Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Neafsey DE, Juraska M, Bedford T, et al. Genetic diversity and protective efficacy of the RTS,S/AS01 malaria vaccine. N Engl J Med 2015;373:2025-37. DOI: 10.1056/NEJMoa1505819

Supplementary Appendix

Genetic Diversity and Protective Efficacy in a Phase 3 Trial of the RTS,S/AS01 Malaria Vaccine

TABLE OF CONTENTS

LIST OF INVESTIGATORS

Albert Schweitzer Hospital, Lambaréné, Gabon/Institute of Tropical Medicine, University of Tübingen, Tübingen, Germany:
Selidji Todagbe Agnandji, M.D., Bertrand Lell, M.D., Peter G. Kremsner, M.D.

Broad Institute of MIT and Harvard, Cambridge, USA:
Scott Anderson, B.S., Bruce W. Birren, Ph.D., Kristen M. Connolly, B.S., Allison D. Griggs, M.Sc., Jonna Grimsby, Ph.D., Niall J. Lennon, Ph.D., Eli L. Moss, B.S., Daniel E. Neafsey, Ph.D., Daniel J. Park, Ph.D., Dana Robbins, B.S., Carsten Russ, Ph.D., Elizabeth M. Ryan, B.S., Brian Sogoloff, B.S., Qing Yu, M.Sc.

Centro de Investigação em Saúde de Manhiça, Manhiça, Mozambique/Barcelona Centre for International Health Research (CRESIB), Universitat de Barcelona, Hospital Clinic, Barcelona, Spain:
Carlota Dobaño, Ph.D., Jahit Sacarlal M.D., Ph.D., Pedro Aide, M.D., Ph.D., John J. Aponte, M.D., Ph.D.

Fred Hutchinson Cancer Research Center, Seattle, USA:
Trevor Bedford, Ph.D., Peter B. Gilbert, Ph.D., Michal Juraska, Ph.D.

GlaxoSmithKline Vaccines, Rixensart, Belgium:
Marc Lievens, M.Sc., Amanda Leach, M.R.C.P.C.H., Didier Lapierre, M.D., Myriam Bruls, M.Sc.

Harvard T.H. Chan School of Public Health, Boston, USA:
Karell Pellé, Ph.D., Clarissa Valim, M.D., Sc.D., Sarah K. Volkman, Sc.D., Dyann F. Wirth, Ph.D.

Ifakara Health Institute, Bagamoyo, Tanzania/Swiss Tropical and Public Health Institute, Basel, Switzerland:
Salim Abdulla, M.D., Ph.D., Jackson T. Molel, Ms.C., Marcel Tanner, Ph.D., M.P.H., Ms.C.

Institut de Recherche en Sciences de la Santé, Nanoro, Burkina Faso/Institute of Tropical Medicine, Antwerp, Belgium:
Halidou Tinto, Ph.D., Hermann Sorgho, Ph.D., Innocent Valea, Pharm.D., M.Sc., Ph.D.

KEMRI/CDC Research and Public Health Collaboration, Kisumu, Kenya:
Mary Hamel, M.D., D.T.M.&H., Simon Kariuki, Ph.D., Kephas Otieno, M.P.H.

KEMRI–Walter Reed Project, Kombewa, Kenya:
Walter Otieno, Ph.D., M.B.Ch.B.,M.Med, Lucas Otieno, M.B.Ch.B, M.P.H., Bernhards R. Ogutu, Ph.D.

KEMRI–Wellcome Trust Research Program, Kilifi, Kenya:
Philip Bejon, Ph.D., M.B.Ch.B., M.Med., Kevin Marsh, M.D., Patricia Njuguna, M.Med.

Kintampo Health Research Centre, Kintampo, Ghana/London School of Hygiene and Tropical Medicine, London, UK:

Seth Owusu-Agyei, M.Sc., Ph.D., Kwaku Poku Asante, M.D., M.P.H., Ph.D, Brian Greenwood, M.D.

Medical Research Council Unit, Banjul, The Gambia:
Umberto D'Alessandro, M.D., Ph.D.

National Institute for Medical Research, Korogwe, Tanzania/University of Copenhagen, Copenhagen, Denmark/London School of Hygiene and Tropical Medicine, London, UK:
John Lusingu, M.D., Ph.D., Samwel Gesase, M.D., M.Sc., Thor G. Theander, M.D., D.Sc.

PATH Malaria Vaccine Initiative, Washington DC, USA:
Ashley J. Birkett, Ph.D., Chris F. Ockenhouse, M.D., Ph.D.

School of Medical Sciences, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana: Tsiri Agbenyega, M.B., Ch.B., Ph.D. Daniel Ansong, M.B.Ch.B., F.W.A.C.P., Samuel Adjei, M.B.Ch.B., D.T.M.

University of North Carolina Project, Lilongwe, Malawi: Francis Martinson, M.D., Ph.D., Irving Hoffman, P.A., M.P.H., Portia Kamthunzi, M.D.

University of Washington, Seattle, USA:
David Benkeser, M.S.

AUTHOR CONTRIBUTIONS

CV, DEN, SKV, DFW, ML, AL, DL, MB, MJ, TB, DCB, and PBG designed the study and analysis plan. SA(bdulla), SA(jei), SA(gnandji), PA, DA, JJA, KPA, PB, SG, BG, TH, IH, PK, SK, PGK, BL, JL, KM, FM, JTM, PN, BOR, WO, LO, KO, SOA, JS, HS, MT, TT, and IV executed the trial and collected the samples. SA(nderson), KC, JG, NJL, DR, EMR, BS, QY generated the genotyping data. ADG, ELM, DEN, DJP, CV, and TB filtered the data. MJ, CV, DCB, TB, DEN, AG, and PBG analyzed the data. BWB, SKV, CR, DFW, DEN, KGP, PBG, CD, MJH, UD, TA, AJB, and CFO provided project oversight and supervision. DEN, MJ, PBG, and DFW vouch for the data and the analysis. PBG and DFW decided to publish the paper. DEN, MJ, PBG, and DFW wrote the first draft.

No agreements regarding confidentiality of the data were made between the sponsors of the study and the authors or their institutions.

SUPPLEMENTARY METHODS

We developed custom workflows at the Broad Institute to extract DNA from filter paper samples, PCR amplify the DNA, and sequence it via the Illumina MiSeq platform (for short amplicons) or the Pacific Biosciences RS II platform (for the large amplicon containing the NANP/NVDP repeat region).

SAMPLE RECEIPT AND REGISTRATION

FTA blood spot cards, aggregated and shipped from the Quintiles office (Lyttelton Manor, South Africa), were received by the Broad Institute Genomics Platform (GP) under CDC import permit 201305069, tagged with GP sample specific barcodes, both GP and GSK barcodes were registered in the GP laboratory information management system (LIMS), and cards were stored in desiccator cabinets at room temperature until further processing. Samples were held in a room requiring a Broad Institute identification card for access.

DNA EXTRACTION AND QUANTITATION

Genomic DNA extraction was performed in batches of 96 Whatman FTA cards, including one blank control FTA card. For each FTA card 7 disks were punched out of the blood spot, using an automated laser guided hole puncher, into a distinct well of 96 well plates. Genomic DNA was extracted from the punches using the automated Chemagen Chemagic bead-based DNA extraction platform using standard protocols. DNA samples were registered in LIMS and stored in barcoded tubes. DNA concentration of each sample was quantified using standard automated PicoGreen quantification. All steps of the process were tracked in the LIMS.

PCR

Four plasmodium PCR amplicons, “CS C-terminus”, “SERA2”, “TRAP” and “NANP/NVDP repeat”, were amplified in 5,160 samples. Two of these targets, CSP C-terminus and NANP/NVDP repeat, are located in the circumsporozoite (*CS*) gene. The CSP C-terminus amplicon captures the polymorphic C terminus T cell epitope region of *CS.* The NANP/NVDP repeat amplicon captures the NANP/NVDP repeat region of *CS*. TRAP and SERA2 are not located within *CS*, but are targets that fall within other polymorphic regions of the *P. falciparum* genome and served as control amplicons for data analysis. The SERA2 locus was used as a control for sieve analyses. Final haplotype calls from the CS C-terminus, NANP/NVDP, SERA2 and TRAP amplicons were all used to estimate complexity of infection (COI) for individual samples. Full amplicon sizes (including adapter sequence, flow cell attachment sequences and indices) were 400, 427, 371 and 742 bp for the CS C-terminus, TRAP, SERA2 and NANP/NVDP repeat amplicons, respectively. Plasmodium portions of these amplicons were 333, 360, 304 and 700 bp for CS C-terminus, TRAP, SERA2 and NANP/NVDP repeat amplicons, respectively. The formal *P. falciparum* 3D7 gene IDs and nucleotide coding sequence (CDS) coordinates for these amplicons are as follows:

CS C-terminus (*CS*; PF3D7_0304600): CDS bp 858-1190

NANP/NVDP repeat (*CS*; PF3D7_0304600): CDS bp 168 - 867

TRAP (*trap*; PF3D7_1335900): CDS bp 1,222 - 1,581

SERA2: (*sera2*: PF3D7_0207900): CDS bp 72 - 357

The CS C-terminus, TRAP and SERA2 amplicons were sequence-ready constructs and did not require further library construction after PCR. These PCRs were carried out in two reactions. Round 1 PCR primers contained *Plasmodium* sequence and Illumina adapter sequences while round 2 PCR primers were “tailing” primers, containing some

overlap of the Illumina adapter sequence, flow cell attachment sequences, and an eight bp index on the reverse primer between the adapter sequence and flow cell attachment sequence (primer sequences below). Amplification of NANP/NVDP repeat was carried out in one reaction with primers containing plasmodium sequence, 10 bp indices on both forward and reverse primers, and a 5 bp buffer sequence (GGTAG).

First-round PCRs for CS C-terminus, TRAP and SERA2 were carried out using the Hot Star Plus DNA Polymerase Kit (Qiagen). Reactions consisted of 5 μl DNA at ~0.5 ng/μl, 10 μl mixed F/R primer (1.0 μM for CS C-terminus and TRAP, 2.0 μM for SERA), 2μl 10X buffer, 0.8 μl 25 mM MgCl2, 0.16 μl dNTPs (100 mM dNTP mix, Agilent Technologies), 0.08 μl HotStar Taq (5U/μl), 3.96 μl nuclease free water. Thermal cycling consisted of 95˚C for 5 min, 30 cycles of [94˚C 30 sec, 60˚C 30 sec, 72˚C 1 min] and 3 min at 72˚C. Second-round PCRs for CS C-terminus, TRAP and SERA2 consisted of 1 μl of PCR1 product, 3.13 μl nuclease free water, 11.72 μl Pfu Buffer, 0.12 μl Pfu DNA polymerase and 10 μl mixed F/R indexed primer (1.6μM). Second-round PCR thermal cycling for CS C-terminus and TRAP consisted of 50˚C for 2 min, 70˚C for 20 min, 95˚C for 10 min, 5 cycles of [95˚C 15 sec, 60˚C 30 sec, 72˚C 1 min], 1 cycle of [95˚C 15 sec, 80˚C 30 sec, 60˚C 30 sec, 72˚C 1 min], 4 cycles of [95˚C 15 sec, 60˚C 30 sec, 72˚C 1 min], 1 cycle of [95˚C 15 sec, 80˚C 30 sec, 60˚C 30 sec, 72˚C 1 min], 4 cycles of [95˚C 15 sec, 60˚C 30 sec, 72˚C 1 min], 5 cycles of [95˚C 15 sec, 80˚C 30 sec, 60˚C 30 sec, 72˚C 1 min]. Second-round PCR thermal cycling for SERA2 consisted of 95˚C for 5 min, 9 cycles of [94˚C 30 sec, 60˚C 30 sec, 72˚C 1 min] and 72˚C for 3 min. PCR for the CSP-repeat amplicon consisted of 5 μl DNA at 0.5 ng/μl, 11.72 μl Pfu buffer, 0.12 μl Pfu DNA polymerase, and 10 μl of mixed F/R NANP/NVDP repeat indexed primers (2.0 μM). Thermal cycling for CSP-repeat consisted of 95˚C for 2 min, 5 cycles of [95˚C 30 sec,

58˚C 30 sec, 72˚C 30 sec], 30 cycles of [95˚C 30 sec, 65˚C 30 sec, 72˚C 30 sec] and 72˚C for 10 min.

Samples were batched into groups of 192 (2 sets of 96), with a negative control sample (originating from a blank blood-spot card) included within each set of 96. For each batch of 192 (28+ batches total), 14 96-well plate PCRs were carried out in total (8 PCR1 [CS C-terminus, TRAP, SERA2, NANP/NVDP repeat set 1, set 2], 6 PCR2 [CS C-terminus, TRAP, SERA2, set 1, set 2]). A sampling of CS C-terminus, TRAP, SERA2 PCR2 and NANP/NVDP repeat products were visually inspected using a Lab Chip GX II Caliper Instrument (Perkin Elmer). Indices for sample identification were assigned during PCR so that, within a batch, the same sample was assigned the same index for CS C-terminus, TRAP and SERA2 amplicons. For each batch of 192 samples, indexed CS C-terminus PCR2 products were pooled by volume, as were TRAP PCR2 and SERA2 PCR2 products. NANP/NVDP amplicons for each batch were also pooled by volume. These 4 amplicon pools were purified using a 0.7X solid-phase reversible immobilization (SPRI) cleanup with Agencourt Ampure XP beads (Beckman Coulter). CS C-terminus, TRAP, SERA2 and NANP/NVDP repeat product pools for each batch were then assessed and quantified on a BioAnalyzer (Agilent Technologies). For automated PCR set-up, pooling, LIMS tracking and messaging, a Bravo Automated Liquid Handling Platform (Agilent Technologies) was used. To avoid PCR contamination, automated setup of PCR2 included tip piercing of PCR1 plate and primer plate covers to avoid amplicon spray going into nested PCR. In addition, PCR workspaces were decontaminated with DNA ZAP (Ambion) and negative control wells were visually inspected on the Lab Chip GX II Caliper Instrument.

PRIMER SEQUENCES

Round 1 PCR primers (CS C-terminus, TRAP, SERA2; *Plasmodium* sequence in bold; X indicates positions of sample-specific barcode sequences):

CS C-terminus_Round_1_Forward:
ACACTCTTTCCCTACACGACGCTCTTCCGATCT**TTAAGGAACAAGAAGGATAATACCA**
CS C-terminus_Round_1_Reverse:
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**AAATGACCCAAACCGAAATG**

TRAP_Round_1_Forward:
ACACTCTTTCCCTACACGACGCTCTTCCGATCT**TCCAGCACATGCGAGTAAAG**
TRAP_Round_1_Reverse:
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**AAACCCGAAAATAAGCACGA**

SERA2_Round_1_Forward:
ACACTCTTTCCCTACACGACGCTCTTCCGATCT**TACTTTCCCTTGCCCTTGTG**
SERA_Round_1_Reverse:
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**CACTACAGATGAATCTGCTACAGGA**

Round 2 PCR Primers (CS C-terminus, TRAP, SERA2):

Round2_Forward:
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
Round 2_Reverse:
CAAGCAGAAGACGGCATACGAGATXXXXXXXXGTGACTGGAGTTCAGACGTGTGCTCTTCC
GATCT

NANP/NVDP repeat Primers:
Forward Primer GGTAGXXXXXXXXXXXXXXXX**TGGGTCATTTGGCATATTGT**
Reverse Primer GGTAGXXXXXXXXXXXXXXXX**TGGGAAACAGGAAAATTGGT**

SEQUENCING

For each batch, CS C-terminus, TRAP and SERA2 PCR2 product pools were normalized and combined into a single pool for MiSeq sequencing. The library was quantified by SYBR green qPCR before cluster generation, and one MiSeq run (2x250 bp paired end with standard sequencing primers) was carried out for each sample batch using standard methods (V2 sequencing chemistry). PhiX library, derived from the well-characterized and small PhiX genome, was spiked in at 15% to add diversity for improved cluster imaging. Sequencing data were processed through the Broad Picard sequencing analysis pipeline generating standard sequencing metrics (e.g. reads counts) and demultiplexed sample specific sequencing read BAM files. BAM files were then screened and filtered for human contaminating sequences, and packaged for submission to the NCBI short read archive database. All steps were tracked in the LIMS.

The pools of indexed CSP-repeat amplicons underwent standard PacBio library construction and 3 SMRT cells were run per library (P4C2 sequencing chemistry, 180-minute movie). Sequencing data were processed through Pacific Biosciences and Broad-developed pipelines generating standard sequencing metrics and demultiplexed sample specific barcode-stripped sequencing read files. Read files were converted to the BAM file format, screened and filtered for human contaminating sequences, and packaged for submission to the NCBI Short Read Archive database (BioProject PRJNA235895). All steps were tracked in a LIMS.

SEQUENCING DATA ERROR FILTRATION

Sequencing is prone to both systematic and random errors, which occur at a frequency dependent on sequencing platform. We examined a panel of 56 mock samples of known genotype (determined via Illumina whole genome sequencing) and complexity level (ranging from one to eleven strains) that was subjected to the described PCR and sequencing protocol in order to characterize the error profiles of the Miseq and PacBio sequencing platforms. The raw Miseq data exhibited an abundance of very low frequency random base call errors, manifested as haplotypes present at very low frequencies in a read set that differ from an expected haplotype by average of one to two base changes. The raw PacBio data exhibited both base call errors and spurious insertions/deletions (indels). PacBio base calling accuracy was approximately 80%, and exhibited a mean of 5 spurious indels per read in the NANP/NVDP repeat amplicon sequences. The circular consensus (CC) reads purge much of the random error through creation of a consensus sequence, but do not result in completely clean data. The following error correction/filtration methods were therefore applied to MiSeq and PacBio data.

For MiSeq data, reads with unexpected barcode sequences (due to mutation) or an excess of low quality bases were removed by the Broad Institute's Picard pipeline[1]. Multiplexed reads were separated into files for each sample based on barcode sequences. Overlapping 250 bp mate pair reads (forward and reverse) were merged into single reads using the FLASH utility[2]. Merged reads were then aligned to the PlasmoDB v9.0 3D7 *P. falciparum* reference genome assembly using BWA version 0.7.4-r385[3] to confirm affiliation with each amplicon. Reads containing one or more uncalled bases were removed, and haplotype sequences represented by fewer than 1% of reads for a given sample were removed. Remaining haplotypes were then clustered, using a rule by which haplotypes exhibiting 1 bp differences were collapsed into a single majority consensus if the a 20-fold or greater difference in abundance was observed between major and minor haplotypes. Minority haplotypes distinguished from nearest neighbors only by differences the length of T nucleotide homopolymers were removed. Haplotypes observed only once in the entire sample set and supported by fewer than 500 reads were removed. Finally, all samples exhibiting more than two haplotypes following clustering and filtering were evaluated for PCR chimerism. If a haplotype could be expressed as a simple combination of two other haplotypes called for the same sample and was supported by fewer than 250 reads, it was eliminated from the final dataset.

