PLOS ONE

RESEARCH ARTICLE

Systematic review of *Plasmodium falciparum* and *Plasmodium vivax* polyclonal infections: Impact of prevalence, study population characteristics, and laboratory procedures

Luis Lopez[1], Cristian Koepfli[1,2]*

1 Department of Biological Sciences, University of Notre Dame, Notre Dame, IN, United States of America,
2 Eck Institute for Global Health, University of Notre Dame, Notre Dame, IN, United States of America

* ckoepfli@nd.edu

Abstract

Multiple infections of genetically distinct clones of the same *Plasmodium* species are common in many malaria endemic settings. Mean multiplicity of infection (MOI) and the proportion of polyclonal infections are often reported as surrogate marker of transmission intensity, yet the relationship with traditional measures such as parasite prevalence is not well understood. We have searched Pubmed for articles on *P. falciparum* and *P. vivax* multiplicity, and compared the proportion of polyclonal infections and mean MOI to population prevalence. The impact of the genotyping method, number of genotyping markers, method for diagnosis (microscopy/RDT vs. PCR), presence of clinical symptoms, age, geographic region, and year of sample collection on multiplicity indices were assessed. For *P. falciparum*, 153 studies met inclusion criteria, yielding 275 individual data points and 33,526 genotyped individuals. The proportion of polyclonal infections ranged from 0–96%, and mean MOI from 1–6.1. For *P. vivax*, 54 studies met inclusion criteria, yielding 115 data points and 13,325 genotyped individuals. The proportion of polyclonal infections ranged from 0–100%, and mean MOI from 1–3.8. For both species, the proportion of polyclonal infections ranged from very low to close to 100% at low prevalence, while at high prevalence it was always high. Each percentage point increase in prevalence resulted in a 0.34% increase in the proportion of polyclonal *P. falciparum* infections (*P*<0.001), and a 0.78% increase in the proportion of polyclonal *P. vivax* infections (*P*<0.001). In multivariable analysis, higher prevalence, typing multiple markers, diagnosis of infections by PCR, and sampling in Africa were found to result in a higher proportion of *P. falciparum* polyclonal infections. For *P. vivax*, prevalence, year of study, typing multiple markers, and geographic region were significant predictors. In conclusion, polyclonal infections are frequently present in all settings, but the association between multiplicity and prevalence is weak.

Introduction

Malaria remains a global public health threat with over 200 million confirmed cases and over 400,000 deaths recorded in 2019 [1]. Reliable measurements of transmission intensity are

needed in order to optimally allocate resources for control and to measure changes response to control. Obtaining such measures is a challenge [2].

Individuals in malaria endemic regions are often concurrently infected with several clones or strains of the same parasite species. Polyclonal infections can be the result of a single mosquito bite transmitting a genetically diverse inoculum (coinfection), or of repeated bites transmitting different clones (superinfection) [3]. In *P. vivax*, relapses of heterologous hypnozoites during ongoing blood-stage infection can also result in polyclonal infections [4, 5].

Multiplicity of infection (MOI, also termed 'complexity of infection', COI), and the proportion of polyclonal infections (also termed 'multiple clone infections') are among the most frequently reported measures of molecular malaria surveillance studies of *Plasmodium falciparum* and *Plasmodium vivax*. A large number of genotyping protocols is used to determine multiplicity. The most common methods include typing of size-polymorphic markers (antigens or microsatellites) and sizing on agarose gel or by capillary electrophoresis (CE) [6, 7], SNP panels [8], and more recently amplicon sequencing [9, 10].

MOI is generally assumed to be higher in high-transmission settings, while in low-transmission settings most individuals might carry single clone infections. Multiplicity has thus been proposed as surrogate marker to measure changes in transmission intensity in time and space. Numerous studies found a correlation between transmission intensity and MOI. For example, in West Papua, Indonesia, clinical incidence of *P. vivax* and *P. falciparum* approximately halved over a 14-year period. The proportion of polyclonal infections fell from 70% to 35% for *P. vivax*, and from 28 to 18% for *P. falciparum* [11]. Along the Thai-Myanmar border, a 10-fold decrease in incidence over a 10-year period was accompanied by a reduction in the proportion of polyclonal *P. falciparum* infections from 63% to 14% [12]. In Senegal, the proportion of polyclonal infections fell from 50% in 2001–2002 to 0% in 2014 along with a decrease in incidence [13]. In Venezuela, multiplicity increased dramatically in parallel to an increase in incidence [14]. In Burkina Faso, MOI mirrored seasonal patterns of incidence [15]. In other studies, however, the association between multiplicity of infection and transmission intensity was weak [16], or absent [17].

Understanding the extent of polyclonal infections is important for molecular malaria surveillance activities beyond measuring transmission intensity. Typing of molecular markers for drug resistance is an important aspect of malaria surveillance [18]. Polyclonal infections can result in estimates of prevalence (i.e. the proportion of all individuals carrying a drug resistant parasite) to differ substantially from the true frequency (i.e. the proportion of all parasite strains in the population carrying the marker of resistance) [19]. Rapid diagnostic tests (RDTs) are key tools for malaria diagnosis. Most RDTs for *P. falciparum* diagnose the HRP2 protein. Increasing numbers of parasites are being reported carrying a deletion of the *hrp2* gene, which results in false negative tests. Molecular surveillance of *hrp2* deletion is a high priority [20]. In case of polyclonal infections, wild type parasites can mask strains carrying the deletion, resulting in an underestimation of the frequency of deletion [21]. Changes in the frequency of polyclonal infections could result in observed changes in the prevalence of deletions and mutation, even if population frequencies remain constant. Data on multiplicity is also required for the reconstruction of haplotypes from whole genome sequencing data. In the presence of polyclonal infections, specific software tools are required [22]. As a further line of research, the impact of polyclonal infection on clinical symptoms has been studied [23–25].

Despite the large number of studies reporting multiplicity data, the relationship between wMOI or the proportion of polyclonal infections and transmission intensity across a wide range of studies remains poorly characterized. Likewise, little effort has been made to systematically assess the impact on multiplicity estimates of characteristics of the study population (e.g. asymptomatic vs. clinical infections) or the method used for genotyping (e.g. gel electrophoresis

vs. capillary electrophoresis for sizing of polymorphic markers). We have conducted a systematic literature search and compared the proportion of polyclonal infections and mean multiplicity to population prevalence of infection as common measure of transmission intensity. Furthermore, we studied the effect of laboratory procedures and characteristics of the study population on multiplicity indices.

Methods

Search strategy

PubMed was searched on 1/10/21 using the searches "Multiplicity of infection *Plasmodium falciparum*", "Complexity of infection *Plasmodium falciparum*", and "Multiple clone infection *Plasmodium falciparum*". Articles were eligible if they reported data on mean MOI and/or the proportion of polyclonal infections. From clinical trials, data was included only when reported separately for the baseline. 265 studies were identified as eligible for assessment out of which 153 were included in the pooled analysis (S1 Fig: PRISMA 2009 Flow Diagram).

The same search strategy for *P. vivax* yielded very few results. Thus, a broader search attempt was made. This included screening of additional studies from authors that were identified in the initial search for *P. falciparum* or *P. vivax* data, and screening of studies that cited widely used methods for *P. vivax* genotyping (e.g. [7, 26, 27]). The final dataset for *P.* vivax included 54 studies.

