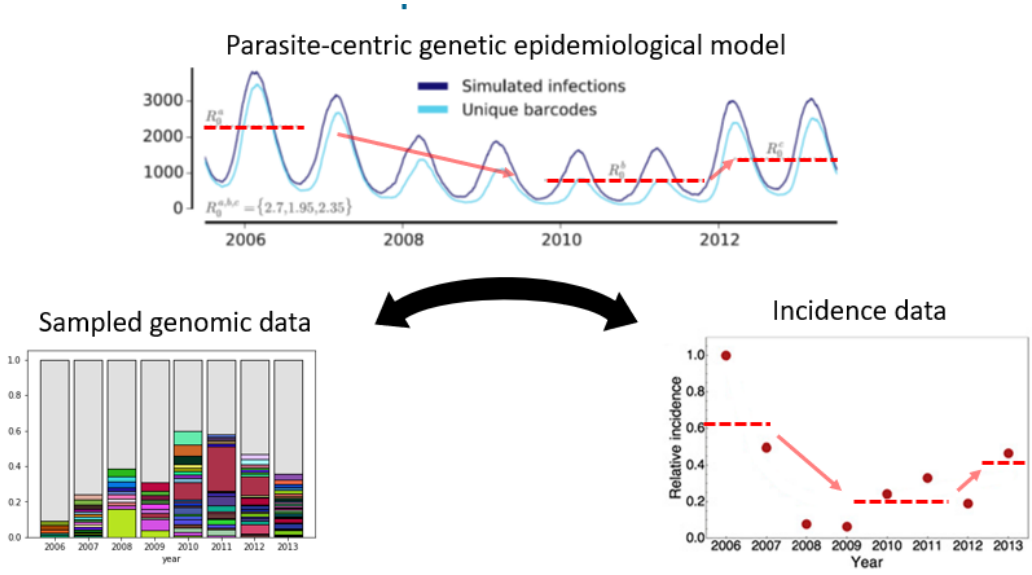
Presenter: Albert Lee

Integrating mathematical modelling can add value to genomic surveillance by bridging gaps when surveillance is limited, unifying information from multiple data sources via a common modelling framework, and pressure-testing the interpretation of genetic signals. Genetic models have been

shown to achieve actionable outputs, for example, linking R_0 and other epidemiological indicators to genomic signals and characterizing spatiotemporal patterns to estimate connectivity.

Transmission can be estimated using dynamic genetic models that build on connections with transmission indicators. For example, COI is an important genetic indicator, and its correlation with R_0 has been assessed. By simulating biological mechanisms, it is then possible to build on an understanding of parasite genetics to find order in complex genetic relationships and test theories against the data to determine where they may be most effective. This means that modelling may simulate trends in transmission without input from incidence data by using sampled genomic data from local parasite populations (Fig. 16). There are still challenges and limitations to this approach, such as the large number of components required to produce dynamic models that must be thoroughly tested, the dependency on priors that must be well understood and data-driven, and uncertainties in inputs that must be propagated to uncertainties in outputs. These issues can be resolved through close collaboration with local experts and a solid mathematical framework for uncertainty quantification.

Fig. 16. Parasite-centric genomic model reproduces trends in transmission without input from incidence data



Genetics improves the differentiation between importation and localized transmission. For example, models can provide a detailed view of local transmission properties by examining spatiotemporal correlations with links between strains from multiple infections. Positions of clonal infections over time enable estimates of dispersal velocity in emergent strains, and a mechanistic transmission model can then link dispersal velocity to spatial connectivity. The limitation, however, is that while models can estimate dispersal velocity to understand local versus imported transmission in relatively small geographic areas, this would require very robust data to ensure accuracy. The models are based on the parasites' clonal expansion, but have yet to incorporate the effects of vector biology and human movement.

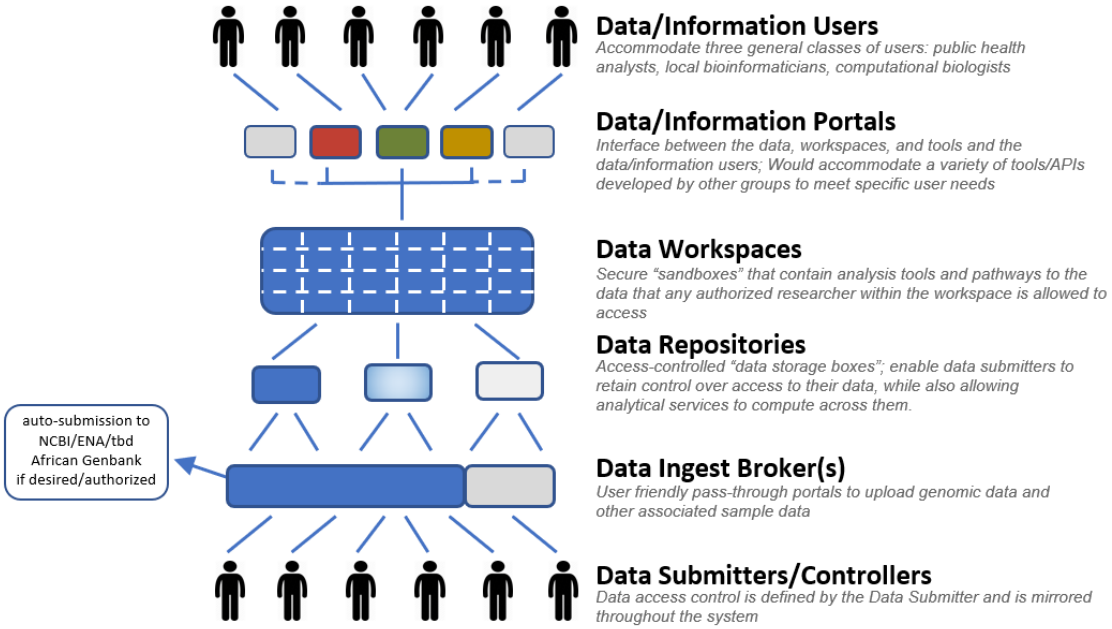
Genetic models can improve the precision and efficiency of surveillance programmes and the understanding of local parasite populations and patterns of transmission. With further development, such methods may allow for predictive modelling to identify hotspots and target interventions. The availability of comprehensive data will enable the training of such models to reduce the number of assumptions that need to be made and improve the accuracy of model projections. This will allow future surveillance efforts to be more effective with sparse data.