Application of this pre-processing pipeline to data deriving from the validation sample panel resulted in the following performance at the whole-amplicon and epitope haplotype level for the CSP-C-terminus, TRAP, and SERA2 amplicons:

Amplicon	COI	1	2	3	4	5	6	7	8	9	10	11
All (combined)	TP	163	162	150	96	117	72	249	0	341	0	62
	FP	3	0	0	0	2	0	0	0	0	0	4
	FN	0	0	0	0	3	0	3	0	10	0	4
	Sensitivity	**1**	**1**	**1**	**1**	**0.978**	**1**	**0.989**	**NaN**	**0.974**	**NaN**	**0.943**
SERA2	COI	1	2	3	4	5	6	7	8	9	10	11
	TP	56	54	51	32	38	24	84	0	117	0	22
	FP	0	0	0	0	2	0	0	0	0	0	2
	FN	0	0	0	0	2	0	0	0	0	0	0
	Sensitivity	**1**	**1**	**1**	**1**	**0.95**	**1**	**1**	**NaN**	**1**	**NaN**	**1**
TRAP	COI	1	2	3	4	5	6	7	8	9	10	11
	TP	56	54	51	32	40	24	84	0	116	0	20
	FP	1	0	0	0	0	0	0	0	0	0	0
	FN	0	0	0	0	0	0	0	0	1	0	2
	Sensitivity	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**NaN**	**0.993**	**NaN**	**0.923**
CS C-term	COI	1	2	3	4	5	6	7	8	9	10	11
	TP	51	54	48	32	39	24	81	0	108	0	20
	FP	2	0	0	0	0	0	0	0	0	0	2
	FN	0	0	0	0	1	0	3	0	9	0	2
	Sensitivity	**1**	**1**	**1**	**1**	**0.975**	**1**	**0.964**	**NaN**	**0.923**	**NaN**	**0.909**

TP=True Positives, FP=False Positives, FN=False Negatives, Sensitivity=TP/(TP+FN), NaN='Not a Number'

For the PacBio data, multiplexed reads were separated into individual files on the basis of barcode sequences. Circular consensus reads were aligned to the 3D7 reference sequence for the NANP/NVDP repeat amplicon sequence using BLAST[4]. Reads that failed to completely span the repetitive region were excluded from further analysis. Tandem Repeats Finder[5] was used to identify and count NANP and NVDP-encoding repeats in the nucleotide sequence using a consensus-based approach to address low base calling accuracy. Repeat unit counts represented by fewer than 10% of the reads associated with a sample, or fewer than 50 reads total, were discarded.

This pipeline was applied to data deriving from the validation sample panel. Due to the difficulty of sequencing the NANP/NVDP repeat region via Sanger sequencing, we were unable to directly evaluate the absolute accuracy of the results obtained, but can evaluate the general concordance in repeat counts resulting from validation samples composed of mixtures of strains that were also amplified and sequenced individually. Using this approach, with expected haplotypes in the strain mixtures determined from

those observed in the single strains, we detected 70 (true positive) out of 83 expected haplotypes across the validation samples, for a sensitivity of 84%. We reported only 2 haplotypes that were not expected (false positives). This pipeline yields at least one repeat count from 3,137 of the clinical samples.

Analysis of repeat counts yielded by the panel of validation samples found no bias with respect to the number of repeat counts. We observed no statistically significant correlation between repeat unit count and sequencing coverage depth within samples containing a single strain (Pearson's product-moment correlation = -0.16, P = 0.48), or within samples containing multiple strains, where the opportunity for competitive amplification could exist (Pearson's product moment-correlation = 0.06, P = 0.62).

Figure S1. Sieve Analysis Schematic.

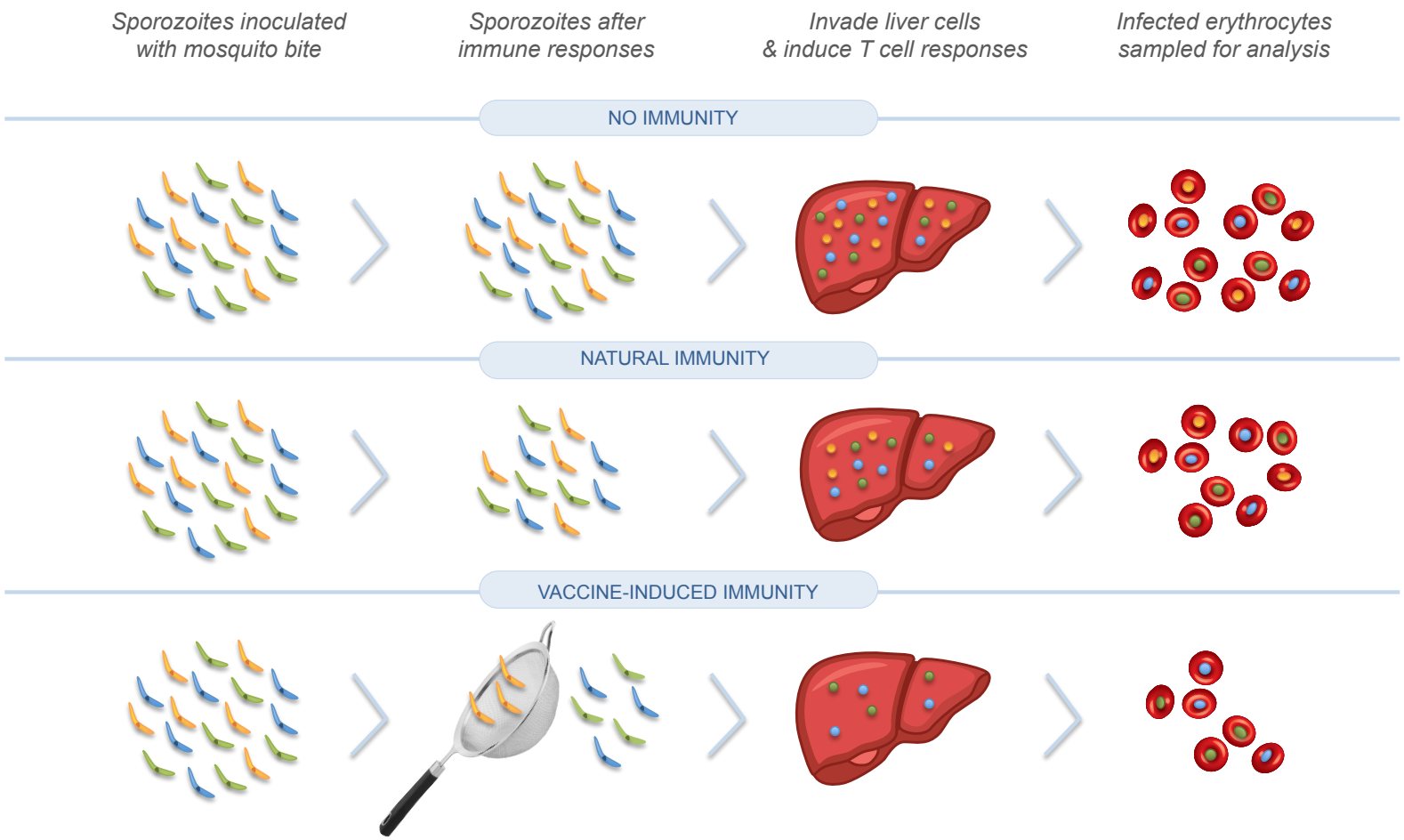


Figure S2. Data Generation and Sample/Data Filtration in Children Aged 5-17 Months for the CS C-terminus, Parasite Positive Endpoint.

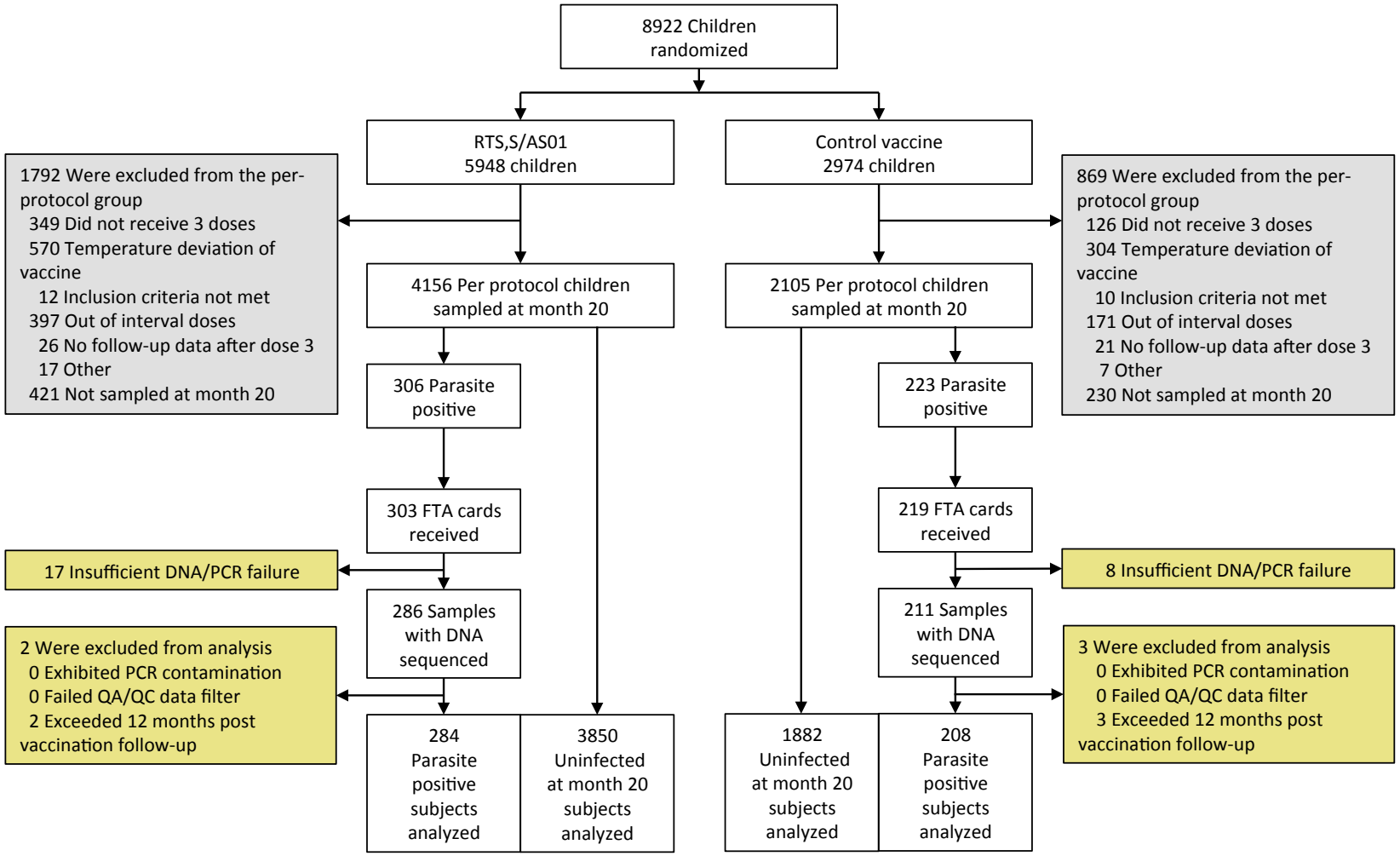


Figure S3. Data Generation and Sample/Data Filtration in Infants Aged 6-12 Weeks for the CS C-terminus, Primary Clinical Malaria Endpoint.

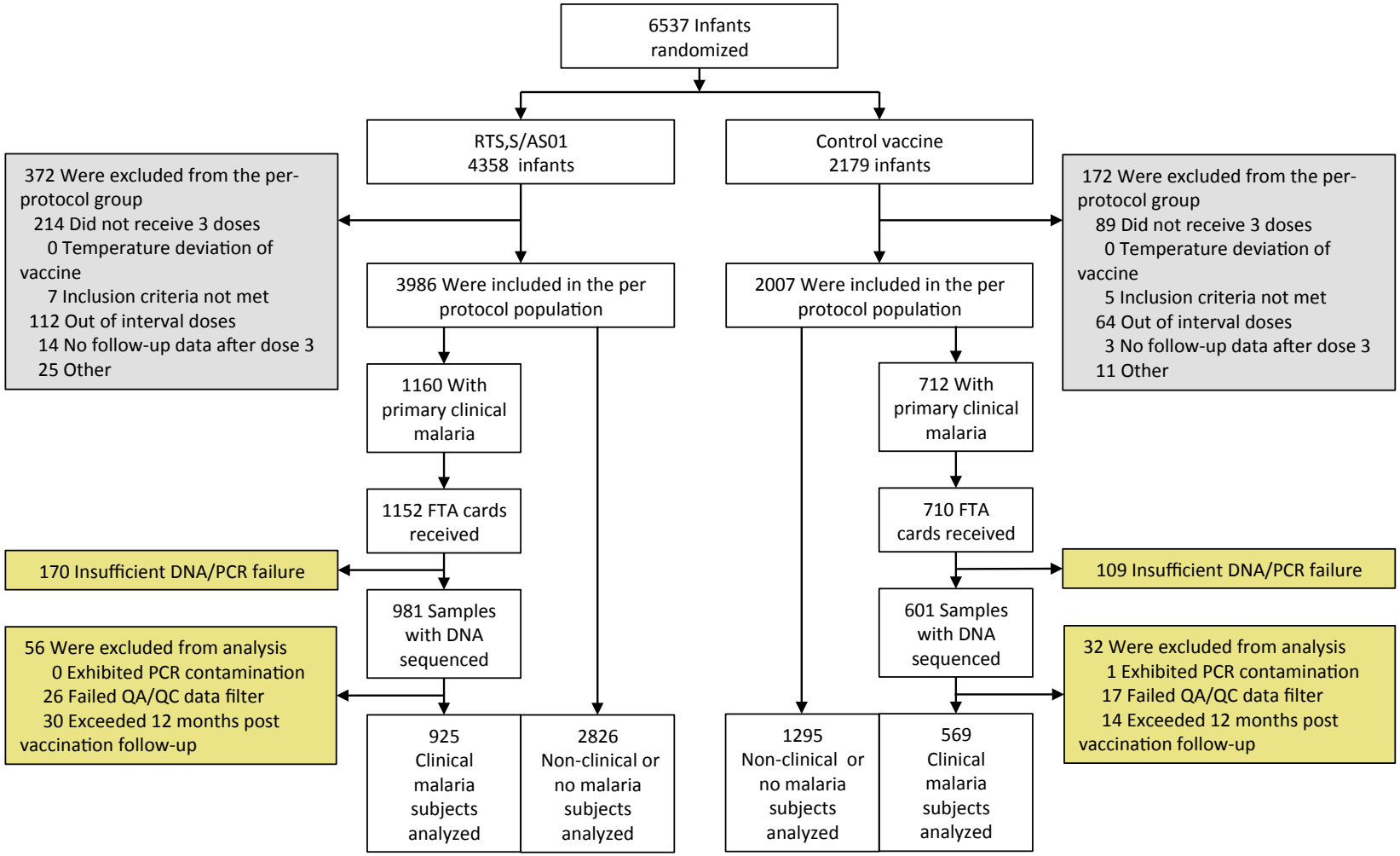


Figure S4. Data Generation and Sample/Data Filtration in Infants Aged 6-12 Weeks for the CS C-terminus, Parasite Positive Endpoint.

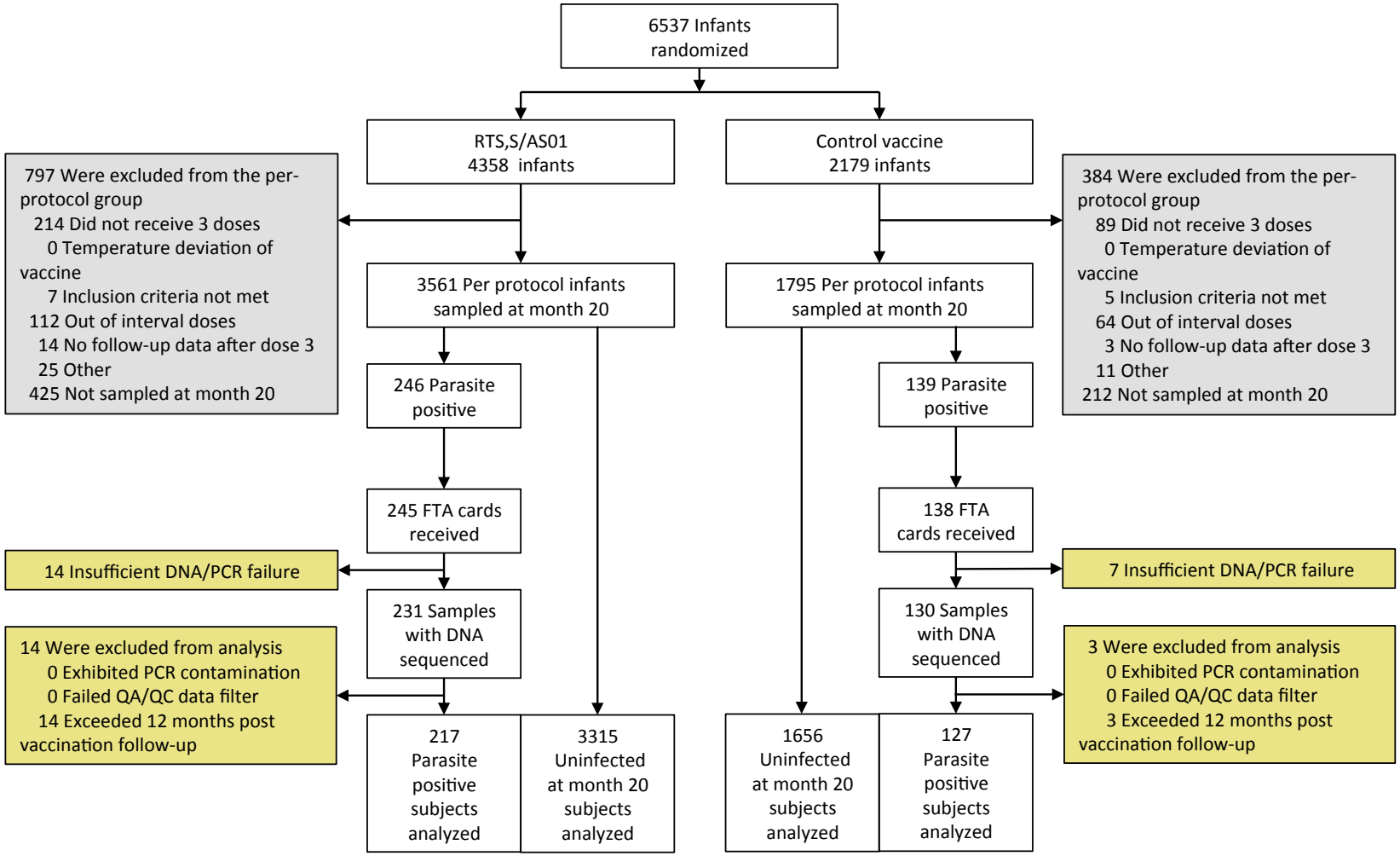


Figure S5. Data Generation and Sample/Data Filtration in Children Aged 5-17 Months for the NANP/NVDP Repeat Region, Primary Clinical Malaria Endpoint.

