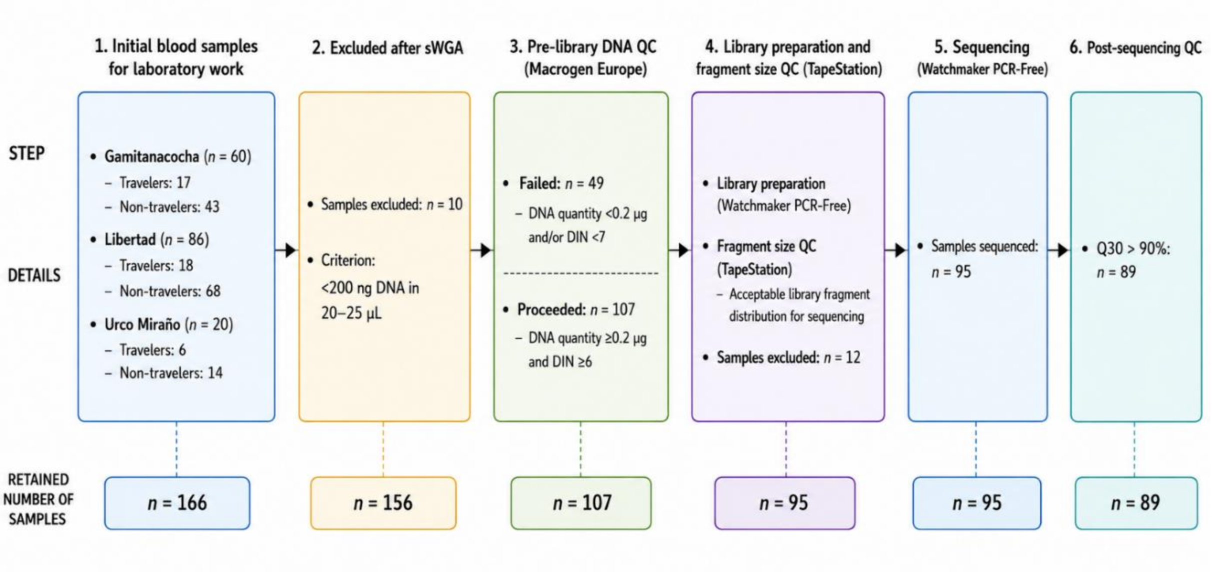

Figure S1: Flowchart summarizing sample selection, quality control, and whole-genome sequencing workflow from initial blood sample collection to post-sequencing quality assessment.

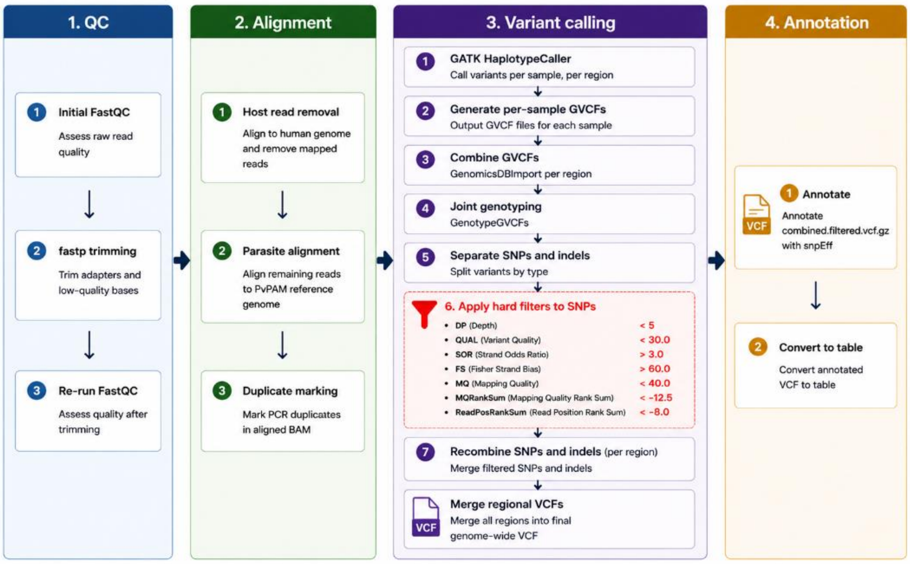

Figure S2: Overview of the bioinformatics pipeline for whole-genome sequencing data analysis