From each study, data on multiplicity of infection and/or the percentage of polyclonal infections were recorded. If available, population prevalence of infection was recorded from the article. If no prevalence data was reported, data was obtained from the Malaria Atlas Project (MAP) website (https://malariaatlas.org). MAP provides modeled prevalence of infection by microscopy in 1-10-year-old children at the country or sub country level. For the current study, data from the lowest administrative level available was recorded. With very few exceptions, data points for which prevalence data was taken from MAP had only genotyped children. Thus, the age range of the individuals genotyped matched the age range of the data in MAP. For studies conducted before the year 2000, data from 2000 was recorded (the earliest available from MAP).

Additional elements of the eligible articles that were analyzed included the country of study, the year of sample collection, whether the study population was clinical or asymptomatic, the method for diagnosis, the method for genotyping, and the age range of the study population. The total number of patients surveyed, and the number of patients genotyped was recorded. A few studies did not report genotyping success rate, and thus the total number of genotyped samples is unknown. In these cases, it was assumed that all infections underwent genotyping, and the number of samples with *P. falciparum* and *P. vivax* infections was included.

Certain articles that met the inclusion criteria yielded more than one data point. Elements within studies that prompted individual treatment included examination of multiple geographic locations, investigation of both clinical and asymptomatic patients separately, inclusion of multiple age ranges, or reporting data from multiple different time periods.

Data analysis

In addition to prevalence of infection, characteristics of the study population and the laboratory methods selected for diagnosis and genotyping can affect MOI and the proportion of polyclonal infections. The impact of the following factors on multiplicity indices was analyzed:

(i) Number of genotyping markers: A larger number of molecular markers increases the probability to detect additional clones. Studies were distinguished between using a single

marker or multiple markers. SNP panels always include multiple markers, but only two alleles can be distinguished for each SNP. Whole genome sequencing (WGS) covers the entire genome, but read depth is generally lower than amplicon sequencing. Thus, SNP panels and WGS were excluded from this analysis. (ii) Diagnosis: In all transmission settings, a large proportion of infections remains below the limit of detection of microscopy or rapid diagnostic test (RDT) and can only be diagnosed by sensitive molecular methods (e.g. PCR) [28, 29]. Detectability of minority clones is higher when parasite density is higher [30]. While individual data on parasite density was not available, data was grouped into infections diagnosed either by PCR (i.e. more low-density infections included) or diagnosed by microscopy or RDT. (iii) Genotyping method: Estimates of multiplicity are heavily impacted by the ability to detect minority clones. Capillary electrophoresis generally increases detectability of minority clones and distinction of fragments of similar size compared to visualization on gel, thus CE detects higher multiplicity [31, 32]. Amplicon deep sequencing further increases the ability to discriminate between clones [9]. Data was thus grouped by genotyping method as follows: a) size polymorphic markers visualized on gel, b) size polymorphic markers visualized by CE, c) SNP panels, d) amplicon sequencing, and e) whole genome sequencing. (iv) Clinical vs. asymptomatic infections: Parasite density in clinical samples is expected to be higher than in asymptomatic samples, affecting the ability to detect minority clones. Thus, the effect of sampling from clinical vs. asymptomatic populations was assessed. (v) Age: Multiplicity of infection has been shown to be affected by age [33]. For the analysis, data was grouped into age ranges as follows: including individuals of all ages, only individuals ≤15 years, and only individuals >15 years included. Some studies that did not fully fit this classification were assigned to the closest matching group (S1 File). In moderate and high transmission settings, clinical disease occurs mostly in children, in particular in the case of *P. vivax*. Thus, it is expected that many studies on clinical patients enrolled mostly children. Nevertheless, unless an age limit was stated, they were classified as including all ages. (vi) Geographic locations: Studies were grouped as conducted either in Africa, the Asia-Pacific, or the Americas. (vii) Year: Studies conducted from 1990 to 2019 were included in the analysis. Over this period, the number of *P. falciparum* and *P. vivax* cases globally has decreased [1]. Associations between the year of sample collection and multiplicity of infection indices were analyzed. If the year of sample collection was not mentioned, the year prior to the year of publication was used for analysis.

Linear regression models were used to analyze the impact of the above-described variables and the proportion of polyclonal infections or mean multiplicity. Multivariable regression was conducted, with step-wise removal of variables with a *P*-value >0.05, until the final model included only variables with *P*≤0.05.

Results

P. falciparum literature search

The literature search yielded 153 studies that met the inclusion criteria, resulting in 275 data points with a total of 33,526 genotyped infections. Summary statistics are presented in Table 1. On average, 122 individuals were genotyped per data point. The proportion of polyclonal infections ranged from 0–99%, and mean MOI from 1–6.1.

P. falciparum proportion of polyclonal infections

In univariate analysis, a significant correlation between the proportion of polyclonal infections and prevalence was observed (Fig 1A). Each percentage point increase in prevalence resulted in an 0.34% increase in the proportion of polyclonal infections (n = 174, *P*<0.001). At low prevalence, the proportion of infections carrying multiple clones ranged from very low to

Table 1. Summary statistics of studies analyzed.

	P. falciparum	*P. vivax*
Number of studies	153	54
Number of data points[1]	275	115
Number of surveyed individuals	105609	39671
Number of genotyped individuals	33526	13325
Method for diagnosis of infection[2]		
Microscopy/RDT	180 (66.4%)	61 (51.7%)
PCR	91 (33.6%)	57 (48.3%)
Number of genotyping markers[2]		
1 marker	133 (49.1%)	11 (9.4%)
>1 marker	138 (50.9%)	106 (90.6%)
Method for genotyping[2]		
Size polymorphic - gel	214 (79.0%)	18 (15.4%)
Size polymorphic - CE	51 (18.8%)	89 (76.1%)
SNP panel	1 (0.4%)	1 (0.9%)
Amplicon sequencing	5 (1.8%)	4 (3.4%)
WGS	0 (0.0%)	5 (4.3%)
Clinical symptoms[2]		
Asymptomatic	120 (44.3%)	26 (22.0%)
Clinical	151 (55.7%)	92 (78.0%)
Age range[2]		
All ages	162 (59.8%)	102 (86.4%)
<15 years	92 (34.0%)	7 (5.9%)
>15 years	17 (6.3%)	9 (7.6%)
Geographic region[2]		
Africa	214 (79.0%)	11 (9.3%)
Asia-Pacific	51 (18.8%)	84 (71.2%)
Latin America	6 (2.2%)	23 (19.5%)
Mean proportion of polyclonal infections [range]	57.0% [0–98.9]	42.8% [0–100]
Mean MOI [range]	2.26 [1–6.1]	1.59 [1–3.8]

Table 1: Characteristics of studies analyzed.

[1] Multiple data points were retrieved from a study if data was reported separately for different locations, age groups, time points, or separately for clinical and asymptomatic individuals.

[2] For each factor, the number and percentage of data points is given.

https://doi.org/10.1371/journal.pone.0249382.t001

almost 100% (Fig 1A). At higher prevalence, very few studies found a low proportion of poly-clonal infections. In only two out of 43 studies reporting prevalence rates >50% was the pro-portions of polyclonal infections <50%. Over the time period analyzed (1990–2019), a pronounced decrease in the number of clinical malaria cases was observed globally. Nevertheless, no significant impact of the year of study on the proportion of polyclonal infections was observed (n = 174, P = 0.869, Fig 2A).