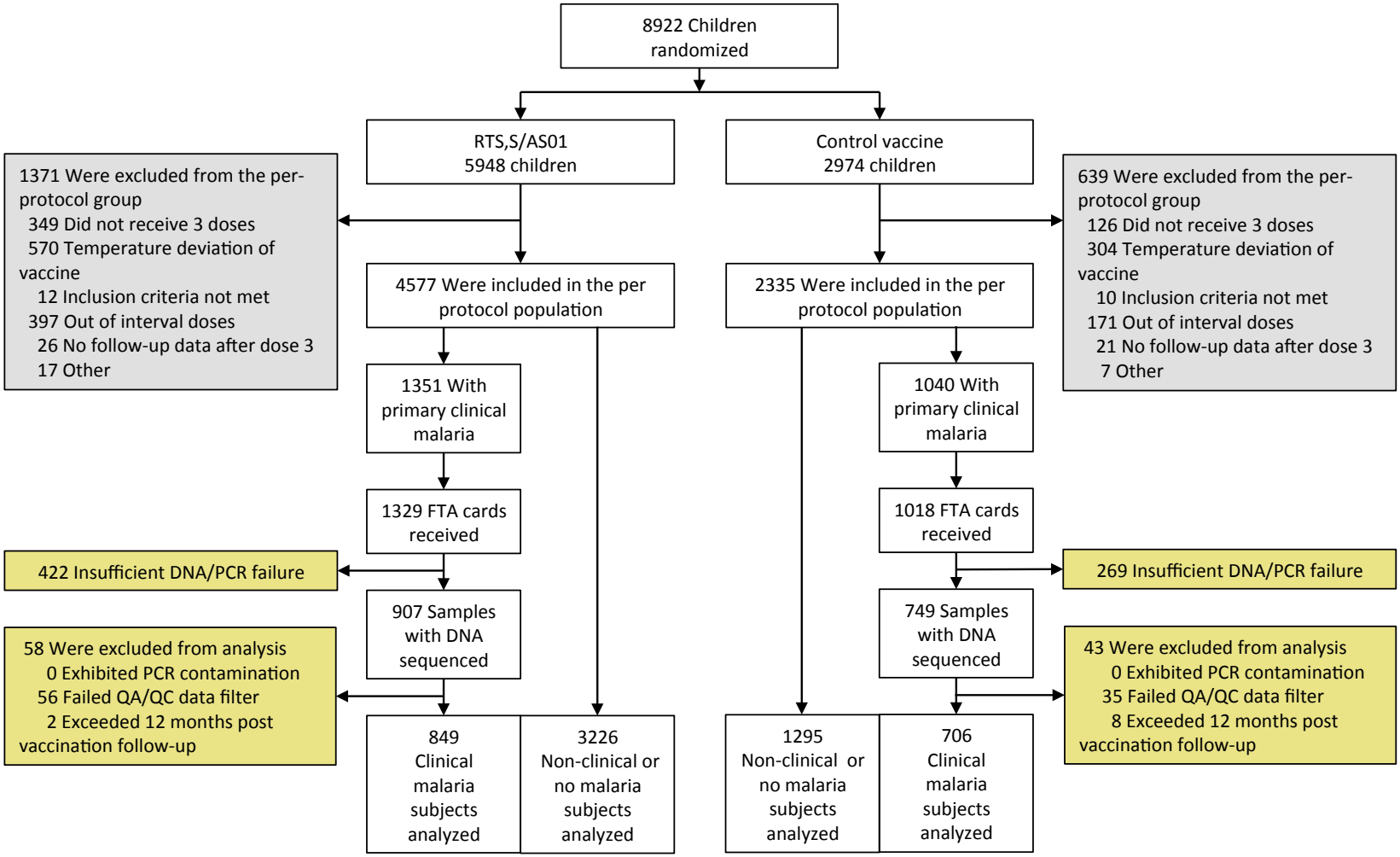


Figure S6. Data Generation and Sample/Data Filtration in Children Aged 5-17 Months for the NANP/NVDP Repeat Region, Parasite Positive Endpoint.

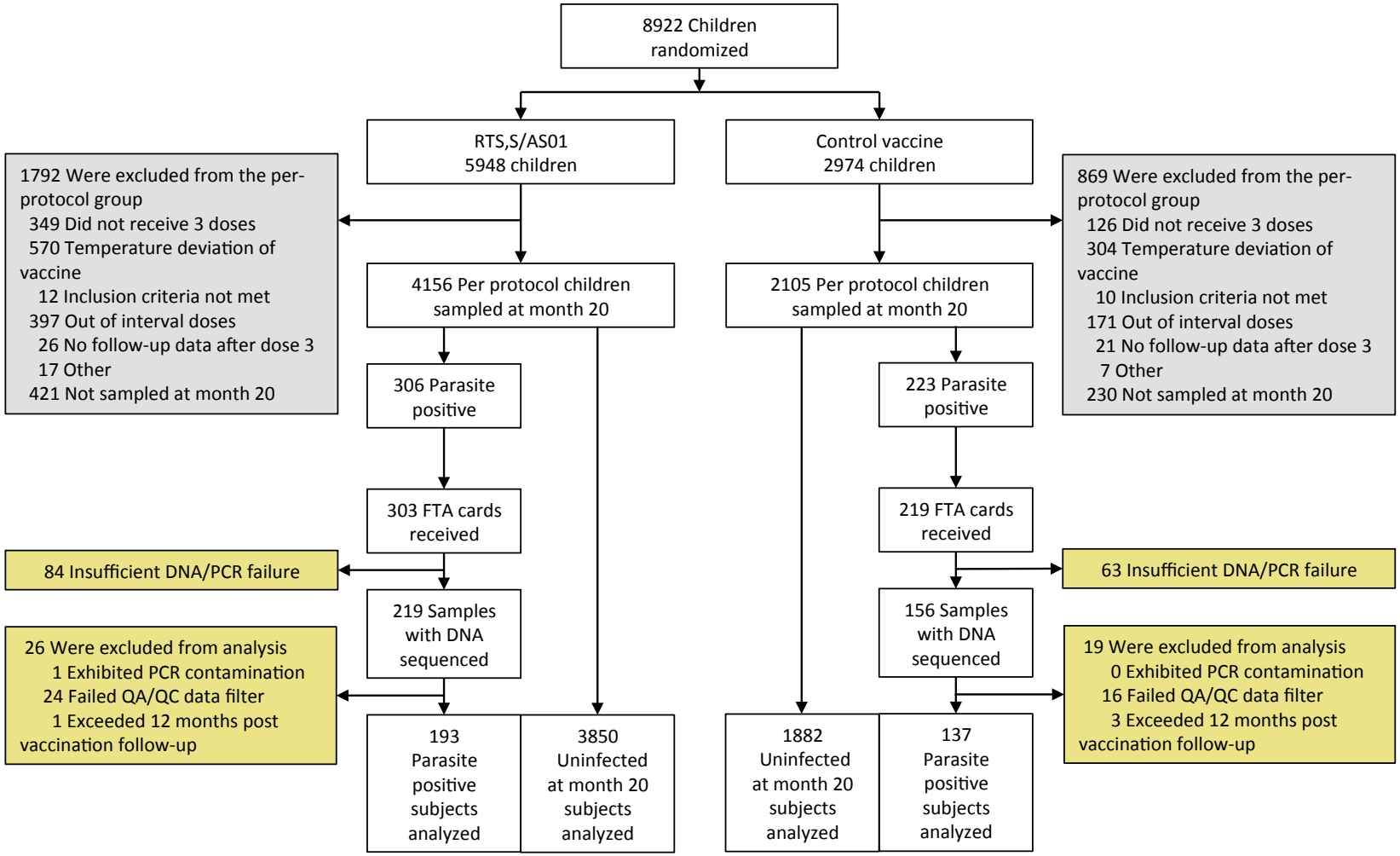


Figure S7. Data Generation and Sample/Data Filtration in Infants Aged 6-12 Weeks for the NANP/NVDP Repeat Region, Primary Clinical Malaria Endpoint.

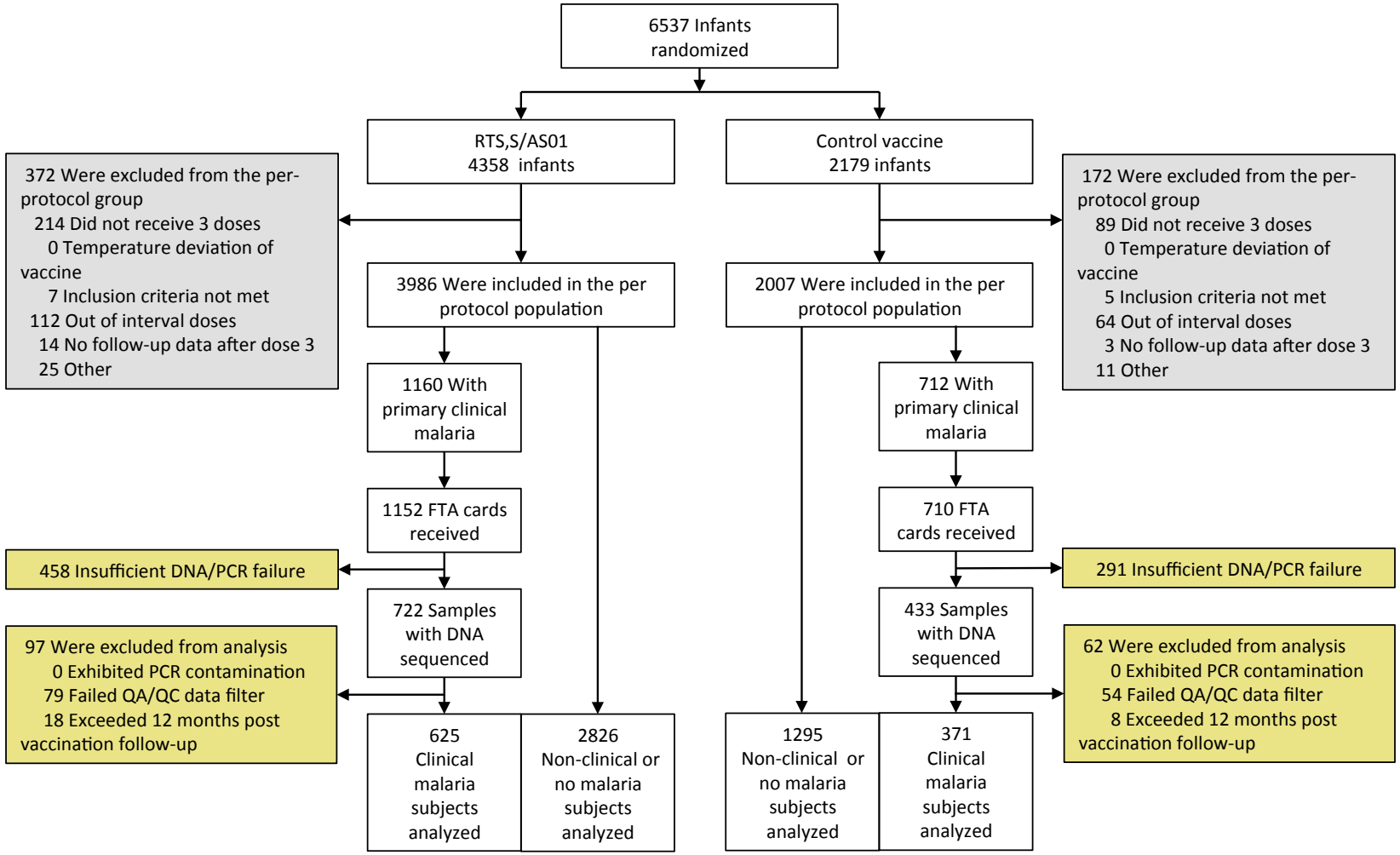


Figure S8. Data Generation and Sample/Data Filtration in Infants Aged 6-12 Weeks for the NANP/NVDP Repeat Region, Parasite Positive Endpoint.

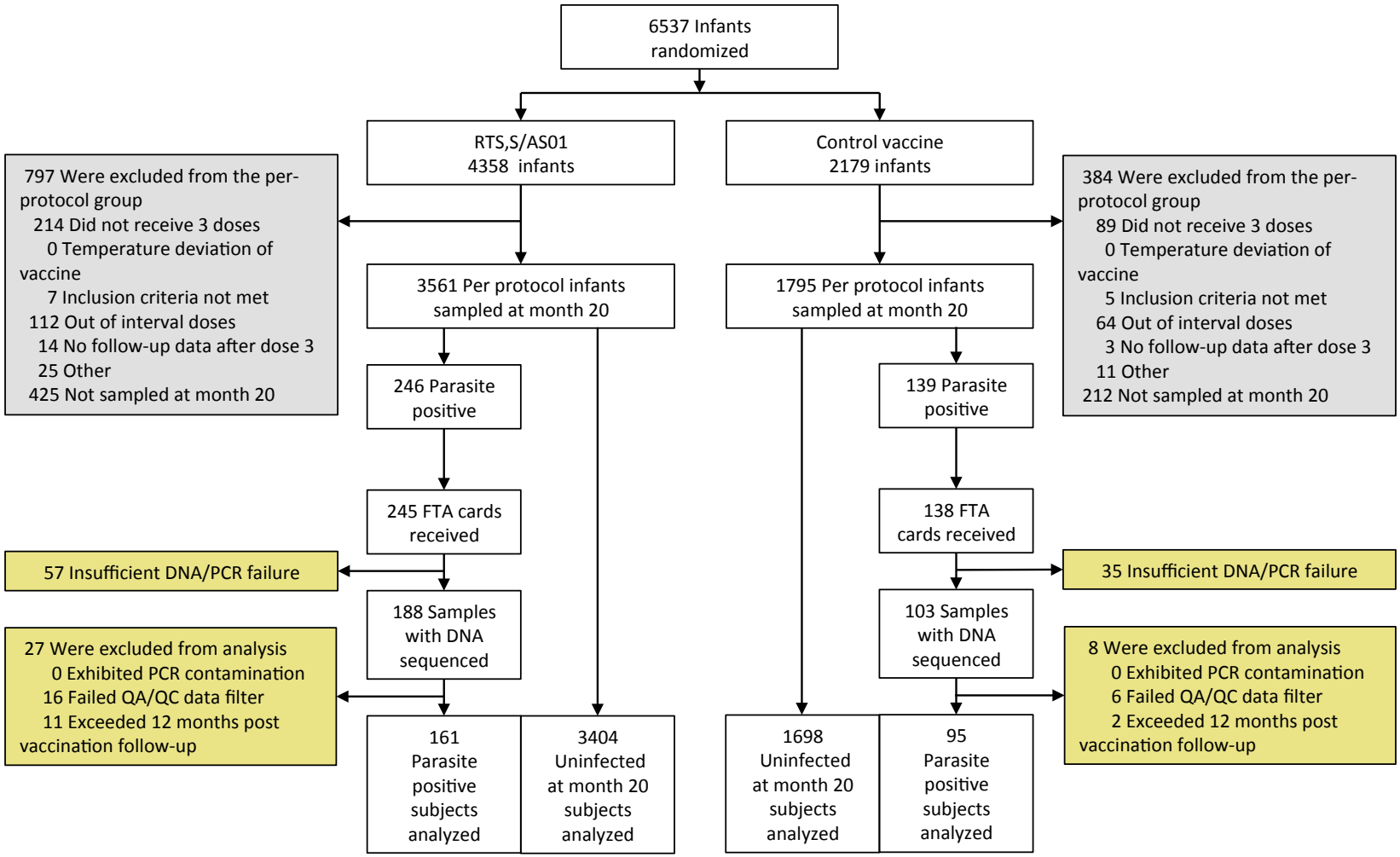


Figure S9. Linkage Disequilibrium in the CS C-terminus in children aged 5-17 months in the control vaccine group.

Calculations of r^2 between polymorphic positions amino acid positions with a minor allele frequency of at least 3% at the five largest study sites. Nucleotide coding sequence positions are indicated in parentheses, and positions comprising the LD haplotype used in sieve analyses are indicated in red. Siaya (Panel A); Nanoro (Panel B); Kombewa (Panel C); Agogo (Panel D); Kintampo (Panel E).

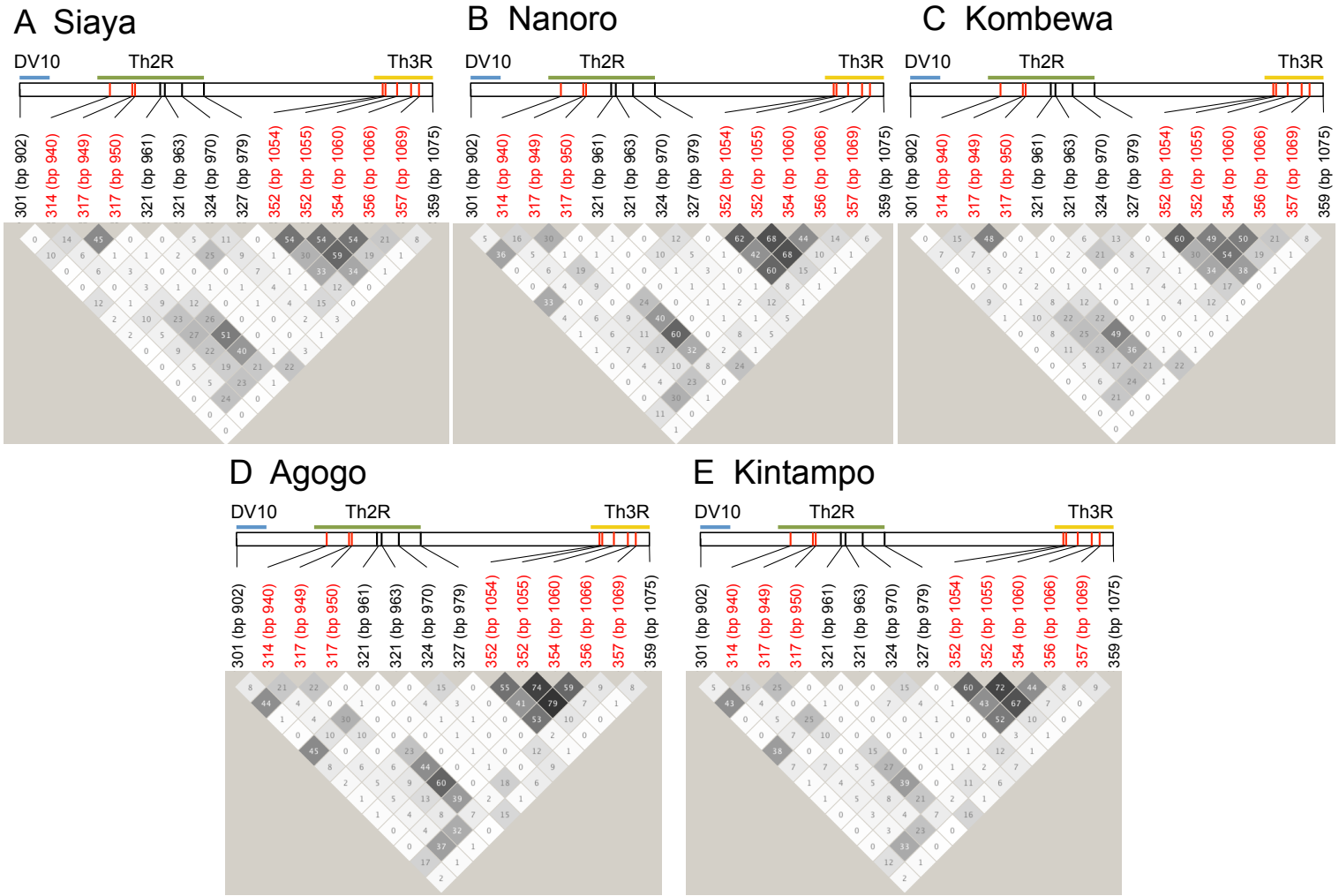


Figure S10. COI and Population Frequencies for CS C-terminus in the Infants Aged 6-12 Weeks for the Primary Clinical Malaria Endpoint.