3.7b. Data standardization and translation for use in routine surveillance – a strategy for scale-up

Presenter: Bronwyn MacInnis

Across all potential use cases for malaria genetic epidemiology, one thing is exceedingly clear: the need for harmonization between data types (WGS, amplicon sequencing and genotyping) to allow for comparisons within and between countries, a common platform for data storage, analysis and reporting (Fig. 16), and agreements on data sharing.

Fig. 16. A Data System Concept for Genomic Pathogen Surveillance and Epidemic Preparedness



Considerations for the types of samples required, whether sequencing is necessary, and whether analysis requires specialized expertise all have implications for whether implementation and scale-up are possible at country level. Scaling up genomic data generation would require a coordinated multi-country effort with international partners and harmonization of core analytical workflows. A cloud-based, access-controlled data system for data storage has been suggested as a possibility to ensure country-level access control that could be deployed locally. This would remove the need for substantial computing infrastructure, downloading, tracking and version control. It is important for any system introduced to be adaptable to future needs as the field progresses.

The purpose of data sharing also needs to be considered – e.g., country-level public health programme needs versus academic research interests – along with the necessity to share data between countries, especially bordering countries. While many use case applications of genomic data require data sharing across country borders, many do not. Within-country applications, or those requiring bilateral data sharing, could be developed while open data sharing terms are considered and agreed. In approaches taken by other programmes, such as the polio laboratory network, it took a decade from proof of concept to establish a data sharing network. WHO has drafted a code of conduct for open and timely sharing of pathogen sequence data during outbreaks of infectious diseases. Simplified data sharing and communication already occur within the genomics community, which implies willingness to engage in informal sharing and the potential to introduce more refined options to solidify sharing mechanisms in a more robust fashion.

A recent Lassa fever outbreak in Nigeria was used as an example to demonstrate how rapid sequencing, analysis and data sharing helped answer the question of whether the outbreak was due to a new variant, due to a more virulent strain of the virus, or due to increased human to human transmission. Comparison of the viruses isolated from patients during the outbreak were compared to genomic sequences from viruses found elsewhere in Nigeria and in other countries. Results showed that there were multiple introductions of genetically independent viruses similar to known lineages in Nigeria, which excluded human to human transmission and the possibility of introduction of a new variant strain. The findings were shared with the Nigerian CDC and Lassa fever clinicians, and genomic data were released openly to the scientific community in real-time.

The key question is: How practical is it to scale up the use of malaria genomics in its current state? Some countries already have routine systems in place, but most high-burden countries do not have capacity to conduct sequencing, analyse data and use molecular epidemiology in the field. NMCPs have the opportunity to influence the approach for a way forward, but this requires consensus among partner organizations and countries to ensure that harmonization in the next steps occurs.

4. Working groups – Resistance, transmission, elimination and data collection: priority use cases for programme implementation

To further discuss potential use cases for malaria genetic epidemiology, two working groups were established: 1) Surveillance for *pfhrp 2/3* deletions and drug and insecticide resistance, and 2) Transmission dynamics across the transmission continuum and elimination. The objectives were to focus on research questions and use cases related to gene flow, including issues relevant to elimination settings, and to develop targets for the next 6–12 months, 1–2 years, 3–5 years and 5–10 years either for implementing the use cases as approaches to malaria control or for identifying areas that require more research. The groups also discussed where genomic surveillance would be most useful for policy, strategy and programme implementation; if there is evidence available for WHO review or a timeline for when information will become available; and what approaches would be best for data collection and interpretation in clinical, public health, surveillance and laboratory settings and what challenges are foreseen. During the group work, participants identified priorities for the application of genetic epidemiology in the detection and control of drug and insecticide resistance and transmission. These priorities are outlined below, while detailed information on operational use, field sampling and laboratory methods, ethics and data sharing, added value over conventional epidemiological methods, and challenges for implementation are provided in the supplementary tables (see Annex 1).

4.1. Surveillance for *pfhrp2/3* deletions and drug and insecticide resistance

In evaluating potential use cases for genomic surveillance of drug and insecticide resistance, there was a need to generate additional evidence and notable challenges for implementation. There were two use cases/applications for the surveillance of *pfhrp2/3* deletions or spread of drug resistance that were deemed ready for immediate action; the remaining use cases will require additional evidence for action in the medium term of 1–2 years or 3–5 years.

4.1.a. For immediate action (6–12 months)

Surveillance of *pfhrp2/3* deletions

- There is sufficient evidence from several countries to show that deletions of *pfhrp2 +/- pfhrp3* can cause false-negative HRP2-RDT results and that these parasites can become dominant in the parasite population. WHO has developed recommendations for investigating suspected false-negative RDTs due to *pfhrp2/3* deletions, as well as indications for conducting surveys, survey templates and criteria for when countries should switch to non-HRP2-exclusive RDTs. To support high-quality and rapid molecular analysis, WHO has also established a network of

reference laboratories experienced in *pfhrp2/3* genotyping and a proficiency testing scheme for malaria NAAT that includes *pfhrp2/3* deleted parasites. Until alternative diagnostic tests that can match the performance, stability and demand of HRP2-RDTs become available, surveillance for *pfhrp2/3* deletions across all epidemiological settings is essential for detecting areas where RDTs are failing and maintaining confidence in HRP2-RDT results.

