Article

Long-term sustained malaria control leads to inbreeding and fragmentation of *Plasmodium vivax* populations

Andreea Waltmann[1,2], Cristian Koepfli[1,2], Natacha Tessier, [1,2] Stephan Karl[1,2], Andrew W Darcy[3], Lyndes Wini[4], G.L. Abby Harrison[1,2], Céline Barnadas[1,2], Charlie Jennison[1,2], Harin Karunajeewa[1,2], Sarah Boyd[1], Maxine Whittaker[5], James Kazura[6], Melanie Bahlo[1, 2], Ivo Mueller[1,2,7*] and Alyssa E. Barry[1,2*]

1. Division of Population Health and Immunity, The Walter & Eliza Hall Institute of Medical Research, Melbourne, Australia
2. Department of Medical Biology, University of Melbourne, Melbourne, Australia
3. The National Health Training and Research Institute, Ministry of Health, Solomon Islands
4. National Vector Borne Disease Control Program, Ministry of Health, Solomon Islands
5. School of Population Health, University of Queensland, Brisbane, Australia
6. Center for Global Health and Diseases, Case Western Reserve University, Cleveland, Ohio, United States
7. Center de Recerca en Salut Internacional de Barcelona, Barcelona, Spain

*Corresponding authors:
barry@wehi.edu.au
mueller@wehi.edu.au

Abstract

Plasmodium vivax populations are more resistant to malaria control strategies than *Plasmodium falciparum*, maintaining high genetic diversity and gene flow even at low transmission. To quantify the impact of declining transmission on *P. vivax* populations, we investigated population genetic structure over time during intensified control efforts and over a wide range of transmission intensities and spatial scales in the Southwest Pacific. Analysis of 887 *P. vivax* microsatellite haplotypes (Papua New Guinea, PNG = 443, Solomon Islands = 420, Vanuatu =24) revealed substantial population structure among countries and modestly declining diversity as transmission decreases over space and time. In the Solomon Islands, which has had sustained control efforts for 20 years, significant population structure was observed on different spatial scales down to the sub-village level. Up to 37% of alleles were partitioned between populations and significant multilocus linkage disequilibrium was observed indicating substantial inbreeding. High levels of haplotype relatedness around households and within a range of 300m are consistent with a focal and clustered infections suggesting that restricted local transmission occurs within the range of vector movement and that subsequent focal inbreeding may be a key factor contributing to the observed population structure. We conclude that unique transmission strategies, including relapse allows *P. vivax* populations to withstand pressure from control efforts for longer than *P. falciparum.* However sustained control efforts do eventually impact parasite population structure and with further control pressure, populations may eventually fragment into clustered foci that could be targeted for elimination.

Introduction

Spatial clustering of infectious diseases is a well-known phenomenon in which micro-epidemiological variations in exposure due to factors controlling disease transmission result in some individuals in the community being disproportionately infected (Cattani, et al. 1986; Bousema, et al. 2012; Cotter, et al. 2013). Malaria is one disease in which spatial clustering of transmission has been frequently reported, with heterogeneity becoming more pronounced as transmission decreases (Woolhouse, et al. 1997; Bousema, et al. 2012). Concerted international efforts over the last 15 years, have reduced the global malaria burden by more than 50% with rapidly declining transmission in many endemic regions (WHO 2015c). As countries aim for elimination, measuring the impact of control efforts and mapping transmission foci will provide data that can guide when to switch from broad ranging to targeted control efforts, and will help to prioritise regions for elimination.

Plasmodium falciparum and *Plasmodium vivax* are the major agents of human malaria, however as malaria transmission declines in co-endemic areas, *P. vivax* becomes the main source of malaria infection and disease because it is more refractory to control efforts (Harris, et al. 2010; Kaneko 2010; Oliveira-Ferreira, et al. 2010; Rodriguez, et al. 2011; Organisation 2013; Kaneko, et al. 2014; Noviyanti, et al. 2015; Waltmann, et al. 2015; WHO 2015a, c). *P. vivax* employs unique life-history strategies including dormant liver-stage infections (hypnozoites), the early development of transmissible forms (gametocytes) and the lower density (and thus detectability) of infections which probably underlie control-driven shifts in species dominance and suggest that *P. vivax* will be the far more challenging species to eliminate (Feachem, et al. 2010; Bousema and Drakeley 2011; White and Imwong 2012; Mueller, et al. 2013). Transmission-reducing interventions originally developed for African *P. falciparum* malaria, may not be sufficient or suitable against *P. vivax* and therefore novel strategies to confront the unique challenges posed by *P. vivax* may need to be developed (Mendis, et al. 2001; Alonso, et al. 2011; Alonso and Tanner 2013; Cotter, et al. 2013; WHO 2015b, a).

Parasite population genetics has been successfully harnessed to understand changes in *P. falciparum* populations in response to sustained control (Daniels, et al. 2015), but this has not yet been applied extensively to *P. vivax* control. As infections reduce in number and transmission becomes more focal, it is expected that effective population size, genetic diversity, gene flow and outcrossing will decrease, eventually leading to inbred, structured populations (Anderson, Haubold, et al. 2000; Markert, et al. 2010; Bousema, et al. 2012). Conversely, in areas of high transmission, high levels of diversity, gene

flow and outcrossing are common (Anderson, Haubold, et al. 2000), resulting in admixed and unstructured populations (Anderson, Haubold, et al. 2000). Whilst *P. falciparum* fits this expectation (Anderson, Haubold, et al. 2000; Markert, et al. 2010; Bousema, et al. 2012), *P. vivax* populations exhibit high levels of diversity and large effective population sizes irrespective of transmission (Van den Eede, et al. 2010; Chenet, et al. 2012; Gray, et al. 2013; Gunawardena, et al. 2014; Barry, et al. 2015). Strong structuring of *P. vivax* populations has been observed among continents indicating long periods of isolation (Koepfli, et al. 2015; Hupalo, et al. 2016; Pearson, et al. 2016), but at regional and local scales sub-structure has been reported only for some areas (Imwong, et al. 2007; Van den Eede, et al. 2010; Abdullah, et al. 2013; Delgado-Ratto, et al. 2016), but not others (Koepfli, et al. 2013; Jennison, et al. 2015; Noviyanti, et al. 2015) with little apparent relationship to transmission intensity. In co-endemic areas where *P. vivax* prevalence is comparable to, or lower than that of *P. falciparum*, *P. vivax* exhibits large panmictic populations (Orjuela-Sanchez, et al. 2013; Jennison, et al. 2015). Regions where *P. vivax* population structure has been observed, such as Peru (Van den Eede, et al. 2010), Colombia (Imwong, et al. 2007) or Malaysia (Abdullah, et al. 2013) tend to have had multiple introductions (Taylor, et al. 2013), historically low *P. vivax* transmission (Van den Eede, et al. 2010,; Delgado-Ratto, et al. 2016), non-overlapping vector species refractory to non-autochthonous strains (Pimenta, et al. 2015) or historically focal transmission combined with recent reductions due to control (Abdullah, et al. 2013). In regions with past hyperendemic *P. vivax* transmission and recent upscaling of malaria control efforts, population structure has not been observed (Noviyanti, et al. 2015).

In comparison to *P. falciparum*, the lack of local population structure of *P. vivax* is consistent with more stable transmission over a long period of time and/or deeper evolutionary roots (Neafsey, et al. 2012) and also reflects the contribution of relapse and multiple infections to outcrossing and gene flow (Jennison 2015). Relapses account for up to 80% of *P. vivax* blood stage infections in highly endemic areas (Robinson, et al. 2015), undoubtedly playing a central role in shaping the complex genetic structure of *P. vivax* populations. At lower prevalence, significant multilocus linkage disequilibrium (LD) in the context of high diversity suggests that *P. vivax* may increasingly undergo clonal transmission and inbreeding as diverse strains in the hypnozoite reservoir are depleted (Chenet, et al. 2012), leading to increasing population structure (Abdullah, et al. 2013). Changes in the *P. vivax* population structure within declining transmission and in the context of long-term intensified control have not been investigated.

Malaria control programs need to measure the effectiveness of control efforts, determine whether their interventions are having an impact and how much control pressure is needed and for how long. For *P. vivax*, current approaches, mostly confined to prevalence surveys and monitoring clinical cases, lack the resolution to discern underlying population processes. Population genetics however, is a powerful approach to determine whether populations are under stress (Markert, et al. 2010). Before these molecular approaches can be effectively utilized, it will be important to understand how declining transmission affects *P. vivax* population structure. Furthermore, in order to stratify interventions for maximum impact, malaria control programs need to know the spatial scales that characterize *P. vivax* populations (Bousema, et al. 2012). Here we define *P. vivax* transmission patterns by measuring population genetic structure at different transmission intensities, spatial scales and in the context of successful long-term malaria control. We analysed almost 900 *P. vivax* microsatellite haplotypes from isolates collected throughout the Southwest Pacific region, which has a natural, gradual decline in malaria endemicity from west to east (high transmission in PNG, moderate-to-high in Solomon Islands and low in Vanuatu), that has been accentuated by recent control efforts. Included were dense spatial and temporal data from areas of residual *P. vivax* transmission in the Solomon Islands (Waltmann, et al. 2015), where over the last two decades malaria incidence has been reduced by approximately 90% (Program 2013). The results suggest that long-term sustained control will eventually impact *P. vivax* populations, highlighting the importance of maintaining control efforts, and the key role that population genetic surveillance can play in malaria control and elimination.

Results

Wide range of *Plasmodium vivax* transmission intensities across the study area

Genotyping of all available *P. vivax* infections using the highly polymorphic markers *MS16* and *msp1*F3 was first done to determine the multiplicity of infection (MOI) and calculate the proportion of polyclonal infections as a surrogate measure of transmission intensity (Nkhoma, et al. 2013). The greatest proportion of polyclonal infections was seen in regions with high *P. vivax* prevalence in PNG and in the Solomon Islands population of Tetere 2004-5 at 72.4% (42/58), consistent with the high level of malaria transmission at the time of sample collection (Figure 1, (Koepfli, et al. 2013; Jennison, et al. 2015)). In Vanuatu, among 30 *P. vivax* isolates in the original study (Boyd et al., unpublished), 25 were successfully genotyped and of these only 10.0% (3/25) were polyclonal and the mean MOI was 1.13 (range 1-2) (Figure 1D).