Distributions of COI for the RTS,S/AS01 vaccine and control vaccine groups (Panel A); Frequencies of full CS C-terminus 3D7 match by study site (Panel B) and for all sites by malaria genomic unit defined by the full CS C-terminus, epitopes, polymorphic region DV-10, and haplotype combining Th2R and Th3R positions in linkage disequilibrium (LD) (Panel C); polymorphic CS C-terminus amino acid positions with frequency between 1% and 99% (Panel D).

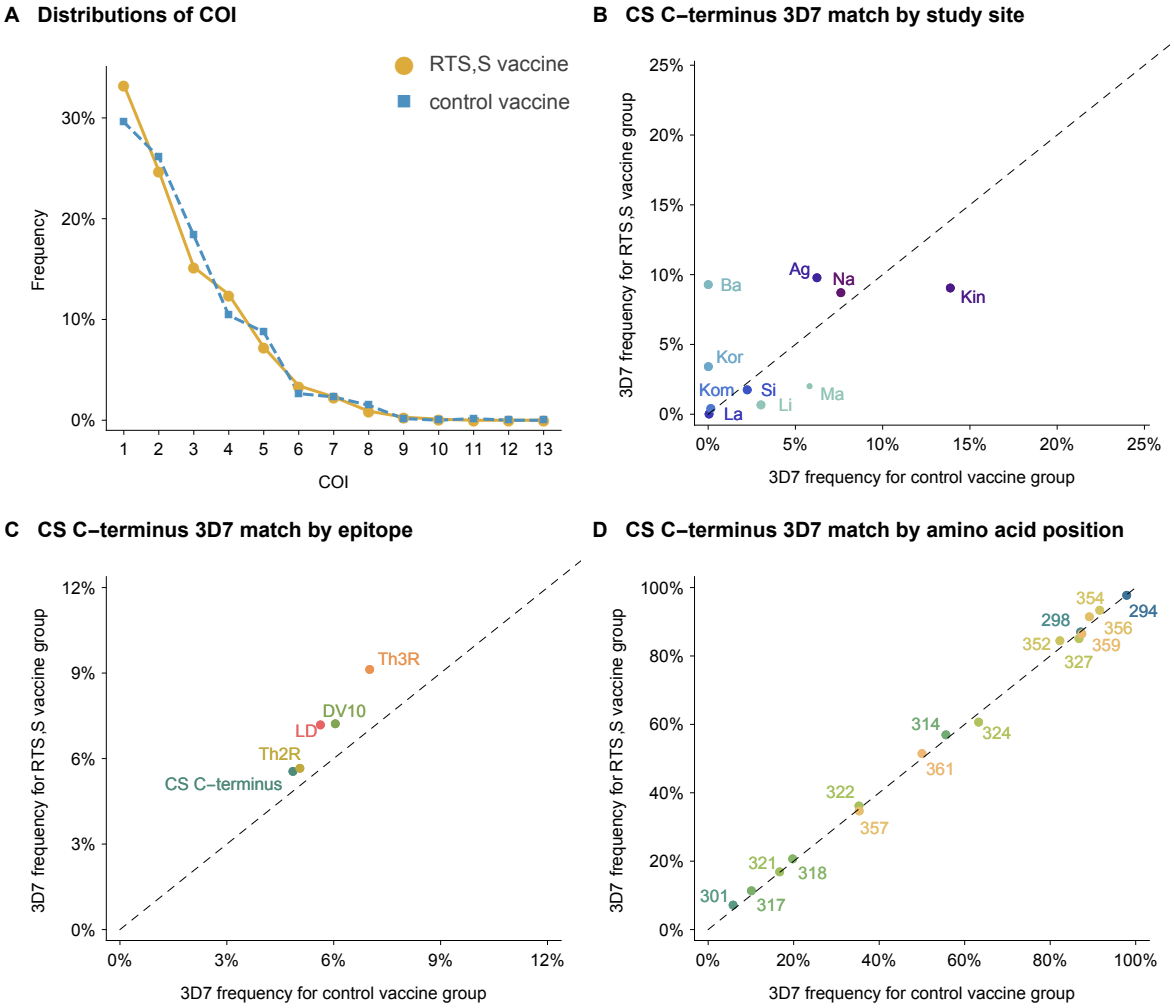


Figure S11. Proportion of Infections Containing a 3D7-Matched Haplotype as a Function of COI in the RTS,S/AS01 Vaccine and Control Vaccine Groups in the Children Aged 5-17 Months for the Primary Clinical Malaria Endpoint.

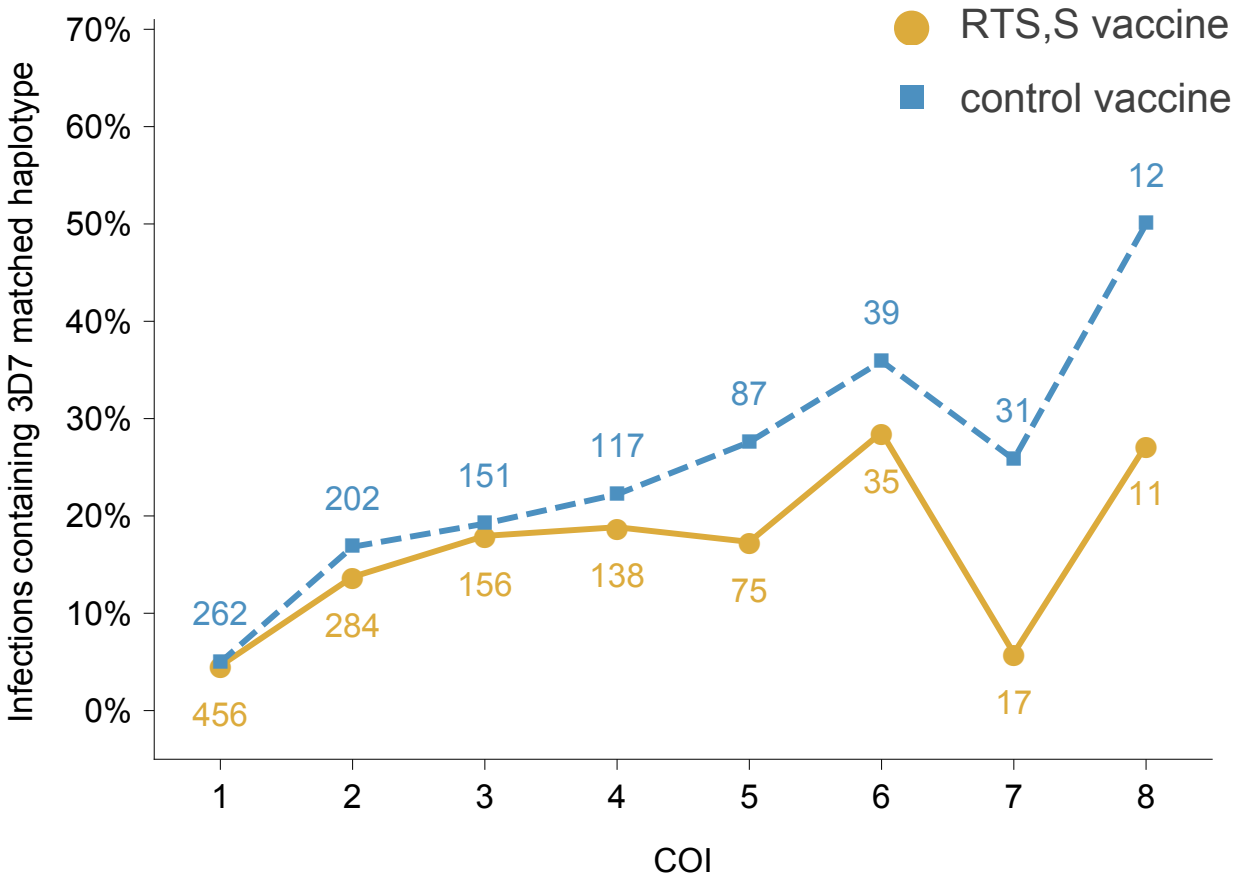
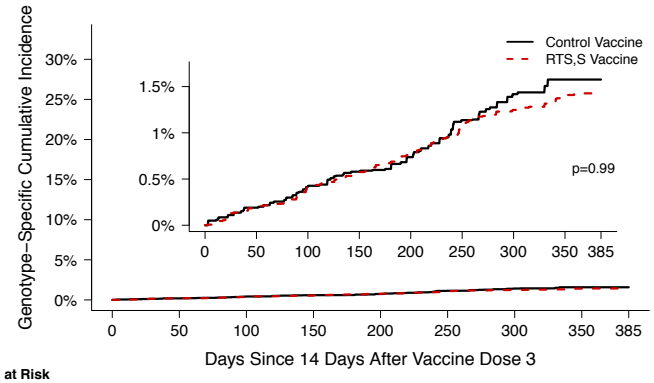


Figure S12. Cumulative Incidences and Vaccine Efficacies (VEs) Against the Primary Clinical Malaria Endpoint in Infants Aged 6-12 Weeks with Parasites Matched and Mismatched to the 3D7 Full CS C-Terminus Haplotype.

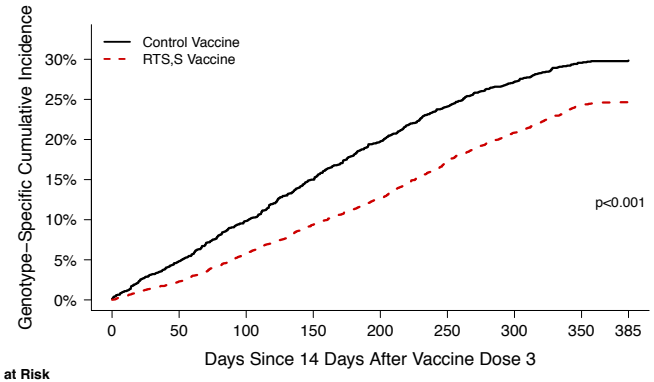
The cumulative incidence during 12 months of post-vaccination follow-up in RTS,S/AS01 and control vaccine recipients in 3D7 matched cases (Panel A), and 3D7 mismatched cases (Panel B). Panel C shows the cumulative VE against 3D7 matched and 3D7 mismatched malaria over the entire post-vaccination follow-up period, and Panel D shows the cumulative and hazard ratio VE against 3D7 matched and 3D7 mismatched malaria at 12 months post-vaccination.

A Cumulative Incidence of 3D7 Matched Malaria



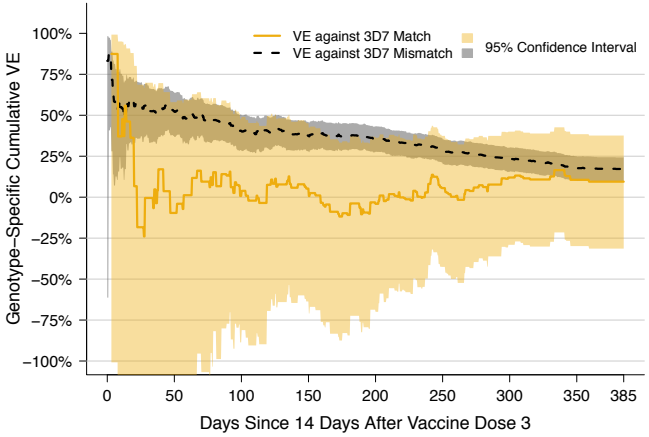
No. at Risk									
Control Vaccine	2007	1863	1720	1585	1458	1342	1241	1165	1160
RTS,S Vaccine	3986	3827	3633	3413	3220	2951	2714	2545	2532
Cumulative No. of Endpoints with 3D7 Match									
Control Vaccine	0	4	8	11	14	21	26	29	29
RTS,S Vaccine	0	8	16	22	30	41	47	52	52

B Cumulative Incidence of 3D7 Mismatched Malaria



No. at Risk									
Control Vaccine	2007	1863	1720	1585	1458	1342	1241	1165	1160
RTS,S Vaccine	3986	3827	3633	3413	3220	2951	2714	2545	2532
Cumulative No. of Endpoints with 3D7 Mismatch									
Control Vaccine	3	96	194	291	380	459	514	550	554
RTS,S Vaccine	1	92	232	367	488	660	785	894	903

C Cumulative Vaccine Efficacy Over Time



D Cumulative and Hazard Ratio Vaccine Efficacy

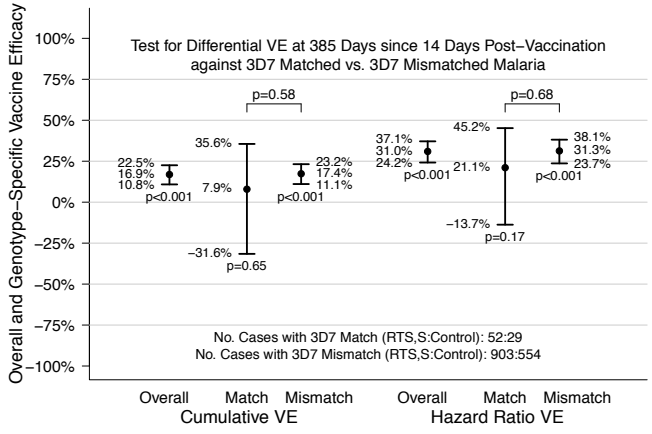
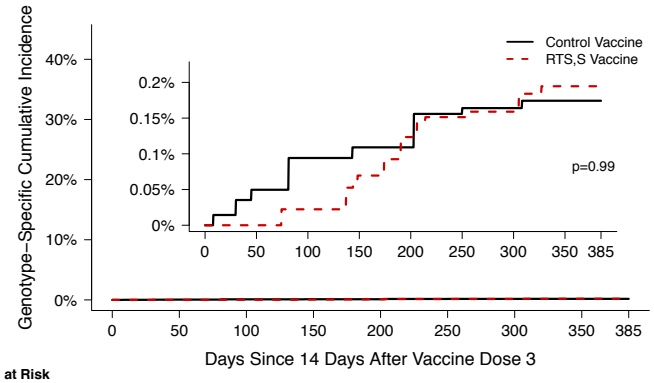


Figure S13. Cumulative Incidences and Vaccine Efficacies (VEs) Against the Primary Clinical Malaria Endpoint in Children Aged 5-17 Months with Parasites Matched and Mismatched to the 3D7 Full SERA2 Haplotype.

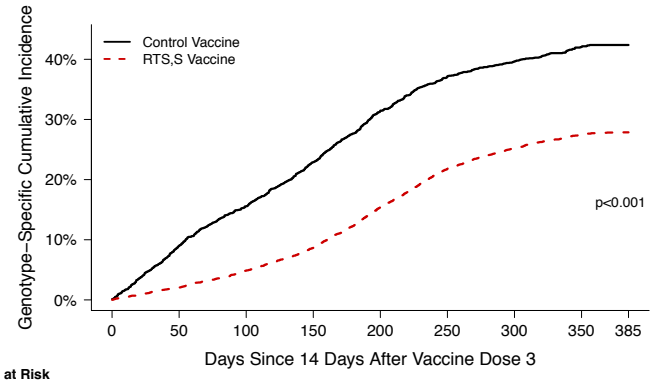
The cumulative incidence during 12 months of post-vaccination follow-up in RTS,S/AS01 and control vaccine recipients in 3D7 matched cases (Panel A), and 3D7 mismatched cases (Panel B). Panel C shows the cumulative VE against 3D7 matched and 3D7 mismatched malaria over the entire post-vaccination follow-up period, and Panel D shows the cumulative and hazard ratio VE against 3D7 matched and 3D7 mismatched malaria at 12 months post-vaccination.