The proportion of infections diagnosed by PCR that were polyclonal was 10.4% less than for infections diagnosed by microscopy or RDT (irrespective of whether individuals showed symptoms) (n = 174, P = 0.010, Fig 3A). Typing multiple markers resulted in 8.2% more poly-clonal infections (n = 174, P = 0.032, Fig 3B). When size-polymorphic markers were typed, typing by CE instead of on gel resulted in detection of 12.8% fewer polyclonal infections (n = 168, P = 0.007, Fig 3C). Notably, the proportion of studies that used CE was 3-fold higher
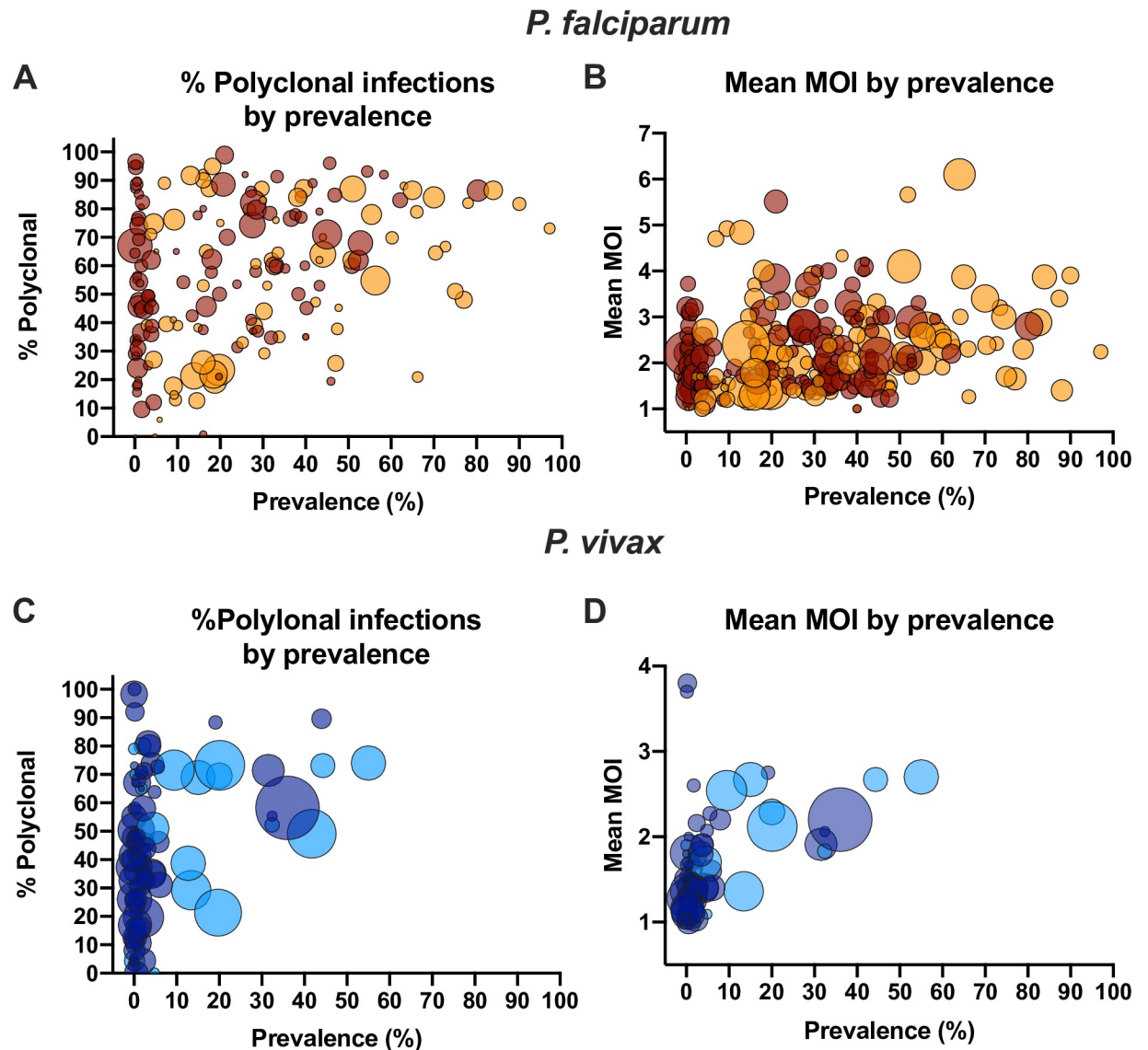
Fig 1. Relationship between prevalence and the proportion P. *falciparum* and *P. vivax* polyclonal infections (A, C) or mean multiplicity (B, D). Dot sizes represent the number of individuals genotyped. Dark red/dark blue = samples collected form clinical patients. Orange/light blue = samples collected form asymptomatic individuals. Overlapping data points result in darker colors.

https://doi.org/10.1371/journal.pone.0249382.g001

in the Asia-Pacific region compared to Africa (42.9 vs. 13.2%, $P<0.001$). In the Asia-Pacific region, the proportion of polyclonal infections was 21.9% lower than the proportion in Africa (n = 168, $P<0.001$, Fig 3F).

No significant difference was observed in the proportion of polyclonal infections between samples collected from clinical and asymptomatic individuals (n = 174, $P = 0.358$, Fig 3D). Age group had a significant impact on the proportion of polyclonal infections, with the highest proportion observed in studies only enrolling children (n = 174, $P = 0.0033$, Fig 3E).

In multivariable analysis, the year the study was conducted, sizing by CE vs. gel, age group, and whether samples were from clinical or asymptomatic individuals did not reach significance and were excluded from the final model (Table 2A). In the reduced model, each percentage point increase in prevalence resulted in an 0.28% increase in the proportion of polyclonal
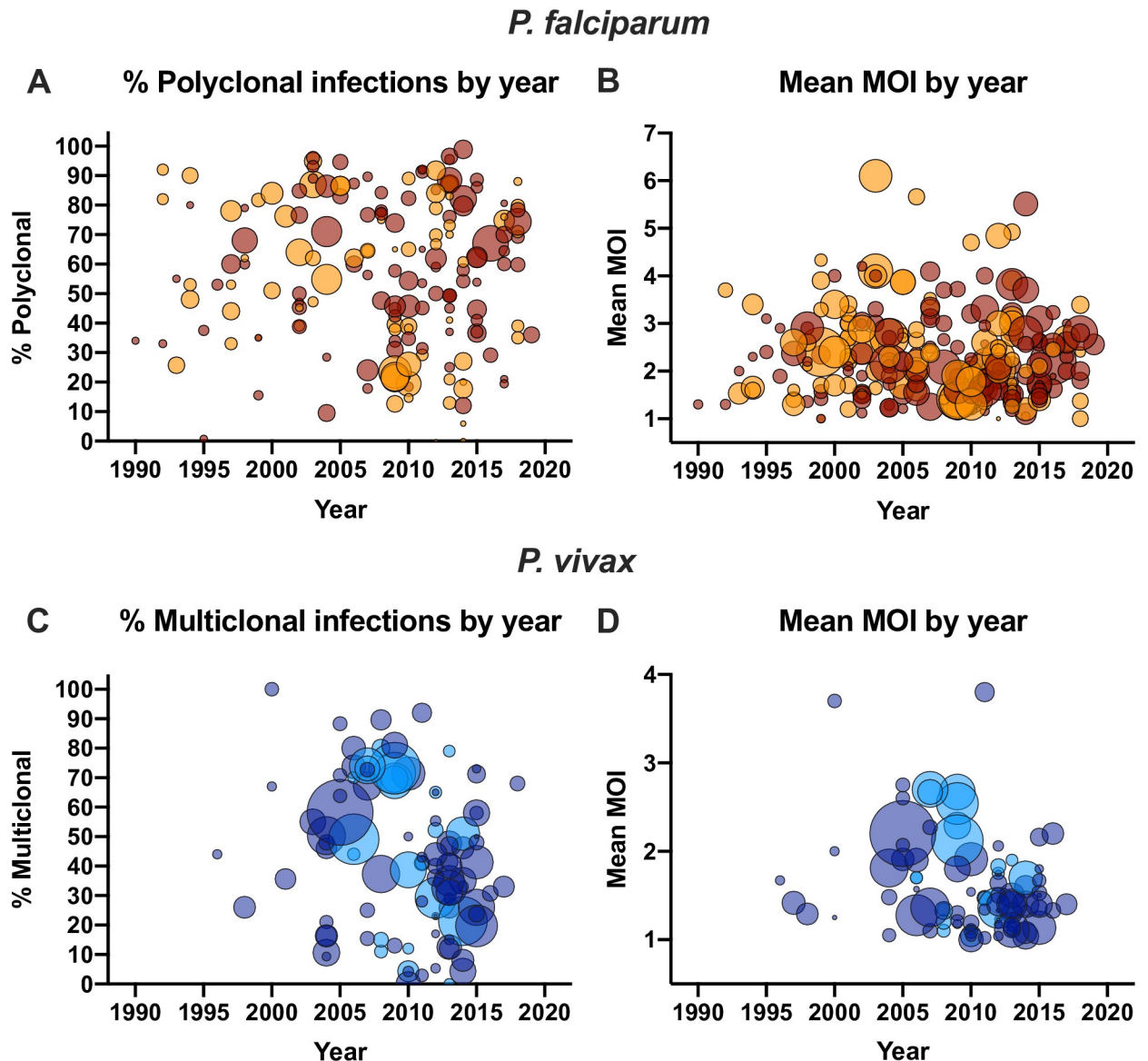
P. falciparum

