

Mapping the global distribution and spread of the *Plasmodium vivax*-associated virus MaRNAV-1

Mary E. Petrone^{1,*}, Justine Charon², Rhys H. Parry³, Matthew J. Grigg^{4,5}, Kim A. Piera⁴, Jacob A.F. Westaway⁴, Kayoko Shioda^{6,7}, Bruce Russell⁸, Ric N. Price^{4,9,10}, Timothy William^{5,11}, Enny Kenangalem¹², James S McCarthy^{13,14}, Bridget E. Barber^{15,16}, Edward C. Holmes¹⁷, Nicholas M. Anstey^{4,5}

¹Sydney Infectious Diseases Institute, School of Medical Sciences, The University of Sydney, Sydney, NSW 2006, Australia

²Fruit Biology and Pathology Unit, University of Bordeaux, INRAE, 71 Av. Edouard Bourlaux, Villenave-d'Ornon, Bordeaux 33140, France

³School of Chemistry and Molecular Biosciences, The University of Queensland, St. Lucia, QLD 4072, Australia

⁴Global and Tropical Health Division, Menzies School of Health Research, Charles Darwin University, John Mathews Building (Bldg 58), Royal Darwin Hospital Campus, Rocklands Drv., Casuarina, Darwin, NT 8010, Australia

⁵Infectious Diseases Society Kota Kinabalu Sabah-Menzies School of Health Research Clinical Research Unit, Kota Kinabalu, Sabah 88200, Malaysia

⁶Department of Global Health, Boston University School of Public Health, Boston, MA 02118, United States of America

⁷Center on Emerging Infectious Diseases, Boston University, 801 Massachusetts Ave, Boston, MA 02118, United States of America

⁸Department of Microbiology and Immunology, University of Otago, Dunedin North, Dunedin 9016, New Zealand

⁹Nuffield Department of Clinical Medicine, Centre for Tropical Medicine and Global Health, University of Oxford, Oxford OX3 7BN, United Kingdom

¹⁰Mahidol-Oxford Tropical Medicine Research Unit (MORU), Faculty of Tropical Medicine, Mahidol University, 420 6 Rajvithi Road, Thung Phaya Thai, Ratchathewi, Bangkok 10400, Thailand

¹¹Subang Jaya Medical Centre, No. 1, Jalan SS 12/1A, Ss 12, Subang Jaya, Selangor 47500, Malaysia

¹²Mimika Regency Health Office, Timika Jaya, Kec. Kuala Kencana, Kabupaten Mimika, Papua Tengah 99963, Indonesia

¹³Walter and Eliza Hall Institute, 1G Royal Parade, Parkville, VIC 3050, Australia

¹⁴Department of Infectious Diseases, University of Melbourne at the Peter Doherty Institute for Infection and Immunity, 792 Elizabeth St, Melbourne, VIC 3000, Australia

¹⁵QIMR Berghofer, 300 Herston Rd, Herston, Queensland 4006, Australia

¹⁶Royal Brisbane and Women's Hospital, 288 Herston Rd, Herston, Queensland 4006, Australia

¹⁷School of Medical Sciences, The University of Sydney, Sydney, NSW 2006, Australia

*Corresponding author. Sydney Infectious Diseases Institute, School of Medical Sciences, The University of Sydney, Sydney, NSW 2006, Australia. E-mail: mary.petrone@sydney.edu.au.

Abstract

Matryoshka RNA virus 1 (MaRNAV-1) is a bi-segmented and single-stranded RNA virus associated with *Plasmodium vivax*, a cause of human malaria. Little has been uncovered about the epidemiology and ecology of this virus since its discovery in 2019. To address this, we used a combination of primary and publicly available metatranscriptomic data to map the geographic distribution and host associations of MaRNAV-1. We detected this virus throughout Southeast Asia, in parts of South America, and, for the first time, in Oceania. Despite its broad distribution, MaRNAV-1 was found exclusively in metatranscriptomes containing *P. vivax*, suggesting that there is a specific virus-host relationship that has shaped the evolutionary history of this virus. We were unable to estimate the emergence date of the MaRNAV-1 lineage; however, phylogeographic mapping analysis suggested that MaRNAV-1 is widely dispersed throughout Southeast Asia. Our findings have both evolutionary and public health implications and can serve as the basis for future investigations in these fields.

Keywords RNA virus, malaria, virus-host co-divergence

Introduction

The first RNA virus to be associated with *Plasmodium*, a single-celled parasite that can cause malaria in humans, was discovered in 2019 (Charon et al. 2019). Matryoshka RNA virus 1 (MaRNAV-1) is a bi-segmented, positive-sense, single-stranded RNA virus from the family *Narnaviridae* (phylum *Lenarviricota*) that, to date, is exclusively associ-

ated with the species *P. vivax*. The longer segment (~3 000 bp) encodes the RNA-dependent RNA polymerase (RdRp). The other (~1 300 bp) encodes two overlapping open reading frames for hypothetical proteins of unknown function. MaRNAV-1 has been identified in *P. vivax*-infected human and mosquito blood samples from Malaysia, Cambodia, Thailand, and Colombia (Charon et al. 2019). It was recently shown

that MaRNAV-1 infects *P. vivax* cells at all stages of its life cycle (Seng et al. 2026). Since its discovery, six additional lineages of MaRNAVs (MaRNAV-2-8) have been identified (Charon et al. 2019, Rodrigues et al. 2022, Duc et al. 2025, Esperanza et al. 2025, Bourke et al. 2026). Collectively, these exhibit broader diversity and host range than MaRNAV-1. Identified in birds sampled in Oceania, North America, and Europe, MaRNAVs 2-8 are presumed to infect species of *Leucocytozoon* and *Haemoproteus* sp. which, along with *Plasmodium*, are members of the order Haemosporida (phylum Apicomplexa) (Valkiunas 2004).

That only one MaRNAV species has been identified in *Plasmodium* compared to four in leucocytozoa and two in Haemoproteus indicates an important gap in our knowledge of virus evolution. The identification of MaRNAVs in multiple genera of Haemosporida suggests that this viral lineage was present in the common ancestor of the order, which is estimated to have originated ~70 million years ago around the time of the Cretaceous–Paleogene (K–Pg) mass extinction event (Pacheco et al. 2018). However, Charon et al., found no evidence of MaRNAVs in the *P. knowlesi* and *P. falciparum* parasites that were co-circulating with the MaRNAV-positive *P. vivax* infected-samples in their cohort (Charon et al. 2019). Similarly, studies of avian parasites did not identify MaRNAV-like viruses in isolates with exclusively *Plasmodium* infection (Rodrigues et al. 2022, Duc et al. 2025, Esperanza et al. 2025). One explanation is that this was the consequence of the small sample size (n = 18) and limited catchment area (Sabah, east Malaysia) of this initial study. Alternatively, the life cycle of *P. vivax*, which involves the establishment of a latent infection in the liver that can reactivate after initial infection and cause relapsing episodes of malaria and ongoing transmission of the parasite, may facilitate infection by MaRNAV-1 or have its relapse frequency modulated by viral infection (Esperanza et al. 2025). *P. ovale* and *P. cynomolgi* are the only other human-infecting *Plasmodium* species that exhibit a latent hepatic hypnozoite stage (Mikolajczak et al. 2015). Hence, such a feature might render *P. vivax* susceptible to MaRNAV infection or less equipped to mount a protective response against it compared to other *Plasmodium* species.