- **Challenges for implementation:** Although not likely to be the only factor, the use of HRP2-RDTs themselves is expected to be driving the selection for *pfhrp2* deletions. The *pfhrp2*-negative parasites in Eritrea and Peru showed distinct haplotypes that strongly suggested de novo development of these parasites in both locations. Such development would imply that all malaria-endemic areas are at risk and that there is an urgent need to map the prevalence of *pfhrp2*-negative parasites to inform case management policy. The key challenge then becomes the mobilization of resources to conduct such mapping.

Monitoring changes in frequencies of molecular markers of drug resistance over time and space

- There is sufficient evidence to show that molecular markers can be used to monitor changes in drug resistance and *pfhrp2/3* deletions in parasite populations over space and time. This is essential for detecting populations at risk of treatment failure or under-detection by RDTs in order to subsequently inform first-line drug policy decisions (ensuring that effective treatment is given to patients) and ensure that patients can be adequately diagnosed. While passive surveillance is acceptable, active sampling biannually or annually using dried blood spots would be desirable in order to rapidly detect changes in drug resistance. Routine monitoring should be implemented at the appropriate administrative level, which is relevant for the implementation of national drug policies. This approach is less expensive and timelier than a therapeutic efficacy study (TES).

- **Challenges for implementation:** Countries require clear guidance, training and capacity-building on the establishment of an appropriate spatial sampling strategy, methods for amplicon sequencing or other genotyping methods, and data generation, analysis and interpretation – all of which could prove costly in the short term. Clear procedures also need to be developed to ensure that policies on first-line drugs can be modified rapidly in response to changing resistance patterns and implemented in the field in a timely manner.

4.1.b. For medium-term action (1–2 years)

Determining the origins of drug resistance

- Determining the origins of drug resistance can facilitate the monitoring of the spread of resistance within and between countries. By monitoring haplotypes associated with drug resistance mutations from samples on a routine basis and comparing them over time and across regions, it is possible to determine if drug resistance is emerging locally or spreading – something that is difficult to infer using standard epidemiological approaches. Identifying populations at risk can inform regional drug policies and ensure interventions are targeted to contain resistance.

- **Challenges for implementation**: More research is needed to identify molecular markers in different geographical settings, along with a database of parasite samples across multiple geographic regions that can be used for comparison in analyses. Outsourcing of WGS may be necessary to generate the data for establishing such a database, as well as data sharing agreements within and between countries. In addition to the challenges outlined for use case 4.1.a, further research is needed, followed by clear guidance on public health responses to emerging resistance versus spread.

Determining the number and spatial distribution of sentinel sites needed to assess insecticide resistance and monitor new interventions

- This is important to improve the timeliness of surveillance of insecticide resistance, allocate adequate resources and monitor the impact of new interventions.
- **Challenges for implementation**: In many countries, there is a lack of entomological capacity, as well as a lack of the geospatial and entomological expertise required to develop a spatial sampling strategy. This means that clear guidance and technical support would be required to support this activity.

4.1.c. For medium-term action (3-5 years)

Detecting changes in parasite population structure or signatures of positive selection

- Detecting changes in parasite population structure to determine whether there is anthropogenic impact from interventions or other selective pressures can help to identify populations at risk for emergence of resistance. It can also lead to early detection of emergence of new resistance mechanisms through identification of new resistance markers. This requires continuous longitudinal spatial sampling of populations over time.
- **Challenges for implementation:** In addition to the challenges outlined for case 4.1a, parasite genomic data from the same region over time or from nearby regions are required for comparison, which may take several years to establish. The current analytical approaches, such as IBD, also have some limitations in terms of determining the origin of resistant alleles. More research is required to determine whether these approaches can be used at the operational level rather than as a research tool, which is how they are currently being used.

Monitoring local species composition and changes over time

- Improved understanding of local species composition and changes over time can i) inform selection of vector control tools by identifying key vectors responsible for transmission, and ii) aid in assessing residual transmission and its implications for the effectiveness of interventions. Cross-sectional sampling over time at sentinel sites can reveal the heterogeneity in the local vector and parasite populations and support the development of other use cases, such as improving the understanding of resistance patterns and transmission dynamics in a region. This activity is simple to implement even if local entomological expertise is available or species complexes have already been formally defined.
- **Challenges for implementation**: There is a lack of reference genomes for many species, as well as a lack of validated spatial density data for decision-making.

Insecticide resistance surveillance

- Monitoring insecticide resistance allows for the targeting of specific interventions (e.g., pyrethroid-PBO nets) and resistance mechanisms (e.g., mixed-function oxidase (MFO) resistance mechanisms) over time. Such monitoring also enables programmes to assess the value of different insecticide resistance management strategies (e.g., IRS rotation, new types of ITNs, attractive toxic sugar baits). Using genotyping to detect drug resistance is simpler to implement than phenotypic assays that require rearing of larvae. With this approach, shifts in allele frequencies may be easier to detect than shifts in phenotype over short time periods.
- **Challenges for implementation**: There are currently no molecular markers for MFO resistance and therefore research is required in this area. The number of sentinel sites required to assess insecticide resistance is also unknown, and there is a lack of reference genomes for many species. In terms of using the information to inform strategies, it is still necessary to assess the value of each strategy in the context of the NMCP's national malaria control strategy. More research is required to assess the effectiveness of the combination of different control

strategies targeting the mosquito and the parasite. As an insecticide resistance management strategy, baits are the least developed and will require extensive research before they can be recommended as a tool in the context of malaria genomic surveillance. To begin to bridge the gaps that exist in the capacity for data generation and analysis, establishment of a network approach would facilitate further progress.

4.2. Transmission

Use cases for malaria genomics in the context of gene flow for transmission to elimination include both parasite and vector dynamics. Generally, the use of genomic data in this context is viewed as confirmatory, or as an augmentation to traditional epidemiologic data, providing more precise information where gaps or discrepancies remain. A key concern, however, is that more evidence is needed to validate genomic data with respect to traditional epidemiology. With the use of different methods at different levels of confidence for interpreting genomic data, quality assurance and control are needed to standardize approaches and establish the evidence base to support policy and programmatic decision-making at a higher level.