Fig 2. Relationship between year of sample collection and the proportion *P. falciparum* and *P. vivax* polyclonal infections (A, C) or mean multiplicity (B, D). Dot sizes represent the number of individuals genotyped. Dark red/dark blue = samples collected form clinical patients. Orange/light blue = samples collected form asymptomatic individuals. Overlapping data points result in darker colors.

https://doi.org/10.1371/journal.pone.0249382.g002

infections (P = 0.001). Typing multiple markers instead of only one marker resulted in detection of 14.1% more polyclonal infections (P<0.001). Diagnosis of infections by PCR instead of microscopy or RDT resulted in 9.4% fewer polyclonal infections detected (P = 0.013). Compared to Africa, the proportion of polyclonal infections in the Asia-Pacific region was 17.8% lower, and in Latin America, it was 28.0% lower (P<0.001).

P. falciparum mean multiplicity

In univariate analysis, each percentage point increase in prevalence resulted in an increase in mean MOI of 0.010 (n = 254, P<0.001, Fig 1B). No impact of the year of study on mean MOI was observed (n = 254, P = 0.241, Fig 2B).
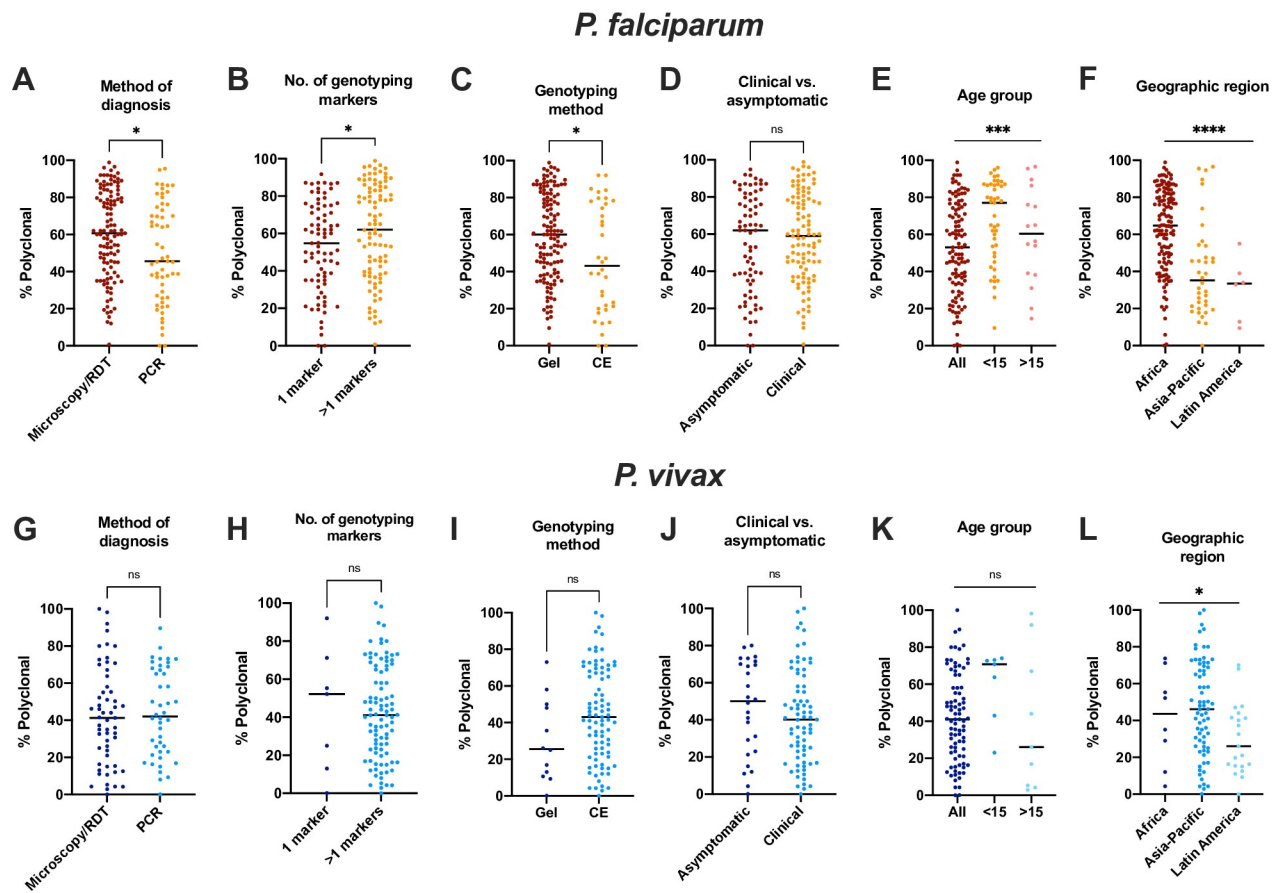
P. falciparum

P. vivax

Fig 3. Impact of presence of laboratory methods (method for diagnosis, number of genotyping markers, method to size size-polymorphic amplicons) and patient characteristics (presence of clinical symptoms, age group, geographic region) on the proportion P. *falciparum* (A-F) and P. *vivax* (G-L) polyclonal infections. Black line denotes median. Groups were compared by Mann-Whitney test (if 2 groups), or regression analysis (if 3 groups). * $P \le 0.05$, **$P \le 0.01$, *** $P \le 0.001$, **** $P \le 0.0001$, ns = not significant.

Mean MOI did not differ significantly between studies on clinical and asymptomatic individuals (n = 254, P = 0.158). A trend towards lower mean MOI was observed in studies using PCR for diagnosis compared to microscopy or RDT (mean MOI 0.219 lower if PCR used, n = 254, P = 0.069). In studies including children <15 years (n = 90), mean MOI was 0.350 higher compared to studies including all ages (n = 150, P = 0.004). No difference was observed between studies including only adults >15 years (n = 14) and all ages (P = 0.988). Typing of size-polymorphic markers by CE instead of gel resulted in a trend toward lower mean MOI (0.270 lower by CE, n = 249, P = 0.073).

In multivariable analysis, higher prevalence and typing multiple markers resulted in higher mean MOI. Diagnosis by PCR and typing clinical samples resulted in lower mean MOI (Table 2B).