The changing ecology of *P. vivax* might also impact the distribution and prevalence of MaRNAV-1. Although *P. falciparum* accounts for > 90% of malaria cases globally (Zekar and Sharman 2025), successful elimination strategies in co-endemic areas are leading to a relative increase in the proportion of malaria due to *P. vivax*. Indeed, *P. vivax* has become the predominant cause of malaria in Southeast Asia, the Horn of Africa, and the Americas (Weiss et al. 2025). *P. vivax* is more difficult to eliminate than *P. falciparum* since it requires radical cure of all stages of the parasite. These challenges are compounded by the emergence and spread of chloroquine resistance and the need for prolonged treatment course of primaquine to kill liver stages (Price et al. 2014a, 2014b, Thriemer et al. 2017, Ferreira et al. 2021, Ngwana-Joseph et al. 2024). Whether the shift in the epidemiology of *P. vivax* corresponds to an increase in MaRNAV-1 has not been assessed, in part because no baseline measurements have been recorded. Specifically, the frequency at which MaRNAV-1 co-occurs with *P. vivax* is not known, nor if this frequency varies by geographic region.

Herein, we aim to help fill this gap by extending our search for MaRNAV-1 to three additional species of human-infecting *Plasmodium* species collected across southeast Asia, Africa, South America, and Oceania. In doing so, we consider three hypotheses that could explain the evolutionary history and phylogenetic distribution of MaRNAV-1: (1) MaRNAV-1 can infect all *Plasmodium* species, and this will be revealed by screening additional species from a broader geographic

range; (2) MaRNAV-1 emerged in *P. vivax* after this species diverged from other members of its genus; (3) MaRNAVs were present in the *Plasmodium* common ancestor, but non-*vivax Plasmodium* species evolved mechanisms to prevent infection by MaRNAV-1. We leverage metatranscriptomic sequencing and phylogenetics to investigate the geographic and putative host range of MaRNAV-1, thereby evaluating the plausibility of each hypothesis.

Results

The distribution of MaRNAV-1 in *P. vivax* is global and varied

To establish a reference for the geographic range and frequency of MaRNAV-1, we began by measuring these metrics in *P. vivax* metatranscriptomes. Accordingly, we compiled a data set of *P. vivax* metatranscriptomes from (i) primary human blood isolates (n = 34, including 7 that were previously described by Charon et al. (Charon et al. 2019); Table S1) and (ii) publicly available RNA libraries from the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) (n = 1 431, as of May 2025). These libraries were generated from human blood isolates (n = 940), laboratory *Plasmodium* strains (n = 102), and mosquitoes (n = 389) (Table S2). Screening these libraries for MaRNAV-like viruses using DIAMOND BLAST (Buchfink et al. 2021) returned 610 putative MaRNAV RNA-dependent RNA polymerase (RdRp) sequences (Supp. Data 1) with minimal sequence diversity (at least 95% amino acid similarity to the original MaRNAV-1 RNA-dependent RNA polymerase (RdRp) [acc. QGV56801] and 97% amino acid similarity to the original MaRNAV-1 hypothetical protein [acc. QGV56796]). Relatives of MaRNAVs-2-7 were not detected. We classified all putative viruses identified as MaRNAV-1.

Our search expanded the known geographic and temporal distribution of MaRNAV-1, revealing the first instances of MaRNAV-1 in *Plasmodium* libraries from Peru (n = 27 of 319), Indonesia (n = 14 of 25), Solomon Islands (n = 1 of 2), and Papua New Guinea (n = 1 of 3) (Fig. 1). The presence of this virus in human blood isolates from Peru suggested that virus-infected *Plasmodium* have been transmitted to humans by mosquito vectors in South America. Previously, MaRNAV-1 virus had only been detected in mosquitoes collected in Colombia (Charon et al. 2019). We also detected MaRNAV-1 in the earliest libraries in our data set (Indonesia, 2004), extending its known circulation in Southeast Asia by 10 years. Fragments of the hypothetical protein were recovered in two of the 27 libraries we screened from Brazil without corresponding RdRp segments. We did not consider this as sufficient evidence that MaRNAV-1 is circulating in the country, but additional sampling could resolve this given that the virus is present in neighbouring Colombia and Peru. We did not identify MaRNAV-1 in the one *P. vivax* library available from Africa (Ethiopia); however, a recent study has shown that it is present in this country (Seng et al. 2026).

The frequency of MaRNAV-1 in *P. vivax* libraries varied considerably. Among the primary human blood isolates, frequency ranged from 0% (India, 0 out of 1) to 77.8% (Papua, Indonesia, 14 out of 18) (Table 1).

The frequency of MaRNAV-1 in libraries available on the SRA was more difficult to estimate because they were derived from a combination of human and mosquito samples, and we could not assume that individual libraries represented individual infections due to pooling. Therefore, frequency estimates from the SRA could not be directly

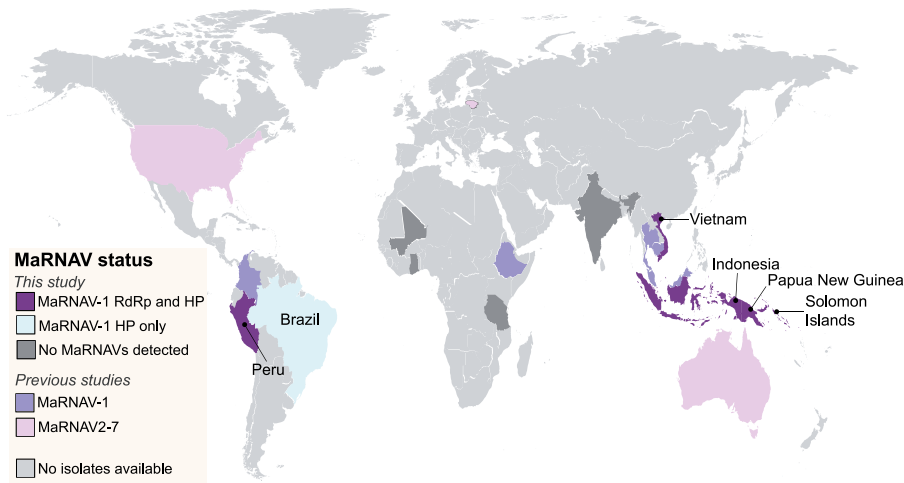


Figure 1 Global distribution of MaRNAVs in Haemosporida transcriptomes. MaRNAVs 2–7 were previously detected in USA (Esperanza et al. 2025), Australia (Charon et al. 2019), and Lithuania (Duc et al. 2025) (pink). MaRNAV-1 was previously detected in Colombia, Malaysia, and Cambodia by Charon et al. (Charon et al. 2019) and in Ethiopia by Seng et al. (Seng et al. 2026). Abbreviations: RdRp (RNA-dependent RNA polymerase), HP (hypothetical protein).