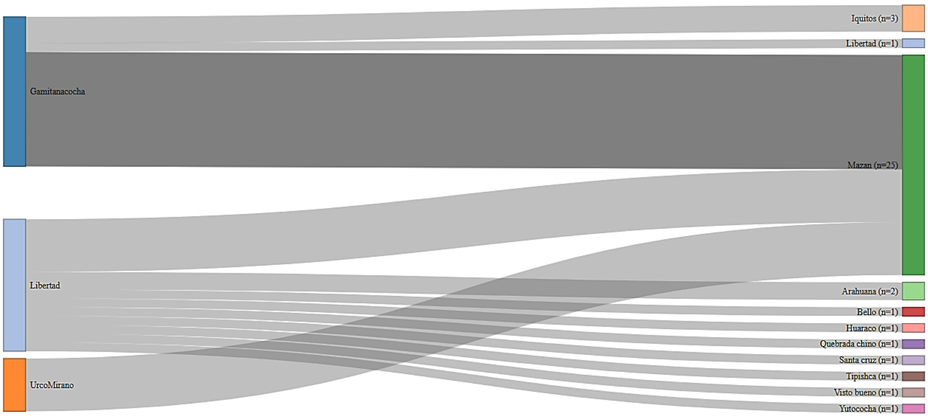

Figure S3: Travel routes of malaria-infected individuals. Sankey diagram shows reported travel routes of malaria-infected individuals from the three study communities in the Mazan district. The width of each flow is proportional to the number of reported travel events between the community of residence and the travel destination.