A Cumulative Incidence of 3D7 Matched Malaria



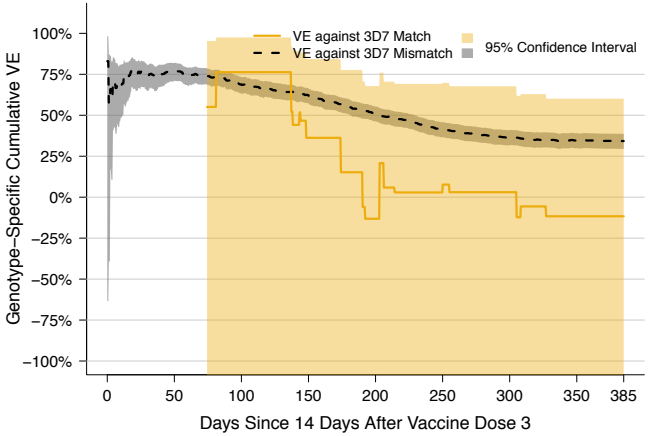
No. at Risk									
Control Vaccine	2335	2067	1881	1685	1467	1317	1221	1131	1126
RTS,S Vaccine	4577	4409	4238	4014	3647	3283	2994	2788	2777
Cumulative No. of Endpoints with 3D7 Match									
Control Vaccine	0	1	2	2	2	4	4	4	4
RTS,S Vaccine	0	0	1	3	5	7	7	8	8

B Cumulative Incidence of 3D7 Mismatched Malaria



No. at Risk									
Control Vaccine	2335	2067	1881	1685	1467	1317	1221	1131	1126
RTS,S Vaccine	4577	4409	4238	4014	3647	3283	2994	2788	2777
Cumulative No. of Endpoints with 3D7 Mismatch									
Control Vaccine	3	209	356	518	703	824	875	916	919
RTS,S Vaccine	1	94	221	394	682	954	1092	1173	1183

C Cumulative Vaccine Efficacy Over Time



D Cumulative and Hazard Ratio Vaccine Efficacy

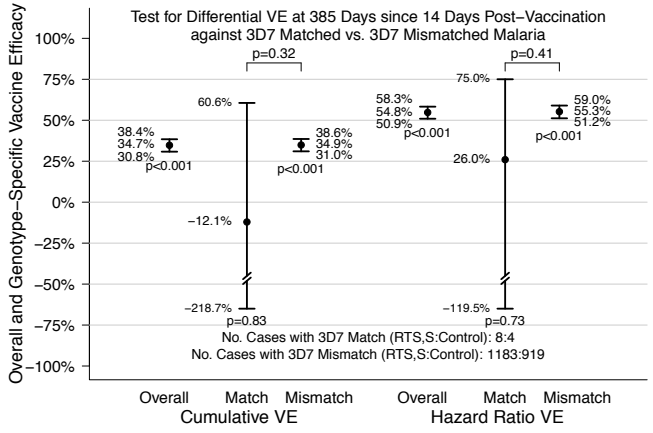
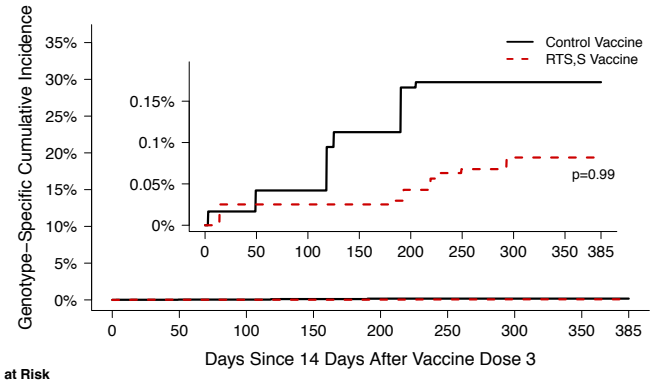


Figure S14. Cumulative Incidences and Vaccine Efficacies (VEs) Against the Primary Clinical Malaria Endpoint in Infants Aged 6-12 Weeks with Parasites Matched and Mismatched to the 3D7 Full SERA2 Haplotype.

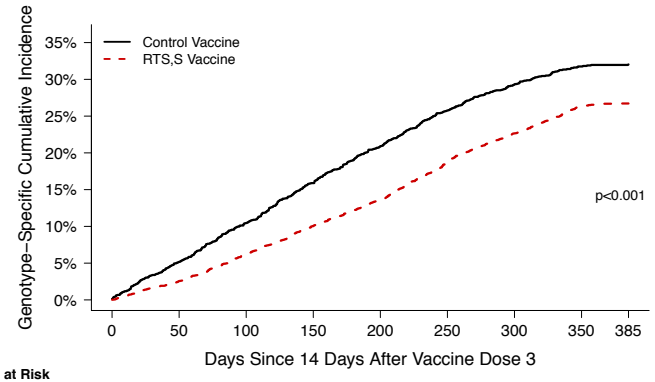
The cumulative incidence during 12 months of post-vaccination follow-up in RTS,S/AS01 and control vaccine recipients in 3D7 matched cases (Panel A), and 3D7 mismatched cases (Panel B). Panel C shows the cumulative VE against 3D7 matched and 3D7 mismatched malaria over the entire post-vaccination follow-up period, and Panel D shows the cumulative and hazard ratio VE against 3D7 matched and 3D7 mismatched malaria at 12 months post-vaccination.

A Cumulative Incidence of 3D7 Matched Malaria



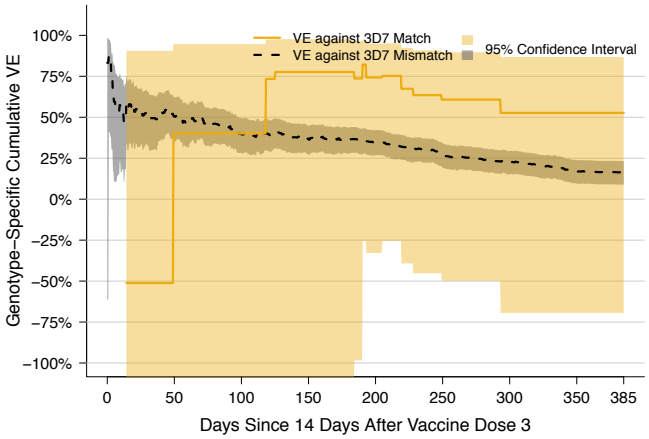
No. at Risk									
Control Vaccine	2007	1863	1720	1585	1458	1342	1241	1165	1160
RTS,S Vaccine	3986	3827	3633	3413	3220	2951	2714	2545	2532
Cumulative No. of Endpoints with 3D7 Match									
Control Vaccine	0	1	1	2	3	3	3	3	3
RTS,S Vaccine	0	1	1	1	2	3	3	3	3

B Cumulative Incidence of 3D7 Mismatched Malaria



No. at Risk									
Control Vaccine	2007	1863	1720	1585	1458	1342	1241	1165	1160
RTS,S Vaccine	3986	3827	3633	3413	3220	2951	2714	2545	2532
Cumulative No. of Endpoints with 3D7 Mismatch									
Control Vaccine	3	103	206	309	403	492	556	594	598
RTS,S Vaccine	1	102	249	394	524	719	855	973	982

C Cumulative Vaccine Efficacy Over Time



D Cumulative and Hazard Ratio Vaccine Efficacy

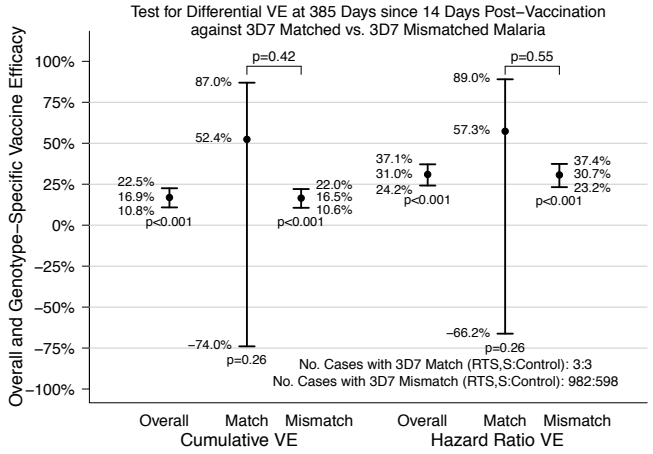


Figure S15. Cumulative Vaccine Efficacies (VE) Against the Primary Clinical Malaria Endpoint in Children Aged 5-17 Months with Parasites Matched and Mismatched to the CS C-terminus 3D7 Vaccine Strain at 1-7 of the Signature Sieve Positions.

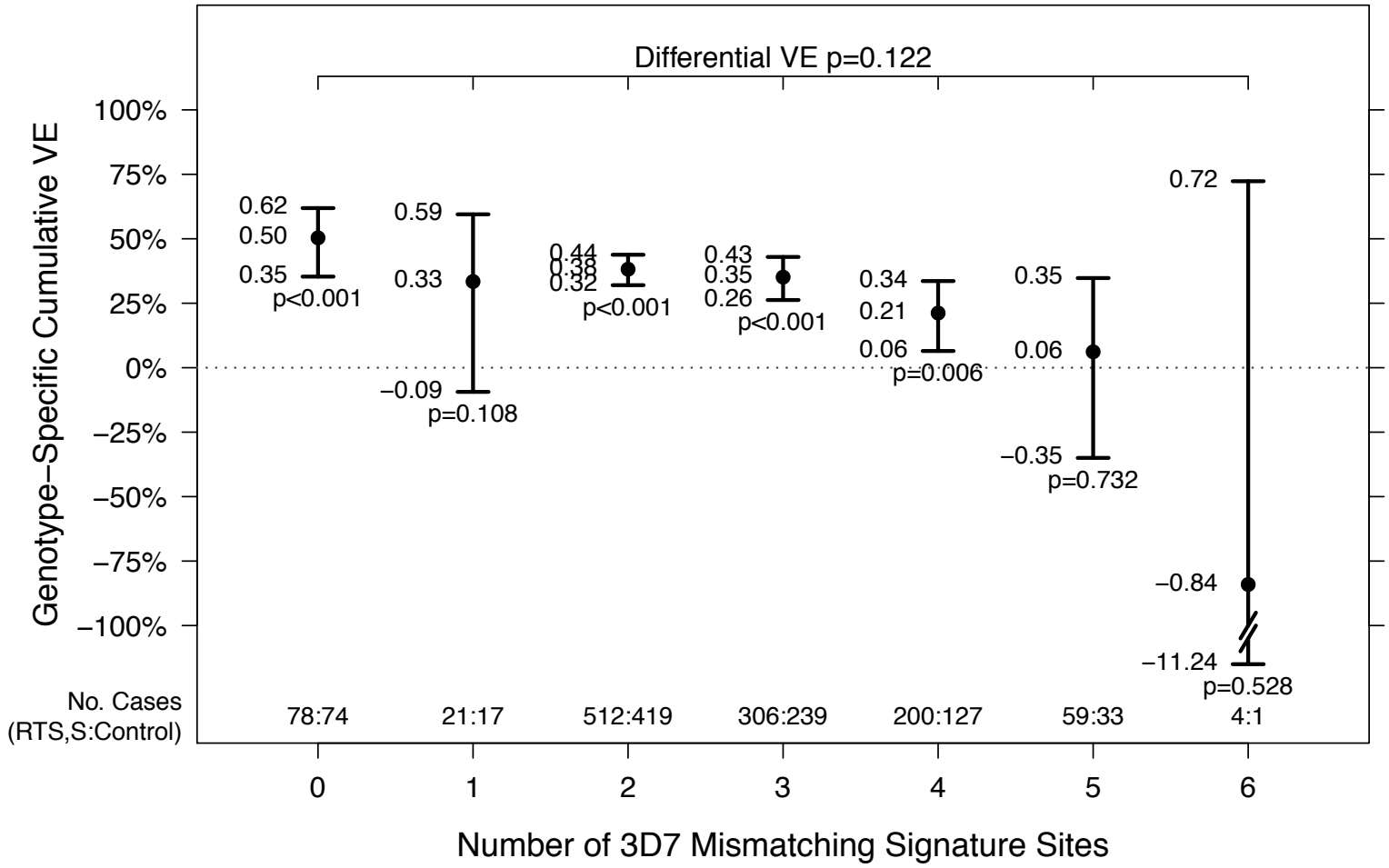


Figure S16. Cumulative Vaccine Efficacies (VE) Against the Primary Clinical Malaria Endpoint in Children Aged 5-17 Months Stratified by the Number of NANP/NVDP repeats.

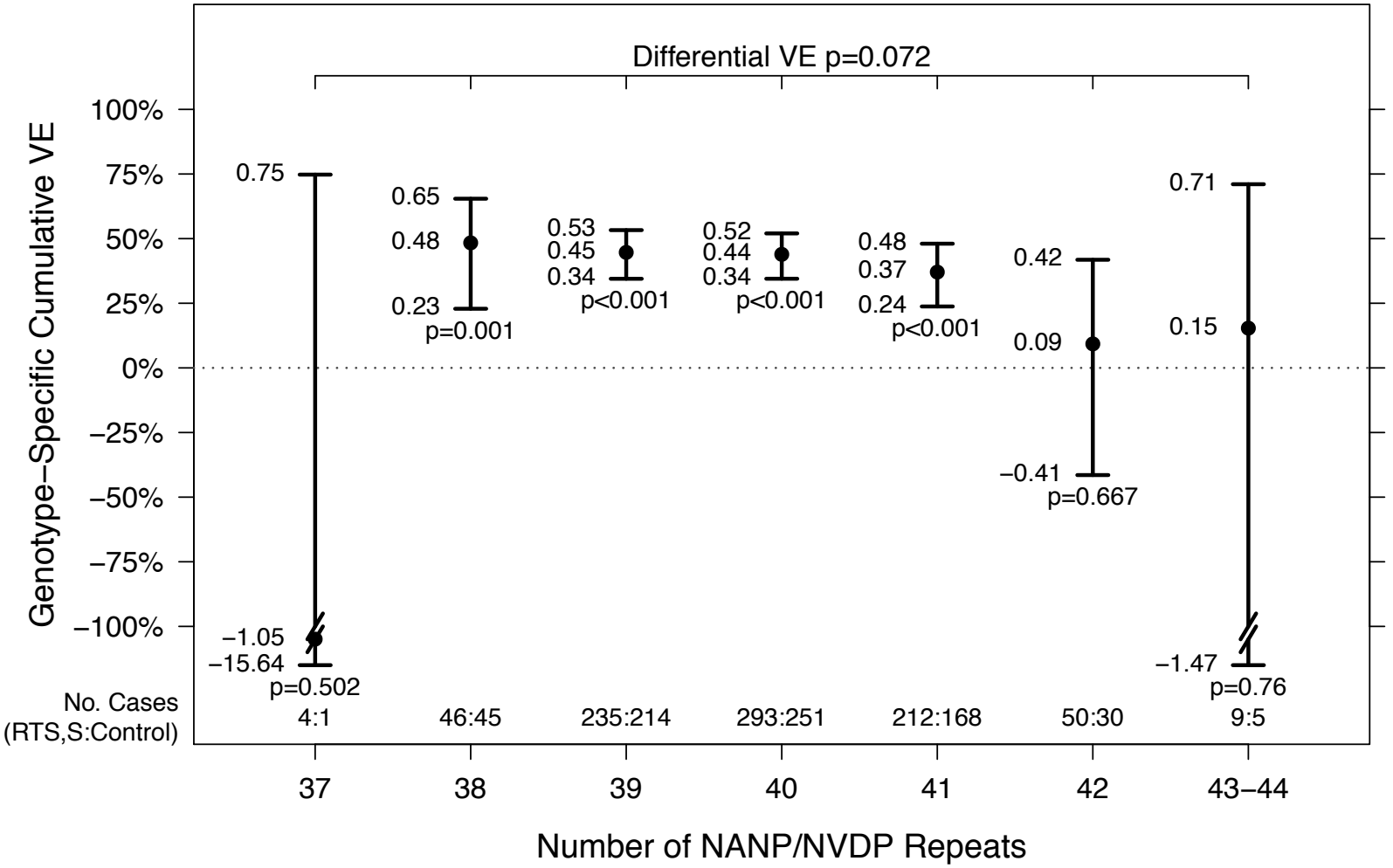


Figure S17. Cumulative Vaccine Efficacies (VE) Against the Primary Clinical Malaria Endpoint in Infants Aged 6-12 Weeks Stratified by the Number of NANP/NVDP repeats.

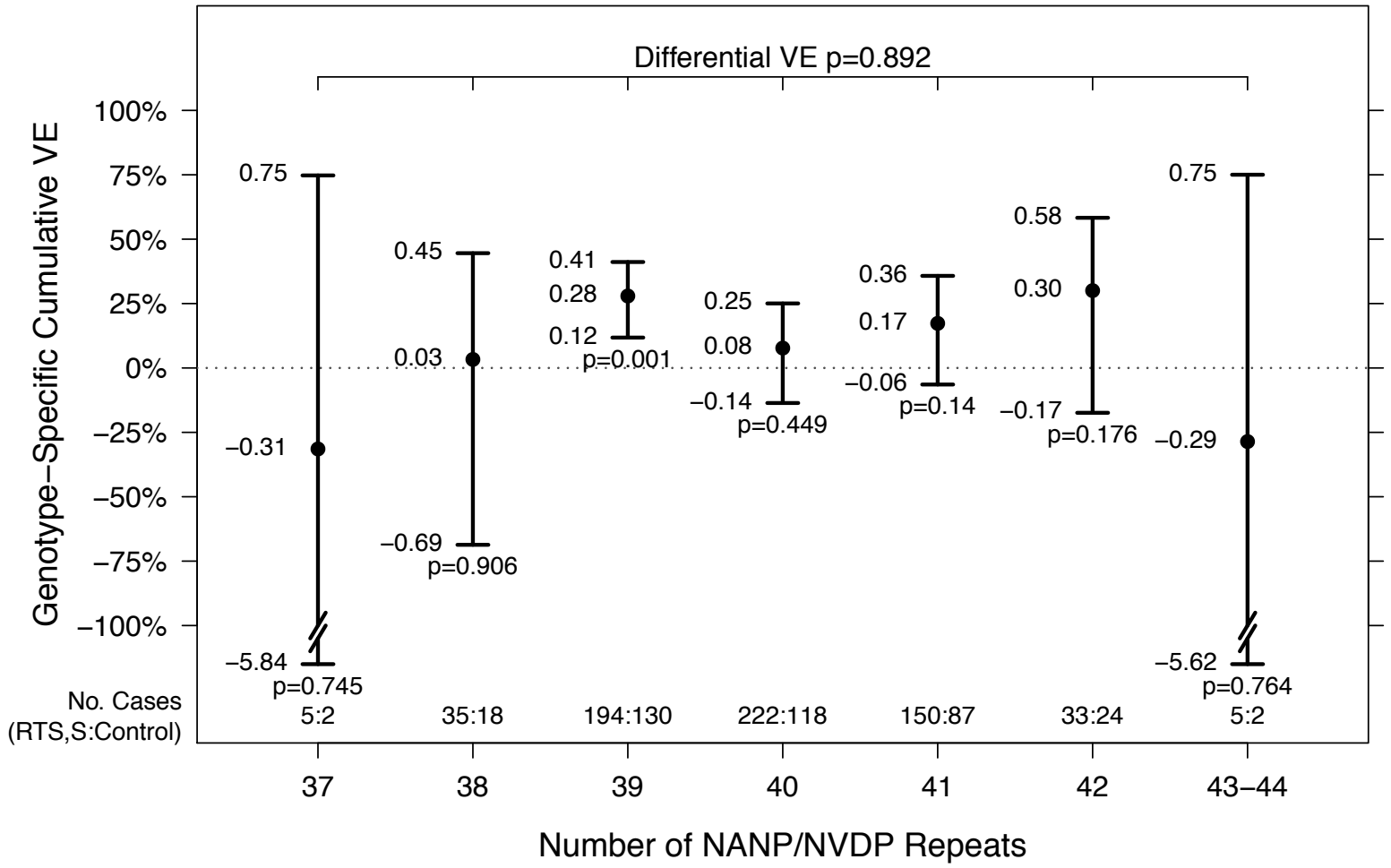


Figure S18. 3D7 Matched CS C-terminus Haplotype Frequency vs. Cumulative Vaccine Efficacy (VE) by Study Site

The 3D7 match haplotype frequency in the control group for each study site vs. vaccine efficacy during 12 months of post-vaccination follow-up. Panels A and B show overall and haplotype-specific cumulative VE, respectively. Panels C and D show overall and haplotype-specific hazard ratio VE, respectively.