Table 1 Frequency of MaRNAV-1 in primary human blood isolates infected with *P. vivax*.

Country	% positive	# positive samples	# total samples
India	0	0	1
Indonesia	77.8	14	18
Malaysia	61.5	8	13
Solomon Islands	50	1	2

Table 2 Frequency of MaRNAV-1 in publicly available *P. vivax* transcriptomes generated from primary isolates of human blood.

Country	% positive	# positive samples	# total samples
Cambodia	9.1	5	55
Colombia	0	0	100
Ethiopia	0	0	1
Indonesia	0	0	7
Malaysia	0	0	1
Papua New Guinea	33.3	1	3
Peru	45.2	14	31
Thailand	43.6	41	94

compared to our findings of the primary human blood isolates. Despite this, we again observed varying frequencies of MaRNAV-1 in clinical samples from individual countries, albeit at a lower rate than what we observed in the primary human blood isolates (Table 2, Table S2). Overall, our observations suggest that MaRNAV-1 is present in human-infecting *P. vivax* on at least three continents.

MaRNAV-1 is specific to *P. vivax*

Having previously shown an absence of MaRNAV-1 in small numbers of other *Plasmodium* species in Sabah, Malaysia (*P. falciparum* and *P. knowlesi*) (Charon et al. 2019), we next assessed whether MaRNAV-1 is

similarly absent in a broader range of non-*vivax Plasmodium* species, including the relapsing species *P. ovale* and *P. cynomolgi*. We tested this by screening primary human blood isolates from patients who had non-*P. vivax* malaria from the same catchment area in Southeast Asia in which MaRNAV-1 was first detected (Charon et al. 2019) and in patients from the Pacific and Africa (Fig. 1). These samples comprised 13 human blood samples with documented *P. falciparum* from Sabah, Malaysia (n = 5), Papua, Indonesia (imported; n = 2), Cambodia (n = 1), Mali (n = 1), Ghana (n = 1) and Tanzania (n = 2); 8 samples with documented *P. malariae* from Sabah (n = 2) and Papua, Indonesia (n = 6); and 7 samples with documented *P. ovale* infections from Papua, Indonesia (n = 6) and Sabah (n = 1) (Table S1). We also screened a laboratory-adapted strain of *P. cynomolgi* Berok (Christensen et al. 2022). Finally, we screened all publicly available *Plasmodium* meta-transcriptomes that were not previously captured by Charon et al (Charon et al. 2019), (n = 7 662 available on NCBI as of October 2025; Tables S3 [*P. falciparum*; n = 5 753], S4 [non-*P. falciparum*; n = 1909]). These included transcriptomes of laboratory strains (e.g. *P. falciparum* 3D7) and laboratory-culture mouse strains (e.g. *P. berghei*) in addition to natural human infections.

This analysis did not identify MaRNAV-1 in any of the non-*P. vivax* transcriptomes of the primary human blood isolates, including those that were co-circulating with infected *P. vivax*. Specifically, we found no evidence of MaRNAV-1 in *P. malariae*, *P. ovale*, or *P. falciparum* parasites that were co-circulating with *P. vivax* in Papua, Indonesia between 2004 and 2016 (including imported cases), a period during which we repeatedly detected MaRNAV-1 in *P. vivax* infections (Fig. 2a). In addition, we did not identify MaRNAV-1 in any of the seven primary human blood isolates from patients infected with *P. ovale*, which can establish latent liver stages and cause relapsing malaria.

Interestingly, however, we detected the MaRNAV-1 RdRp in 120 *P. falciparum* libraries (8.3%), including 9 collected in Vietnam (Table S3). To assess whether their presence could be *bona fide* evidence of MaRNAV-1 association with *P. falciparum*, we first considered whether these libraries were co-infected with *P. falciparum* and *P. vivax*. This was plausible because all MaRNAV-1-positive *P. falciparum* libraries were generated as part of studies for which co-infection was not

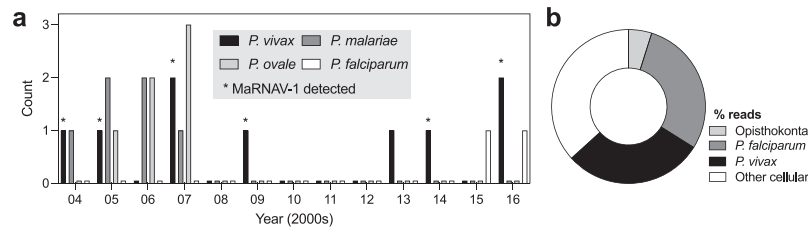


Figure 2 Evidence for MaRNAV-1 specificity to *P. vivax*. (a) Number of primary human-infecting *Plasmodium* species collected per year originating in Papua, Indonesia (including imported cases). The MaRNAV-1 RdRp segment was only detected in *P. vivax* samples (asterisks); (b) Representative output of Kraken2 analysis showing the composition of MaRNAV-1-positive *P. falciparum* library (SRA ID: SRR25922176).

an exclusion criterion (BioProjects PRJNA716963, PRJNA1013216, and PRJDB2573) (Zhu et al. 2022, Nayak et al. 2024), and all three studies were carried out in Southeast Asia, where both *P. falciparum* and *P. vivax* co-circulate (Noviyanti et al. 2015). Additionally, no MaRNAV-1 reads were detected in libraries from samples from the African continent (where *P. vivax* transmission is low and spatially heterogeneous) or in laboratory-adapted strains, which would be known to not contain mixed species. To test for co-infection, we screened the 120 *P. falciparum* libraries containing MaRNAV-1 RdRp sequences for reads aligning to the *P. vivax* genome (Supp. Data 2). We found evidence of *P. vivax* in all 120 *P. falciparum* libraries, including 23 instances in which *P. vivax* was equal to or more abundant in the library compared to *P. falciparum* (e.g. Fig. 2b). We also calculated the probability of finding MaRNAV-1 in 8.3% or less of human-infecting *P. falciparum* if the true prevalence was consistent with that of MaRNAVs in publicly available *P. vivax* transcriptomes (weighted mean = 20.9%, Table 2). To do this, we applied a binomial distribution function (Eq. 1), which returned a one-sided *P*-value of <.0001: the probability of observing 120 or fewer positives in 1453 samples if the true prevalence was 20.89% was effectively zero. From these two lines of evidence, we concluded that only one MaRNAV species (MaRNAV-1) is associated with human-infecting *Plasmodium*, and it is specific to *P. vivax*.