4.2.a. For immediate action (6–12 months) to medium-term action (1–2 years)

Vector species dynamics

- For understanding vector dynamics, use cases include understanding vectorial capacity and vector competence to inform surveillance and control measures surrounding imported cases. This use case is also of importance for imported case management in countries with low transmission or in malaria-free countries with high receptivity risk for sustained introduced transmission. Understanding the local vector competence for imported malaria species can help to define risk and inform response strategies for outbreak prevention.
- **Challenges for implementation:** A basic molecular biology laboratory with trained personnel and a map of local vector species would be required and could take some time to establish.

4.2.b. For long-term action (5–10 years)

Changes in transmission

- Genomic data can help to shed light on other changes or fluctuations in population dynamics that are not always clear, e.g., due to natural phenomena. Understanding changing transmission and being able to distinguish between natural fluctuations in parasite populations and the impact of interventions are important for future strategic planning.
- **Challenges for implementation:** There are challenges with interpretation using current methods; for example, do changes detected reflect those of the broader population? Extensive research using well designed studies is required.

Transmission intensity

- Understanding the levels of transmission intensity and transmission patterns with accuracy can inform stratification and malaria control strategies, detect persistent local transmission and help to establish a baseline of variation for future parasite population-genetics studies.
- **Challenges for implementation:** See challenges under elimination use cases.

Gene drive

- With increasing research on gene drive as a control strategy, it is necessary to map implementation of research and assess impact on local mosquito and parasite populations.

Determining the necessary spatial resolution of gene drive in the context of natural selective pressures in the field would improve the precision and future applications of this approach.

4.3. Elimination

The applications for the use of genomic data in elimination settings are either immediately actionable or actionable within the next 1 to 2 years.

4.3.a. For immediate action (6–12 months)

Elimination and low transmission settings: case classification of local, introduced or imported cases

- In low transmission settings, accurate case classification is crucial to certify a country as malaria-free (certification). The use of genomic data can add precision to case classification (indigenous vs imported), providing a country with evidence demonstrating zero indigenous cases of malaria.

4.3.b. For medium-term action (1–2 years)

Elimination and low transmission settings: risk factors for local transmission and outbreak investigations

- In low transmission settings, genomics can also help to identify active foci, provide information on the origin of imported cases, identify high-risk groups for infection and for sustaining transmission ("hotpops"), and assess their contribution to onward transmission. Accurate data on whether local transmission is occurring, and identification of associated risk factors enable high-risk groups to be targeted with screening/awareness campaigns.
- Genomic data can help to determine how geographical areas may be linked through regular travel/importations. In considering progress towards elimination, it is important to generate data that help to elucidate parasite boundaries in a region, regardless of administrative borders, so that determination of origin and control measures can be implemented in relation to the parasite boundary rather than administrative borders. Genomic data could serve as supplemental to conventional epidemiologic data in understanding the movement of people. Determining cross-border connectivity of parasites will allow for a coordinated response between bordering countries, across artificial or porous borders, and inform relevant decision-making in a regional context.
- In outbreak investigations, genomic data can be used in conjunction with conventional epidemiology to confirm linkages between locally transmitted cases. This information can be used to direct public health resources appropriately and prevent unnecessary investigations or interventions.
- **Challenges for implementation:** There is a need for a local and global repository of genetic parasite sequences that can be queried and ideally integrated into existing databases within the malaria community in order to aid in the identification of parasite origins. This will take some time to establish. Standardization is necessary across data, genotyping and analysis types to enable comparison, along with established mechanisms for quality assurance and control. Capacity-building for *timely* genotyping, analysis, interpretation and use of data in countries is required. Guidance will be needed on translating genetic data into information that can easily be used by control and elimination programmes, particularly as part of outbreak investigations.

5. Next steps

Several next steps were identified:

1. The table of research priority areas (Table A1) identified during this meeting should be made available online and updated on an annual basis by WHO with help from research networks and individuals.
2. A database of researchers and institutions involved in policy-relevant malaria genetic epidemiology studies should be developed by WHO, and this database should be updated annually.
3. Use cases share several overlapping themes across the spectrum of transmission in terms of understanding gene flow in insecticide and drug resistance. Studies should maximize these linkages so that common data generation platforms and samples can be used, wherever possible.
4. In addition to research studies, there are opportunities to explore drug and insecticide resistance monitoring sites: collecting genetic samples during case detection and investigations in elimination settings, and, in burden reduction settings, passive case detection systems and household surveys could become the mainstay for genomic surveillance. A structured approach that will not add any unnecessary burden on the health system is needed.
5. Stakeholders should work with researchers to ensure that study protocols are designed to generate evidence in formats relevant to policy and programmes. For example, studies exploring the relevance of genomic surveillance metrics must include a comparison to metrics currently recommended by WHO and used by countries in terms of their relevance, reliability, accuracy, precision, cost and sustainability. WHO to work with network of research during study design stage.
6. Established global databases should be harnessed to develop information products relevant for policy and country operations. WHO to work with groups such as Sanger Institute and BROAD on appropriate information products once policy relevance is established.
7. Investment in regional and national capacities for genetic epidemiology should be sought. WHO to work with researchers and funders such as BMGF on pathways to increased national capacity.

