Commentary

Tracing the geographic origins of *Plasmodium falciparum* malaria parasites

Xin-zhuan Su

Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA

Invited Commentary on 'A barcode of organellar genome polymorphisms identifies the geographic origin of Plasmodium falciparum *strains', Preston* et al.

The genome of *Plasmodium falciparum* 3D7 parasite has been sequenced, and genes from the parasite genome have been used extensively in vaccine development and other studies. One of the questions frequently asked in the malaria research community is where does the 3D7 parasite come from? The parasite was derived from the NF54 isolate initially obtained from a patient living near Schiphol Airport, Amsterdam, who had never left the Netherlands,[1] and the origin of this parasite has been a mystery. A recent publication by Preston *et al.* in *Nature Communications*[2] suggests that the 3D7 parasite originated from Africa.

In the study, the authors analyzed the mitochondrion (*mt*) and apicoplast (*apico*) genomes of 711 *P. falciparum* isolates from 14 countries in West Africa, East Africa, Southeast Asia, Oceania, and South America, including the 3D7 parasites, and identified 151 and 488 single nucleotide polymorphisms (SNPs) from the *mt* and *apico* genomes, respectively. Clustering analyses of the SNPs revealed grouping of parasites by their geographic origin, similar to results using genome-wide microsatellites,[3] genome-wide SNPs[4] or *mt* polymorphisms.[5] The dataset used in this study was larger than the dataset used in the previous analysis of the *mt* genome, and the lack of recombination in *mt* and *apico* genomes provided greater power for identifying the parasite origins than genetic makers from the nuclear genome. Among the 290 distinct haplotypes generated from the combined 639 SNPs, 282 (97.2%) were observed in one region only, which allowed prediction of parasite origins with ~95% accuracy. The authors then identified 23 SNPs that had high prediction value and developed a SNP barcode that had a predictive accuracy of 92%. Haplotype analysis using the barcoded SNPs clearly placed the 3D7 parasite with African parasites. The absence of parasites from India, Central America, Southern Africa, and the Caribbean in this study indicates that these regions cannot be unequivocally ruled out as a source of 3D7; however, previous studies clearly showed that 3D7 is markedly different from parasites of Central America or South Africa.[3,5] The haplotype analysis also supports a hypothesis of *P. falciparum* radiation from Africa.[5]

In addition to the barcode for parasite identification, the large dataset also allowed testing of recombination of the maternally inherited elements. Interestingly, the two organelles were found to be mostly co-inherited, supporting previous theoretical prediction based on parasite gametocytogenesis. Intriguingly, the two co-inherited organelles appeared to experience different selection pressures. The *mt* genome had more synonymous (S) than non-synonymous (NS) substitutions, whereas the *apico* had more NS than S substitutions. Pressures from anti-malarial drugs may play a role in the high NS/S ratio of the *apico* SNPs, although these differences may also point to other unknown mechanisms.

This SNP barcode will be useful for tracing parasite origins. Building on the promising results from this study, it would be useful to collect and analyze more samples from additional endemic regions such as India and Central America to establish a global database for tracing parasite origins of future outbreaks. For discriminating local isolates, additional markers of high variability such as microsatellites may be required.

Correspondence to: X.-z. Su, Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA. Email: xsu@niaid.nih.gov

Acknowledgements

This work was supported by the Division of Intramural Research at the National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health, USA. The author thanks Dr Deirdre Joy for comments and Cindy Clark of the NIH Library for editing.

References

1 Walliker D, Quakyi IA, Wellems TE, McCutchan TF, Szarfman A, London WT, *et al.* Genetic analysis of the human malaria parasite *Plasmodium falciparum*. Science. 1987;236:1661–6.
2 Preston MD, Campino S, Assefa SA, Echeverry DF, Ocholla H, Amambua-Ngwa A, *et al.* A barcode of organellar genome polymorphisms identifies the geographic origin of *Plasmodium falciparum* strains. Nat Commun. 2014;5:4052.

DOI 10.1179/2047772414Z.000000000225

3 Wootton JC, Feng X, Ferdig MT, Cooper RA, Mu J, Baruch DI, *et al*. Genetic diversity and chloroquine selective sweeps in *Plasmodium falciparum*. Nature. 2002;418:320–3.
4 Miotto O, Almagro-Garcia J, Manske M, Macinnis B, Campino S, Rockett KA, *et al*. Multiple populations of artemisinin-resistant *Plasmodium falciparum* in Cambodia. Nat Genet. 2013;45:648–55.
5 Joy DA, Feng X, Mu J, Furuya T, Chotivanich K, Krettli AU, *et al*. Early origin and recent expansion of *Plasmodium falciparum*. Science. 2003;300:318–21.