Polyphyletic distribution of MaRNAV-1 in Southeast Asia

We next sought to determine when and from where MaRNAV-1 emerged in *P. vivax*. A recent emergence date would be consistent with the hypothesis that MaRNAV-1 evolved after the divergence of *P. vivax* from a common *Plasmodium* ancestor.

We began by placing the phylogenetic diversity of the novel MaRNAV-1 viruses discovered in this study within all previously identified MaRNAVs and the 100 most closely related RdRp sequences to the MaRNAV-1 RdRp. This analysis placed MaRNAV-4 and MaRNAV-5 in a separate lineage with non-MaRNAV members of the *Narnaviridae* (Fig. S2). Consistent with Esperanza et al., who could not detect a second segment in the corresponding libraries (Esperanza et al. 2025), we concluded that these viruses should not be classified as MaRNAVs and excluded them from further analysis of this group. We applied this conclusion to MaRNAV-8, which forms a sister group to MaRNAVs 4 and 5 (Bourke et al. 2026). However, this does not rule out their association with *Haemoproteus*.

To help determine from where MaRNAV-1 may have emerged, we performed a phylogeographic mapping analysis. Viruses from Southeast Asia were polyphyletic with no defined country-specific clustering (Fig. 3). Although some country-specific clades were present, particularly for Thailand, sequences from Cambodia and Vietnam were

dispersed throughout the phylogeny, likely a reflection of their close geographic relationships. Similarly, viruses from Oceania (Solomon Islands and Papua New Guinea) fell in different parts of the phylogeny, yet among viruses from Southeast Asia, consistent with separate introductions from this region. In contrast, viruses from South America formed a monophyletic group among Southeast Asian viruses, indicative of a single introduction (Fig. 3). This analysis also supports the conclusion that the MaRNAV-1 sequences detected with putative *P. falciparum* libraries were associated with *P. vivax* because MaRNAV-1 from both host species were distributed throughout the MaRNAV-1 clade rather than forming distinct lineages as has been observed in leucocytozoan-associated MaRNAVs (Fig. S2) (Charon et al. 2019).

We were similarly unable to estimate the precise date of emergence of MaRNAV-1 relative to the divergence of *P. vivax*. An analysis of root-to-tip genetic distances against year of sampling using a refined data set of tips with known collection dates did not return a detectable clock-like structure (*R*-squared = 2.55E-2 with heuristic residual mean squared function for the best-fitting root; Fig. S3), likely reflecting the short sampling period (~15 years). Sampling of *Plasmodium* collected prior to the 21st century may be needed to resolve this limitation.

Discussion

Our study presents additional evidence for a specific association between MaRNAV-1 and *P. vivax*. Our data also suggest that MaRNAVs are not associated with other human-infecting *Plasmodium* species, including *P. falciparum* and relapsing species such as *P. ovale*, suggesting that MaRNAV-1 may not be able to infect all *Plasmodium* species. This apparent specificity contrasts with the four distinct MaRNAV lineages that have been reported in leucocytozoans and two related viruses in *Haemoproteus*. Our *P. vivax* metatranscriptome data set contained a disproportionate number of samples from Southeast Asia, but we were still able to conclude that this region is a likely driver of spread of MaRNAV-1-infected *P. vivax*, as evidenced by the monophyly of viruses from South America and the two introductions of MaRNAV-1 into Oceania from Southeast Asia (Fig. 3).

More granular assessment of historical MaRNAV-1 circulation patterns will require robust historical sampling and better representation of the African continent. The earliest samples in our data set were from 2004, and this limited time span was insufficient to allow the emergence of a temporal signal in our root-to-tip analysis. We could not determine whether the arrival of MaRNAV-1 in South America coincided with the first introduction of *P. vivax* to the continent (the timing of which is still debated (Rougeron et al. 2022)). A broader sampling is therefore necessary to determine when MaRNAV-1 emerged in *P. vivax* relative to the spread of this cellular pathogen around the world.

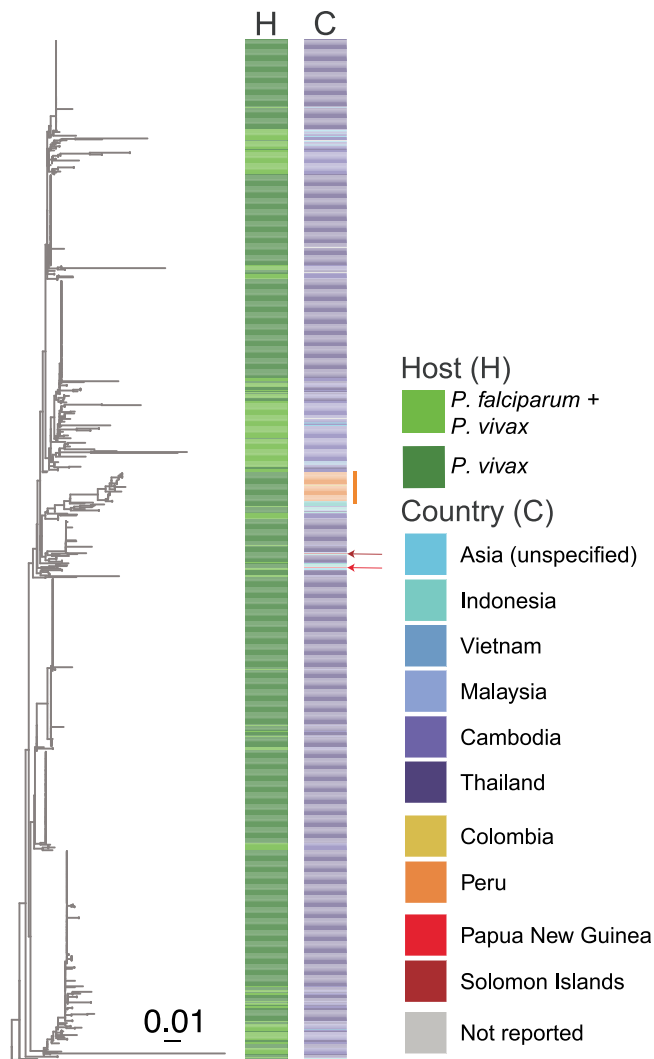


Figure 3 Phylogenetic distribution of MaRNAV-1 by country. Mid-point rooted maximum likelihood phylogenetic tree of the MaRNAV-1 RdRp. Arrows indicate placement of MaRNAV-1 RdRp sequences identified in Oceania. The vertical orange bar corresponds to samples collected in South America. Branches are scaled according to number of amino acid substitutions.