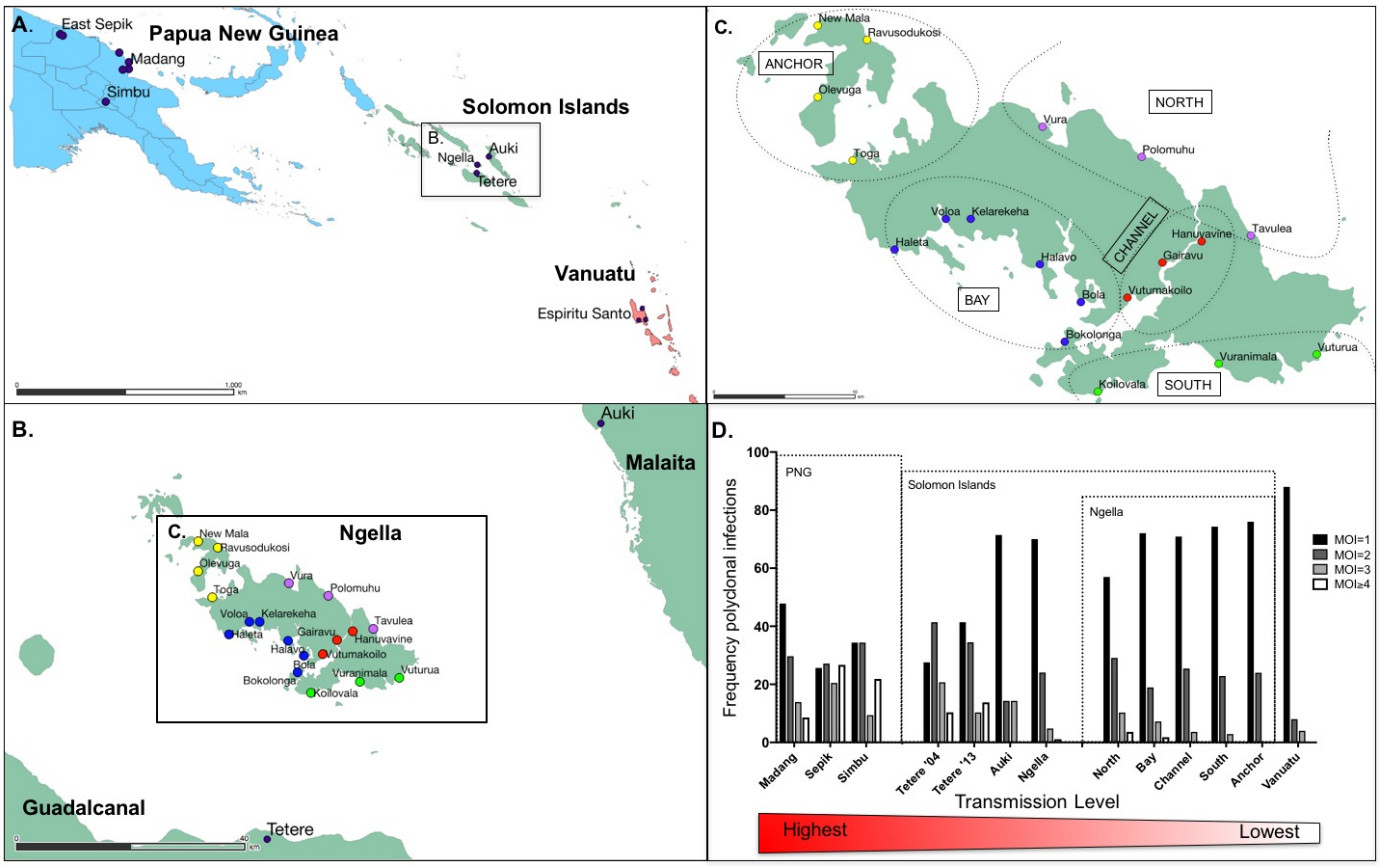

Figure 1. **Map of the study areas and transmission intensity**
(A) Southwest Pacific sampling locations showing Papua New Guinea in blue, Solomon Islands in green and Vanuatu in red. (B) Central Solomon Islands. (C) Ngella, including 19 villages from five distinct geographical / ecological regions. Anchor villages are indicated in yellow, Bay in blue, South Coast in green, Channel in red and North Coast in purple). (D) The frequency of monoclonal and polyclonal infections is shown for each sampling location, as an indicator of transmission intensity.

Within Solomon Islands, variation in the proportion of polyclonal infections was observed over time and space (Figure 1D). By 2013, Tetere had a lower proportion that in 2004-5 at 57.1% (32/56) indicating lower transmission than in 2004-5, but remaining at moderately high levels similar to the Madang Province of PNG. The mean MOI was 1.73 (range 1-5). In the other Solomon Islands populations of Auki and Ngella, the majority of infections were monoclonal consistent with much lower transmission. Of 18 Auki *P. vivax* infections, only 27.8% (5/18) were polyclonal and mean MOI was 1.33 (range 1-4). Within Ngella, an area of dense sampling, the smallest proportion of polyclonal infections was found in Anchor (the zone with the lowest prevalence) and the greatest proportion was found on the North Coast (Figure 1D).

Definition of high quality microsatellite haplotypes

A total of 889 high quality haplotypes with data for at least five out of nine loci were available for analysis (Figure S1). These included 557 confirmed single (MOI =1) and 332 "dominant" haplotypes from samples with MOI=2, which comprised the dominant allele calls (highest peaks) for all markers. However, two haplotypes were identified as outliers (i.e. those that do not conform to the expected distribution), due to rare singleton alleles at the MS2 locus (one from PNG and one from Tetere 2004-5). These were discarded for subsequent analyses leaving a final dataset of 887 haplotypes (Table 1). No significant genetic differentiation was observed between haplotypes constructed from dominant alleles from multiple infections and those from confirmed single infections (Table S2) thus the haplotypes were combined for further analyses. The 887 haplotypes were distributed across all catchment areas. Smaller sample sizes were available for lower prevalence regions (Table 1, Table S1). The allele frequencies for each of the populations are summarized in Table S3.

Temporal changes in *Plasmodium vivax* population structure after sustained control

Although sustained control has been ongoing in the Solomon Islands since 1996, from 2003 with support from the Global Fund for combatting AIDS, Tuberculosis and Malaria, there have been several interventions in more recent years including indoor residual spraying (IRS) in 2006, introduction of artemisinin combination therapy (ACT) in 2008, and widespread distribution of long lasting insecticide treated bednets (LLIN) in 2009 (WHO 2015c). Data was available for two time points, 2004-5 and 2013 for Tetere, a village on the north coast of the main island of Guadalcanal. In Tetere 2013, diversity (H_E and R_S) was lower and effective population sizes were half that of the values observed for the 2004-5 population, consistent with a significant impact on the *P. vivax* population over that period (Table 1). Furthermore, in 2004-5 there were no identical haplotypes and no significant LD (Koepfli, et al. 2013) (Table 2), indicating high levels of outcrossing due to high transmission. By 2013, multilocus LD had increased to significant levels consistent with an increase in inbreeding (Table 2), and the populations from the two time points showed low but significant levels of genetic differentiation (Jost D=0.195, G_{ST}=0.021, F_{ST}=0.029, Figure 2, Table S4). The Tetere 2004-5 population was also genetically differentiated from other two Solomon Islands populations (Auki and Ngella), which included samples collected in 2012 and 2013, respectively (Figure 2). This clearly demonstrates increasing LD and population structure in the Tetere *P. vivax* population in the period from 2004-2013, most likely as a result of declining transmission due to intensified malaria control.

SOUTHWEST PACIFIC

	PNG	*Solomon Is.*	*Vanuatu*
PNG	0	0.216	0.408
Solomon Is.		0	0.417
n	443	420	24

SOLOMON ISLANDS

	Tetere 2004-5	*Tetere 2013*	*Bay*	*South*	*Channel*	*North*	*Anchor*	*Auki*
Tetere 2004-5	0	0.197	0.305	0.307	0.275	0.292	0.278	0.265
Tetere 2013		0	0.209	0.193	0.265	0.189	0.201	0.233
Bay			0	0.098	0.212	0.156	0.177	0.322
South				0	0.320	0.158	0.155	0.396
Channel					0	0.251	0.352	0.372
North						0	0.132	0.307
Anchor							0	0.323
n	45	39	83	35	46	136	23	13

NGELLA: NORTH

	Vura	*Polohomu*	*Tavulea*
Vura	0	0.184	0.236
Polohomu		0	0.192
n	58	46	33

NGELLA: CHANNEL

	Vutumakoilo	*Huanavavine*
Vutumakoilo	0	0.488
n	14	29

Figure 2. Estimates of genetic differentiation for *P. vivax* populations of the Southwest Pacific. Genetic differentiation values are shown for populations at different spatial scales, and are based on Jost's *D* (Jost 2008). Darker shading indicates higher values. Values for F_{ST} and G_{ST} are available in the supporting materials (Tables S4 and S5).

Spatially variable diversity and effective population sizes for *Plasmodium vivax* according to level of transmission

The study area included a wide range of transmission intensities and spatial scales (Figure 1). Diversity based on the mean expected heterozygosity was high for all populations with marginally lower values in the lowest transmission site of Espiritu Santo in Vanuatu (0.72) compared to the PNG and Solomon Islands populations (range 0.79 - 0.85, Table 1) demonstrating the ability of *P. vivax* populations to maintain high levels of diversity even at very low transmission. In Ngella, the lowest H_E was found in the Channel and Anchor populations (0.79) and the highest on the North Coast (0.85). The mean allelic richness (R_S) displays a similar pattern but broader range of values, with the lowest in Vanuatu (5.45 alleles/locus) and the highest in Solomon Islands and PNG, on the North Coast of Ngella (9.73 alleles/locus) and Madang Province (9.62 alleles/locus) respectively. Within Solomon Islands, Anchor, an area of Ngella, and the area with the lowest *P. vivax* prevalence, had the lowest mean number of estimated alleles/locus (6.51) (Table 1)

Effective population sizes (N_e) were also variable across the different parasite populations, with Vanuatu having the lowest values. Solomon Islands and PNG populations showed comparably

moderate to high N_e (Table 1) even though transmission in Solomon Islands was lower than that of PNG (Figure 1). Among the Ngella populations, Channel had the smallest effective population size. These relative patterns in N_e (effectively another measure of diversity) show that very low transmission is needed (e.g. Vanuatu) for N_e to be impacted substantially, particularly when the diversity of microsatellite markers are used for its calculation.