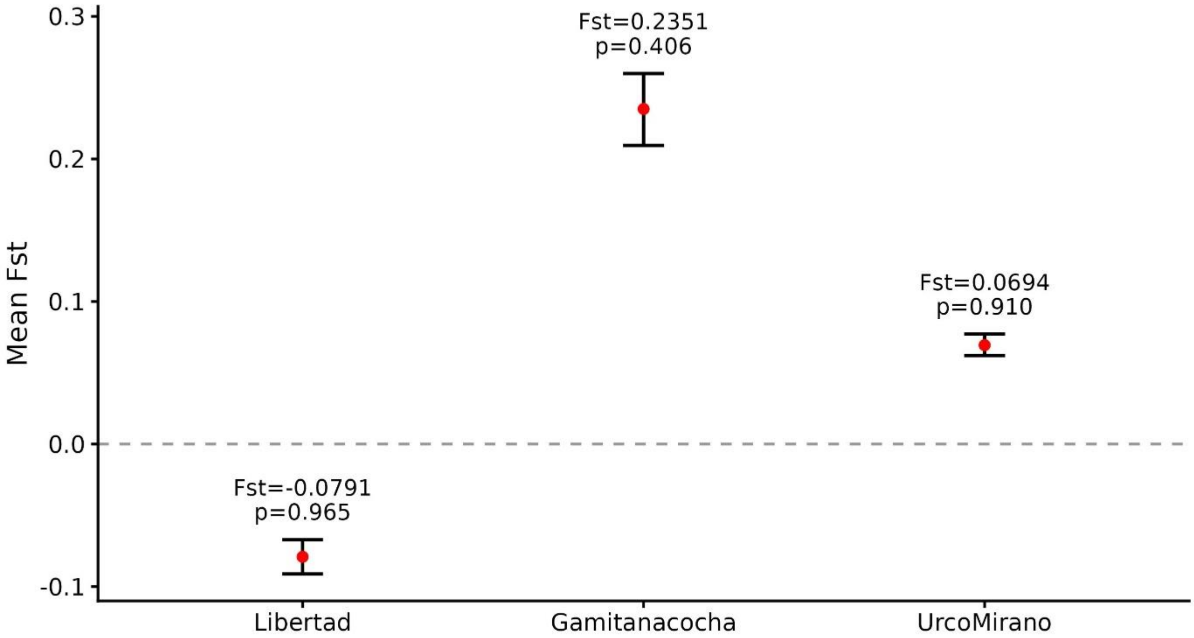

Figure S4: Within-community Fst between travelers and non-travelers. Red circles indicate point estimates, and black bars show 95% bootstrap confidence intervals across loci. The dashed horizontal line indicates Fst = 0.

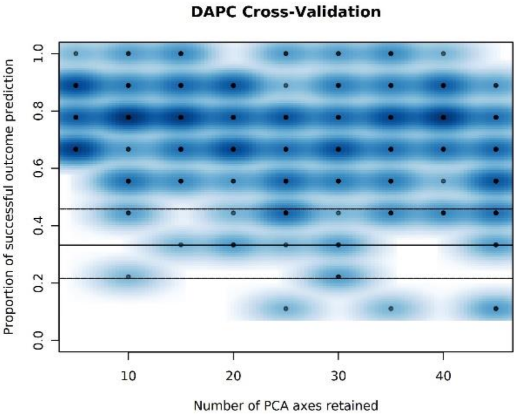

Figure S5: Cross-validation results for discriminant analysis of principal components (DAPC).

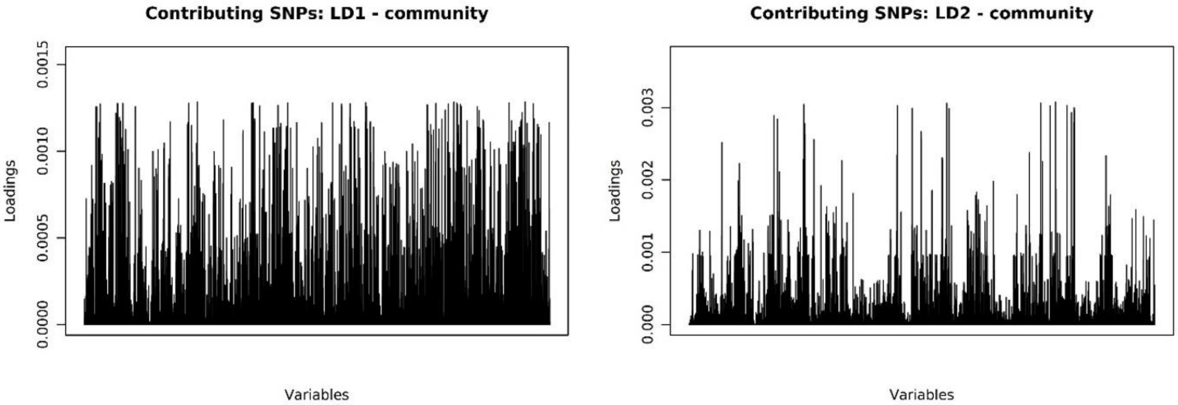

Figure S6: NP loading plot for the first and second discriminant function (LD1 and LD2) from the DAPC analysis.

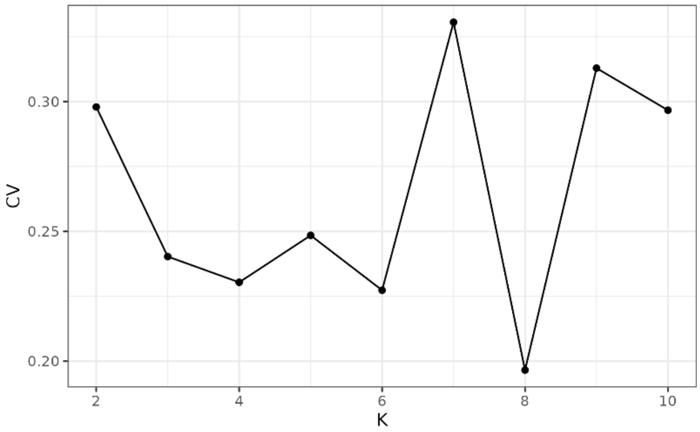

Figure S7: Cross-validation (CV) error plot from ADMIXTURE analysis showing the relationship between the number of ancestral populations (K) and model CV error.