Despite these limitations, the heterogeneity of MaRNAV-1 status with *P. vivax* populations suggests that MaRNAV-1 may be transmissible between *P. vivax* hosts, which would be a unique feature among the *Narnaviridae*. Canonical members of the *Narnaviridae* were so named because their genomes comprise a single gene encoding the RdRp (hence the name ‘naked RNA’—narna). The lack of capsid is likely to limit the possibility of horizontal transmission. Indeed, it has been hypothesised that capsid-less RNA viruses are transmitted vertically, persisting exclusively in the cytoplasm of the cells they infect (Esteban and Fujimura 2003, Richaud et al. 2019). Interestingly, however, MaRNAVs are bi-segmented, encoding ‘hypothetical proteins’ that could have a structural purpose. If the hypothetical proteins are indeed capsid-like and facilitate the transmission of MaRNAVs between protozoan hosts, the acquisition would represent a significant event in the evolution of the MaRNAV viral lineage, possibly even marking its emergence. Identifying the source of this segment and when it was acquired would generate insights into the evolution

of the *Narnaviridae* and potentially virus-protozoa interactions if the second segment was derived from a host gene. It could also mark an evolutionary commitment to a specific host type, just as the acquisition of the E1/E2 glycoprotein in some members of the *Flaviviridae* appears to be linked to commitment to vertebrate hosts (Mifsud et al. 2024). Determining the precise mode of MaRNAV transmission is therefore an important step in disentangling these possibilities. In the absence of a robust *in vitro* system, an investigation of the genetic relatedness of *P. vivax* sub-populations that harbour this virus could be used to further understand both *Plasmodium* and viral transmission dynamics between mosquitoes and humans.

The absence of MaRNAV-1 in *P. cynomolgi* and *P. knowlesi* may reflect distinct evolutionary trajectories between these Asian primate parasites and human *P. vivax*. Phylogenetic analyses position *P. vivax* either as a sister taxon to, or basal to, the Asian primate parasite clade (Arisue et al. 2019), with recent genomic evidence demonstrating that extant human *P. vivax* represents a bottlenecked lineage that emerged from more diverse African ape parasites (Loy et al. 2018). The specificity of MaRNAV-1 to *P. vivax* could reflect either acquisition of this viral association after the divergence from Asian primate *Plasmodium* lineages, or differential retention/loss dynamics during independent host adaptations. African ape *P. vivax*-like parasites exhibit approximately 10-fold greater genetic diversity than human *P. vivax*, consistent with a severe population bottleneck during the establishment of the human parasite lineage (Loy et al. 2018).

In contrast, the absence of MaRNAV-1 infection of the other human-infecting *Plasmodium* species that relapse from latent hepatic hypnozoite parasite stages, *P. ovale* and *P. cynomolgi*, suggests that MaRNAV-1 may not be essential in enabling relapse from hypnozoites in *Plasmodium* spp. Nevertheless, the role of MaRNAV-1 in the biology, pathogenesis and transmission of *P. vivax* in humans remains unclear.

Together, our findings expand our understanding of the Matryoshka RNA virus lineage while raising new questions about the virus-parasite relationship. These questions, such as why *P. vivax* is the only known *Plasmodium* species to harbour MaRNAVs and its role in *P. vivax* pathobiology and transmission in humans, have broader implications for both evolutionary biology and public health. Resolving them through expanded sampling and experimental analyses will become increasingly pertinent if global strategies for the elimination of *P. vivax* falter.

Methods

Collection of primary human blood isolates

Whole blood was collected from patients presenting with *P. vivax* malaria in Sabah, Malaysia (n = 6) (Auburn et al. 2018), Papua, Indonesia (Marfurt et al. 2021), including malaria imported to Darwin from this region after 2008 (Auburn et al. 2016) (n = 18), Solomon Islands (n = 2) (McCarthy et al. 2013, Auburn et al. 2016) and India (n = 1) (Collins et al. 2020); *P. falciparum* detected in Sabah (n = 5) (Barber et al. 2013), Tanzania (n = 2) (Anstey et al. 1996), and imported from Papua, Indonesia (n = 2), Mali (n = 1), Ghana (n = 2), and Cambodia (n = 1) (Auburn et al. 2016); *P. ovale* detected in Sabah, Malaysia (n = 1) (Barber et al. 2013) and Papua, Indonesia (n = 6) (Siswanto et al. 2011); and *P. malariae* detected in Sabah, Malaysia (n = 2) (Grigg et al. 2018) and Papua, Indonesia (n = 6) (Siswanto et al. 2011). Ethics approval was obtained for blood collections from patients with malaria as listed in each of the studies cited above. We also tested a laboratory-adapted

Table 3 Reference genomes used to remove *Plasmodium* reads from primary libraries.

Host	Reference genome
<i>P. ovale</i>	GCA_900090025.2_PowCR01
<i>P. malariae</i>	GCF_900090045.1_PmUG01
<i>P. falciparum</i>	GCF_000002765.4_ASM276v2
<i>P. vivax</i>	Pv_GCF_000002415.2_ASM241v2
<i>P. knowlesi</i>	GCF_000006355.1

strain of *P. cynomolgi Berok* (Chua et al. 2019) for the presence of RNA viruses including MaRNAVs.

Total RNA sequencing of primary human blood isolates

Total RNA was extracted from the primary human blood samples using the RNeasy Universal kit (QIAGEN). Subsequent gDNA removal was performed on total RNAs using DNase I treatment (QIAGEN). Pooled RNA was prepared for total RNA sequencing with the TruSeq stranded library (Illumina) and Globin-Zero gold kit (Illumina). Libraries were sequenced as paired-end 150 bp using either the Illumina NovaSeq 6000 S4 or SP lane (Australian Genome Research Facility, Melbourne).

Bioinformatic processing

SRA screening

All publicly available *Plasmodium* RNA SRA libraries generated through RNA-Seq were downloaded from NCBI according to the following criteria: (i) not included in the original Charon et al., study (Charon et al. 2019); (ii) at least 0.4GB in size (Table S4). These libraries were then screened for the presence of MaRNAV-1-7. The *P. falciparum* SRA libraries were downloaded using SRA Toolkit v3.0.3.