Evidence of Significant Inbreeding in *Plasmodium vivax* parasite populations of Solomon Islands and Vanuatu

In previously published data from PNG and Solomon Islands there were no identical haplotypes and no significant LD was observed (Koepfli, et al. 2013; Jennison, et al. 2015). In the new data from Solomon Islands (samples collected almost a decade later to the previous study) and Vanuatu, seven haplotypes were found repeatedly. All of these shared haplotypes were observed within Ngella, which may in part be due to lower transmission as well as the greater depth of sampling in this region. The seven repeated haplotypes were found in four Ngella sub-regions, not including Bay (Table S6, Figure S2). One haplotype was found in five isolates, one in four isolates, three were found in three isolates each and two in two isolates each. Three identical haplotypes were restricted to the same village, and the other four were shared among villages of the same or different region (Table S6, Figure S2). For those haplotypes shared among villages, at least one of the infected individuals reported travel other parts of Ngella or Guadalcanal.

Based on the complete microsatellite haplotypes (n=248), significant multilocus LD was observed for Solomon Islands and Vanuatu populations, with the exception of Tetere 2004-5 as mentioned above (Table 2) (Koepfli, et al. 2013). Strong LD was also observed within villages (Table 2). The pattern of strong LD was retained when only single infections were considered (n=93, Table 2) (any differences in I_A^S estimates or increases in *p*-values are due to the sample size reductions), as well as when only one locus per chromosome was analyzed to confirm that LD was not the result of physical linkage (Table S7). As previously reported, significant multilocus LD was not found in PNG, indicating high levels of outcrossing in that population (Jennison, et al. 2015).

Geographic Population Structure of *Plasmodium vivax* in the Southwest Pacific

To investigate population structure across the study area, genetic differentiation (Jost's *D*) was calculated and clusters defined from the haplotype data. Substantial population structure was observed

across the region with low to moderate G_{ST} (0.20-0.54) and F_{ST} values (0.038-0.085), and Jost's D values among countries indicating that between 21.6-41.7% of the alleles are private (Figure 2, Tables S4, S5). Higher values of genetic differentiation between parasite populations of different regions and villages in Solomon Islands were observed (Figure 2, Tables S4, S5), but not among regions or villages in PNG (Koepfli, et al. 2013; Jennison 2015), while Vanuatu samples were from only one region and thus within country structure could not be determined.

STRUCTURE analysis showed sub-structuring at various spatial scales down to the village level (Figures 3, S3). Analysis of all haplotypes showed that Southwest Pacific parasites optimally clustered into three genetically distinct populations associated with each of the countries (Figures 3, S3). Within Solomon Islands, further substructure was observed (K=4, Figures 2, 3). Moderate to high levels of genetic differentiation were observed with the highest values observed for all populations against Auki, albeit the small sample size (n=13) may have elevated these values (Figure 2). Tetere 2013 and Ngella infections were the least differentiated indicating greater gene flow between these two populations (Figure 2). STRUCTURE analyses confirmed this additional sub-population structure within Solomon Islands showing four genetically distinct groups of haplotypes with weak clustering between the Tetere, Ngella and Auki haplotypes (Figures 3, S3).

Within Ngella, genetic differentiation and clustering was observed amongst the five defined geographic regions and even among villages in the same region (Figure 2). Differentiation values were highest for the Channel population against all other populations, with the strongest differentiation as compared to the Anchor and South Coast regions (Figure 2). As sample sizes were substantial, we also compared three villages on the North Coast located approximately 6-15km apart (Vura n=58, Polomuhu n=46, and Tavulea n=33) and two villages in the Channel region (Hanuvavine n=29 and Vutumakoilo n=14) which are approximately 6km apart. Moderate genetic differentiation was observed among the North Coast villages and high levels between the two Channel villages (Jost's D=0.488) (Figure 2). In addition, clustering was evident among the five defined regions. STRUCTURE analysis containing only Ngella haplotypes gave an optimal K of 4 (Figures 3, S3). With this analysis, further clustering was also discernible within the Channel region, with the Hanuvavine and Vutumakoilo village isolates forming distinct clusters (Figures 3, S3) consistent with the genetic differentiation results (Figure 2). A separate analysis of the three North Coast villages (Vura, Polomuhu and Tavulea), which were between

6-15km apart also confirmed that clustering in Ngella is present even at small geographical scale (Figure 3, S3).

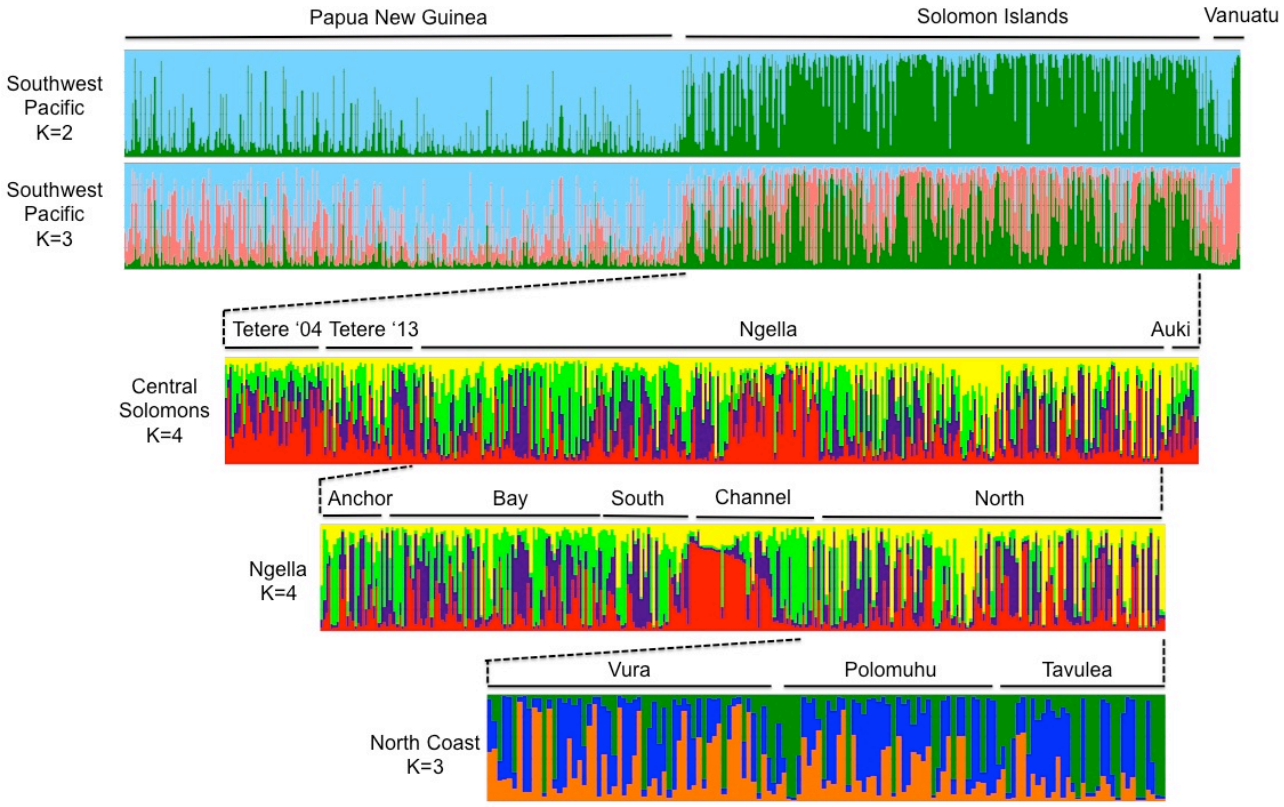

Figure 3. Clustering patterns of *P. vivax* microsatellite haplotypes. Results of STRUCTURE analysis are shown for different geographic strata including (A) Southwest Pacific, (B) Solomon Islands, (C) Ngella and (D) Ngella, North Coast. The analysis assigns *P. vivax* haplotypes to a defined number of genetic clusters (*K*) based on genetic distance. Vertical bars indicate individual *P. vivax* haplotype and colours represent the ancestry co-efficient (membership) within each cluster.

Fine-scale clustering of *Plasmodium vivax* infections in Solomon Islands

To investigate whether clustering of infections could be observed on a very fine scale (sub-village) we also investigated genetic relatedness of infections within and between households (Figure 4A). The analyses made use of two datasets of pairwise haplotype comparisons with only high quality haplotypes with at least six of the nine genetic loci successfully typed (Table S4). These included an intra-household comparison (n comparisons with ≥ 6 markers assessed = 164), where pairwise allele sharing was calculated only between each haplotypes of the same household and included 208 Ngella haplotypes from 86 households with ≥2 infections (Figures 4A and B); and an inter-household comparison (n comparisons=46,479; n haplotypes = 315; n households with ≥1 infection = 187), where pairwise allele sharing was calculated between all haplotypes.