6. References

1. Global Technical Strategy for malaria 2016–2030. Geneva: World Health Organization; 2015 (https://www.who.int/malaria/publications/atoz/9789241564991/en/).
2. Walker TM, Merker M, Knoblauch AM, Helbling P, Schoch OD, van der Werf MJ, et al. A cluster of multidrug-resistant Mycobacterium tuberculosis among patients arriving in Europe from the Horn of Africa: a molecular epidemiological study. Lancet Infect Dis. 2018;18(4):431–40. doi:10.1016/S1473-3099(18)300004-5.
3. World Health Organization, Foundation for Innovative New Diagnostics. The use of next-generation sequencing technologies for the detection of mutations associated with drug resistance in *Mycobacterium tuberculosis* complex: technical guide. Geneva: World Health Organization; 2018 (WHO/CDS/TB/2018.19; https://apps.who.int/iris/bitstream/handle/10665/274443/WHO-CDS-TB-2018.19-eng.pdf).
4. Anderson TJ, Nair S, McDew-White M, Cheeseman IH, Nkhoma S, Bilgic F, et al. Population parameters underlying an ongoing soft sweep in Southeast Asian malaria parasites. Mol Biol Evol. 2017;34(1):131–44. doi:10.1093/molbev/msw228.
5. Miotto O, Amato R, Ashley EA, MacInnis B, Almagro-Garcia J, Amaratunga C, et al. Genetic architecture of artemisinin-resistant Plasmodium falciparum. Nat Genet. 2015;47(3):226–34. doi:10.1038/ng.3189.
6. Amato R, Pearson RD, Almagro-Garcia J, Amaratunga C, Lim P, Suon S, et al. Origins of the current outbreak of multidrug-resistant malaria in southeast Asia: a retrospective genetic study. Lancet Infect Dis. 2018;18(3):337–45. doi:10.1016/S1473-3099(18)30068-9.
7. Takala-Harrison S, Jacob CG, Arze C, Cummings MP, Silva JC, Dondorp AM, et al. Independent emergence of artemisinin resistance mutations among Plasmodium falciparum in Southeast Asia. J Infect Dis. 2015;211(5):670–9. doi:10.1093/infdis/jiu491.
8. Neafsey DE, Waterhouse RM, Abai MR, Aganezov SS, Alekseyev MA, Allen JE, et al. Mosquito genomics. Highly evolvable malaria vectors: the genomes of 16 Anopheles mosquitoes. Science. 2015;347(6217):1258522. doi:10.1126/science.1258522.
9. Hamilton WL, Amato R, van der Pluijm RW, Jacob CG, Quang HH, Thuy-Nhien NT, et al. Evolution and expansion of multidrug-resistant malaria in southeast Asia: a genomic epidemiology study. Lancet Infect Dis. 2019;19(9):943–51. doi:10.1016/S1473-3099(19)30392-5.
10. Shetty AC, Jacob CG, Huang F, Li Y, Agrawal S, Saunders DL, et al. Genomic structure and diversity of Plasmodium falciparum in Southeast Asia reveal recent parasite migration patterns. Nat Commun. 2019;10(1):2665. doi:10.1038/s41467-019-10121-3.
11. A framework for malaria elimination. Geneva: World Health Organization; 2017 (https://www.who.int/malaria/publications/atoz/9789241511988/en/).
12. Patel JC, Taylor SM, Juliao PC, Parobek CM, Janko M, Gonzalez LD, et al. Genetic evidence of importation of drug-resistant Plasmodium falciparum to Guatemala from the Democratic Republic of the Congo. Emerg Infect Dis. 2014;20(6):932–40. doi:10.3201/eid2006.131204.
13. Molina-Cruz A, Canepa GE, Kamath N, Pavlovic NV, Mu J, Ramphul UN, et al. Plasmodium evasion of mosquito immunity and global malaria transmission: the lock-and-key theory. Proc Natl Acad Sci U S A. 2015;112(49):15178–83. doi:10.1073/pnas.1520426112.
14. Molecular surveillance for HRP2 and HRP3 gene expression in *Plasmodium falciparum* parasites from South and Central America. Atlanta: Centers for Disease Control and Prevention; 2012 (https://www.paho.org/hq/dmdocuments/2013/CDC-Molecular-Surv-SCAmerica-Eng.pdf).

Annex 1

Table A1 Summary of priority research areas and questions

SPREAD OF DRUG RESISTANCE AND *PFHRP 2/3* DELETIONS							
Evidence/use cases: SpotMALARIA, Plasmodium Diversity Network Africa (PDNA) *(1–3)*, MalariaGen Network *(4)*, GenRe-Mekong (Genetic reconnaissance to support malaria elimination in the GMS) *(5–7)*, Malawi *(8)*, Bangladesh, Mali *(9)*, Cambodia *(10–12)*, Thailand *(13,14)*, Lao People's Democratic Republic *(15)*, Myanmar, Viet Nam, China *(16)*, French Guiana *(17), Peru (18), Eritrea (19)*							
Use case/ application	**Operational component**	**Field sampling (methods, data source, spatial scale, frequency)**	**Laboratory (methods, standardization, expected advances)**	**Ethics and data sharing**	**Added value**	**Challenges for implementation**	**Immediate, medium- or long-term action**
1) Monitoring changes in frequencies of molecular markers of drug resistance over time and space	First-line drug policy decisions Identify populations at risk of treatment failure	Passive case detection Active sampling desirable Desired frequency: annual or semi-annual Dried blood spots Spatial sampling strategy should be relevant to implementation of national drug policies (e.g., district, province, or country level).	Amplicon sequencing or other genotyping methods	Data ownership: country owns primary data Aggregate data shared with the malaria community	Less expensive than TES Early warning of clinical failure Ability to genotype from dried blood spots Allows more dense sampling in time and space and at epidemiological scales	Unbiased population sampling – including establishment of appropriate spatial sampling strategy Nagoya protocol Countries need technical support and capacity-building to generate, store and analyse the data. Procurement and access to reagents. May need to rely on regional reference laboratories Costs Timeliness for modifying policies and implementation in the field	Immediate Evidence ready for submission to WHO for review within six months to one year