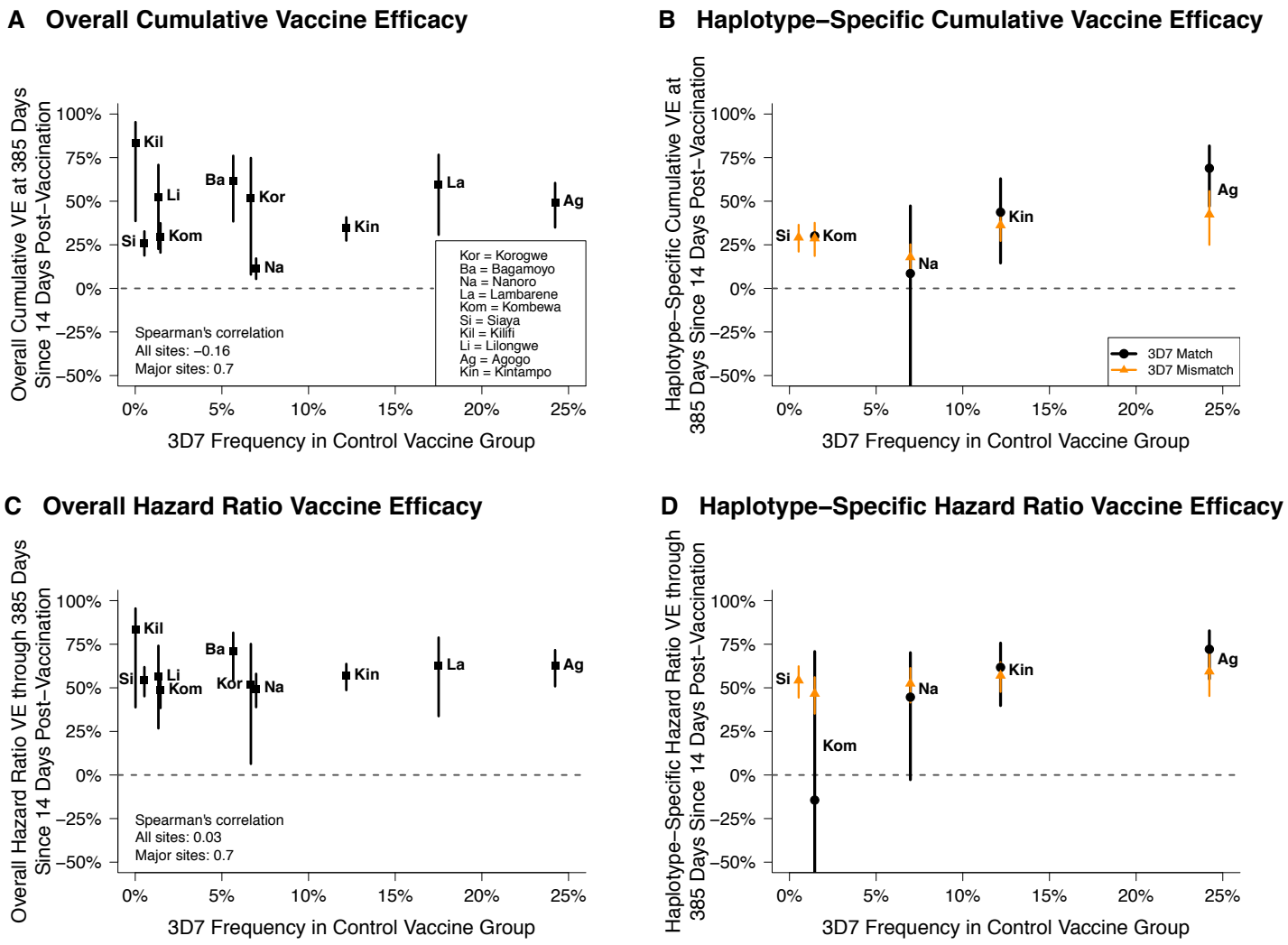


Figure S19. 3D7 Matched CS C-terminus Haplotype Cumulative Incidence vs. Vaccine Efficacy (VE) by Study Site

The 3D7 match haplotype cumulative incidence in the control group for each study site vs. vaccine efficacy during 12 months of post-vaccination follow-up. Panels A and B show overall and haplotype-specific cumulative VE, respectively. Panels C and D show overall and haplotype-specific hazard ratio VE, respectively.

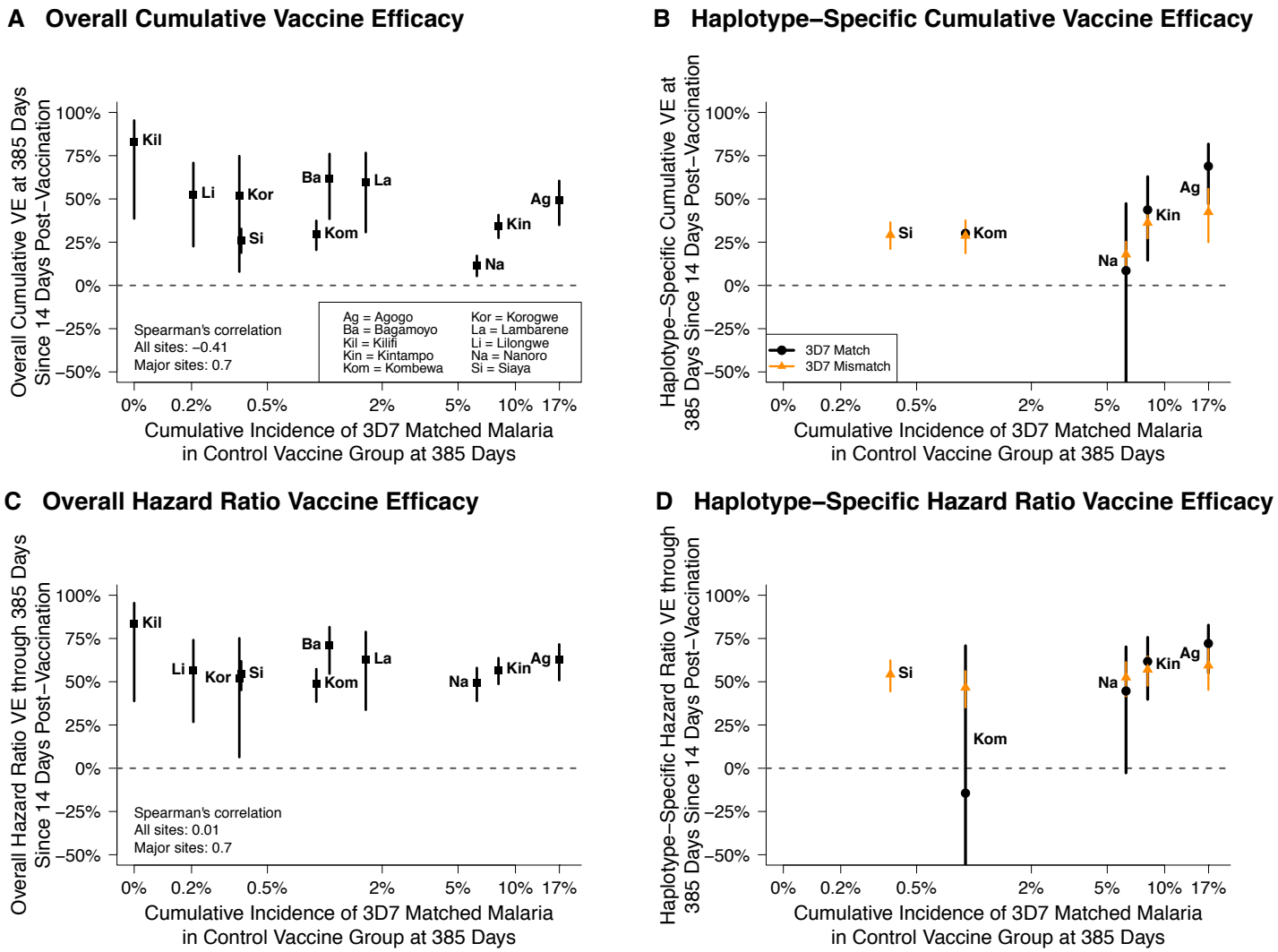


Table S1. Description of the Study Population for the Primary Clinical Malaria Endpoint by Study Site and PCR Amplicon for Children Aged 5-17 Months.

Study Site/ Vaccine Group	Endpoint: Primary Clinical Malaria	CS C-terminus		NANP/NVDP		SERA2	
		Samples with DNA Sequenced	Subjects with Analyzable Sequence	Samples with DNA Sequenced	Subjects with Analyzable Sequence	Samples with DNA Sequenced	Subjects with Analyzable Sequence
Agogo							
RTS,S/AS01 (N = 371)	102	101 (99%)	95 (93%)	74 (73%)	69 (68%)	98 (96%)	96 (94%)
Control (N = 191)	103	100 (97%)	95 (92%)	86 (83%)	79 (77%)	97 (94%)	91 (88%)
Bagamoyo							
RTS,S/AS01 (N = 470)	31	29 (94%)	28 (90%)	20 (65%)	20 (65%)	30 (97%)	30 (97%)
Control (N = 236)	48	48 (100%)	46 (96%)	35 (73%)	31 (65%)	44 (92%)	43 (90%)
Kilifi							
RTS,S/AS01 (N = 335)	3	3 (100%)	3 (100%)	2 (67%)	2 (67%)	3 (100%)	3 (100%)
Control (N = 172)	9	6 (67%)	6 (67%)	5 (56%)	5 (56%)	6 (67%)	6 (67%)
Kintampo							
RTS,S/AS01 (N = 609)	298	259 (87%)	246 (83%)	202 (68%)	191 (64%)	247 (83%)	246 (83%)
Control (N = 301)	228	197 (86%)	191 (84%)	166 (73%)	158 (69%)	192 (84%)	192 (84%)
Kombewa							
RTS,S/AS01 (N = 613)	269	254 (94%)	249 (93%)	204 (76%)	189 (70%)	247 (92%)	247 (92%)
Control (N = 311)	199	184 (92%)	178 (89%)	161 (81%)	152 (76%)	181 (91%)	180 (90%)
Korogwe							
RTS,S/AS01 (N = 568)	17	15 (88%)	15 (88%)	9 (53%)	8 (47%)	15 (88%)	15 (88%)
Control (N = 293)	18	16 (89%)	15 (83%)	13 (72%)	12 (67%)	15 (83%)	15 (83%)
Lambaréne							
RTS,S/AS01 (N = 382)	21	21 (100%)	21 (100%)	14 (67%)	13 (62%)	21 (100%)	21 (100%)
Control (N = 196)	27	22 (81%)	19 (70%)	17 (63%)	16 (59%)	21 (78%)	20 (74%)
Lilongwe							
RTS,S/AS01 (N = 358)	27	20 (74%)	19 (70%)	11 (41%)	10 (37%)	20 (74%)	19 (70%)
Control (N = 185)	30	28 (93%)	27 (90%)	14 (47%)	14 (47%)	27 (90%)	27 (90%)
Nanoro							
RTS,S/AS01 (N = 389)	301	262 (87%)	255 (85%)	178 (59%)	165 (55%)	263 (87%)	263 (87%)
Control (N = 198)	176	161 (91%)	156 (89%)	118 (67%)	108 (61%)	162 (92%)	162 (92%)
Siaya							
RTS,S/AS01 (N = 482)	282	265 (94%)	250 (89%)	193 (68%)	182 (65%)	253 (90%)	253 (90%)
Control (N = 252)	202	185 (92%)	176 (87%)	134 (66%)	131 (65%)	181 (90%)	181 (90%)

Table S2. Description of the Study Population for the Primary Clinical Malaria Endpoint by Study Site and PCR Amplicon for Infants Aged 6-12 Weeks.

Study Site/ Vaccine Group	Endpoint: Primary Clinical Malaria	CS C-terminus		NANP/NVDP		SERA2	
		Samples with DNA Sequenced	Subjects with Analyzable Sequence	Samples with DNA Sequenced	Subjects with Analyzable Sequence	Samples with DNA Sequenced	Subjects with Analyzable Sequence
Agogo							
RTS,S/AS01 (N = 417)	149	119 (80%)	90 (60%)	86 (58%)	65 (44%)	120 (81%)	94 (63%)
Control (N = 221)	90	69 (77%)	56 (62%)	49 (54%)	34 (38%)	72 (80%)	59 (66%)
Bagamoyo							
RTS,S/AS01 (N = 501)	30	25 (83%)	25 (83%)	21 (70%)	21 (70%)	26 (87%)	26 (87%)
Control (N = 245)	20	19 (95%)	17 (85%)	15 (75%)	12 (60%)	17 (85%)	17 (85%)
Kilifi							
RTS,S/AS01 (N = 184)	2	1 (50%)	1 (50%)	1 (50%)	1 (50%)	1 (50%)	1 (50%)
Control (N = 102)	1	0 (0%)	0 (0%)	(%)	0 (0%)	0 (0%)	0 (0%)
Kintampo							
RTS,S/AS01 (N = 199)	125	101 (81%)	96 (77%)	82 (66%)	71 (57%)	104 (83%)	104 (83%)
Control (N = 100)	58	51 (88%)	48 (83%)	34 (59%)	27 (47%)	52 (90%)	52 (90%)
Kombewa							
RTS,S/AS01 (N = 387)	152	130 (86%)	127 (84%)	97 (64%)	88 (58%)	129 (85%)	129 (85%)
Control (N = 196)	102	85 (83%)	80 (78%)	61 (60%)	53 (52%)	79 (77%)	79 (77%)
Korogwe							
RTS,S/AS01 (N = 382)	11	10 (91%)	10 (91%)	7 (64%)	6 (55%)	10 (91%)	10 (91%)
Control (N = 183)	10	9 (90%)	9 (90%)	6 (60%)	5 (50%)	9 (90%)	9 (90%)
Lambaréne							
RTS,S/AS01 (N = 147)	10	9 (90%)	9 (90%)	8 (80%)	8 (80%)	10 (100%)	10 (100%)
Control (N = 62)	4	2 (50%)	2 (50%)	2 (50%)	2 (50%)	2 (50%)	2 (50%)
Lilongwe							
RTS,S/AS01 (N = 497)	59	50 (85%)	48 (81%)	32 (54%)	26 (44%)	49 (83%)	49 (83%)
Control (N = 257)	55	44 (80%)	43 (78%)	28 (51%)	21 (38%)	44 (80%)	44 (80%)
Manhiça							
RTS,S/AS01 (N = 380)	21	18 (86%)	16 (76%)	15 (71%)	13 (62%)	16 (76%)	16 (76%)
Control (N = 188)	11	10 (91%)	10 (91%)	6 (55%)	6 (55%)	10 (91%)	10 (91%)
Nanoro							
RTS,S/AS01 (N = 441)	327	287 (88%)	286 (87%)	216 (66%)	195 (60%)	292 (89%)	292 (89%)
Control (N = 224)	186	171 (92%)	171 (92%)	127 (68%)	115 (62%)	170 (91%)	170 (91%)
Siaya							
RTS,S/AS01 (N = 451)	274	231 (84%)	217 (79%)	157 (57%)	131 (48%)	253 (92%)	224 (82%)
Control (N = 229)	175	141 (81%)	133 (76%)	105 (60%)	96 (55%)	146 (83%)	145 (83%)

Table S3. Description of the Study Population for the Parasite Positive Endpoint by Study Site and PCR Amplicon for Children Aged 5-17 Months.

Study Site/ Vaccine Group	Endpoint: Parasite Positive	CS C-terminus		NANP/NVDP		SERA2	
		Samples with DNA Sequenced	Subjects with Analyzable Sequence	Samples with DNA Sequenced	Subjects with Analyzable Sequence	Samples with DNA Sequenced	Subjects with Analyzable Sequence
Agogo							
RTS,S/AS01 (N = 371)	27	27 (100%)	26 (96%)	23 (85%)	19 (70%)	27 (100%)	27 (100%)
Control (N = 191)	24	24 (100%)	21 (88%)	24 (100%)	17 (71%)	24 (100%)	24 (100%)
Bagamoyo							
RTS,S/AS01 (N = 470)	4	3 (75%)	3 (75%)	0 (0%)	0 (0%)	3 (75%)	3 (75%)
Control (N = 236)	3	3 (100%)	3 (100%)	0 (0%)	0 (0%)	3 (100%)	3 (100%)
Kilifi							
RTS,S/AS01 (N = 335)	1	1 (100%)	1 (100%)	0 (0%)	0 (0%)	1 (100%)	1 (100%)
Control (N = 172)	1	1 (100%)	1 (100%)	0 (0%)	0 (0%)	1 (100%)	1 (100%)
Kintampo							
RTS,S/AS01 (N = 609)	79	76 (96%)	76 (96%)	69 (87%)	59 (75%)	76 (96%)	76 (96%)
Control (N = 301)	57	55 (96%)	55 (96%)	46 (81%)	42 (74%)	56 (98%)	56 (98%)
Kombewa							
RTS,S/AS01 (N = 613)	63	56 (89%)	56 (89%)	18 (29%)	16 (25%)	56 (89%)	56 (89%)
Control (N = 311)	45	42 (93%)	42 (93%)	13 (29%)	12 (27%)	42 (93%)	42 (93%)
Korogwe							
RTS,S/AS01 (N = 568)	1	1 (100%)	1 (100%)	0 (0%)	0 (0%)	1 (100%)	1 (100%)
Control (N = 293)	1	1 (100%)	1 (100%)	0 (0%)	0 (0%)	1 (100%)	1 (100%)
Lambaréne							
RTS,S/AS01 (N = 382)	7	7 (100%)	7 (100%)	7 (100%)	6 (86%)	7 (100%)	7 (100%)
Control (N = 196)	6	6 (100%)	6 (100%)	6 (100%)	6 (100%)	6 (100%)	6 (100%)
Lilongwe							
RTS,S/AS01 (N = 358)	10	9 (90%)	8 (80%)	9 (90%)	7 (70%)	9 (90%)	9 (90%)
Control (N = 185)	7	7 (100%)	7 (100%)	5 (71%)	3 (43%)	7 (100%)	7 (100%)
Nanoro							
RTS,S/AS01 (N = 389)	35	33 (94%)	33 (94%)	29 (83%)	27 (77%)	33 (94%)	33 (94%)
Control (N = 198)	40	36 (90%)	36 (90%)	33 (83%)	30 (75%)	37 (93%)	37 (93%)
Siaya							
RTS,S/AS01 (N = 482)	79	73 (92%)	73 (92%)	64 (81%)	59 (75%)	75 (95%)	75 (95%)
Control (N = 252)	39	36 (92%)	36 (92%)	29 (74%)	27 (69%)	37 (95%)	36 (92%)

Table S4. Description of the Study Population for the Parasite Positive Endpoint by Study Site and PCR Amplicon for Infants Aged 6-12 Weeks.