P. falciparum marker *msp2*

While a wide variety of genotyping markers was used for *P. falciparum* genotyping, *msp2* was the most frequently used marker. 97 data sets reported the proportion of polyclonal infections based on *msp2*, and 172 data sets reported mean multiplicity based on *msp2*. To study whether typing the same marker impacts the relationship between multiplicity and prevalence, the analysis was repeated using *msp2* only.

Table 2. Multivariable predictors of A) *P. falciparum* **proportion of polyclonal infections, and B) mean multiple clone infections.** 95%CI = 95% confidence interval.

A) *P. falciparum* proportion of polyclonal infections					
	Full model			Reduced model	
	Coefficient [95%CI]	*P*		Coefficient [95%CI]	*P*
Prevalence	0.284 [0.105, 0.464]	0.002		0.285 [0.122, 0.449]	0.001
Year	0.031 [-0.512, 0.574]	0.911			
Sizing by CE	-4.234 [-13.838, 5.370]	0.385			
>1 genotyping marker	13.505 [6.264, 20.746]	<0.001		14.13 [7.180, 21.079]	<0.001
PCR for diagnosis	-7.621 [-15.829, 0.587]	0.069		-9.377 [-16.761, -1.993]	0.013
Geographic Region					
Africa	[reference]			[reference]	
Asia-Pacific	-16.143 [-25.903, -6.383]	0.001		-17.787 [-26.683, -8.891]	<0.001
Latin America	-26.805 [-46.174, -7.436]			-28.008 [-46.291, -9.725]	
Age group					
All	[reference]				
<15	6.032 [-1.988, 14.052]	0.1602			
>15	9.086 [-3.734, 21.907]				
Presence of clinical symptoms	0.992 [-7.207, 9.190]	0.812			
B) *P. falciparum* mean multiplicity of Infection					
	Full model			Reduced model	
	Coefficient [95%CI]	*P*		Coefficient [95%CI]	*P*
Prevalence	0.009 [0.003, 0.015]	0.003		0.012 [0.007, 0.017]	<0.001
Year	0 [-0.019, 0.018]	0.96			
Sizing by CE	-0.0997 [-0.413, 0.213]	0.531			
>1 genotyping marker	0.331 [0.101, 0.561]	0.005		0.319 [0.092, 0.546]	0.006
PCR for diagnosis	-0.347 [-0.619, -0.075]	0.013		-0.421 [-0.686, -0.156]	0.002
Geographic Region					
Africa	[reference]				
Asia-Pacific	-0.286 [-0.608, 0.036]	0.1689			
Latin America	0.229 [-0.67, 1.127]				
Age group					
All	[reference]				
<15	0.149 [-0.096, 0.394]	0.4827			
>15	0.012 [-0.492, 0.516]				
Presence of clinical symptoms	-0.314 [-0.585, -0.044]	0.023		-0.273 [-0.538, -0.008]	0.044

The proportion of polyclonal infections increased by 0.32% per percentage point increase in prevalence (n = 97, *P* = 0.002, Fig 4A). As for the entire dataset, at low prevalence, the proportion of polyclonal infections varied from 0 to almost 100%. At high prevalence, in most studies, the proportion was >50%. Mean MOI based on *msp2* increased by 0.012 per percentage point increase in prevalence (n = 172, *P*<0.001, Fig 4B).

P. vivax literature search

Using a broad set of search criteria, 54 studies were identified that met inclusion criteria, yielding 115 data points that included 13,325 genotyped individuals (Table 1). On average 118 individuals were genotyped per study. The proportion of polyclonal infections ranged from 0–100%, and the mean multiplicity ranged from 1–3.8.
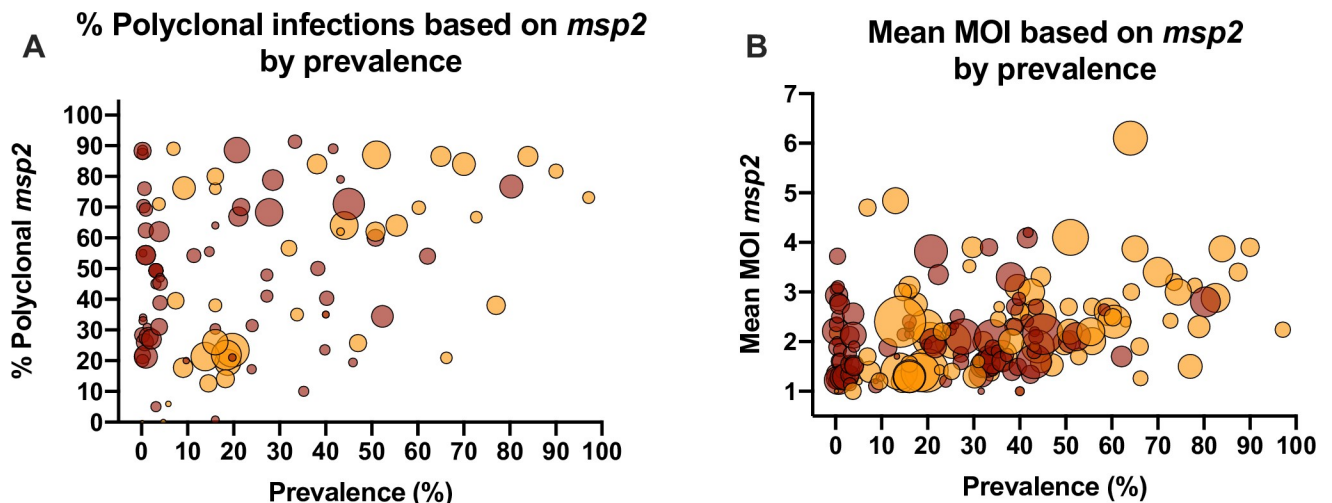
Fig 4. Relationship between the proportion *P. falciparum* polyclonal infections (A) or mean multiplicity (B) and prevalence based on typing of marker *msp2*. Dot sizes represent the number of individuals genotyped. Dark red = samples collected form clinical patients. Orange = samples collected form asymptomatic individuals. Overlapping data points result in darker colors.

https://doi.org/10.1371/journal.pone.0249382.g004

P. vivax proportion of polyclonal infections

As for *P. falciparum*, at low *P. vivax* prevalence, the proportion of polyclonal infections varied widely from almost 0% to close to 100% (Fig 1C). Few studies were available where *P. vivax* prevalence was high, and the proportion of polyclonal infections was high in all of these studies. In univariate analysis, each percentage point increase in prevalence resulted in an 0.78% increase in the proportion of polyclonal infections (n = 101, $P<0.001$). Over the time period analyzed (1996–2018), a trend toward a lower proportion of polyclonal infections was observed (1.02% reduction per year, n = 99, $P = 0.070$, Fig 1C).

In univariate analysis, no difference was observed in the proportion of polyclonal infections whether diagnosis was by microscopy/RDT or PCR (n = 101, $P = 0.764$, Fig 3G), using multiple genotyping markers (n = 10, $P = 0.917$, Fig 3H), typing by CE vs. gel (n = 90, $P = 0.082$, Fig 3J), between asymptomatic and clinical infections (n = 101, $P = 0.418$, Fig 3J), or among age groups (n = 101, $P = 0.169$, Fig 3K). A significant difference was observed among geographic regions (n = 101, $P = 0.028$, Fig 3L).

In multivariable analysis, each percentage point increase in prevalence resulted in 0.48% more polyclonal infections ($P<0.001$, Table 3A). Compared to Africa, in the Asia-Pacific the proportion of polyclonal infections was 9.6% higher, and in Latin America 7.5% lower ($P = 0.0141$, Table 3A). Few studies typed a single marker, which resulted in 35% fewer polyclonal infections detected ($P = 0.006$, Table 3A). Further significant predictors were year of sample collection and age group (Table 3A).