2) Identifying and monitoring changes in frequencies of *pfhrp2/3* deletions	Directly informs RDT selection for national programmes	Prospective surveys of symptomatic patients presenting to health facilities: survey templates available Parallel testing using HRP2 and pf-LDH RDTs or microscopy and collection of dried blood spots Prioritize HRP2 negative/pf-LDH or microscopy positive for *pfhrp2/3* genotyping Target countries with reports of *pfhrp2/3* deletions and neighbouring countries If >5% *Pf* cases are missed due to *pfhrp2/3* deletions, replace RDTs in the country; if <5%, repeat survey in 1–2 years	PCR to confirm *Pf* infection; *pfhrp2* and *pfhrp3* and at least two other single copy genes Flanking genes, serology, whole genome sequencing, next-generation sequencing optional	Country owns primary data and should establish MTA with international reference laboratory Aggregate data shared with the malaria community through WHO Malaria Threat Maps	As post-market surveillance and complaint reporting are weak in endemic countries, and confidence in RDTs remains fragile in many places, surveillance across all settings where HRP2 RDTs are in use is necessary to allow for early warning of *pfhrp2/3* deletions causing false-negative RDTs.	Financial resources to implement baseline surveys and monitoring	Immediate
3) Determining origins of drug resistance (independent emergence vs spread)	To guide targeting of interventions for containment of resistance & inform regional drug policies	Passive case detection Active sampling desirable Desired frequency: annual or semi-annual Dried blood spots	Amplicon sequencing or other genotyping methods followed by whole genome sequencing	Country owns primary data Aggregate data shared with the malaria community	Ease of field implementation Shared haplotypes associated with drug resistance mutations provide evidence of origins that may be difficult to infer using standard epidemiological approaches. Continuous sampling possible	Access to whole genome sequencing – outsourcing may be necessary A large database of parasites from multiple geographic regions needs to be available for comparison. More research to identify molecular markers in different geographical settings	Medium-term Evidence for WHO review likely to be ready within the next 1–2 years

4) Detecting changes in parasite population structure or signatures of positive selection	Identifying populations at risk for emergence of resistance Early detection of emergence of new resistance mechanisms through identification of new resistance markers	Passive case detection Active sampling desirable Sampling of populations over time, ideally annually or semi-annually Dried blood spots Spatial sampling strategy should be relevant to implementation of national drug policies (e.g., district, province, or country level).	Whole genome sequencing or genome-wide genotyping	Country owns primary data Aggregate data shared with the malaria community	Detection of emerging new resistance mechanisms Ease of field implementation Early warning of populations at risk for emergence of resistance Continuous sampling possible. Can make use of historical samples	Access to whole genome sequencing – outsourcing may be necessary Parasite genomic data from the same region over time or nearby regions need to be available for comparison.	Medium-term Evidence for WHO review likely to be ready within the next 3–5 years

SPREAD OF INSECTICIDE RESISTANCE

Evidence/use cases: *Anopheles gambiae* 1000 Genomes Project *(20)*

Use case/application	Operational component	Field sampling (methods, data source, spatial scale, frequency)	Laboratory (methods, standardization, expected advances)	Ethics and data sharing	Added value	Challenges for implementation	Immediate, medium or long-term action
1) Determine the number and spatial distribution of sentinel sites needed to assess insecticide resistance and monitor new interventions	Improved surveillance of insecticide resistance and impact of interventions	Larval sampling and adult sampling (traps, human landing catches)	Amplicon sequencing or other genotyping methods for known resistance markers; whole genome sequencing on a subset	Country owns primary data Aggregate data shared with the malaria community	More timely monitoring and adequate resources allocated Inform gene drive development and deployment	Lack of entomological capacity in-country Need for spatial sampling strategy (i.e., need geospatial expertise along with entomological expertise)	Medium-term 1–2 years

2) Monitoring local species composition and changes over time	Incrimination of key vectors Informing selection of vector control tools Assessing residual transmission and its implications for intervention effectiveness	To be informed by surveillance work (see above) Cross-sectional sampling over time Spatial density Sampling to be conducted at sentinel sites Human bloodmeal index Sporozoite infection rate	Amplicon sequencing or other genotyping methods for known resistance markers; whole genome sequencing on a subset	Country owns primary data Aggregate data shared with the malaria community	Inform gene drive development and deployment Simple to implement when local entomological expertise is lacking or species complexes have not been formally defined	Lack of reference genomes for many species Lack of validated spatial density data for decision-making	Medium term 3–5 years
3) Insecticide resistance surveillance	Targeting of specific interventions (pyrethroid-PBO nets) and resistance mechanisms (e.g., MFO resistance mechanisms) over time Assessing the value of different insecticide resistance management strategies (e.g., IRS rotation, new types of ITNs, attractive toxic sugar baits)	Sampling over time Spatial density dependent on spatial insecticide exposure Larval sampling, adult sampling (inside and outside buildings to detect behavioural resistance)	Amplicon sequencing or other genotyping methods for known resistance markers; whole genome sequencing on a subset	Country owns primary data Aggregate data shared with the malaria community	Simpler to implement than phenotypic assays requiring rearing of larvae Shifts in allele frequencies may be easier to detect than shifts in phenotype over short time periods.	No known markers for MFO-mediated resistance Feasibility of outsourcing whole genome sequencing Lack of reference genomes for many species	Medium-term 3–5 years to long-term 5–10 years