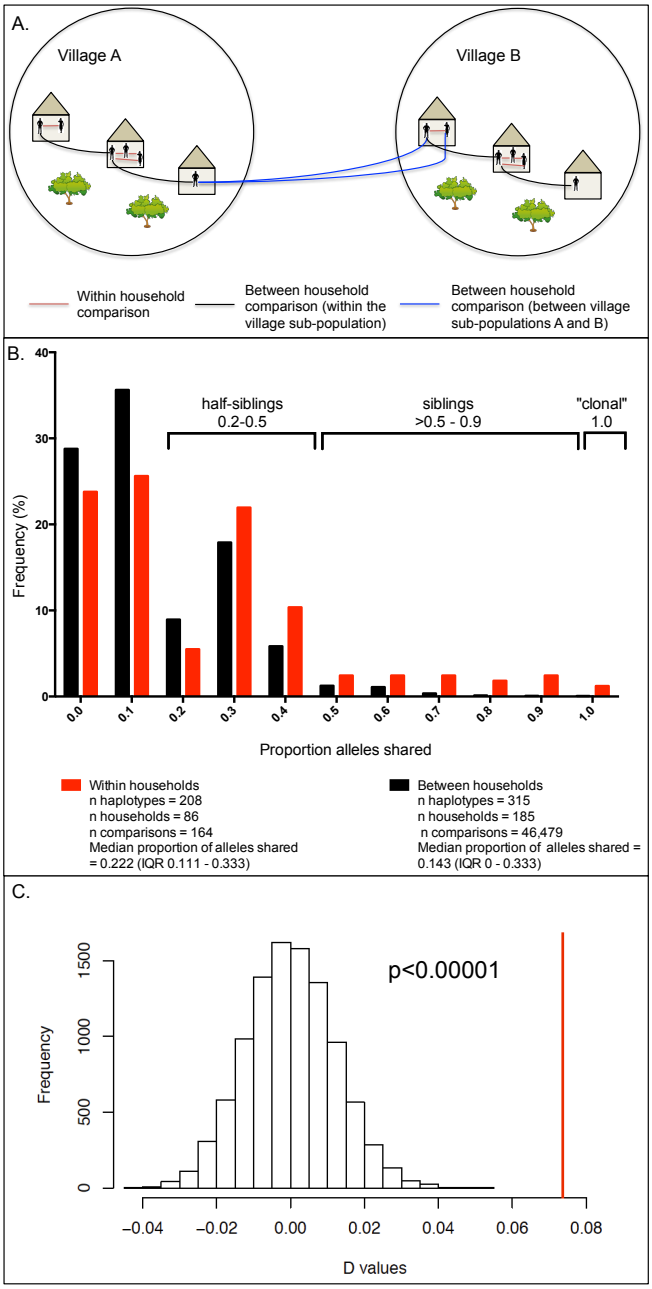

Figure 4. Fine scale genetic relatedness (allele sharing) of *P. vivax* haplotypes in Ngella, Solomon Islands. (A) Schematic of the types of pairwise allele sharing comparisons within and between households with *P. vivax* carriers of two theoretical villages, A and B. Intra-household comparisons, red lines; inter-household comparisons within a village, black lines; inter-household comparisons between villages, blue lines). (B) Proportion of allele sharing (P_S) within and between Ngella households. (C) Distribution of the *D* statistic (proportion alleles shared within households – proportion of alleles shared between households). A total of 10,000 permutations were used. The observed D value (0.074) is shown in red. Under the null hypothesis, the 10,000 *D* values permuted never reach the observed *D* value. Hence, the distribution of the proportion of alleles shared within household compared to that between households is statistically different (p<0.00001).

Comparisons of allele sharing (P_S) showed that clonal pairs, which had identical alleles at all markers, were found within households (2/164, 1.2%) more frequently than between households (27/46479, 0.06%, p<0.001) (Figure 4B, S5). There was also a greater proportion of full siblings, sharing 51-90% of their alleles, within household pairs (19/164, 11.6%) compared to that between households (1338/46479, 2.88%, p<0.001) (Figure 4B). Half-siblings, which share 20-50% of their alleles, were also more common within households (62/164, 37.8%) compared to inter-household comparisons but the difference did not reach statistical significance (15177/46479, 32.7%, p=0.160) (Figure 3B). Overall, there was 7.4% more allele sharing within-households (*D*=0.074), which is well beyond the

range of normal variation, as none of the *D* values of the 10,000 permutations reached the observed *D* ($p<0.00001$, Figure 4C). The assessment of P_S at individual microsatellite loci was also consistent with a strong gradient of spatial variation in relatedness with a mean P_S across the loci of 0.34 within households compared to 0.29 between households in the same village ($p<0.00001$), 0.26 between households of the same region ($p<0.0001$) and 0.23 between households in the entire Ngella comparison set ($p<0.00001$, Figure S6).

In order to assess the spatial scale of clustering, we then assessed associations between the extent of allele sharing and geographic distance of the households. The median physical distance between pairs of haplotypes sharing less than five alleles was 10 km with significant differences in the distance distributions of haplotypes with zero to four alleles shared. For clonal and sibling haplotypes sharing more than five alleles, there was a highly significant decrease in the median distance to between 100-300m (Figure 5).

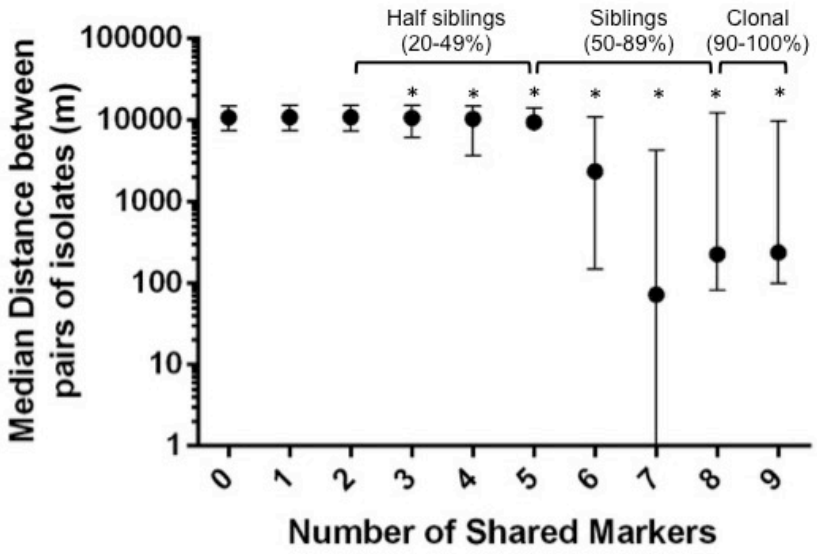

Figure 5. Relationship between geographic distance and genetic relatedness of *P. vivax* haplotypes from Ngella, Solomon Islands. Geographic distance in metres for each pair of haplotypes was calculated based on GPS co-ordinates of households from which the haplotypes originated. Genetic relatedness was measured using the pairwise allele sharing (P_s) of haplotypes. The distribution of geographic distances was then plotted for all pairs as a function of genetic relatedness. Circles represent the median geographic distance for a given genetic distance value, bars represent the 5-95th percentiles. Asterisks indicate highly significant ($p<0.000001$) differences between 0 shared alleles and a given number of shared alleles.

Discussion

The spatial scales that define malaria parasite populations and clustering of foci becomes particularly important at low transmission (Bousema, et al. 2012) to map the distribution of infections and aid in the spatial stratification of interventions for maximum impact. However, this may be challenging for *P. vivax* given high levels of outcrossing and complex patterns of gene flow that threaten to undermine control efforts (Jennison 2015; Noviyanti, et al. 2015; Hupalo, et al. 2016; Pearson, et al. 2016). Using the most spatially dense dataset of geo-positioned *P. vivax* genotypes to date, our results reveal decreasing diversity and increasing multilocus LD over time as well as fragmented *P. vivax* populations with declining transmission in space in the context of sustained long-term malaria control. In the central study area of Ngella, Solomon Islands, *P. falciparum* has almost disappeared due to ongoing control interventions, but significant *P. vivax* residual, mostly sub-clinical transmission remains (Waltmann, et al. 2015). In this region, *P. vivax* parasite populations were spatially structured among sub-regions, villages and households. A substantial decrease in diversity and an increase in LD and population structure over time on a neighbouring island (Guadalcanal) indicate that the patterns observed are predominantly the result of sustained malaria control, which has been ongoing in Solomon Islands for more than 20 yrs. The results show that while *P. vivax* may be overall more resistant to control efforts than *P. falciparum* (Feachem, et al. 2010; Bousema and Drakeley 2011; White and Imwong 2012; Mueller, et al. 2013), long-term sustained malaria control will put parasite populations under substantial stress and may lead to at least partial fragmentation of parasite subpopulations. While human movement is a major factor for the spread of infections at large scales and will also counteract population differentiation in the Solomon Islands, at the microepidemiological scale, the predominant clustering of infections is between household and most sibling and clonal parasites were found within 100-300 m i.e. within the usual flight distance of vectors in the Pacific (Charlwood, et al. 1988). Given the highly heterogeneous nature of mosquito-borne disease transmission (Perkins, et al. 2013; Stoddard, et al. 2013), this fine scale spatial clustering thus indicates that most infections persist and spread locally.

Across the Southwest Pacific, diversity amongst *P. vivax* populations was predominantly partitioned by country of origin reflecting the limited mixing of these populations. Southwest Pacific *P. vivax* populations have also been shown to be genetically distinct from other worldwide populations (Koepfli, et al. 2015; Hupalo, et al. 2016; Pearson, et al. 2016). Only minimal population structure between PNG and Solomon Islands was previously reported for the small number of samples collected in Tetere

between 2004-5 (Koepfli, et al. 2013; Jennison, et al. 2015). The differences with these earlier studies can be attributed to a much larger sample size from multiple Solomon Islands locations, and the intervening intensification of antimalarial interventions in the region. This is supported by the comparison of two time points (Tetere 2004-5 and 2013) that revealed a decline in polyclonal infections, corresponding with decreasing diversity and effective population size and an increase in multilocus LD and population structure. Although no pre-2013 samples were available from Ngella, evidence from malaria surveys indicate a 90% reduction in cases from 1992 to 2013 (Solomon Islands National Vector Borne Diseases Control Program, unpublished data), consistent with the pattern of population structure observed being a direct result of malaria control. Thus, sustained intervention has likely resulted in the inbred and fragmenting parasite populations observed.

The genetic structure of malaria parasite populations in relation to variable transmission has previously been investigated primarily with *P. falciparum* or with *P. vivax* populations over large spatial scales (e.g. between countries or distant locations within countries) (Ferreira, et al. 2007; Imwong, et al. 2007; Arnott, et al. 2013; Koepfli, et al. 2013; Abdullah et al. 2013; Jennison, et al. 2015; Koepfli, et al. 2015). The high-resolution analyses of *P. vivax* population structure in the central zone of Solomon Islands, a region spanning around 100 km, revealed population structure among different island provinces. This region, which contains the most populous provinces, has historically had the highest malaria transmission in the country (Avery 1977) and continues to have the highest API nationally (National Vector Borne Diseases Control Program 2013). Ngella *P. vivax* populations were found to have moderate levels of differentiation from populations of the other island provinces. Ngella is well connected to the higher endemicity areas of the Central Solomon Islands zone, as a direct and popular shipping route exists between Guadalcanal and Malaita Provinces via Ngella. This suggests that despite a significant level of human movement among these three provinces, importation of *P. vivax* cases into Ngella is sufficiently reduced to impact *P. vivax* gene flow. Whilst within country population structure has not been observed previously in the Southwest Pacific, it has been found in Malaysia (Abdullah, et al. 2013), where prevalence of *P. vivax* has traditionally been described as focal, and has recently reduced; and in South America (e.g. Peru (Van den Eede, et al. 2010; Delgado-Ratto, et al. 2016), Venezuela (Chenet, et al. 2012), Colombia (Imwong, et al. 2007) and Brazil (Ferreira, et al. 2007) where *P. vivax* has likely been introduced multiple times with adaptation to local vectors likely to have resulted in founder effects and influenced gene flow (Taylor, et al. 2013). At reduced transmission in an African setting, *P. falciparum* populations were shown to be more inbred and with genetic

relatedness rapidly increasing within the first year of intensified control as a result of inbreeding, however this would be dependent on the structure and effective population size at baseline (Daniels, et al. 2015). Notably, it has taken at least 20 years for Solomon Islands *P. vivax* populations to show signs of instability. Structured parasite populations within Ngella (20-50km) were subdivided into four genetic clusters distributed unevenly among Anchor/Bay/South (combined), North Coast and the two villages in the Channel region. The Channel villages lay in an extensive mangrove system on both sides of a channel, however, the area has comparable prevalence and proportions of polyclonal infections to other Ngella areas, showing that the population structure is likely to be influenced by the ecology and isolation of this region. Population structure was also observed among neighbouring villages of the North Coast. Thus, *P. vivax* population structure in Ngella seems to be organized as a hierarchical island model, consisting of a metapopulation of several sub-populations (Slatkin and Voelm 1991).