P. vivax mean multiplicity

In univariate analysis, each percentage point increase in prevalence resulted in an increase in mean MOI of 0.025 (n = 91, $P<0.001$, Fig 1D). The year of study was not a significant predictor for mean MOI (n = 91, $P = 0.207$, Fig 2D). Presence of clinical symptoms (n = 91, $P = 0.169$) and method of diagnosis (n = 91, $P = 0.999$) did not impact mean MOI. Mean MOI was higher in studies enrolling only children (n = 77, mean MOI = 2.26) compared to all ages (n = 6, mean MOI = 1.54, $P = 0.002$). In multivariable analysis, the same predictors as for the

Table 3. Multivariable predictors of A) *P. vivax* proportion of polyclonal infections, and B) mean multiple clone infections. 95%CI = 95% confidence interval.

A) *P. vivax* proportion of polyclonal infections				
	Full model		Reduced model	
	Coef [95%CI]	P	Coef [95%CI]	P
Prevalence	0.476 [-0.004, 0.955]	0.052	0.479 [0.037, 0.921]	0.034
Year	-1.375 [-2.483, -0.266]	0.016	-1.274 [-2.345, -0.202]	0.02
Sizing by CE	6.301 [-10.692, 23.294]	0.463		
>1 genotyping marker	29.408 [0.018, 58.798]	0.05	35.421 [10.294, 60.548]	0.006
PCR for diagnosis	3.81 [-6.633, 14.253]	0.47		
Geographic Region				
Africa	[reference]		[reference]	
Asia-Pacific	10.301 [-15.537, 36.139]	0.0135	9.584 [-15.463, 34.631]	0.0141
Latin America	-7.299 [-33.845, 19.248]		-7.45 [-33.437, 18.537]	
Age group				
All	[reference]		[reference]	
<15	1.845 [-16.029, 19.719]		2.658 [-14.916, 20.231]	0.0134
>15	-30.677 [-50.749, -10.606]	0.0109	-29.072 [-48.705, -9.439]	
Presence of clinical symptoms	5.521 [-6.186, 17.228]	0.351		
B) *P. vivax* mean multiplicity of infections				
	Full model		Reduced model	
	Coef [95%CI]	P>t	Coef [95%CI]	P>t
Prevalence	0.019 [0.009, 0.029]	<0.001	0.019 [0.009, 0.029]	<0.001
Year	-0.04 [-0.061, -0.019]	<0.001	-0.039 [-0.059, -0.019]	<0.001
Sizing by CE	-0.045 [-0.39, 0.301]	0.798		
>1 genotyping marker	0.434 [0.004, 0.865]	0.048	0.403 [0.096, 0.711]	0.011
PCR for diagnosis	-0.027 [-0.216, 0.162]	0.775		
Geographic Region				
Africa	[reference]		[reference]	
Asia-Pacific	0.165 [-0.203, 0.534]	0.0108	0.173 [-0.182, 0.527]	0.0068
Latin America	-0.238 [-0.653, 0.177]		-0.225 [-0.62, 0.17]	
Age group				
All	[reference]		[reference]	
<15	0.193 [-0.171, 0.556]	0.0167	0.193 [-0.162, 0.548]	0.0136
>15	-0.482 [-0.85, -0.115]		-0.477 [-0.832, -0.122]	
Presence of clinical symptoms	-0.018 [-0.239, 0.203]	0.869		

https://doi.org/10.1371/journal.pone.0249382.t003

proportion of polyclonal infections were significant, namely prevalence, year of sample collection, typing multiple markers, geographic region, and age group (Table 3B).

Discussion

MOI and the proportion of polyclonal infections are among the most frequently reported measures of malaria genomic epidemiology studies and have been proposed as surrogate markers for transmission intensity. Analysis of data from a large number of studies showed that polyclonal infections were frequent in almost all studied populations, even in low transmission settings. A weak correlation between multiplicity indices and prevalence was observed. For both species, once prevalence reached 40%, the majority of infections were polyclonal in almost all studies. At lower prevalence, some studies found a low proportion of polyclonal infections, while in others it was very high.

Multiplicity indices were lower in *P. falciparum* infections diagnosed by PCR instead of microscopy or RDT. Detectability of clones decreases as parasite densities decrease [30]. While parasite densities were not reported by most studies and thus not analyzed, the lower multiplicity when PCR was used for diagnosis can likely be explained by lower mean parasite densities in these infections. However, despite generally higher mean parasite densities in clinical cases compared to asymptomatic individuals, no difference in the proportion of polyclonal infections was observed between the two groups. Asymptomatic infections can persist for many months. Superinfections during this time can result in high multiplicity [3, 34]. A previous study comparing asymptomatic and clinical infections in West Papua from the same site indeed found higher multiplicity in asymptomatic individuals [24]. Possibly, in the data set analyzed here, higher detectability of clones in high-density clinical infections was balanced by higher multiplicity in asymptomatic individuals, resulting in overall similar estimates of the proportion of polyclonal infections in clinical and asymptomatic infections.

Estimates of multiplicity are heavily impacted by the ability to detect minority clones and by the discrimination power of the genotyping method selected. Size polymorphic markers were the preferred method for both species, but pronounced differences in the methods used for *P. falciparum* and *P. vivax* were observed. For *P. falciparum*, sizing of polymorphic markers on gel remains the primary method for genotyping. 79% of all studies sized PCR amplicons on gel, with little change in more recent studies (74% in studies conducted since 2015). For *P. vivax*, 76% of studies used CE for sizing. The choice of gel vs. CE can affect multiplicity estimates for two reasons. First, capillary electrophoresis increases detectability of minority clones. Second, it allows distinction of fragments of similar size compared to visualization on gel. By CE, fragments differing as little as 3 base pairs in size can be distinguished. In contrast, when binning using gels, resolution might be as low as 50 base pairs [35]. Previous studies comparing the two methods directly showed that CE yields higher multiplicity [31, 32]. In the current study, however, in univariate analysis typing by CE resulted in a lower proportion of *P. falciparum* polyclonal infections, and the method was not a significant predictor in multivariable analysis. The result from the univariate analysis could be explained by the more widespread use of CE in the Asia-Pacific compared to Africa, where transmission and multiplicity are generally lower than in Africa.

For *P. falciparum* genotyping, 49% of studies typed a single marker. In multivariable analysis, typing multiple *P. falciparum* markers increased the proportion of polyclonal infections detected by 14%. An even stronger increase of 35% was observed for *P. vivax*, where only 9% of studies typed a single marker.

Pronounced differences were observed among geographic regions. For *P. falciparum*, the proportion of polyclonal infections was highest in Africa, and for *P. vivax*, in the Asia-Pacific. In both cases, this mirrors high transmission intensity.

Marker *msp2* has been widely used for *P. falciparum* genotyping, allowing to repeat the analysis on this marker only. Results mirrored those of the entire dataset. The relationship between the proportion of polyclonal infections and prevalence remained poor. Despite its wide use, inclusion criteria for minority clones when typing *msp2* differ among laboratories, and are often not reported. When using CE, the criteria applied include an arbitrary cut off of 33% of the height of the main peak [6], a cut off based on the signal intensity of the size standard [36], or inclusion of all peaks above a set fluorescent signal threshold [17]. This variation in criteria to include minority clones likely adds to the weak relationship between multiplicity estimates and transmission intensity.