TRANSMISSION							
Change in transmission							
Evidence/use cases: *P. falciparum* Community Project, MalariaGEN Network - Pf3k, SpotMALARIA							
Use case/application	**Operational component**	**Field sampling (methods, data source, spatial scale, frequency)**	**Laboratory (methods, standardization, expected advances)**	**Ethics and data sharing**	**Added value**	**Challenges for implementation**	**Immediate, medium- or long-term action**
1) Is transmission changing and how do we interpret natural fluctuation?	Can remove need for some active case detection to save costs	Epidemiological framework to benchmark	Whole genome sequencing or amplicon sequencing; requires measure of relatedness	Country information is sufficient	Independent, orthogonal metric of transmission, helps to triangulate current metrics. At low transmission, current measures lose dynamic range/sensitivity. Genomic metrics may be more sensitive. Complementary data to help understand other changes/fluctuations that are not always clear, e.g., due to natural phenomena	No validated framework for calibrating epidemiologically relevant metrics or for deriving them. How to design this experiment? Challenges with interpretation; do changes detected reflect those of the broader population?	Long-term Evidence for WHO review likely to be ready within the next 5–10 years
2) Are interventions making an impact?	Signals whether interventions are impacting parasite populations in target regions Orthogonal data type to measure incidence May be an earlier or more sensitive indicator (tbd)	As above	Whole genome sequencing or amplicon sequencing	As above	As above	As above	

TRANSMISSION INTENSITY							
Evidence/use cases: Senegal, Panama, Malawi							
Use case/application	**Operational component**	**Field sampling (methods, data source, spatial scale, frequency)**	**Laboratory (methods, standardization, expected advances)**	**Ethics and data sharing**	**Added value**	**Challenges for implementation**	**Immediate, medium- or long-term action**
1) What are the levels of transmission intensity? Is it possible to develop a standalone genomic metric of transmission?	Inform stratification and malaria control strategies	A local and global repository of genetic sequences that can be queried and ideally integrated with existing databases within the malaria community to inform parasite origins	Whole genome sequencing and amplicon sequencing	Country information can be sufficient for identification of local transmission, but requires shared data to identify source of imported infections.	Understand transmission intensity with more accuracy and better target malaria control	Standardization across data, genotyping and analysis types for comparison Quality assurance/quality control How to harmonize, "what tool/approach" should I use? Establishing a global repository for sharing of parasite genetic sequences Capacity for *timely* genotyping, analysis, interpretation and use of data in country Translation of genetic data into information that can easily be used for control and elimination programmes	Long-term
2) What is R_0 in different populations?	Predict spread and identify risk groups	As above	As above	As above		As above	

VECTOR SPECIES DYNAMICS							
Evidence/use cases: The MalariaGen Vector Observatory, Ag1000G Consortium							
Use case/application	**Operational component**	**Field sampling (methods, data source, spatial scale, frequency)**	**Laboratory (methods, standardization, expected advances)**	**Ethics and data sharing**	**Added value**	**Challenges for implementation**	**Immediate, medium- or long-term action**
1) What are *An. stephensi*'s population dynamics in Africa and Sri Lanka?	Guide vector control interventions If single introduction and low levels of genetic diversity, then these mosquitos could be eradicated	Entomological collections in Djibouti and Ethiopia	Molecular taxonomic identification and genotyping (PCR of ITS2, microsatellite and mitochondrial markers)			Requires a basic molecular biology lab with trained personnel	Immediate
2) How does it compare with its parent population?	As above	As above	As above	As above	As above	As above	
3) Understanding local vectoral infectivity for imported parasites	As above	As above	As above	As above	As above	Need map of local vector species	Medium-term 1–2 years
GENE DRIVE							
Evidence/use cases:							
1) Map vector species to then track changes from gene drive							Long-term

ELIMINATION (AND SOURCE OF INFECTION MORE BROADLY)							
Evidence/use cases: Greece, China, Guatemala, Peru, Colombia, Ecuador, Panama, Kingdom of Eswatini, northern Senegal, northern Namibia, Bangladesh. Pfs47 as a potential candidate. Could Greece be a good retrospective benchmarking proof of concept?							
Use case/application	**Operational component**	**Field sampling (methods, data source, spatial scale, frequency)**	**Laboratory (methods, standardization, expected advances)**	**Ethics and data sharing**	**Added value**	**Challenges for implementation**	**Immediate, medium- or long-term action**
1) Are new cases locally transmitted, introduced or imported?	Improve classification of cases as indigenous or imported (over travel history) Identify transmission foci/sources Gives additional information about parasite origin for imported cases Can we calculate a rate of importation? Identify high-risk groups for infection and for sustaining transmission (hotpops) Certification of malaria-free/demonstration of zero indigenous cases of malaria	A local and global repository of genetic sequences that can be queried and ideally integrated with existing databases within the malaria community to inform parasite origins Travel history data Routine malaria surveillance data for elimination settings, including case investigation data	Whole genome sequencing and amplicon sequencing Dried blood spot samples	Country information can be sufficient for identification of local transmission, but requires shared data to identify source of imported infections from other regions in the country or other countries. May require data sharing between administrative boundaries within the country. Should respect governance that protects patient confidentiality	More accurate data/ higher resolution for case classification and identification of foci Allows targeting of high-risk groups with screening/awareness	Standardization across data, genotyping and analysis types for comparison Quality assurance/quality control Establishing a global repository for sharing of parasite genetic sequences Capacity for *timely* genotyping, analysis, interpretation and use of data in country Translation of genetic data into information that can easily be used for control and elimination programmes	Immediate (in some contexts)

2) Is there ongoing local transmission?	Possible information about linkages between locally transmitted cases Stratification of interventions, i.e., to deploy vector control if local transmission is ongoing	As above	As above	As above	Determine whether transmission is occurring with higher accuracy Determine the source/source region of imported infections	Standardization across data, genotyping and analysis types for comparison Quality assurance/quality control Establishing a global repository for sharing of parasite genetic sequences Capacity for *timely* genotyping, analysis, interpretation and use of data in country Translation of genetic data into information that can easily be used for control and elimination programmes	Medium-term 1–2 years
3) Mapping transmission chains and better defining risk of onward infection	Assess contribution to onward transmission Determine how geographical areas may be linked through regular travel/importations	As above	As above	As above	Predict spread and carry out resource planning/targeted interventions Determine cross-border connectivity of parasites allowing a coordinated response between bordering countries	As above	
4) Outbreak/cluster investigation	Validation of epi linkages (or not) Determine the source of outbreak (local vs imported)	As above	As above	As above	Direct public health resources appropriately or prevent unnecessary investigations or interventions	As above	