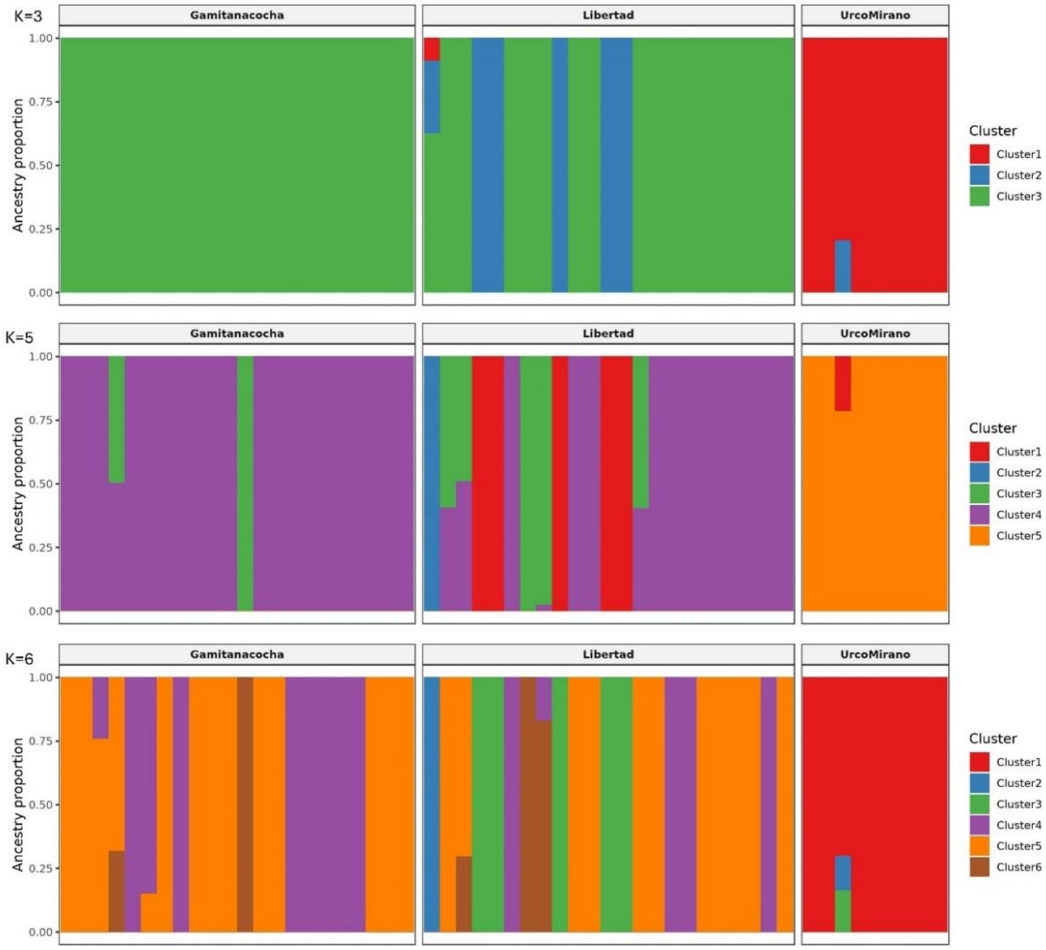

Figure S8: ADMIXTURE analysis for the unpruned dataset at K = 3, K=5, and K = 6. Each vertical bar represents one *P. vivax* isolate, and colors indicate the estimated ancestry proportion assigned to each inferred genetic cluster.