Despite marked reductions over time in one population (Tetere), relatively high genetic diversity and high effective population sizes remain in Solomon Islands *P. vivax* populations in the context of inbreeding and population structure. High *P. vivax* genetic diversity at low transmission was first recognized in Sri Lanka (Gunawardena, et al. 2014) and has also been observed together with significant LD in Peru (Chenet, et al. 2012), Malaysia (Abdullah, et al. 2013)and Indonesia (Noviyanti, et al. 2015). Despite a substantial range of transmission intensities, the genetic diversity observed for PNG and Solomon Islands were similar while hypoendemic Vanuatu had much lower levels of diversity, indicating that *P. vivax* transmission must reach very low levels before genetic diversity is impacted. LD and population structure can however signal changes in transmission intensity much earlier. The presence of identical haplotypes shared among Ngella parasites and significant multilocus LD is consistent with considerable levels of inbreeding due to increasingly clustered transmission of clonal or highly related, perhaps relapsing infections. In most endemic regions, identical *P. vivax* haplotypes are rare and were seen only at very low transmission in Central Asia where the *P. vivax* population is nearly clonal or at low transmission in the Amazon (Koepfli, et al. 2015). With decreasing transmission and polyclonality, opportunities for recombination between diverse strains are reduced. Focal inbreeding in and around households can explain the presence of LD in the context of high diversity, which is measured at the village level.

Relapse, which has been shown to account for up to 80% of *P. vivax* infections in the high transmission setting of PNG (Robinson, et al. 2015), is undoubtedly a major contributor to the observed population

structure. For some time after a reduction in transmission, the re-activation of parasites from a pool of genetically diverse hypnozoites from numerous past infections provides opportunities for the exchange and dissemination of alleles, thus sustaining genetic diversity in the population. However, as the hypnozoite reservoir is depleted, focal clusters will be composed of more recent infections and subsequent relapses with highly related strains (Bright, et al. 2014). If diversity is measured at larger scales (i.e. village or region) this could explain high diversity in the context of significant LD. In addition, as transmission declines to very low levels, imported infections can become an important source of new, inbred foci. Thus relapse is likely to sustain residual transmission and maintain diverse meta-populations with high evolutionary potential. Other biological characteristics of *P. vivax* that are likely to sustain transmission and resilience to intervention include the rapid and continuous gametocyte production coupled with efficient transmissibility at lower gametocyte loads that drives high rates of human-to-vector transmission (Boyd and Kitchen 1937; Jeffery and Eyles 1955) and the rapid acquisition of clinical immunity early in life and low density of infection (Mueller, et al. 2013) that would lead to a larger population reservoir of asymptomatic carriers (Harris, et al. 2010; Waltmann, et al. 2015). However, unlike relapse, these do not fully explain the patterns of population structure that we have observed in the context of declining transmission.

As national malaria control programs switch from control to elimination strategies, widespread application of control interventions eventually becomes unfeasible and spatially targeted interventions, more cost effective. In order to optimally plan intervention such as reactive case detection (van Eijk, et al. 2016) or focal mass drug administration (Gerardin, et al. 2016), it is essential to know the spatial scales required to deploy interventions for maximum impact. In line with a recent review (van Eijk, et al. 2016) we have previously shown that risk of *P. vivax* infection was enhanced by approximately 40% if an individual was living in a household with at least one other infected co-inhabitant, suggesting that within-household transmission may be important (Waltmann, et al. 2015). Spatial studies of other mosquito-borne infections, such as dengue virus (Harrington, et al. 2005; Mammen, et al. 2008; Stoddard, et al. 2013) or filarial worms (Michael, et al. 2001), have demonstrated that transmission occurs within communities, often around homes. These spatial studies have employed various data types and approaches, including profiling of human DNA in mosquito blood meals (Michael, et al. 2001; De Benedictis, et al. 2003), cluster analysis around index cases (Mammen, et al. 2008) or index case contact tracing (Stoddard, et al. 2013). Spatial characteristics of *P. vivax* malaria transmission have not been previously investigated using population genetics data. Within villages, the results show

significantly more genetic relatedness between parasites of the same household than between parasites of different households. The finding of more highly related parasites among people living in the same household than among the general population, indicates that co-inhabitants may be infected by more inbred strains either due to spatial clustering of transmission or by the bites of the same, infected mosquito(es).

Despite the principal vector in Solomon Islands, *Anopheles farauti*, feeding outdoors (Russell, et al. 2016), much of the exposure to infective bites remains highly clustered around homes. This vector behaviour is considered to be a major challenge for elimination, but our data suggests that interventions focused on index households (e.g. reactive case detection or focal mass drug administration) could make a substantial impact. Whilst not all sibling or "clonal" parasites were found in the same household, they circulate in close proximity since high genetic relatedness was observed within approximately 100 - 300 meters, after which point it substantially decreased. These village "neighbourhoods" of parasite lineages appear to emanate from within-household co-transmission of highly related parasites. This radius of high genetic relatedness is consistent with the mosquito flight path (Charlwood, et al. 1988; Russell, et al. 2016). Thus, the fine scale patterns of population structure detected are likely to be driven by mosquito movement, rather than that of the human host. This data provides a basis to identify and attack residual pockets of transmission. The findings highlight that improved malaria surveillance and intervention can be local in nature, an approach previously recommended (Greenwood 1989; Stoddard, et al. 2013). Spatial decision support systems have been already proposed for the elimination provinces of Temotu and Isabel (Kelly, et al. 2013). The data suggests that a 300 meter response perimeter around index households could be included as part of a reactive, hypnozoite-targeting intervention against *P. vivax*.

In summary, the results demonstrate *P. vivax* population structure at all spatial scales with hampered gene flow and inbreeding within parasite populations after long term sustained malaria control. These findings have significant public health implications showing that albeit more resistant to control efforts than *P. falciparum* (Alonso and Tanner 2013; WHO 2015a), *P. vivax* populations eventually will become increasingly inbred and fragmented if control pressure is maintained over an extended period. These results emphasize the need for interventions to be sustained for very long periods, well beyond the time frame required for *P. falciparum*. Given the proposal to eliminate malaria from the Asia-Pacific by 2030 (APLMA 2014), intensive control pressure must be maintained to capitalize on these

successes and avoid rebound. Enhanced control efforts including targeted control in and around hotspots of transmission will help to reach these goals.

Materials and Methods

Study sites and *Plasmodium vivax* isolates

Historically, the Southwest Pacific region, in particular PNG and Solomon Islands, has endured some of the highest *P. vivax* transmission anywhere in the world and a *P. falciparum* incidence comparable to that of Sub-Saharan Africa (Gething, et al. 2011; Gething, et al. 2012; Organisation 2013). Sustained control efforts in the Solomon Islands over the past 20 years have reduced transmission to very low levels (Harris, et al. 2010; PacMISC 2010; Waltmann, et al. 2015) not seen since the end of the last malaria elimination program in the mid 1970s. At the time of sampling, transmission in this region ranged from moderate to high in PNG, low in Solomon Islands and very low in Vanuatu (Gething, et al. 2012).

A total of 887 isolates from PNG (n=443), Solomon Islands (n=420) and Vanuatu were used for the current study (n=24) (Table S1). These included previously published genotyping data from PNG collected in 2005-6 (n=486) and Solomon Islands in 2004 (Tetere 2004-5, n=45) (Koepfli, et al. 2013; Jennison, et al. 2015) in addition to 398 newly typed *P. vivax* isolates from multiple sites in the Solomon Islands collected in 2012-2013 (Waltmann, et al. 2015) and 24 from Espiritu Santo, Vanuatu collected in 2013 (Figure 1A). Importantly, the dense sampling of the central region of the Solomon Islands allowed analyses at a wide range of spatial scales. The 2012-13 Solomon Islands samples are from three neighbouring island provinces, namely Malaita (Auki, n=13), Guadalcanal (Tetere, n=39) and Central Province (Ngella, n= 323, Figure 1B). The 323 Ngella haplotypes spanned all three islands, including five distinct geographical areas: Bay (n=83), South (n=35), Channel (n=46), North (n=136) and Anchor (n=23, Figure 1C) comprising 19 villages and 190 households. Of all the households included in the Ngella survey, 93 had only one *P. vivax*-infected member, 69 had two, 23 had three, five households had four infections and one household had five members infected. The Tetere 2013 (Wini et al., unpublished), Auki (Wini et al., unpublished), and Vanuatu (Boyd et al., unpublished) samples were collected as part of antimalarial drug efficacy trials.

Further details of the samples and study sites are summarised in Table S1 and Text S1. The study was approved by The Walter and Eliza Hall Institute Human Research Ethics Committee (12/01, 11/12 and

13/02), the Papua New Guinea Institute of Medical Research Institutional Review Board (11-05), the Papua New Guinea Medical Research Advisory Committee (11-06), the Solomon Islands National Health Research Ethics Committee (12/022) and the Vanuatu Ministry of Health (19-02-2013).

Multiplicity of Infection (MOI)

To determine multiplicity of infection (MOI) in each population and to allow the selection of low complexity infections (MOI = 1 or 2) for the population genetics analyses, MS16 and *msp1*F3 genotyping data were used (Koepfli, et al. 2013; Jennison, et al. 2015). These data were previously available for the PNG (Koepfli, et al. 2013; Jennison 2015), Tetere 2004-5 (Koepfli, et al. 2013), and Ngella datasets (Waltmann, et al. 2015). The MOI in the Tetere 2013, Auki 2013, and Vanuatu *P. vivax* populations was assessed for this study, according to protocols previously described (Karunaweera, et al. 2008; Koepfli, et al. 2011).