References

1. Roper C, Pearce R, Nair S, Sharp B, Nosten F, Anderson T. Intercontinental spread of pyrimethamine-resistant malaria. Science. 2004;205(5687):1124.
2. Wooton JC, Feng X, Ferdig MT, Cooper RA, Mu J, Baruch DI, et al. Genetic diversity and chloroquine selective sweeps in Plasmodium falciparum. Nature. 2002;418(6895):320–3.
3. Ghansah A, Amenga-Etego L, Amambua-Ngwa A, Andagalu B, Apinjoh T, Bouyou-Akotet M, et al. Monitoring parasite diversity for malaria elimination in sub-Saharan Africa. Science. 2014;345(6202):1297–8. doi:10.1126/science.1259423.
4. Hamilton WL, Amato R, van der Pluijm RW, Jacob CG, Quang HH, Thuy-Nhien NT, et al. Evolution and expansion of multidrug-resistant malaria in southeast Asia: a genomic epidemiology study. Lancet Infect Dis. 2019;19(9):943–51. doi:10.1016/S1473-3099(19)30392-5.
5. Zhu L, Tripathi J, Rocamora FM, Miotto O, van der Pluijm R, Voss TS, et al. The origins of malaria artemisinin resistance defined by a genetic and transcriptomic background. Nat Commun. 2018;9(1):5158. doi:10.1038/s41467-018-07588-x.
6. Anderson TJ, Nair S, McDew-White M, Cheeseman IH, Nkhoma S, Bilgic F, et al. Population parameters underlying an ongoing soft sweep in Southeast Asian malaria parasites. Mol Biol Evol. 2017;34(1):131–44. doi:10.1093/molbev/msw228.
7. Miotto O, Amato R, Ashley EA, MacInnis B, Almagro-Garcia J, Amaratunga C, et al. Genetic architecture of artemisinin-resistant Plasmodium falciparum. Nat Genet. 2015;47(3):226–34. doi:10.1038/ng.3189.
8. Laufer MK, Takala-Harrison S, Dzinjalamala FK, Stine OC, Taylor TE, Plowe CV. Return of chloroquine-susceptible falciparum malaria in Malawi was a reexpansion of diverse susceptible parasites. J Infect Dis. 2010;202(5):801–8. doi:10.1086/655659.
9. Djimdé AA, Dolo A, Ouattara A, Diakité S, Plowe CV, Doumbo OK. Molecular diagnosis of resistance to antimalarial drugs during epidemics and in war zones. J Infect Dis. 2004;190(4):853–5.
10. Amato R, Pearson RD, Almagro-Garcia J, Amaratunga C, Lim P, Suon S, et al. Origins of the current outbreak of multidrug-resistant malaria in southeast Asia: a retrospective genetic study. Lancet Infect Dis. 2018;18(3):337–45. doi:10.1016/S1473-3099(18)30068-9.
11. Agrawal S, Moser KA, Morton L, Cummings MP, Parihar A, Dwivedi A, et al. J Infect Dis. 2017;216(4):468–76. doi:10.1093/infdis/jix334.
12. Takala-Harrison S, Jacob CG, Arze C, Cummings MP, Silva JC, Dondorp AM, et al. Independent emergence of artemisinin resistance mutations among Plasmodium falciparum in Southeast Asia. J Infect Dis. 2015;211(5):670–9. doi:10.1093/infdis/jiu491.
13. Nkhoma SC, Nair S, Al-Saai S, Ashley E, McGready R, Phyo AP, et al. Population genetic correlates of declining transmission in a human pathogen. Mol Ecol. 2013;22(2):273–85. doi:10.1111/mec.12099.
14. Cerquiera GC, Cheeseman IH, Schaffner SF, Nair S, McDew-White M, Phyo AP, et al. Longitudinal genomic surveillance of Plasmodium falciparum malaria parasites reveals complex genomic architecture of emerging artemisinin resistance. Genome Biol. 2017;18(1):78. doi:10.1186/s13059-017-1204-4.
15. Nash D, Nair S, Mayxay M, Newton PN, Guthmann JP, Nosten F, et al. Selection strength and hitchhiking around two anti-malarial resistance genes. Proc Biol Sci. 2005;272(1568):1153–61.

16. Huang F, Takala-Harrison S, Jacob CG, Liu H, Sun X, Yang H, et al. A single mutation in K13 predominates in southern China and is associated with delayed clearance of Plasmodium falciparum following artemisinin treatment. J Infect Dis. 2015;212(10):1629–35. doi:10.1093/infdis/jiv249.

17. Pelleau S, Moss EL, Dhingra SK, Volney B, Casteras J, Gabryszewski SJ, et al. Adaptive evolution of malaria parasites in French Guiana: reversal of chloroquine resistance by acquisition of a mutation in pfcrt. Proc Natl Acad Sci U S A. 2015;112(37):11672-7. doi:10.1073/pnas.1507142112.

18. Gamboa D, Ho MF, Bendezu J, Torres K, Chiodini PL, Barnwell JW, et al. A large proportion of *P. falciparum* isolates in the Amazon region of Peru lack *pfhrp2* and *pfhrp3*: implications for malaria rapid diagnostic tests. PLoS One. 2010;5(1):e8091. doi:10.1371/journal.pone.0008091.

19. Berhane A, Anderson K, Mihreteab S, Gresty K, Rogier E, Mohamed S, et al. Major threat to malaria control programs by *Plasmodium falciparum* lacking histidine-rich protein 2, Eritrea. Emerg Infect Dis. 2018;24(3):462–70. doi:10.3201/eid2403.171723.

20. The Anopheles gambiae 1000 Genomes Consortium. Genetic diversity of the African malaria vector Anopheles gambiae. Nature. 2017;552:96–100. https://www.nature.com/articles/nature24995.