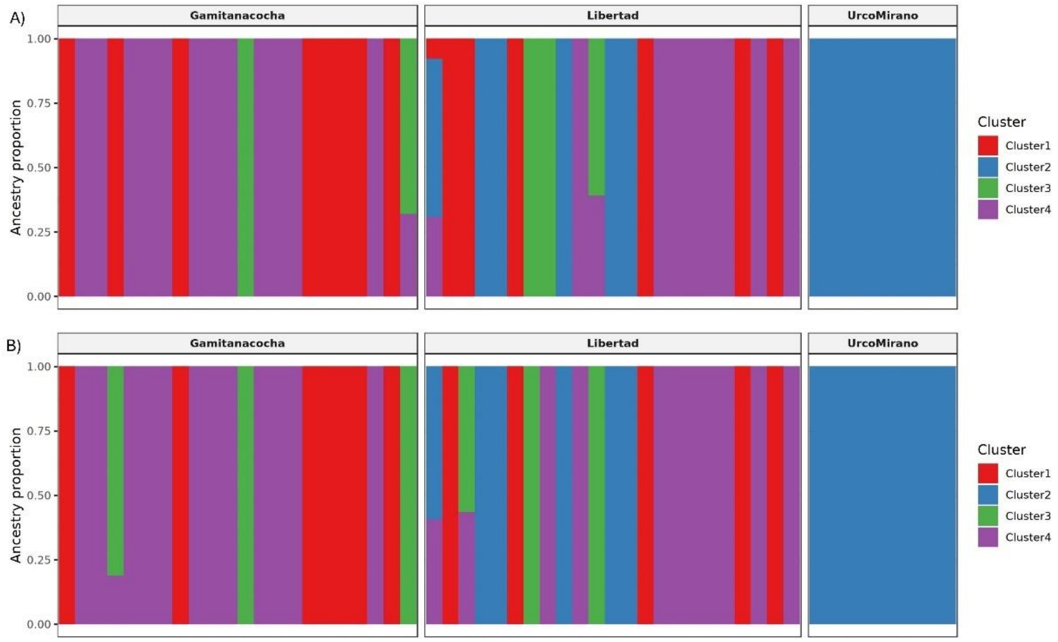

Figure S9: ADMIXTURE analysis for the pruned dataset at K = 4. (A) LD-pruned dataset with $r^2 = 0.1$ and (B) LD-pruned dataset with $r^2 = 0.2$. Each vertical bar represents one *P. vivax* isolate, and colors indicate the estimated ancestry proportion assigned to each inferred genetic cluster.