Multilocus microsatellite genotyping

All confirmed low complexity infections (MOI = 1 and 2) were then genotyped with nine genome-wide and putatively neutral microsatellites loci (MS1, MS2, MS5, MS6, MS7, MS9, MS10, MS12 and MS15) (Karunaweera, et al. 2008). A semi-nested PCR was employed, whereby a multiplex primary PCR was followed by nine individual secondary reactions, with a fluorescently labelled forward primer, as previously described (Koepfli, et al. 2013; Jennison, et al. 2015). PCR products were sent to a commercial facility for GeneScan™ fragment analysis on an ABI3730xl capillary electrophoresis platform (Applied Biosystems) using the size standard LIZ500.

Data analysis

Electropherograms resulting from the fragment analysis were visually inspected and the sizes of the fluorescently labeled PCR products were scored with Genemapper V4.0 software (Applied Biosystems), with the peak calling strategy done as previously described (Jennison, et al. 2015). Raw data from the published dataset was added to the new dataset and binned together to obtain consistent allele calls. Automatic binning (i.e. rounding of fragment length to specific allele sizes) was performed with Tandem (Matschiner and Salzburger 2009). After binning, quality control for individual *P. vivax* haplotypes and microsatellite markers was conducted to confirm the markers were not in linkage disequilibrium (LD) and to identify outlier haplotypes and/or markers (i.e. haplotypes or markers which are disproportionately driving variance in the dataset). For isolates with an MOI=2, the dominant

alleles were used to construct dominant clone haplotypes as previously described (Jennison, et al. 2015).

Allele frequencies and input files for the various population genetics software programs were created using CONVERT version 1.31. Allele frequencies and genetic diversity parameters (number of alleles (*A*), expected heterozygosity (H_E) and allelic richness (R_S)) were calculated in *FSTAT* version 2.9.3.2. Effective Population Size (N_e) was calculated using the stepwise mutation model (SMM) and infinite alleles model (IAM), as previously described (Anderson, Haubold, et al. 2000). Mutation rates for *P. vivax* were not available and thus the *P. falciparum* mutation rate was used (Anderson, Su, et al. 2000).

For SMM, N_e was calculated as follows:

$$N_e = \frac{\frac{1}{8}\, x \left\{ \left[\frac{1}{1-H_{E_{mean}}} \right]^2 - 1 \right\}}{\mu}$$

where $H_{E\ mean}$ is the expected heterozygosity averaged across all loci.

For the IAM, N_e was calculated using the formula:

$$N_e = \left(\frac{H_{E_{mean}}}{4\left(1 - H_{E_{mean}}\right)} \right) x \frac{1}{\mu}$$

As a measure of inbreeding in the populations studied, multilocus LD (non-random associations between alleles of all pairs of markers) was estimated using the standardized index of association (I_A^S) in LIAN version 3.6. I_A^S compares the observed variance in the number of shared alleles between parasites with that expected under equilibrium, when alleles at different loci are not in association (Haubold and Hudson 2000). The measure was followed by a formal test of the null hypothesis of LD and *p*-values were derived. Only unique haplotypes with complete genotypes were used and Monte Carlo tests with 100,000 re-samplings were applied (Haubold and Hudson 2000). The number of unique haplotypes was assessed using DROPOUT (McKelvey and Schwartz 2005). To confirm that LD was not artificially reduced by false reconstruction of dominant haplotypes, the analysis was also performed for the combined dataset of dominant and single haplotypes, and for single infections only.

MS2 and MS5 both localize to chromosome 6 and MS12 and MS15 to chromosome 5 thus, analyses were repeated on datasets where MS5 and MS15 were excluded (chosen due to a greater degree of missing data) using the remaining seven loci spanning seven chromosomes. Where sample size permitted (n > 5), multilocus LD was also estimated at village level.

To investigate geographic population structure, we first calculated three measures of genetic differentiation, namely F_{ST}, G_{ST} and Jost's D, for all pairwise comparisons of the predefined populations. F_{ST} was estimated using *FSTAT*. G_{ST} (Nei and Chesser 1983) and Jost's D (Jost 2008) were estimated using the *R* package *DEMEtics*, as previously described (Gerlach, et al. 2010). Population structure was further confirmed by Bayesian clustering of haplotypes implemented in the software STRUCTURE version 2.3.4 (Pritchard, et al. 2000), which was used to investigate whether haplotypes cluster into distinct genetic populations (K) among the defined geographic areas. The analyses were run for K=1-20, with 20 independent stochastic simulations for each K and 100,000 MCMCs, after an initial burn-in period of 10,000 MCMCs using the admixture model and correlated allele frequencies. The results were processed using STRUCTURE Harvester (Earl and Vonholdt 2012), to calculate the optimal number of clusters as indicated by a peak in ΔK according to the method of Evanno *et al*. (Evanno, et al. 2005). The programs CLUMPP version 1.1.2 (Jakobsson and Rosenberg 2007) and DISTRUCT 1.1 (Rosenberg 2004) were used to display the results.

As our dataset comprised substantial numbers of infections from the same household it was possible to investigate fine-scale (within and between households) clustering of infections. To do this we assessed the extent of allele sharing, P_S among *P. vivax* haplotypes, calculated as the number of alleles shared between a pair of haplotypes divided by the number of loci for which data was available for that pair of isolates. The formula is as follows:

$$Ps_{i,j} = \frac{\sum_{k=1}^{9}(k_i = k_j)}{k}$$

Where:

i and j = the two haplotypes compared

k = the number of markers

Note: the number of missing markers is subtracted from the denominator

P_S was also measured for each individual microsatellite locus to confirm the patterns. First, we computed the number of identical alleles observed between two pairs of infections. Next, the minimum number of alleles available in each of the pairwise comparisons is considered as the denominator. For example, if for a given marker *x*, infection *i* has three detected alleles and infection *j* has two detected alleles and they have in common one allele, the proportion of alleles shared is 1/2 (50%). Therefore the formula for P_S, for each marker, is as follows:

$$Ps_{i,j} = \frac{|A \cap B|}{\min(|A|, |B|)}$$

The analyses per haplotype included dominant and single haplotypes. For the per marker analyses, households with only one infected individual were excluded and the dataset included all observed alleles for each *P. vivax* infection. Permutation tests were used to formally assess the difference in the allele sharing within households compared to that among households.

Parasites which shared between 20-49% of their alleles were considered half-siblings, those which shared 50-89% of alleles were classified as full-siblings and (nearly) clonal parasites were those which shared 90-100% of their alleles.

For the haplotype data, the test statistic *D,* which is the difference between the mean P_S within households and the mean P_S between households, was calculated. The sampling distribution of *D* under the null hypothesis (allele sharing within households is equal to the allele sharing between households, H_0: *D=0*) was computed using 10,000 permutations and compared to the observed *D* and the *p*-value (the proportion of statistics, including the original, that are larger than the observed *D*) derived. For the per marker analyses*,* P_S was calculated by using the minimum number of alleles available in the pairwise comparison for each marker individually as the denominator. For example, if for a given marker *x*, individual *i* has three detected alleles and individual *j* has two detected alleles and they have in common one allele, the proportion p_{ij} of alleles shared is 1/2 (50%). The dataset for this analysis excluded households with only one infected individual. Permutation tests (10,000 re-samplings) were employed to test for the observed difference in pairwise allele sharing (*D*) within and between households.

To investigate the spatial scale of haplotype clustering, the physical distance (in metres) was calculated for all pairs of isolates. Distance distributions were then calculated for each level of relatedness (i.e. 0-9 shared markers). Significant differences in the distance distributions were then compared using a Mann Whitney U test. All statistical analyses were done using Mathematica and GraphPad Prism.

Acknowledgements

The authors wish to acknowledge the support of the communities and field workers at all study sites. This study was supported by the International Centers of Excellence in Malaria Research (ICEMR, NIH grant U19AI089686 "Research to control and eliminate malaria in the Southwest Pacific") and by the National Health and Medical Research Council of Australia (NHMRC, #1021544). IM is supported by a NHMRC Senior Research Fellowship (#1043345) and AW was supported by an NHMRC Postgraduate Scholarship (1056511). The authors acknowledge the Victorian State Government Operational Infrastructure Support and Australian Government National Health and Medical Research Council Independent Research Institute Infrastructure Support Scheme.

References

Abdullah NR, Barber BE, William T, Norahmad NA, Satsu UR, Muniandy PK, Ismail Z, Grigg MJ, Jelip J, Piera K, et al. 2013. Plasmodium vivax population structure and transmission dynamics in Sabah Malaysia. PLoS One 8:e82553.

Alonso PL, Brown G, Arevalo-Herrera M, Binka F, Chitnis C, Collins F, Doumbo OK, Greenwood B, Hall BF, Levine MM, et al. 2011. A research agenda to underpin malaria eradication. PLoS Med 8:e1000406.

Alonso PL, Tanner M. 2013. Public health challenges and prospects for malaria control and elimination. Nat Med 19:150-155.

Anderson TJ, Haubold B, Williams JT, Estrada-Franco JG, Richardson L, Mollinedo R, Bockarie M, Mokili J, Mharakurwa S, French N, et al. 2000. Microsatellite markers reveal a spectrum of population structures in the malaria parasite Plasmodium falciparum. Mol Biol Evol 17:1467-1482.

Anderson TJ, Su XZ, Roddam A, Day KP. 2000. Complex mutations in a high proportion of microsatellite loci from the protozoan parasite Plasmodium falciparum. Mol Ecol 9:1599-1608.

APLMA. 2014. Task Force Progress Report 2014. In.

Arnott A, Barnadas C, Senn N, Siba P, Mueller I, Reeder JC, Barry AE. 2013. High genetic diversity of Plasmodium vivax on the north coast of Papua New Guinea. Am J Trop Med Hyg 89:188-194.

Avery J. 1977. The Epidemiology of Disappearing Malaria in the Solomon Islands. [[Sheffield, United Kingdom]: University of Sheffield.

Barry AE, Waltmann A, Koepfli C, Barnadas C, Mueller I. 2015. Uncovering the transmission dynamics of Plasmodium vivax using population genetics. Pathog Glob Health 109:142-152.

Bousema T, Drakeley C. 2011. Epidemiology and infectivity of Plasmodium falciparum and Plasmodium vivax gametocytes in relation to malaria control and elimination. Clin Microbiol Rev 24:377-410.

Bousema T, Griffin JT, Sauerwein RW, Smith DL, Churcher TS, Takken W, Ghani A, Drakeley C, Gosling R. 2012. Hitting hotspots: spatial targeting of malaria for control and elimination. PLoS Med 9:e1001165.