Trimming and assembly

Primary libraries were trimmed using TRIMMOMATIC v0.38 (Bolger et al. 2014) and TruSeq3 paired-end adapters (minimum length = 25). Libraries from the SRA were trimmed using Cutadapt (Martin 2011) v1.8.3 (-u 5 -U 5 -q 25 -m 25). *Plasmodium* reads were removed from primary libraries using Bowtie2 (Langmead and Salzberg 2012) (Table 3). All trimmed and filtered reads were assembled using MEGAHIT (Li et al. 2015) v1.2.9 with default parameters.

Targeted virus discovery

We used two approaches to search for MaRNAV-like contigs in the libraries. For the *P. falciparum* SRA libraries, we downloaded raw reads using SRA Toolkit v3.0.3 and assembled raw reads using MEGAHIT v1.2.9 (Li et al. 2015) with default parameters.

For the libraries generated from the primary human blood isolates and the *P. vivax* SRA, we created a custom DIAMOND database comprising all known MaRNAVs (both RdRp and hypothetical protein segments) and screened our libraries against this database using DIAMOND v2.1.6 (Buchfink et al. 2021). All contigs that returned a significant hit to this database (either e-value < 0.001 and default sensitivity for *P. falciparum* SRA libraries or e-value < 1e-5, --ultra-sensitive for all others) were checked against the NCBI non-redundant (nr) protein

database (as of March 2025) again using DIAMOND BLASTx, and false positives (i.e. contigs with detectable sequence similarity to cellular genes, e-value < 1e-5, --very-sensitive) were removed.

RT-qPCR confirmation of MaRNAV-1 in primary human blood isolates

The primary human blood isolates were pooled for total RNA sequencing. We used RT-qPCR as reported previously (Charon et al. 2019) to test for the presence of MaRNAV-1 in individual isolates that comprised pools in which MaRNAV-1 was detected with RNASeq. Briefly, cDNA was synthesised using the SuperScript IV VILO reverse transcriptase (Invitrogen) and amplification was performed using virus-specific primers (Table S5) and the Platinum SuperFi DNA polymerase (Invitrogen) (35 cycles, annealing temperature 64C). The human gene RSP18 was used as the positive control, and a water blank was used as the negative control (Fig. S1). One PCR reaction was run per sample.

Assessment of host composition

We analysed the composition of *P. falciparum* libraries in which we had detected the MaRNAV-1 RdRp segment using Kraken2 v2.1.3 (Wood et al. 2019). Read counts for Opisthokonta (i.e. animals and fungi), *Plasmodium (Laverania)*, *P. vivax*, and remaining cellular organisms were extracted from the report file in R. The ratio of *P. vivax* to *P. falciparum* reads was calculated using these read counts.

MaRNAV-1 frequency estimates

Percent positivity

To estimate the frequency of association between MaRNAV-1 and *P. vivax*, we filtered the complete data set of *Plasmodium* libraries that we screened to exclude all samples except those known to have been collected from human infections. This data set was further filtered to remove instances of repeated sampling from the same individual for longitudinal study, retaining samples taken prior to antimalarial treatment where possible.

We first tabulated the samples in this refined data set by country. We then measured frequency as the number of libraries (SRA) or samples (primary) with MaRNAV-1 RdRp per sample count per country and reported this figure as a percentage.

Binomial probability function

To measure the probability of observing 120 or fewer MaRNAV-1-positive samples among the 1453 human-infecting *P. falciparum* libraries screened here (8.26% frequency), we used the binomial distribution function implemented in R (pbinom, Eq. 1). We assumed a true frequency of 20.89%, as this was the weighted average in our data set.

$$\Pr(X \leq 120) = \sum_{x=0}^{120} \binom{1453}{x} (0.2089)^x (1 - 0.2089)^{1453-x} \quad (1)$$

Translation and ORF identification

Nucleotide sequences of the MaRNAV-1 RdRp genes identified in this study were translated using the EMBOSS getorf tool (<https://www.bioinformatics.nl/cgi-bin/emboss/getorf>). Open reading frame (ORFs) detection was enabled between stop codons and were limited to

a minimum length of 300 bps. For contigs that returned multiple putative ORFs that met these parameters, the longest ORF was selected. To assess the accuracy of this selection, a preliminary multiple sequence alignment was performed using MAFFT (Katoh and Standley 2013) v7.490 with default parameters implemented in Geneious Prime. Incorrect ORFs were identified, removed, and replaced manually.

Phylogenetic analysis

Phylogenetic position of MaRNAVs within the Narnaviridae

An initial reference data set was compiled by screening the RdRp segment of MaRNAV-1 against the NCBI nr database and downloading the first 100 hits. The amino acid sequences of all published MaRNAVs (Charon et al. 2019, Rodrigues et al. 2022, Duc et al. 2025, Esperanza et al. 2025) were also downloaded. Sequences were aligned with MAFFT (Katoh and Standley 2013) v7.490 using default parameters and implemented in Geneious Prime. Ambiguous sites were removed using trimAl (Capella-Gutiérrez et al. 2009) v1.4.1 ('gappyout'). Phylogenetic analysis (Fig. S2) was performed using the maximum likelihood method implemented in IQ-TREE (Minh et al. 2020) v1.6.12 with the ModelFinder limited to LG models with 1 000 ultrafast bootstraps and SH-aLRT test utilised to measure support. LG + RF + R7 was selected as the best-fit model according to the Bayesian Information Criterion (BIC). The resulting phylogenetic tree was rooted on the non-MaRNAV clade.

Phylogeographic mapping

A data set of all MaRNAV-1 RdRp sequences was compiled (n = 1 110; one of which was identified previously (Charon et al. 2019)). We again used MAFFT, trimAl, and IQ-TREE to infer a phylogenetic tree. All parameters for these methods were retained except that we did not restrict the ModelFinder. LG + R4 was selected as the best-fit model according to BIC. The geographic (i.e. country of sampling) of each MaRNAV was then mapped onto the tips of the phylogenetic tree.

To visualise the phylogeny, we used the R libraries ape (Paradis et al. 2004) and gtree (Yu et al. 2017) implemented in Rv4.4.0. We applied midpoint rooting. We assigned heatmap colours according to the designated country (if known) and microbial host obtained from our primary sample or SRA metadata. We refined the visualisation with Adobe Illustrator.

Root-to-tip analysis of evolutionary rates

To estimate the evolutionary rate of MaRNAV-1 in this data set, we retained only MaRNAV-1 RdRp sequences for which at least the year of sampling was known (n = 794). When only the year was known, we assigned June 1 of that year as the sampling date. All sequences were aligned with MAFFT v7.490 using default parameters (final alignment: 6564 nucleotides in length, 93.2% pairwise identity). We inferred a maximum likelihood phylogenetic tree using IQ-TREE v1.6.12 (1 000 ultrafast bootstraps and sh-aLRT). ModelFinder was unrestricted and selected TPM2u + F + R5 according to the BIC score. Evolutionary rate (i.e. the number of nucleotide changes per site per year) was estimated by plotting root-to-tip genetic distances against the year of sampling with TempEst v1.5.3 (Rambaut et al. 2016), employing the heuristic residual mean squared function to find the best-fitting root (Fig. S3).