In many studies, in particular those typing samples from clinical patients, no data on transmission intensity was given, such as population prevalence, clinical incidence, or test positive rate. In order to conduct more detailed analysis of the relationship between multiplicity and

transmission intensity, it is recommended to report such measures along with genotyping data.

In conclusion, the weak correlation between mean MOI or the proportion of polyclonal infections and prevalence indicates these measures do not accurately stratify transmission intensity when different studies are combined. Further studies will be required to understand the factors affecting the presence of polyclonal infections when prevalence is low. The possibility of a high proportion of polyclonal infections even in low transmission sites should be considered when typing markers of drug resistance or *hrp2* deletion. In polyclonal infections the presence of deletions can be masked by wild-type infections [21], and the frequency of mutations conferring drug resistance observed by genotyping might not accurately reflect the prevalence among all clones [19].

Supporting information

S1 File. Data analyzed for this study.
(XLSX)

S1 Fig. PRISMA 2009 flow diagram.
(DOC)

Author Contributions

Conceptualization: Luis Lopez, Cristian Koepfli.

Data curation: Luis Lopez, Cristian Koepfli.

Formal analysis: Cristian Koepfli.

Funding acquisition: Cristian Koepfli.

Investigation: Luis Lopez, Cristian Koepfli.

Project administration: Cristian Koepfli.

Supervision: Cristian Koepfli.

Validation: Cristian Koepfli.

Visualization: Cristian Koepfli.

Writing – original draft: Luis Lopez, Cristian Koepfli.

Writing – review & editing: Cristian Koepfli.

References

1. World Health Organisation. World Malaria Report 2020. 2020.

2. Tusting LS, Bousema T, Smith DL, Drakeley C. Measuring changes in Plasmodium falciparum transmission: precision, accuracy and costs of metrics. Adv Parasitol. 2014; 84:151–208. Epub 2014/02/01. https://doi.org/10.1016/B978-0-12-800099-1.00003-X PMID: 24480314; PubMed Central PMCID: PMC4847140.

3. Nkhoma SC, Banda RL, Khoswe S, Dzoole-Mwale TJ, Ward SA. Intra-host dynamics of co-infecting parasite genotypes in asymptomatic malaria patients. Infect Genet Evol. 2018; 65:414–24. Epub 2018/08/27. https://doi.org/10.1016/j.meegid.2018.08.018 PMID: 30145390; PubMed Central PMCID: PMC6219893.

4. Koepfli C, Mueller I. Malaria Epidemiology at the Clone Level. Trends Parasitol. 2017; 33(12):974–85. Epub 2017/10/03. https://doi.org/10.1016/j.pt.2017.08.013 PMID: 28966050; PubMed Central PMCID: PMC5777529.

5. Volkman SK, Neafsey DE, Schaffner SF, Park DJ, Wirth DF. Harnessing genomics and genome biology to understand malaria biology. Nat Rev Genet. 2012; 13(5):315–28. Epub 2012/04/13. https://doi.org/10.1038/nrg3187 PMID: 22495435.

6. Anderson TJ, Haubold B, Williams JT, Estrada-Franco JG, Richardson L, Mollinedo R, et al. Microsatellite markers reveal a spectrum of population structures in the malaria parasite Plasmodium falciparum. Mol Biol Evol. 2000; 17(10):1467–82. Epub 2000/10/06. https://doi.org/10.1093/oxfordjournals.molbev.a026247 PMID: 11018154.

7. Karunaweera ND, Ferreira MU, Hartl DL, Wirth DF. Fourteen polymorphic microsatellite DNA markers for the human malaria parasite Plasmodium vivax. Molecular Ecology Notes. 2006; 7(1):172–5. https://doi.org/10.1111/j.1471-8286.2006.01534.x

8. Daniels R, Volkman SK, Milner DA, Mahesh N, Neafse–y DE, Park DJ, et al. A general SNP-based molecular barcode for Plasmodium falciparum identification and tracking. Malar J. 2008; 7:223. Epub 2008/10/31. https://doi.org/10.1186/1475-2875-7-223 PMID: 18959790; PubMed Central PMCID: PMC2584654.

9. Lerch A, Koepfli C, Hofmann NE, Messerli C, Wilcox S, Kattenberg JH, et al. Development of amplicon deep sequencing markers and data analysis pipeline for genotyping multi-clonal malaria infections. BMC Genomics. 2017; 18(1):864. Epub 2017/11/15. https://doi.org/10.1186/s12864-017-4260-y PMID: 29132317; PubMed Central PMCID: PMC5682641.

10. Lin JT, Hathaway NJ, Saunders DL, Lon C, Balasubramanian S, Kharabora O, et al. Using Amplicon Deep Sequencing to Detect Genetic Signatures of Plasmodium vivax Relapse. J Infect Dis. 2015; 212 (6):999–1008. Epub 2015/03/10. https://doi.org/10.1093/infdis/jiv142 PMID: 25748326; PubMed Central PMCID: PMC4548463.

11. Pava Z, Puspitasari AM, Rumaseb A, Handayuni I, Trianty L, Utami RAS, et al. Molecular surveillance over 14 years confirms reduction of Plasmodium vivax and falciparum transmission after implementation of Artemisinin-based combination therapy in Papua, Indonesia. PLoS Negl Trop Dis. 2020; 14(5): e0008295. Epub 2020/05/08. https://doi.org/10.1371/journal.pntd.0008295 PMID: 32379762; PubMed Central PMCID: PMC7237043.

12. 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. Epub 2012/11/06. https://doi.org/10.1111/mec.12099 PMID: 23121253; PubMed Central PMCID: PMC3537863.

13. Bei AK, Niang M, Deme AB, Daniels RF, Sarr FD, Sokhna C, et al. Dramatic Changes in Malaria Population Genetic Complexity in Dielmo and Ndiop, Senegal, Revealed Using Genomic Surveillance. J Infect Dis. 2018; 217(4):622–7. Epub 2018/01/13. https://doi.org/10.1093/infdis/jix580 PMID: 29325146; PubMed Central PMCID: PMC6279132.

14. Pacheco MA, Forero-Pena DA, Schneider KA, Chavero M, Gamardo A, Figuera L, et al. Malaria in Venezuela: changes in the complexity of infection reflects the increment in transmission intensity. Malar J. 2020; 19(1):176. Epub 2020/05/10. https://doi.org/10.1186/s12936-020-03247-z PMID: 32380999; PubMed Central PMCID: PMC7206825.

15. Sondo P, Derra K, Rouamba T, Nakanabo Diallo S, Taconet P, Kazienga A, et al. Determinants of Plasmodium falciparum multiplicity of infection and genetic diversity in Burkina Faso. Parasit Vectors. 2020; 13(1):427. Epub 2020/08/21. https://doi.org/10.1186/s13071-020-04302-z PMID: 32819420; PubMed Central PMCID: PMC7441709.

16. Fola AA, Harrison GLA, Hazairin MH, Barnadas C, Hetzel MW, Iga J, et al. Higher Complexity of Infection and Genetic Diversity of Plasmodium vivax Than Plasmodium falciparum Across All Malaria Transmission Zones of Papua New Guinea. Am J Trop Med Hyg. 2017; 96(3):630–41. Epub 2017/01/11. https://doi.org/10.4269/ajtmh.16-0716 PMID: 28070005; PubMed Central PMCID: PMC5361537.

17. Koepfli C, Waltmann A, Ome-Kaius M, Robinson LJ, Mueller I. Multiplicity of Infection Is a Poor Predictor of Village-Level Plasmodium vivax and P. falciparum Population Prevalence in the Southwest Pacific. Open Forum Infect Dis. 2018; 5(11):ofy240. Epub 2018/11/07. https://doi.org/10.1093/ofid/ofy240 PMID: 30397622; PubMed Central PMCID: PMC6210381.