Boyd M, Kitchen S. 1937. On the infectiousness of patients infected with Plasmodium vivax and Plasmodium falciparum. Am J Trop Med Hyg s1-17:253-262.

Bright AT, Manary MJ, Tewhey R, Arango EM, Wang T, Schork NJ, Yanow SK, Winzeler EA. 2014. A high resolution case study of a patient with recurrent Plasmodium vivax infections shows that relapses were caused by meiotic siblings. PLoS Negl Trop Dis 8:e2882.

Cattani JA, Tulloch JL, Vrbova H, Jolley D, Gibson FD, Moir JS, Heywood PF, Alpers MP, Stevenson A, Clancy R. 1986. The epidemiology of malaria in a population surrounding Madang, Papua New Guinea. Am J Trop Med Hyg 35:3-15.

Charlwood JD, Graves PM, Marshall TF. 1988. Evidence for a 'memorized' home range in Anopheles farauti females from Papua New Guinea. Med Vet Entomol 2:101-108.

Chenet SM, Schneider KA, Villegas L, Escalante AA. 2012. Local population structure of Plasmodium: impact on malaria control and elimination. Malar J 11:412.

Cotter C, Sturrock HJ, Hsiang MS, Liu J, Phillips AA, Hwang J, Gueye CS, Fullman N, Gosling RD, Feachem RG. 2013. The changing epidemiology of malaria elimination: new strategies for new challenges. Lancet 382:900-911.

Daniels RF, Schaffner SF, Wenger EA, Proctor JL, Chang HH, Wong W, Baro N, Ndiaye D, Fall FB, Ndiop M, et al. 2015. Modeling malaria genomics reveals transmission decline and rebound in Senegal. Proc Natl Acad Sci U S A 112:7067-7072.

De Benedictis J, Chow-Shaffer E, Costero A, Clark GG, Edman JD, Scott TW. 2003. Identification of the people from whom engorged Aedes aegypti took blood meals in Florida, Puerto Rico, using polymerase chain reaction-based DNA profiling. Am J Trop Med Hyg 68:437-446.

Delgado-Ratto C, Gamboa D, Soto-Calle VE, Van den Eede P, Torres E, Sanchez-Martinez L, Contreras-Mancilla J, Rosanas-Urgell A, Rodriguez Ferrucci H, Llanos-Cuentas A, et al. 2016. Population Genetics of Plasmodium vivax in the Peruvian Amazon. PLoS Negl Trop Dis 10:e0004376.

Earl DA, Vonholdt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4:359-361.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611-2620.

Feachem RG, Phillips AA, Hwang J, Cotter C, Wielgosz B, Greenwood BM, Sabot O, Rodriguez MH, Abeyasinghe RR, Ghebreyesus TA, et al. 2010. Shrinking the malaria map: progress and prospects. Lancet 376:1566-1578.

Ferreira MU, Karunaweera ND, da Silva-Nunes M, da Silva NS, Wirth DF, Hartl DL. 2007. Population structure and transmission dynamics of Plasmodium vivax in rural Amazonia. J Infect Dis 195:1218-1226.

Gerardin J, Bever CA, Hamainza B, Miller JM, Eckhoff PA, Wenger EA. 2016. Optimal Population-Level Infection Detection Strategies for Malaria Control and Elimination in a Spatial Model of Malaria Transmission. PLoS Comput Biol 12:e1004707.

Gerlach G, Jueterbock A, Kraemer P, Deppermann J, Harmand P. 2010. Calculations of population differentiation based on GST and D: forget GST but not all of statistics! Mol Ecol 19:3845-3852.

Gething PW, Elyazar IR, Moyes CL, Smith DL, Battle KE, Guerra CA, Patil AP, Tatem AJ, Howes RE, Myers MF, et al. 2012. A long neglected world malaria map: Plasmodium vivax endemicity in 2010. PLoS Negl Trop Dis 6:e1814.

Gething PW, Patil AP, Smith DL, Guerra CA, Elyazar IR, Johnston GL, Tatem AJ, Hay SI. 2011. A new world malaria map: Plasmodium falciparum endemicity in 2010. Malar J 10:378.

Gray KA, Dowd S, Bain L, Bobogare A, Wini L, Shanks GD, Cheng Q. 2013. Population genetics of Plasmodium falciparum and Plasmodium vivax and asymptomatic malaria in Temotu Province, Solomon Islands. Malar J 12:429.

Greenwood BM. 1989. The microepidemiology of malaria and its importance to malaria control. Trans R Soc Trop Med Hyg 83 Suppl:25-29.

Gunawardena S, Ferreira MU, Kapilananda GM, Wirth DF, Karunaweera ND. 2014. The Sri Lankan paradox: high genetic diversity in Plasmodium vivax populations despite decreasing levels of malaria transmission. Parasitology 141:880-890.

Harrington LC, Scott TW, Lerdthusnee K, Coleman RC, Costero A, Clark GG, Jones JJ, Kitthawee S, Kittayapong P, Sithiprasasna R, et al. 2005. Dispersal of the dengue vector Aedes aegypti within and between rural communities. Am J Trop Med Hyg 72:209-220.

Harris I, Sharrock WW, Bain LM, Gray KA, Bobogare A, Boaz L, Lilley K, Krause D, Vallely A, Johnson ML, et al. 2010. A large proportion of asymptomatic Plasmodium infections with low and sub-microscopic parasite densities in the low transmission setting of Temotu Province, Solomon Islands: challenges for malaria diagnostics in an elimination setting. Malar J 9:254.

Haubold B, Hudson RR. 2000. LIAN 3.0: detecting linkage disequilibrium in multilocus data. Linkage Analysis. Bioinformatics 16:847-848.

Hupalo DN, Luo Z, Melnikov A, Sutton PL, Rogov P, Escalante A, Vallejo AF, Herrera S, Arevalo-Herrera M, Fan Q, et al. 2016. Population genomics studies identify signatures of global dispersal and drug resistance in Plasmodium vivax. Nature genetics 48:953-958.

Imwong M, Nair S, Pukrittayakamee S, Sudimack D, Williams JT, Mayxay M, Newton PN, Kim JR, Nandy A, Osorio L, et al. 2007. Contrasting genetic structure in Plasmodium vivax populations from Asia and South America. Int J Parasitol 37:1013-1022.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801-1806.

Jeffery GM, Eyles DE. 1955. Infectivity to Mosquitoes of Plasmodium Falciparum as Related to Gametocyte Density and Duration of Infection. Am J Trop Med Hyg 4:781-789.

Jennison C, Arnott A, Tessier N, Tavul L, Koepfli C, Felger I, Siba PM, Reeder JC, Bahlo M, Mueller I, et al. 2015. Plasmodium vivax populations are more genetically diverse and less structured than sympatric Plasmodium falciparum populations. PLoS Negl Trop Dis 9:e0003634.

Jennison C, Arnott, A., Tessier, N., Tavul, L., Koepfli, C., Felger, I., Siba, P.M., Reeder, J.C., Bahlo, M., Mueller, I., Barry, A.E. 2015. Plasmodium vivax populations are more genetically diverse and less structured than sympatric Plasmodium falciparum populations. PLoS Negl Trop Dis 9:e0003634.

Jost L. 2008. G(ST) and its relatives do not measure differentiation. Molecular ecology 17:4015-4026.

Kaneko A. 2010. A community-directed strategy for sustainable malaria elimination on islands: short-term MDA integrated with ITNs and robust surveillance. Acta Trop 114:177-183.

Kaneko A, Chaves LF, Taleo G, Kalkoa M, Isozumi R, Wickremasinghe R, Perlmann H, Takeo S, Tsuboi T, Tachibana S, et al. 2014. Characteristic age distribution of Plasmodium vivax infections after malaria elimination on Aneityum Island, Vanuatu. Infect Immun 82:243-252.

Karunaweera ND, Ferreira MU, Munasinghe A, Barnwell JW, Collins WE, King CL, Kawamoto F, Hartl DL, Wirth DF. 2008. Extensive microsatellite diversity in the human malaria parasite Plasmodium vivax. Gene 410:105-112.

Kelly GC, Hale E, Donald W, Batarii W, Bugoro H, Nausien J, Smale J, Palmer K, Bobogare A, Taleo G, et al. 2013. A high-resolution geospatial surveillance-response system for malaria elimination in Solomon Islands and Vanuatu. Malar J 12:108.

Koepfli C, Rodrigues PT, Antao T, Orjuela-Sanchez P, Van den Eede P, Gamboa D, van Hong N, Bendezu J, Erhart A, Barnadas C, et al. 2015. Plasmodium vivax Diversity and Population Structure across Four Continents. PLoS Negl Trop Dis 9:e0003872.

Koepfli C, Ross A, Kiniboro B, Smith TA, Zimmerman PA, Siba P, Mueller I, Felger I. 2011. Multiplicity and diversity of Plasmodium vivax infections in a highly endemic region in Papua New Guinea. PLoS Negl Trop Dis 5:e1424.

Koepfli C, Timinao L, Antao T, Barry AE, Siba P, Mueller I, Felger I. 2013. A Large Reservoir and Little Population Structure in the South Pacific. PLoS One 8:e66041.

Mammen MP, Pimgate C, Koenraadt CJ, Rothman AL, Aldstadt J, Nisalak A, Jarman RG, Jones JW, Srikiatkhachorn A, Ypil-Butac CA, et al. 2008. Spatial and temporal clustering of dengue virus transmission in Thai villages. PLoS Med 5:e205.

Markert JA, Champlin DM, Gutjahr-Gobell R, Grear JS, Kuhn A, McGreevy TJ, Jr., Roth A, Bagley MJ, Nacci DE. 2010. Population genetic diversity and fitness in multiple environments. BMC Evol Biol 10:205.

Matschiner M, Salzburger W. 2009. TANDEM: integrating automated allele binning into genetics and genomics workflows. Bioinformatics (Oxford, England) 25:1982-1983.

McKelvey KS, Schwartz MK. 2005. DROPOUT: a program to identify problem loci and samples for noninvasive genetic samples in a capture-mark-recapture framework. Molecular Ecology Notes 5:716-718.

Mendis K, Sina BJ, Marchesini P, Carter R. 2001. The neglected burden of Plasmodium vivax malaria. Am J Trop Med Hyg 64:97-106.

Michael E, Ramaiah KD, Hoti SL, Barker G, Paul MR, Yuvaraj J, Das PK, Grenfell BT, Bundy DA. 2001. Quantifying mosquito biting patterns on humans by DNA fingerprinting of bloodmeals. Am J Trop Med Hyg 65:722-728.