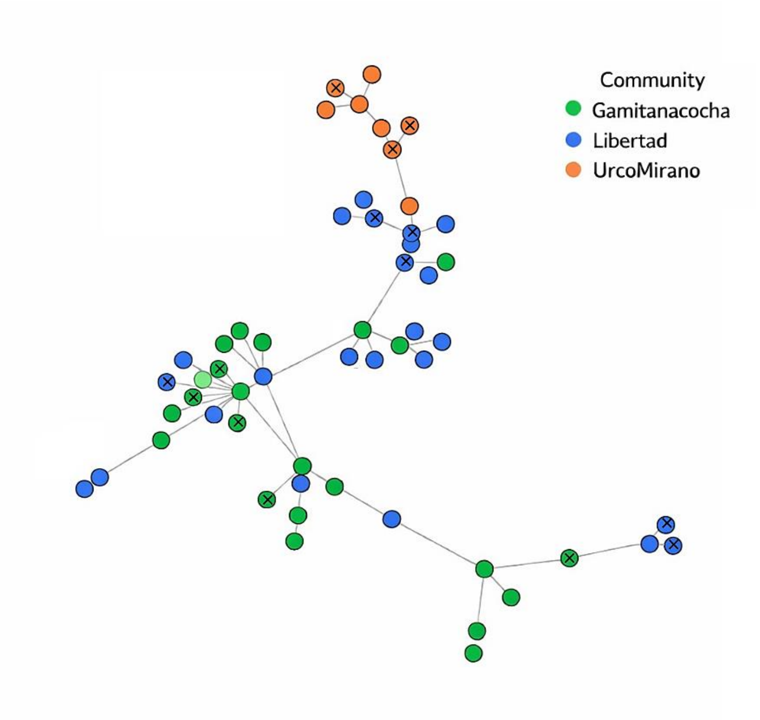

Figure S10. Minimum spanning tree (MST) showing genetic relationships among *Plasmodium vivax* infections based on SNP data. Nodes represent individual samples, and edges connect samples with the smallest pairwise genetic distance (1−PS), where shorter edges indicate higher genetic similarity. Nodes are colored by community of origin (Gamitanacocha, Libertad, and Urco MirañoUrco MirañoUrco Miraño), and crosses (x) denote traveler-associated infections.

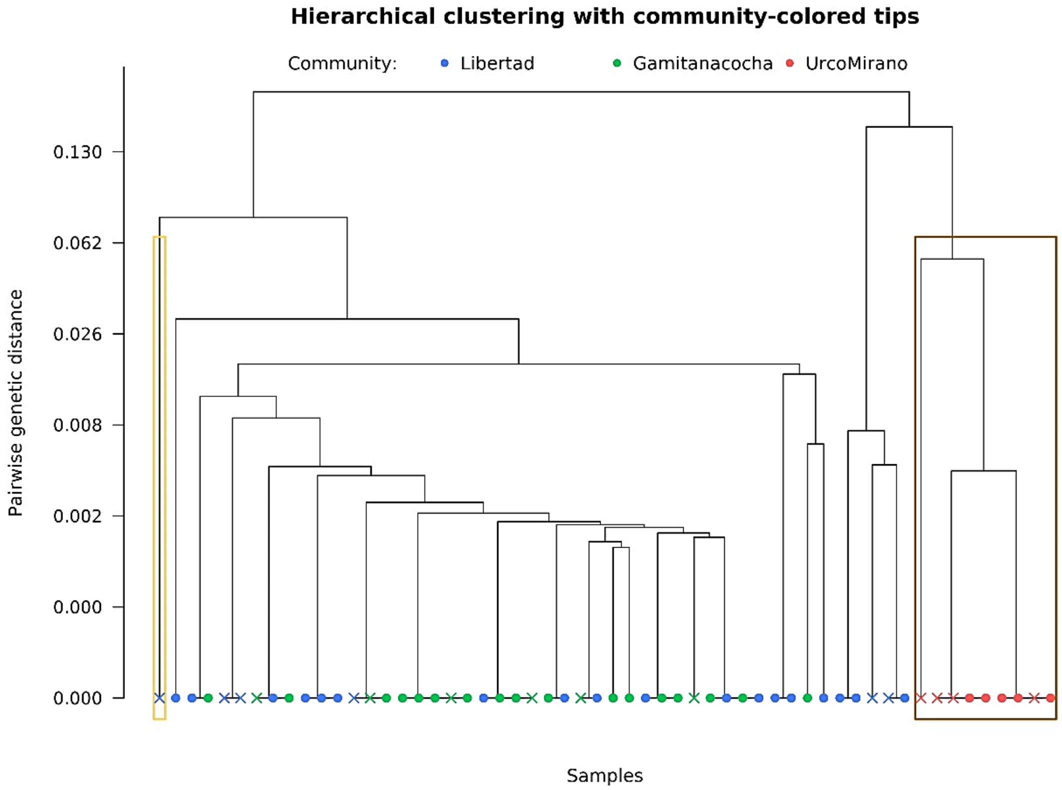

Figure S11: Hierarchical clustering of samples based on pairwise genetic distance. Tips are colored by community (Libertad, Gamitanacocha, and Urco Miraño), and crosses (x) indicate traveler samples. Clusters are defined using an optimal K = 4. The highlighted cluster on the left corresponds to a traveler from Libertad forming a distinct isolated cluster, whereas the highlighted cluster on the right consists only Urco Miraño samples.