Acknowledgements

This work was funded by the National Health & Medical Research Council (NHMRC) [Investigator grant (GNT2017197) to E.C.H., Program Grant (1132975) to N.M.A, R.N.P and J.S.M.] and the Australian Research Council (ARC) [Discovery Project grant (DP240101313) to E.C.H. and M.E.P and Early Career Researcher Discovery Project (DE260100558) to M.E.P.]. We thank Ethan Mandojana for assistance with data cleaning and our colleagues at the National Institute of Health Research and Development and Royal Darwin Hospital for assistance with the studies enrolling patients with malaria.

Author contributions

M.E.P., J.C., M.G., J.W., E.C.H., and N. M. A. conceived this study. M.G., J.W., B.R., R.P., T.W., E.K., J.M., and B.E.B. collected the primary human blood isolates. M.E.P., J.C., R.H.P., and K.S. performed the analyses. M.E.P., E.C.H., and N.M.A. drafted the manuscript. All authors reviewed and edited the manuscript.

Supplementary material

Supplementary material is available at *VEVOLUTION* Journal online.

Conflicts of interest

None declared.

Funding

None declared.

Data availability

The MaRNAV-1 sequences identified in this study are available on our Github repository (https://github.com/mary-petrone/Global_MaRNAV_2026). The RNASeq libraries generated from primary human blood isolates in this study are available on NCBI (BioSamples SAMN57402577—SAMN57402586).

References

- Anstey NM, Weinberg JB, Hassanali MY *et al.* Nitric oxide in Tanzanian children with malaria: inverse relationship between malaria severity and nitric oxide production/nitric oxide synthase type 2 expression. *J Exp Med* 1996;**184**:557–67. <https://doi.org/10.1084/jem.184.2.557>
- Arisue N, Hashimoto T, Kawai S *et al.* Apicoplast phylogeny reveals the position of *Plasmodium vivax* basal to the Asian primate malaria parasite clade. *Sci Rep* 2019;**9**:7274. <https://doi.org/10.1038/s41598-019-43831-1>
- Auburn S, Bohme U, Steinbiss S *et al.* A new *Plasmodium vivax* reference sequence with improved assembly of the subtelomeres reveals an abundance of pir genes. *Wellcome Open Res* 2016;**1**:4. <https://doi.org/10.12688/wellcomeopenres.9876.1>
- Auburn S, Benavente ED, Miotto O *et al.* Genomic analysis of a pre-elimination Malaysian *Plasmodium vivax* population reveals selective pressures and changing transmission dynamics. *Nat Commun* 2018;**9**:2585. <https://doi.org/10.1038/s41467-018-04965-4>
- Barber BE, William T, Grigg MJ *et al.* A prospective comparative study of knowlesi, falciparum, and vivax malaria in Sabah, Malaysia: high