Mueller I, Galinski MR, Tsuboi T, Arevalo-Herrera M, Collins WE, King CL. 2013. Natural acquisition of immunity to Plasmodium vivax: epidemiological observations and potential targets. Adv Parasitol 81:77-131.

Neafsey DE, Galinsky K, Jiang RHY, Young L, Sykes SM, Saif S, Gujja S, Goldberg JM, Young S, Zeng Q, et al. 2012. The malaria parasite Plasmodium vivax exhibits greater genetic diversity than Plasmodium falciparum. Nature genetics 44:1046-1050.

Nei M, Chesser RK. 1983. Estimation of fixation indices and gene diversities. Annals of human genetics 47:253-259.

Nkhoma SC, Nair S, Al-Saai S, Ashley E, McGready R, Phyo AP, Nosten F, Anderson TJ. 2013. Population genetic correlates of declining transmission in a human pathogen. Mol Ecol 22:273-285.

Noviyanti R, Coutrier F, Utami RA, Trimarsanto H, Tirta YK, Trianty L, Kusuma A, Sutanto I, Kosasih A, Kusriastuti R, et al. 2015. Contrasting Transmission Dynamics of Co-endemic Plasmodium vivax and P. falciparum: Implications for Malaria Control and Elimination. PLoS Negl Trop Dis 9:e0003739.

Oliveira-Ferreira J, Lacerda MV, Brasil P, Ladislau JL, Tauil PL, Daniel-Ribeiro CT. 2010. Malaria in Brazil: an overview. Malar J 9:115.

Organisation WH. 2013. World Malaria Report 2013. In.

Orjuela-Sanchez P, Sa JM, Brandi MCC, Rodrigues PT, Bastos MS, Amaratunga C, Duong S, Fairhurst RM, Ferreira MU. 2013. Higher microsatellite diversity in Plasmodium vivax than in sympatric Plasmodium falciparum populations in Pursat, Western Cambodia. Experimental parasitology 134:318-326.

PacMISC. 2010. Malaria on isolated Melanesian islands prior to the initiation of malaria elimination activities. Malar J 9:218.

Pearson RD, Amato R, Auburn S, Miotto O, Almagro-Garcia J, Amaratunga C, Suon S, Mao S, Noviyanti R, Trimarsanto H, et al. 2016. Genomic analysis of local variation and recent evolution in Plasmodium vivax. Nature genetics 48:959-964.

Perkins TA, Scott TW, Le Menach A, Smith DL. 2013. Heterogeneity, mixing, and the spatial scales of mosquito-borne pathogen transmission. PLoS Comput Biol 9:e1003327.

Pimenta PF, Orfano AS, Bahia AC, Duarte AP, Rios-Velasquez CM, Melo FF, Pessoa FA, Oliveira GA, Campos KM, Villegas LM, et al. 2015. An overview of malaria transmission from the perspective of Amazon Anopheles vectors. Mem Inst Oswaldo Cruz 110:23-47.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Robinson LJ, Wampfler R, Betuela I, Karl S, White MT, Li Wai Suen CS, Hofmann NE, Kinboro B, Waltmann A, Brewster J, et al. 2015. Strategies for Understanding and Reducing the Plasmodium vivax and Plasmodium ovale Hypnozoite Reservoir in Papua New Guinean Children: A Randomised Placebo-Controlled Trial and Mathematical Model. PLoS Med 12:e1001891.

Rodriguez JC, Uribe GA, Araujo RM, Narvaez PC, Valencia SH. 2011. Epidemiology and control of malaria in Colombia. Mem Inst Oswaldo Cruz 106 Suppl 1:114-122.

Rosenberg NA. 2004. DISTRUCT: a program for the graphical display of population structure. Molecular Ecology Notes 4:137-138.

Russell TL, Beebe NW, Bugoro H, Apairamo A, Chow WK, Cooper RD, Collins FH, Lobo NF, Burkot TR. 2016. Frequent blood feeding enables insecticide-treated nets to reduce transmission by mosquitoes that bite predominately outdoors. Malar J 15:156.

Slatkin M, Voelm L. 1991. FST in a hierarchical island model. Genetics 127:627-629.

Solomon Islands National Vector Borne Diseases Control Program (NVBDCP). 2013. Annual Malaria Report 2012. In. Honiara, Solomon Islands: Ministry of Health.

Stoddard ST, Forshey BM, Morrison AC, Paz-Soldan VA, Vazquez-Prokopec GM, Astete H, Reiner RC, Jr., Vilcarromero S, Elder JP, Halsey ES, et al. 2013. House-to-house human movement drives dengue virus transmission. Proc Natl Acad Sci U S A 110:994-999.

Taylor JE, Pacheco MA, Bacon DJ, Beg MA, Machado RL, Fairhurst RM, Herrera S, Kim JY, Menard D, Povoa MM, et al. 2013. The evolutionary history of Plasmodium vivax as inferred from mitochondrial genomes: parasite genetic diversity in the Americas. Mol Biol Evol 30:2050-2064.

Van den Eede P, Van der Auwera G, Delgado C, Huyse T, Soto-Calle VE, Gamboa D, Grande T, Rodriguez H, Llanos A, Anne J, et al. 2010. Multilocus genotyping reveals high heterogeneity and strong local population structure of the Plasmodium vivax population in the Peruvian Amazon. Malar J 9:151.

van Eijk AM, Ramanathapuram L, Sutton PL, Kanagaraj D, Sri Lakshmi Priya G, Ravishankaran S, Asokan A, Tandel N, Patel A, Desai N, et al. 2016. What is the value of reactive case detection in malaria control? A case-study in India and a systematic review. Malar J 15:67.

Waltmann A, Darcy AW, Harris I, Koepfli C, Lodo J, Vahi V, Piziki D, Shanks GD, Barry AE, Whittaker M, et al. 2015. High Rates of Asymptomatic, Sub-microscopic Plasmodium vivax Infection and Disappearing Plasmodium falciparum Malaria in an Area of Low Transmission in Solomon Islands. PLoS Negl Trop Dis 9:e0003758.

White NJ, Imwong M. 2012. Relapse. Adv Parasitol 80:113-150.

WHO. 2015a. Confronting Plasmodium vivax malaria. In. Geneva, Switzerland: World Health Organization.

WHO. 2015b. Global Technical Strategy for Malaria 2016-2030. In. Geneva: World Health Organization.

WHO. 2015c. World Malaria Report 2015. In. Geneva, Switzerland: World Health Organization.

Woolhouse ME, Dye C, Etard JF, Smith T, Charlwood JD, Garnett GP, Hagan P, Hii JL, Ndhlovu PD, Quinnell RJ, et al. 1997. Heterogeneities in the transmission of infectious agents: implications for the design of control programs. Proc Natl Acad Sci U S A 94:338-342.

Tables

Table 1: Genetic diversity of *P. vivax* populations in the Southwest Pacific

Country	Province	Population	n haplotypes	*A*±SEM	H_E±SEM	R_S±SEM	N_e (95% CI) SMM	N_e (95% CI) IAM
Papua New Guinea	East Sepik		229	13.44±0.38	0.81±0.006	8.75±0.20	16059 (6901, 36582)	6811 (2927, 15515)
	Madang		175	15±0.39	0.84±0.005	9.62±0.20	24293 (10439, 55337)	8234 (3539, 18757)
	Simbu		39	8±0.46	0.81±0.011	7.37±0.38	15539 (6678, 35397)	6713 (2885, 15291)
Solomon Islands	Guadalcanal	Tetere 2004-5	45	9.44±0.53	0.84±0.009	8.44±0.43	23220 (9978, 52893)	8061 (3464, 18362)
		Tetere 2013	39	7.78±0.3	0.79±0.009	7.05±0.24	11766 (5056, 26802)	5963 (2562, 13583)
	Central Islands (Ngella)	Bay	83	12.33±0.27	0.81±0.006	9.20±0.15	16112 (6924, 36703)	6821 (2931, 15538)
		South	35	9.11±0.34	0.82±0.012	8.50±0.32	16599 (7133, 37811)	6911 (2970, 15744)
		Channel	46	9.33±0.35	0.79±0.009	8.10±0.28	11500 (4942, 26196)	5907 (2538, 13456)
		North	136	13.89±0.28	0.85±0.004	9.73±0.17	26387 (11339, 60107)	8564 (3680, 19509)
		Anchor	23	6.56±0.38	0.79±0.019	6.51±0.37	10771 (4629, 24535)	5752 (2472, 13103)
	Malaita	Auki	13	5.33±0.54	0.80±0.026	NA	13170 (5659, 30000)	6250 (2686, 14238)
Vanuatu	Sanma	Espiritu Santo	24	5.56±0.34	0.72±0.022	5.45±0.33	4056 (1743, 9239)	4131 (1775, 9411)

Table 2. Estimates of inbreeding in *P. vivax* populations of Solomon Islands and Vanuatu. Data from PNG and Tetere 2004-5 were previously published (Koepfli, et al. 2013; Jennison 2015).

Population	**All haplotypes, all loci**			**MOI = 1, all loci**		
	n	I_A^S	*p*	n	I_A^S	*p*
Tetere 2004-5	21	0.012	0.1800	0	NA[§]	
Tetere 2013	31	0.022	0.0042	16	0.033	0.0183
Auki	9	0.081	0.0050	6	0.054	0.1300
Ngella	165	0.026	<0.00001	61	0.043	<0.00001
Bay	32	0.041	0.0008	9	0.092	0.0052
Halavo	21	0.073	0.0002	6	0.146	0.0014
Bokolonga	5	0.315	0.00002	1	NA[§]	
Channel	29	0.074	<0.00001	9	0.087	0.0053
Hanuvavine	21	0.101	<0.00001	5	0.189	0.0020
South	17	0.068	0.0003	7	0.163	0.0005
Koilovala	10	0.049	0.0474	5	0.100	0.0592
Vuranimala	5	0.238	0.0014	2	NA[§]	
Anchor	14	0.069	0.0009	8	0.046	0.0948
North	73	0.038	<0.00001	28	0.076	<0.00001
Vura	34	0.058	<0.00001	14	0.067	0.0010
Polomuhu	23	0.066	<0.00001	10	0.199	<0.00001
Tavulea	16	0.097	<0.00001	1	NA[§]	
Vanuatu	22	0.157	<0.00001	10	0.1690	<0.00001

[§]NA = analysis not possible due to sample size constraints.