Study Site/ Vaccine Group	Endpoint: Parasite Positive	CS C-terminus		NANP/NVDP		SERA2	
		Samples with DNA Sequenced	Subjects with Analyzable Sequence	Samples with DNA Sequenced	Subjects with Analyzable Sequence	Samples with DNA Sequenced	Subjects with Analyzable Sequence
Agogo							
RTS,S/AS01 (N = 417)	34	33 (97%)	20 (59%)	28 (82%)	16 (47%)	33 (97%)	33 (97%)
Control (N = 221)	21	21 (100%)	19 (90%)	16 (76%)	15 (71%)	21 (100%)	21 (100%)
Bagamoyo							
RTS,S/AS01 (N = 501)	3	3 (100%)	3 (100%)	1 (33%)	1 (33%)	2 (67%)	2 (67%)
Control (N = 245)	3	3 (100%)	3 (100%)	0 (0%)	0 (0%)	3 (100%)	3 (100%)
Kilifi							
RTS,S/AS01 (N = 184)	1	1 (100%)	1 (100%)	1 (100%)	1 (100%)	1 (100%)	1 (100%)
Control (N = 102)	3	2 (67%)	2 (67%)	1 (33%)	0 (0%)	2 (67%)	2 (67%)
Kintampo							
RTS,S/AS01 (N = 199)	26	25 (96%)	25 (96%)	23 (88%)	22 (85%)	25 (96%)	25 (96%)
Control (N = 100)	13	12 (92%)	12 (92%)	9 (69%)	8 (62%)	12 (92%)	12 (92%)
Kombewa							
RTS,S/AS01 (N = 387)	42	35 (83%)	35 (83%)	31 (74%)	28 (67%)	36 (86%)	36 (86%)
Control (N = 196)	18	15 (83%)	15 (83%)	12 (67%)	12 (67%)	15 (83%)	15 (83%)
Korogwe							
RTS,S/AS01 (N = 382)	0	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)
Control (N = 183)	1	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (100%)	1 (100%)
Lambaréne							
RTS,S/AS01 (N = 147)	6	5 (83%)	5 (83%)	4 (67%)	4 (67%)	5 (83%)	5 (83%)
Control (N = 62)	2	2 (100%)	2 (100%)	2 (100%)	2 (100%)	2 (100%)	2 (100%)
Lilongwe							
RTS,S/AS01 (N = 497)	38	34 (89%)	34 (89%)	25 (66%)	22 (58%)	34 (89%)	34 (89%)
Control (N = 257)	15	13 (87%)	13 (87%)	11 (73%)	11 (73%)	14 (93%)	14 (93%)
Manhiça							
RTS,S/AS01 (N = 380)	5	5 (100%)	5 (100%)	3 (60%)	3 (60%)	5 (100%)	5 (100%)
Control (N = 188)	3	3 (100%)	3 (100%)	1 (33%)	1 (33%)	2 (67%)	2 (67%)
Nanoro							
RTS,S/AS01 (N = 441)	39	39 (100%)	39 (100%)	29 (74%)	28 (72%)	39 (100%)	39 (100%)
Control (N = 224)	20	20 (100%)	20 (100%)	19 (95%)	18 (90%)	20 (100%)	20 (100%)
Siaya							
RTS,S/AS01 (N = 451)	52	51 (98%)	50 (96%)	43 (83%)	36 (69%)	50 (96%)	50 (96%)
Control (N = 229)	40	39 (98%)	38 (95%)	32 (80%)	28 (70%)	40 (100%)	40 (100%)

Table S5. Cumulative Vaccine Efficacy (VE) Against 3D7 Matched and Mismatched Primary Clinical Malaria Through 12 Months Post Vaccination Follow-Up by CS C-terminus Haplotype and Position in Children Aged 5-17 Months.

Estimates are based on cross-validated model selection with candidate models adjusting for geophysical and biological covariates.[0]

Haplotype	Haplotype-Matched Efficacy[†]			Haplotype-Mismatched Efficacy[†]			Differential Efficacy		
Locus[‡]	VE (%)	95% CI	P-value	VE (%)	95% CI	P-value	P-value	FWER P-value	Q-value
CS C-Terminus*	50.0	(35.8, 60.4)	< 0.001	33.5	(29.7, 37.1)	< 0.001	0.03	–	–
DV10*	50.1	(38.0, 60.0)	< 0.001	33.0	(29.2, 36.7)	< 0.001	0.01	0.32	0.07
294	34.9	(31.2, 38.4)	< 0.001	28.3	(-10.4, 53.4)	0.13	0.67	1.00	0.79
295	34.9	(31.3, 38.4)	< 0.001	4.3	(-100.2, 54.2)	0.91	0.31	1.00	0.54
296	34.8	(31.2, 38.2)	< 0.001	2.2	(-335.0, 78.0)	0.98	0.59	1.00	0.79
298	33.6	(29.4, 37.4)	< 0.001	42.3	(31.4, 51.5)	< 0.001	0.15	1.00	0.32
299*	35.5	(31.9, 38.9)	< 0.001	-52.2	(-164.9, 12.6)	0.14	0.0026	0.07	0.07
301*	48.9	(36.6, 58.7)	< 0.001	33.2	(29.3, 36.9)	< 0.001	0.02	0.49	0.07
302	34.9	(31.2, 38.4)	< 0.001	19.3	(-108.0, 68.7)	< 0.001	0.66	1.00	0.79
303	34.7	(31.1, 38.1)	< 0.001	66.2	(-110.1, 94.6)	0.24	0.48	1.00	0.75
LD[§]*	50.8	(37.9, 61.0)	< 0.001	33.2	(29.4, 36.8)	< 0.001	0.01	0.36	0.07
Th2R*	50.3	(36.7, 61.0)	< 0.001	33.4	(29.5, 36.8)	< 0.001	0.02	0.49	0.07
314^{LD}	32.8	(27.1, 37.9)	< 0.001	37.4	(31.4, 42.8)	< 0.001	0.30	1.00	0.54
317^{LD}*	45.7	(34.2, 55.2)	< 0.001	33.3	(29.3, 37.0)	< 0.001	0.05	0.99	0.14
318	36.9	(26.8, 45.6)	< 0.001	34.0	(30.1, 38.2)	< 0.001	0.64	1.00	0.79
320	34.4	(30.8, 37.9)	< 0.001	62.4	(18.6, 82.6)	0.01	0.16	1.00	0.32
321	35.6	(25.3, 44.5)	< 0.001	34.7	(30.6, 38.6)	< 0.001	0.87	1.00	0.93
322	34.8	(27.8, 41.1)	< 0.001	34.9	(29.9, 39.5)	< 0.001	0.98	1.00	0.98
324	37.4	(32.5, 41.9)	< 0.001	30.7	(23.7, 37.1)	< 0.001	0.14	1.00	0.32
327	35.6	(31.6, 39.3)	< 0.001	29.5	(16.4, 40.1)	< 0.001	0.34	1.00	0.56
Th3R*	45.8	(33.2, 56.1)	< 0.001	33.3	(29.4, 37.0)	< 0.001	0.07	1.00	0.18
349	34.8	(31.2, 38.2)	< 0.001	17.6	(-99.2, 65.9)	0.67	0.60	1.00	0.79
352^{LD}	35.2	(31.2, 39.1)	< 0.001	32.6	(20.5, 42.9)	< 0.001	0.68	1.00	0.79
354^{LD}*	36.0	(32.3, 39.5)	< 0.001	8.6	(-23.1, 32.1)	0.56	0.02	0.53	0.07
355	34.7	(31.1, 38.2)	< 0.001	24.8	(-608.6, 92.0)	0.80	0.90	1.00	0.93
356^{LD}*	36.3	(32.5, 39.8)	< 0.001	13.8	(-8.9, 31.8)	0.21	0.02	0.44	0.07
357^{LD}	35.4	(28.0, 42.0)	< 0.001	34.5	(29.7, 39.0)	< 0.001	0.84	1.00	0.93
359*	36.2	(32.4, 40.0)	< 0.001	17.3	(0.0, 31.7)	0.05	0.01	0.36	0.07
361*	39.3	(34.0, 44.2)	< 0.001	29.5	(23.2, 35.4)	< 0.001	0.03	0.56	0.09

CI denotes confidence interval, and FWER p-value adjusted p-value controlling the familywise error rate.

° The models included in the cross-validation selection algorithm adjusted for study site, gender, weight-for-age Z-score, height-for-age Z-score, weight-for-height Z-score, arm circumference Z-score, hemoglobin, distance from nearest inpatient clinic, distance from nearest outpatient clinic, and month of enrollment (rainy vs. dry season based on site). More details on model selection can be found in the statistical analysis plan.

† For each haplotype locus (CS C-terminus, epitope, amino acid residue), haplotype-matched (mismatched) VE was computed only including clinical malaria endpoint events with 3D7 matched (mismatched) malaria in terms of the haplotype locus.

‡ Only amino acid positions with sufficiently high minor allele frequency were included in the analysis.

§ Linkage disequilibrium haplotype includes Th2R and Th3R amino acid positions marked as ^{LD}.

* Statistically significant differential efficacy was defined as q-value ≤ 0.2 for all multiply compared haplotype loci and as unadjusted p-value ≤ 0.05 for the full CS C-terminus amplicon.

Table S6. Cumulative Vaccine Efficacy (VE) Against 3D7 Matched and Mismatched Primary Clinical Malaria Through 12 Months Post Vaccination Follow-Up by CS C-terminus Full Amplicon Haplotype in Children Aged 5-17 Months at Each of the 5 Major Study Sites.

	Haplotype-Matched Efficacy					Haplotype-Mismatched Efficacy					
Study Site	No. of Events	Control Vaccine Cum. Inc. (%)	VE (%)	95% CI	P-value	No. of Events	Control Vaccine Cum. Inc. (%)	VE (%)	95% CI	P-value	Differential VE P-value
Agogo	40	17.0	68.9	(47.1, 81.7)	< 0.001	156	43.1	42.4	(25.0, 55.8)	< 0.001	0.04
Kintampo	48	8.1	43.7	(14.6, 62.9)	0.007	390	61.8	36.2	(27.2, 44.1)	< 0.001	0.57
Kombewa	6	0.9	30.1	(-121.8, 77.9)	0.54	416	60.9	28.7	(18.6, 37.6)	< 0.001	0.98
Nanoro	31	6.3	8.5	(-58.7, 47.3)	0.75	378	82.2	17.9	(9.9, 25.2)	< 0.001	0.74
Siaya	4	0.4	-108.1	(-1431.8, 71.7)	0.47	419	79.0	29.2	(21.1, 36.5)	< 0.001	0.41

CI denotes confidence interval, and cum. inc. cumulative incidence.

Table S7. Hazard Ratio Vaccine Efficacy (VE) Against 3D7 Matched and Mismatched Primary Clinical Malaria Through 12 Months Post Vaccination Follow-Up by CS C-terminus Amino Acid Position in Children Aged 5-17 Months.

Estimates were stratified by study site.

Amino Acid	Haplotype-Matched Efficacy[†]			Haplotype-Mismatched Efficacy[†]			Differential Efficacy		
Position[‡]	VE (%)	95% CI	P-value	VE (%)	95% CI	P-value	P-value	FWER P-value	Q-value
DV10									
294	54.9	(50.7, 58.7)	< 0.001	52.6	(25.2, 70.0)	< 0.001	0.92	1.00	0.98
295	54.9	(50.8, 58.7)	< 0.001	38.3	(-30.8, 70.9)	0.03	0.89	1.00	0.98
296	54.8	(50.7, 58.6)	< 0.001	37.9	(-178.5, 86.1)	0.77	0.92	1.00	0.98
298	53.7	(49.3, 57.8)	< 0.001	60.9	(52.2, 68.0)	< 0.001	0.50	1.00	0.98
299	55.4	(51.3, 59.1)	< 0.001	-21.5	(-115.2, 31.4)	0.71	0.45	1.00	0.97
301	62.5	(52.8, 70.3)	< 0.001	54.0	(49.7, 57.9)	< 0.001	0.03	0.77	0.31
302	54.9	(50.7, 58.6)	< 0.001	46.9	(-44.7, 80.5)	0.19	0.91	1.00	0.98
303	54.8	(50.6, 58.5)	< 0.001	77.0	(-51.3, 96.5)	0.15	0.98	1.00	0.98
Th2R									
314[LD]	53.1	(47.8, 57.9)	< 0.001	56.8	(51.4, 61.6)	< 0.001	0.45	1.00	0.97
317[LD]	60.0	(50.4, 67.8)	< 0.001	54.1	(49.7, 58.1)	< 0.001	0.14	1.00	0.65
318	55.2	(46.9, 62.2)	< 0.001	54.7	(50.3, 58.7)	< 0.001	0.80	1.00	0.98
320	54.6	(50.4, 58.4)	< 0.001	73.2	(40.4, 88.0)	0.002	0.84	1.00	0.98
321	54.2	(45.8, 61.4)	< 0.001	54.9	(50.5, 58.9)	< 0.001	0.78	1.00	0.98
322	54.5	(48.4, 59.9)	< 0.001	54.9	(50.0, 59.3)	< 0.001	0.80	1.00	0.98
324	56.2	(51.4, 60.6)	< 0.001	52.6	(46.3, 58.1)	< 0.001	0.41	1.00	0.97
327	55.4	(51.1, 59.3)	< 0.001	51.4	(41.4, 59.7)	< 0.001	0.77	1.00	0.98
Th3R									
349	54.8	(50.7, 58.6)	< 0.001	52.0	(-20.0, 80.8)	0.06	0.93	1.00	0.98
352[LD]	55.1	(50.8, 59.1)	< 0.001	53.0	(43.1, 61.2)	< 0.001	0.86	1.00	0.98
354[LD]	55.6	(51.4, 59.3)	< 0.001	40.6	(17.9, 57.0)	0.001	0.31	1.00	0.86
355	54.8	(50.7, 58.6)	< 0.001	73.9	(-120.1, 96.9)	0.22	0.98	1.00	0.98
356[LD]	55.8	(51.6, 59.6)	< 0.001	42.7	(25.7, 55.8)	< 0.001	0.26	1.00	0.80
357[LD]	54.9	(48.5, 60.6)	< 0.001	54.7	(49.9, 59.1)	< 0.001	0.80	1.00	0.98
359	55.7	(51.5, 59.6)	< 0.001	46.9	(33.7, 57.4)	< 0.001	0.26	1.00	0.80
361	57.7	(52.7, 62.1)	< 0.001	51.4	(45.4, 56.7)	< 0.001	0.06	1.00	0.33

CI denotes confidence interval, and FWER p-value adjusted p-value controlling the familywise error rate.

[†] For each amino acid position, haplotype-matched (mismatched) VE was computed only including clinical malaria endpoint events with a 3D7 matched (mismatched) amino acid at the given position.

[‡] Only amino acid positions with sufficiently high minor allele frequency were included in the analysis.

[LD] Linkage disequilibrium haplotype includes these Th2R and Th3R amino acid positions.

Table S8. Cumulative Vaccine Efficacy (VE) against "Any 3D7 Matched" and "No 3D7 Matche" Primary Clinical Malaria at 12 Months After Vaccination by Haplotype Locus in Children Aged 5-17 Months.

Estimates were stratified by study site.

Haplotype	"Any 3D7 Matched" Efficacy[†]			"No 3D7 Matched" Efficacy[†]			Differential VE
Locus	VE (%)	95% CI	P-value	VE (%)	95% CI	P-value	P-value
CS C-Terminus	54.1	(43.0, 63.1)	< 0.001	30.6	(25.8, 35.1)	< 0.001	< 0.001
DV10	53.3	(43.2, 61.5)	< 0.001	29.7	(24.7, 34.4)	< 0.001	< 0.001
LD[§]	53.8	(43.1, 62.5)	< 0.001	30.3	(25.4, 34.9)	< 0.001	< 0.001
Th2R	54.1	(43.0, 63.1)	< 0.001	30.6	(25.7, 35.1)	< 0.001	< 0.001
Th3R	48.9	(38.2, 57.8)	< 0.001	31.0	(26.0, 35.6)	< 0.001	0.007

CI denotes confidence interval.

[†] For each haplotype locus (CS C-terminus and epitope), "any matched" ("no matched") VE was computed only including clinical malaria endpoint events with any (no) 3D7 matched malaria in terms of the haplotype locus.

[§] Linkage disequilibrium haplotype includes Th2R and Th3R amino acid positions 314, 317, 352, 354, 356, and 357.

Table S9. Hazard Ratio Vaccine Efficacy (VE) against "Any 3D7 Matched" and "No 3D7 Matche" Primary Clinical Malaria at 12 Months After Vaccination by Haplotype Locus in Children Aged 5-17 Months.

Estimates were stratified by study site.

Haplotype Locus	"Any 3D7 Matched" Efficacy[†] VE (%)	95% CI	P-value	"No 3D7 Matched" Efficacy[†] VE (%)	95% CI	P-value	Differential VE P-value
CS C-Terminus	67.9	(59.6, 74.4)	< 0.001	52.1	(47.3, 56.4)	< 0.001	0.002
DV10	67.2	(59.7, 73.3)	< 0.001	51.5	(46.5, 55.9)	< 0.001	< 0.001
LD[§]	67.5	(59.5, 73.9)	< 0.001	51.9	(47.1, 56.3)	< 0.001	0.001
Th2R	67.8	(59.6, 74.4)	< 0.001	52.0	(47.3, 56.4)	< 0.001	0.002
Th3R	64.3	(56.2, 70.8)	< 0.001	52.3	(47.5, 56.7)	< 0.001	0.017

CI denotes confidence interval.

[†] For each haplotype locus (CS C-terminus and epitope), "any matched" ("no matched") VE was computed only including clinical malaria endpoint events with any (no) 3D7 matched malaria in terms of the haplotype locus.

[§] Linkage disequilibrium haplotype includes Th2R and Th3R amino acid positions 314, 317, 352, 354, 356, and 357.

Table S10. Cumulative Vaccine Efficacy (VE) Against 3D7 Matched and Mismatched Primary Clinical Malaria Through 12 Months Post Vaccination Follow-Up by CS C-terminus Haplotype and Position in Infants Aged 6-12 Weeks.
Estimates were stratified by study site.