18. Cowell AN, Winzeler EA. The genomic architecture of antimalarial drug resistance. Brief Funct Genomics. 2019; 18(5):314–28. Epub 2019/05/24. https://doi.org/10.1093/bfgp/elz008 PMID: 31119263; PubMed Central PMCID: PMC6859814.

19. Hastings IM, Nsanzabana C, Smith TA. A comparison of methods to detect and quantify the markers of antimalarial drug resistance. Am J Trop Med Hyg. 2010; 83(3):489–95. Epub 2010/09/03. https://doi.org/10.4269/ajtmh.2010.10-0072 PMID: 20810808; PubMed Central PMCID: PMC2929039.

20. Cheng Q, Gatton ML, Barnwell J, Chiodini P, McCarthy J, Bell D, et al. Plasmodium falciparum parasites lacking histidine-rich protein 2 and 3: a review and recommendations for accurate reporting. Malar J. 2014; 13:283. Epub 2014/07/24. https://doi.org/10.1186/1475-2875-13-283 PMID: 25052298; PubMed Central PMCID: PMC4115471.

21. Watson OJ, Verity R, Ghani AC, Garske T, Cunningham J, Tshefu A, et al. Impact of seasonal variations in Plasmodium falciparum malaria transmission on the surveillance of pfhrp2 gene deletions. Elife. 2019;8. Epub 2019/05/03. https://doi.org/10.7554/eLife.40339 PMID: 31045490; PubMed Central PMCID: PMC6533063.

22. Zhu SJ, Almagro-Garcia J, McVean G. Deconvolution of multiple infections in Plasmodium falciparum from high throughput sequencing data. Bioinformatics. 2018; 34(1):9–15. Epub 2017/09/30. https://doi.org/10.1093/bioinformatics/btx530 PMID: 28961721; PubMed Central PMCID: PMC5870807.

23. Eldh M, Hammar U, Arnot D, Beck HP, Garcia A, Liljander A, et al. Multiplicity of Asymptomatic Plasmodium falciparum Infections and Risk of Clinical Malaria: A Systematic Review and Pooled Analysis of Individual Participant Data. J Infect Dis. 2020; 221(5):775–85. Epub 2019/10/05. https://doi.org/10.1093/infdis/jiz510 PMID: 31585009; PubMed Central PMCID: PMC7026891.

24. Pava Z, Handayuni I, Trianty L, Utami RAS, Tirta YK, Puspitasari AM, et al. Passively versus Actively Detected Malaria: Similar Genetic Diversity but Different Complexity of Infection. Am J Trop Med Hyg. 2017; 97(6):1788–96. Epub 2017/10/11. https://doi.org/10.4269/ajtmh.17-0364 PMID: 29016343; PubMed Central PMCID: PMC5790166.

25. Mohammed H, Hassen K, Assefa A, Mekete K, Tadesse G, Taye G, et al. Genetic diversity of Plasmodium falciparum isolates from patients with uncomplicated and severe malaria based on msp-1 and msp-2 genes in Gublak, North West Ethiopia. Malar J. 2019; 18(1):413. Epub 2019/12/12. https://doi.org/10.1186/s12936-019-3039-9 PMID: 31823778; PubMed Central PMCID: PMC6905089.

26. Imwong M, Nair S, Pukrittayakamee S, Sudimack D, Williams JT, Mayxay M, et al. Contrasting genetic structure in Plasmodium vivax populations from Asia and South America. Int J Parasitol. 2007; 37(8–9):1013–22. Epub 2007/04/20. https://doi.org/10.1016/j.ijpara.2007.02.010 PMID: 17442318.

27. Koepfli C, Mueller I, Marfurt J, Goroti M, Sie A, Oa O, et al. Evaluation of Plasmodium vivax genotyping markers for molecular monitoring in clinical trials. J Infect Dis. 2009; 199(7):1074–80. Epub 2009/03/12. https://doi.org/10.1086/597303 PMID: 19275476.

28. Okell LC, Bousema T, Griffin JT, Ouedraogo AL, Ghani AC, Drakeley CJ. Factors determining the occurrence of submicroscopic malaria infections and their relevance for control. Nat Commun. 2012; 3:1237. Epub 2012/12/06. https://doi.org/10.1038/ncomms2241 PMID: 23212366; PubMed Central PMCID: PMC3535331.

29. Cheng Q, Cunningham J, Gatton ML. Systematic review of sub-microscopic P. vivax infections: prevalence and determining factors. PLoS Negl Trop Dis. 2015; 9(1):e3413. Epub 2015/01/09. https://doi.org/10.1371/journal.pntd.0003413 PMID: 25569135; PubMed Central PMCID: PMC4288718.

30. Koepfli C, Schoepflin S, Bretscher M, Lin E, Kiniboro B, Zimmerman PA, et al. How much remains undetected? Probability of molecular detection of human Plasmodia in the field. PLoS One. 2011; 6(4):e19010. Epub 2011/05/10. https://doi.org/10.1371/journal.pone.0019010 PMID: 21552561; PubMed Central PMCID: PMC3084249.

31. Falk N, Maire N, Sama W, Owusu-Agyei S, Smith T, Beck HP, et al. Comparison of PCR-RFLP and Genescan-based genotyping for analyzing infection dynamics of Plasmodium falciparum. Am J Trop Med Hyg. 2006; 74(6):944–50. Epub 2006/06/09. PMID: 16760501.

32. Gupta V, Dorsey G, Hubbard AE, Rosenthal PJ, Greenhouse B. Gel versus capillary electrophoresis genotyping for categorizing treatment outcomes in two anti-malarial trials in Uganda. Malar J. 2010; 9:19. Epub 2010/01/16. https://doi.org/10.1186/1475-2875-9-19 PMID: 20074380; PubMed Central PMCID: PMC2817701.

33. Smith T, Beck HP, Kitua A, Mwankusye S, Felger I, Fraser-Hurt N, et al. Age dependence of the multiplicity of Plasmodium falciparum infections and of other malariological indices in an area of high endemicity. Trans R Soc Trop Med Hyg. 1999; 93 Suppl 1:15–20. Epub 1999/08/18. https://doi.org/10.1016/s0035-9203(99)90163-3 PMID: 10492779.

34. Lerch A, Koepfli C, Hofmann NE, Kattenberg JH, Rosanas-Urgell A, Betuela I, et al. Longitudinal tracking and quantification of individual Plasmodium falciparum clones in complex infections. Sci Rep. 2019; 9(1):3333. Epub 2019/03/06. https://doi.org/10.1038/s41598-019-39656-7 PMID: 30833657; PubMed Central PMCID: PMC6399284.

35. Nguetse CN, Ojo JA, Nchotebah C, Ikegbunam MN, Meyer CG, Thomas BN, et al. Genetic Diversity of the Plasmodium falciparum Glutamate-Rich Protein R2 Region Before and Twelve Years after Introduction of Artemisinin Combination Therapies among Febrile Children in Nigeria. American Journal of Tropical Medicine and Hygiene. 2018; 98(3):667–76. https://doi.org/10.4269/ajtmh.17-0621 WOS:000430952800008. PMID: 29363449

36. Mueller I, Schoepflin S, Smith TA, Benton KL, Bretscher MT, Lin E, et al. Force of infection is key to understanding the epidemiology of Plasmodium falciparum malaria in Papua New Guinean children. Proc Natl Acad Sci U S A. 2012; 109(25):10030–5. Epub 2012/06/06. https://doi.org/10.1073/pnas.1200841109 PMID: 22665809; PubMed Central PMCID: PMC3382533.