- proportion with severe disease from *Plasmodium knowlesi* and *Plasmodium vivax* but no mortality with early referral and artesunate therapy. *Clin Infect Dis* 2013;**56**:383–97. <https://doi.org/10.1093/cid/cis902>
- Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;**30**:2114–20. <https://doi.org/10.1093/bioinformatics/btu170>
- Bourke BP, Drovetski SV, Ergunay K *et al*. Diverse novel and avian-associated viruses in the ileal viromes of northern mockingbird (*Mimus polyglottos*). *Arch Virol* 2026;**171**:108. <https://doi.org/10.1007/s00705-026-06575-8>
- Buchfink B, Reuter K, Drost H-G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat Methods* 2021;**18**:366–8. <https://doi.org/10.1038/s41592-021-01101-x>
- Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 2009;**25**:1972–3. <https://doi.org/10.1093/bioinformatics/btp348>
- Charon J, Grigg MJ, Eden JS *et al*. Novel RNA viruses associated with *Plasmodium vivax* in human malaria and Leucocytozoon parasites in avian disease. *PLoS Pathog* 2019;**15**:e1008216. <https://doi.org/10.1371/journal.ppat.1008216>
- Christensen P, Racklyeft A, Ward KE *et al*. Improving in vitro continuous cultivation of *Plasmodium cynomolgi*, a model for *P. vivax*. *Parasitol Int* 2022;**89**:102589. <https://doi.org/10.1016/j.parint.2022.102589>
- Chua ACY, Ong JY, Malleret B *et al*. Robust continuous in vitro culture of the *Plasmodium cynomolgi* erythrocytic stages. *Nat Commun* 2019;**10**:3635. <https://doi.org/10.1038/s41467-019-11332-4>
- Collins KA, Wang CY, Adams M *et al*. A *Plasmodium vivax* experimental human infection model for evaluating efficacy of interventions. *J Clin Invest* 2020;**130**:2920–7. <https://doi.org/10.1172/JCI134923>
- Duc M, Esperanza C, Chagas CRF *et al*. First report of Matryoshka RNA virus in an African-European migrant bird. *PLoS One* 2025;**20**:e0319395. <https://doi.org/10.1371/journal.pone.0319395>
- Esperanza CW, Faircloth CE, Roy SW *et al*. Prevalence and diversity of Haemosporidian-associated Matryoshka RNA viruses in a natural population of wild birds. *Ecol Evol* 2025;**15**:e71239. <https://doi.org/10.1002/ece3.71239>
- Esteban R, Fujimura T. Launching the yeast 23S RNA Narnavirus shows 5' and 3' cis-acting signals for replication. *Proc Natl Acad Sci* 2003;**100**:2568–73. <https://doi.org/10.1073/pnas.0530167100>
- Ferreira MU, Nobrega de Sousa T, Rangel GW *et al*. Monitoring *Plasmodium vivax* resistance to antimalarials: persisting challenges and future directions. *Int J Parasitol Drugs Drug Resist* 2021;**15**:9–24. <https://doi.org/10.1016/j.ijpddr.2020.12.001>
- Grigg MJ, William T, Barber BE *et al*. Age-related clinical Spectrum of *Plasmodium knowlesi* malaria and predictors of severity. *Clin Infect Dis* 2018;**67**:350–9. <https://doi.org/10.1093/cid/ciy065>
- Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013;**30**:772–80. <https://doi.org/10.1093/molbev/mst010>
- Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. *Nat Methods* 2012;**9**:357–9. <https://doi.org/10.1038/nmeth.1923>
- Li D, Liu CM, Luo R *et al*. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 2015;**31**:1674–6. <https://doi.org/10.1093/bioinformatics/btv033>
- Loy DE, Plenderleith LJ, Sundararaman SA *et al*. Evolutionary history of human *Plasmodium vivax* revealed by genome-wide analyses of related ape parasites. *Proc Natl Acad Sci* 2018;**115**:E8450–9. <https://doi.org/10.1073/pnas.1810053115>
- Marfurt J, Wirjanata G, Prayoga P *et al*. Longitudinal ex vivo and molecular trends of chloroquine and piperazine activity against *Plasmodium falciparum* and *P. vivax* before and after introduction of artemisinin-based combination therapy in Papua, Indonesia. *Int J Parasitol Drugs Drug Resist* 2021;**17**:46–56. <https://doi.org/10.1016/j.ijpddr.2021.06.002>
- Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal* 2011;**17**:3.
- McCarthy JS, Griffin PM, Sekuloski S *et al*. Experimentally induced blood-stage *Plasmodium vivax* infection in healthy volunteers. *J Infect Dis* 2013;**208**:1688–94. <https://doi.org/10.1093/infdis/jit394>
- Mifsud JCO, Lytras S, Oliver MR *et al*. Mapping glycoprotein structure reveals defining events in the evolution of the Flaviviridae. *bioRxiv* 2024;2024.2002.2006.579159.
- Mikolajczak SA, Vaughan AM, Kangwanransan N *et al*. *Plasmodium vivax* liver stage development and hypnozoite persistence in human liver-chimeric mice. *Cell Host Microbe* 2015;**17**:526–35. <https://doi.org/10.1016/j.chom.2015.02.011>
- Minh BQ, Schmidt HA, Chernomor O *et al*. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 2020;**37**:1530–4. <https://doi.org/10.1093/molbev/msaa015>
- Nayak S, Peto TJ, Kucharski M *et al*. Population genomics and transcriptomics of *Plasmodium falciparum* in Cambodia and Vietnam uncover key components of the artemisinin resistance genetic background. *Nat Commun* 2024;**15**:10625. <https://doi.org/10.1038/s41467-024-54915-6>
- Ngwana-Joseph GC, Phelan JE, Manko E *et al*. Genomic analysis of global *Plasmodium vivax* populations reveals insights into the evolution of drug resistance. *Nat Commun* 2024;**15**:10771. <https://doi.org/10.1038/s41467-024-54964-x>
- Noviyanti R, Coutrier F, Utami RAS *et al*. Contrasting transmission dynamics of Co-endemic *Plasmodium vivax* and *P. falciparum*: implications for malaria control and elimination. *PLoS Negl Trop Dis* 2015;**9**:e0003739. <https://doi.org/10.1371/journal.pntd.0003739>
- Pacheco MA, Matta NE, Valkiunas G *et al*. Mode and rate of evolution of Haemosporidian mitochondrial genomes: timing the radiation of avian parasites. *Mol Biol Evol* 2018;**35**:383–403. <https://doi.org/10.1093/molbev/msx285>
- Paradis E, Claude J, Strimmer K. APE: analyses of Phylogenetics and evolution in R language. *Bioinformatics* 2004;**20**:289–90. <https://doi.org/10.1093/bioinformatics/btg412>
- Price RN, von Seidlein L, Valecha N *et al*. Global extent of chloroquine-resistant *Plasmodium vivax*: a systematic review and meta-analysis. *Lancet Infect Dis* 2014;**14**:982–91. [https://doi.org/10.1016/S1473-3099\(14\)70855-2](https://doi.org/10.1016/S1473-3099(14)70855-2)
- Rambaut A, Lam TT, Max Carvalho L *et al*. Exploring the temporal structure of heterochronous sequences using TempEst (formerly path-O-gen). *Virus Evol* 2016;**2**. <https://doi.org/10.1093/ve/vew007>
- Richaud A, Frézal L, Tahan S *et al*. Vertical transmission in Caenorhabditis nematodes of RNA molecules encoding a viral RNA-dependent RNA polymerase. *Proc Natl Acad Sci* 2019;**116**:24738–47. <https://doi.org/10.1073/pnas.1903903116>
- Rodrigues JR, Roy SW, Sehgal RNM. Novel RNA viruses associated with avian haemosporidian parasites. *PLoS One* 2022;**17**:e0269881. <https://doi.org/10.1371/journal.pone.0269881>
- Rougeron V, Daron J, Fontaine MC *et al*. Evolutionary history of *pPlasmodium vivax* and *Plasmodium simium* in the Americas. *Malar J* 2022;**21**:141. <https://doi.org/10.1186/s12936-022-04132-7>
- Seng D, Ko K, Orban A *et al*. MaNAV-1 infection of *Plasmodium vivax* is associated with increased parasite transmission

- and host inflammatory responses. *bioRxiv*, 2026.2003.2022.713337. 2026.
- Siswanto H, Russell B, Ratcliff A *et al*. In vivo and in vitro efficacy of chloroquine against *Plasmodium malariae* and *P. Ovale* in Papua, Indonesia. *Antimicrob Agents Chemother* 2011;**55**:197–202. <https://doi.org/10.1128/AAC.01122-10>
- Thriemer K, Ley B, Bobogare A *et al*. Challenges for achieving safe and effective radical cure of *Plasmodium vivax*: a round table discussion of the APMEN Vivax working group. *Malar J* 2017;**16**:141. <https://doi.org/10.1186/s12936-017-1784-1>
- Valkiunas G. *Avian Malaria Parasites and Other Haemosporidia*. Boca Raton, FL: CRC Press, 2004. [10.1201/9780203643792](https://doi.org/10.1201/9780203643792).
- Weiss DJ, Dzionach PA, Saddler A *et al*. Mapping the global prevalence, incidence, and mortality of *Plasmodium falciparum* and *Plasmodium vivax* malaria, 2000–22: a spatial and temporal modelling study. *Lancet* 2025;**405**:979–90. [https://doi.org/10.1016/S0140-6736\(25\)00038-8](https://doi.org/10.1016/S0140-6736(25)00038-8)
- Wood DE, Lu J, Langmead B. Improved metagenomic analysis with kraken 2. *Genome Biol* 2019;**20**:257. <https://doi.org/10.1186/s13059-019-1891-0>
- Yu G, Smith DK, Zhu H *et al*. Ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol* 2017;**8**:28–36. <https://doi.org/10.1111/2041-210X.12628>
- Zekar L, Sharman T. *Plasmodium Falciparum Malaria*. Treasure Island (FL): StatPearls, 2025.
- Zhu L, van der Pluijm RW, Kucharski M *et al*. Artemisinin resistance in the malaria parasite, *Plasmodium falciparum*, originates from its initial transcriptional response. *Commun Biol* 2022;**5**:274. <https://doi.org/10.1038/s42003-022-03215-0>