Haplotype	Haplotype-Matched Efficacy[†]			Haplotype-Mismatched Efficacy[†]			Differential Efficacy		
Locus[‡]	VE (%)	95% CI	P-value	VE (%)	95% CI	P-value	P-value	FWER P-value	Q-value
CS C-Terminus	7.9	(-31.6, 35.6)	0.65	17.4	(11.1, 23.2)	< 0.001	0.58	–	–
DV10	6.5	(-28.6, 31.9)	0.68	17.6	(11.2, 23.5)	< 0.001	0.46	1.00	0.77
294	16.8	(10.7, 22.6)	< 0.001	19.5	(-42.3, 54.5)	0.46	0.89	1.00	0.93
295	16.2	(10.1, 21.9)	< 0.001	73.0	(28.8, 89.8)	0.01	0.03	0.84	0.39
296	17.0	(10.9, 22.6)	< 0.001	-90.8	(-493.0, 38.6)	0.72	0.57	1.00	0.79
298	17.4	(10.7, 23.5)	< 0.001	13.5	(-9.0, 31.3)	0.22	0.72	1.00	0.88
299	17.5	(11.4, 23.1)	< 0.001	-59.0	(-266.3, 31.0)	0.28	0.13	1.00	0.39
301	4.4	(-31.5, 30.4)	0.78	17.7	(11.4, 23.6)	< 0.001	0.38	1.00	0.77
302	16.2	(10.0, 22.0)	< 0.001	64.6	(14.1, 85.4)	0.02	0.06	1.00	0.39
303	16.7	(10.6, 22.4)	< 0.001	43.4	(-69.2, 81.1)	0.31	0.49	1.00	0.77
305	16.9	(10.8, 22.5)	< 0.001	55.3	(NA, NA)	NA	0.81	1.00	0.92
LD[§]	-6.6	(-47.7, 23)	0.70	18.3	(12, 24.1)	< 0.001	0.13	1.00	0.39
Th2R	5.8	(-34.3, 33.9)	0.74	17.5	(11.2, 23.3)	< 0.001	0.49	1.00	0.77
314^{LD}	15.8	(7.0, 23.8)	< 0.001	18.2	(8.2, 27.2)	< 0.001	0.73	1.00	0.88
317^{LD}	7.8	(-18.2, 28.1)	0.52	17.9	(11.4, 24)	< 0.001	0.40	1.00	0.77
318	16.7	(0.2, 30.5)	0.05	16.9	(9.9, 23.4)	< 0.001	0.98	1.00	0.98
320	16.1	(9.9, 21.8)	< 0.001	69.7	(23.2, 88)	0.01	0.03	0.89	0.39
321	15.1	(-3.4, 30.2)	0.10	17.2	(10.4, 23.6)	< 0.001	0.82	1.00	0.92
322	14.1	(2.2, 24.6)	0.02	18.4	(10.3, 25.7)	< 0.001	0.57	1.00	0.79
324	20.9	(13.1, 28.1)	< 0.001	9.9	(-1.9, 20.3)	0.10	0.12	1.00	0.39
327	18.1	(11.5, 24.2)	< 0.001	8.9	(-13.2, 26.6)	0.40	0.38	1.00	0.77
Th3R	-6.1	(-41.9, 20.7)	0.69	18.6	(12.3, 24.5)	< 0.001	0.09	1.00	0.39
349	16.7	(10.6, 22.4)	< 0.001	45.0	(-60.8, 81.2)	0.28	0.41	1.00	0.77
352^{LD}	14.2	(7.1, 20.8)	< 0.001	28.9	(13.2, 41.9)	< 0.001	0.10	1.00	0.39
354^{LD}	15.2	(8.6, 21.3)	< 0.001	35.0	(11.8, 52.2)	0.01	0.11	1.00	0.39
355	17.0	(10.9, 22.7)	< 0.001	-17.9	(-415.3, 73.0)	0.83	0.66	1.00	0.87
356^{LD}	14.5	(7.8, 20.7)	< 0.001	36.2	(16.9, 51.0)	< 0.001	0.04	1.00	0.39
357^{LD}	17.5	(5.8, 27.7)	0.004	16.6	(8.3, 24.0)	< 0.001	0.90	1.00	0.93
359	17.8	(11.2, 24)	< 0.001	10.5	(-11.9, 28.5)	0.33	0.50	1.00	0.77
361	14.5	(4.9, 23.2)	0.004	19.3	(10.0, 27.6)	< 0.001	0.49	1.00	0.77

CI denotes confidence interval, FWER p-value adjusted p-value controlling the familywise error rate, and NA not available result due to insufficient data.

[†] For each haplotype locus (CS C-terminus, epitope, amino acid position), haplotype-matched (mismatched) VE was computed only including clinical malaria endpoint events with 3D7 matched (mismatched) malaria in terms of the haplotype locus. For CS C-terminus, haplotype-matched VE was based on 52 full amplicon 3D7 matched events and 3137.3 person-years in RTS,S vaccine recipients with event rate of 0.017, and 29 events and 1468.3 person-years in control vaccine recipients with event rate of 0.020. Haplotype-mismatched VE was based on 903 full amplicon 3D7 mismatched events and 3137.3 person-years in RTS,S vaccine recipients with event rate of 0.29, and 554 events and 1468.3 person-years in control vaccine recipients with event rate of 0.38.

[‡] Only amino acid positions with sufficiently high minor allele frequency were included in the analysis.

[§] Linkage disequilibrium haplotype includes Th2R and Th3R amino acid positions marked as ^{LD}.

Table S11. Hazard Ratio Vaccine Efficacy (VE) Against 3D7 Matched and Mismatched Primary Clinical Malaria Through 12 Months Post Vaccination Follow-Up by CS C-terminus Haplotype and Position in Infants Aged 6-12 Weeks.
Estimates were stratified by study site.

Haplotype	Haplotype-Matched Efficacy[†]			Haplotype-Mismatched Efficacy[†]			Differential Efficacy		
Locus[‡]	VE (%)	95% CI	P-value	VE (%)	95% CI	P-value	P-value	FWER P-value	Q-value
CS C-Terminus	21.1	(-13.7, 45.2)	0.17	31.3	(23.7, 38.1)	< 0.001	0.68	–	–
DV10	19.5	(-12.3, 42.3)	0.19	31.5	(23.9, 38.4)	< 0.001	0.53	1.00	0.91
294	30.7	(23.1, 37.6)	< 0.001	33.0	(-20.9, 62.9)	0.16	0.84	1.00	0.91
295	30.2	(22.6, 37.1)	< 0.001	77.4	(39.9, 91.5)	0.01	0.53	1.00	0.91
296	30.9	(23.4, 37.7)	< 0.001	NA	(NA, NA)	NA	NA	NA	NA
298	31.1	(23.2, 38.1)	< 0.001	29.0	(9.6, 44.3)	0.004	0.88	1.00	0.91
299	31.3	(23.8, 38)	< 0.001	-28.7	(-201.9, 45.1)	0.72	0.96	1.00	0.96
301	16.5	(-15.8, 39.9)	0.31	31.7	(24.1, 38.6)	< 0.001	0.35	1.00	0.91
302	30.3	(22.6, 37.1)	< 0.001	69.9	(24.4, 88)	0.01	0.55	1.00	0.91
303	30.6	(23, 37.4)	< 0.001	57.9	(-33.7, 86.7)	0.13	0.72	1.00	0.91
305	30.8	(23.2, 37.6)	< 0.001	NA	(NA, NA)	NA	NA	NA	NA
LD[§]	6.8	(-31.2, 33.8)	0.78	32.2	(24.6, 39)	< 0.001	0.10	1.00	0.91
Th2R	19.9	(-15.5, 44.4)	0.22	31.4	(23.7, 38.2)	< 0.001	0.59	1.00	0.91
314^{LD}	29.1	(19.4, 37.5)	< 0.001	33.0	(22.8, 41.8)	< 0.001	0.64	1.00	0.91
317^{LD}	20.6	(-3.2, 38.9)	0.07	32.0	(24.2, 38.9)	< 0.001	0.32	1.00	0.91
318	30.1	(14.6, 42.9)	< 0.001	31.0	(22.8, 38.3)	< 0.001	0.79	1.00	0.91
320	30.2	(22.5, 37.1)	< 0.001	73.0	(31.9, 89.3)	0.01	0.50	1.00	0.91
321	28.4	(11.1, 42.3)	0.002	31.3	(23.3, 38.4)	< 0.001	0.82	1.00	0.91
322	28.6	(16.8, 38.8)	< 0.001	32.0	(23.2, 39.7)	< 0.001	0.75	1.00	0.91
324	34.0	(25.4, 41.7)	< 0.001	25.3	(13.2, 35.6)	< 0.001	0.43	1.00	0.91
327	32.2	(24.4, 39.2)	< 0.001	21.8	(0.8, 38.3)	0.04	0.77	1.00	0.91
Th3R	8.0	(-25.1, 32.4)	0.71	32.5	(24.9, 39.3)	< 0.001	0.08	1.00	0.91
349	30.6	(23.1, 37.5)	< 0.001	54.9	(-38.3, 85.3)	0.07	0.72	1.00	0.91
352^{LD}	28.4	(20, 35.9)	< 0.001	41.5	(27.2, 53)	< 0.001	0.24	1.00	0.91
354^{LD}	29.4	(21.4, 36.5)	< 0.001	46.3	(26.1, 61.1)	< 0.001	0.35	1.00	0.91
355	30.9	(23.3, 37.7)	< 0.001	7.4	(-318.9, 79.5)	0.85	0.85	1.00	0.91
356^{LD}	28.8	(20.7, 36)	< 0.001	47.2	(30.2, 60.1)	< 0.001	0.23	1.00	0.91
357^{LD}	29.9	(18, 40)	< 0.001	31.3	(22.5, 39.1)	< 0.001	0.81	1.00	0.91
359	31.6	(23.7, 38.6)	< 0.001	25.7	(5.4, 41.6)	0.01	0.86	1.00	0.91
361	28.9	(18.9, 37.7)	< 0.001	32.7	(23.1, 41.1)	< 0.001	0.71	1.00	0.91

CI denotes confidence interval, FWER p-value adjusted p-value controlling the familywise error rate, and NA not available result due to insufficient data.

[†] For each haplotype locus (CS C-terminus, epitope, amino acid position), haplotype-matched (mismatched) VE was computed only including clinical malaria endpoint events with 3D7 matched (mismatched) malaria in terms of the haplotype locus. For CS C-terminus, haplotype-matched VE was based on 52 full amplicon 3D7 matched events and 3137.3 person-years in RTS,S vaccine recipients with event rate of 0.017, and 29 events and 1468.3 person-years in control vaccine recipients with event rate of 0.020. Haplotype-mismatched VE was based on 903 full amplicon 3D7 mismatched events and 3137.3 person-years in RTS,S vaccine recipients with event rate of 0.29, and 554 events and 1468.3 person-years in control vaccine recipients with event rate of 0.38.

[‡] Only amino acid positions with sufficiently high minor allele frequency were included in the analysis.

[§] Linkage disequilibrium haplotype includes Th2R and Th3R amino acid positions marked as ^{LD}.

Table S12. Cumulative Vaccine Efficacy (VE) Against 3D7 Matched and Mismatched Primary Clinical Malaria Through 12 Months Post Vaccination Follow-Up by SERA2 Haplotype and Position in Children Aged 5-17 Months.

Estimates were stratified by study site.

Haplotype	Haplotype-Matched Efficacy[†]			Haplotype-Mismatched Efficacy[†]			Differential Efficacy		
Locus[‡]	VE (%)	95% CI	P-value	VE (%)	95% CI	P-value	P-value	FWER P-value	Q-value
SERA-2	-12.1	(-218.7, 60.6)	0.83	34.9	(31.0, 38.6)	< 0.001	0.32	–	–
49	35.9	(31.9, 39.6)	< 0.001	13.0	(-15.5, 34.5)	0.33	0.05	0.72	0.57
50	34.6	(30.6, 38.3)	< 0.001	39.1	(8.8, 59.3)	0.02	0.73	1.00	0.95
51	34.2	(30.0, 38.1)	< 0.001	42.2	(25.1, 55.4)	< 0.001	0.35	1.00	0.91
52	34.1	(28.5, 39.2)	< 0.001	35.7	(28.7, 42.0)	< 0.001	0.73	1.00	0.95
53	35.2	(31.3, 38.9)	< 0.001	7.0	(-47.7, 41.5)	0.76	0.14	1.00	0.57
54	34.7	(30.7, 38.4)	< 0.001	36.3	(4.1, 57.7)	0.03	0.90	1.00	0.95
66	34.4	(29.5, 38.9)	< 0.001	35.7	(26.7, 43.7)	< 0.001	0.80	1.00	0.95
85	34.8	(30.9, 38.4)	< 0.001	15.5	(-225.8, 78.1)	0.81	0.70	1.00	0.95
86	33.9	(29.8, 37.8)	< 0.001	46.9	(30.3, 59.6)	< 0.001	0.13	1.00	0.57
87	35.0	(31.1, 38.6)	< 0.001	-349.7	(-6364.3, 68.7)	0.27	0.09	1.00	0.57
88	31.8	(23.3, 39.3)	< 0.001	36.1	(31.1, 40.7)	< 0.001	0.40	1.00	0.91
89	34.9	(30.0, 39.4)	< 0.001	34.3	(25.2, 42.3)	< 0.001	0.92	1.00	0.95
90	34.7	(30.9, 38.4)	< 0.001	33.4	(-27.7, 65.2)	0.22	0.95	1.00	0.95
91	35.8	(22.8, 46.7)	< 0.001	34.5	(30.2, 38.6)	< 0.001	0.85	1.00	0.95
92	34.9	(31.0, 38.5)	< 0.001	-30.9	(-377.7, 64.1)	0.68	0.31	1.00	0.91
93	34.8	(30.8, 38.5)	< 0.001	33.3	(-7.8, 58.8)	0.10	0.94	1.00	0.95

CI denotes confidence interval, and FWER p-value adjusted p-value controlling the familywise error rate.

[†] For each haplotype locus (SERA-2, amino acid position), haplotype-matched (mismatched) VE was computed only including clinical malaria endpoint events with 3D7 matched (mismatched) malaria in terms of the haplotype locus. For SERA-2, haplotype-matched VE was based on 8 full amplicon 3D7 matched events and 3537.4 person-years in RTS,S vaccine recipients with event rate of 0.0023, and 4 events and 1529.6 person-years in control vaccine recipients with event rate of 0.0025. Haplotype-mismatched VE was based on 1183 full amplicon 3D7 mismatched events and 3537.4 person-years in RTS,S vaccine recipients with event rate of 0.33, and 919 events and 1529.6 person-years in control vaccine recipients with event rate of 0.60.

[‡] Only amino acid positions with sufficiently high minor allele frequency were included in the analysis.

Table S13. Hazard Ratio Vaccine Efficacy (VE) Against 3D7 Matched and Mismatched Primary Clinical Malaria Through 12 Months Post Vaccination Follow-Up by SERA2 Haplotype and Position in Children Aged 5-17 Months.

Estimates were stratified by study site.

Haplotype	Haplotype-Matched Efficacy[†]			Haplotype-Mismatched Efficacy[†]			Differential Efficacy		
Locus[‡]	VE (%)	95% CI	P-value	VE (%)	95% CI	P-value	P-value	FWER P-value	Q-value
SERA-2	26.0	(-119.5, 75.0)	0.73	55.3	(51.2, 59.0)	< 0.001	0.41	–	–
49	55.9	(51.8, 59.6)	< 0.001	40.9	(19.9, 56.4)	< 0.001	0.16	1.00	0.88
50	55.0	(50.8, 58.7)	< 0.001	60.3	(38.9, 74.2)	< 0.001	0.81	1.00	0.88
51	54.6	(50.4, 58.5)	< 0.001	61.7	(49.6, 70.9)	< 0.001	0.88	1.00	0.88
52	54.4	(49.3, 58.9)	< 0.001	56.1	(50.5, 61.2)	< 0.001	0.82	1.00	0.88
53	55.4	(51.4, 59.2)	< 0.001	39.3	(1.2, 62.7)	0.02	0.34	1.00	0.88
54	55.1	(51.0, 58.9)	< 0.001	55.3	(32.2, 70.5)	< 0.001	0.69	1.00	0.88
55	45.2	(18.7, 63.1)	0.002	55.5	(51.4, 59.2)	< 0.001	0.28	1.00	0.88
66	54.6	(50.0, 58.8)	< 0.001	56.5	(49.5, 62.4)	< 0.001	0.84	1.00	0.88
85	55.2	(51.2, 58.9)	< 0.001	31.5	(-167.8, 82.5)	0.77	0.56	1.00	0.88
86	54.6	(50.4, 58.4)	< 0.001	63.0	(50.6, 72.3)	< 0.001	0.86	1.00	0.88
87	55.3	(51.3, 59.0)	< 0.001	NA	(NA, NA)	NA	NA	NA	NA
88	53.4	(46.7, 59.3)	< 0.001	55.9	(51.3, 60.1)	< 0.001	0.79	1.00	0.88
89	55.1	(50.5, 59.2)	< 0.001	55.3	(48.2, 61.3)	< 0.001	0.78	1.00	0.88
90	55.2	(51.1, 58.9)	< 0.001	54.0	(9.2, 76.7)	0.02	0.64	1.00	0.88
91	53.5	(43.8, 61.6)	< 0.001	55.4	(51.1, 59.3)	< 0.001	0.80	1.00	0.88
92	55.2	(51.2, 58.9)	< 0.001	24.0	(-192.4, 80.2)	0.74	0.50	1.00	0.88
93	55.2	(51.1, 59.0)	< 0.001	51.3	(19.4, 70.6)	0.003	0.61	1.00	0.88

CI denotes confidence interval, FWER p-value adjusted p-value controlling the familywise error rate, and NA not available result due to insufficient data.

[†] For each haplotype locus (SERA-2, amino acid position), haplotype-matched (mismatched) VE was computed only including clinical malaria endpoint events with 3D7 matched (mismatched) malaria in terms of the haplotype locus. For SERA-2, haplotype-matched VE was based on 8 full amplicon 3D7 matched events and 3537.4 person-years in RTS,S vaccine recipients with event rate of 0.0023, and 4 events and 1529.6 person-years in control vaccine recipients with event rate of 0.0025. Haplotype-mismatched VE was based on 1183 full amplicon 3D7 mismatched events and 3537.4 person-years in RTS,S vaccine recipients with event rate of 0.33, and 919 events and 1529.6 person-years in control vaccine recipients with event rate of 0.60.

[‡] Only amino acid positions with sufficiently high minor allele frequency were included in the analysis.

Table S14. Cumulative Vaccine Efficacy (VE) Against 3D7 Matched and Mismatched Primary Clinical Malaria Through 12 Months Post Vaccination Follow-Up by SERA2 Haplotype and Position in Infants Aged 6-12 Weeks.

Estimates were stratified by study site.

Haplotype	Haplotype-Matched Efficacy[†]			Haplotype-Mismatched Efficacy[†]			Differential Efficacy		
Locus[‡]	VE (%)	95% CI	P-value	VE (%)	95% CI	P-value	P-value	FWER P-value	Q-value
SERA-2	52.4	(-74.0, 87.0)	0.26	16.5	(10.6, 22.0)	< 0.001	0.42	–	–
46	16.8	(10.9, 22.3)	< 0.001	-88.4	(NA, NA)	NA	0.58	1.00	0.72
49	16.9	(10.8, 22.5)	< 0.001	12.1	(-33.0, 41.9)	0.54	0.80	1.00	0.85
50	15.8	(9.6, 21.5)	< 0.001	38.4	(3.7, 60.6)	0.03	0.18	1.00	0.68
51	17.6	(11.3, 23.3)	< 0.001	5.5	(-25.9, 29.0)	0.70	0.38	1.00	0.68
52	18.7	(10.7, 25.9)	< 0.001	13.5	(2.4, 23.4)	0.02	0.46	1.00	0.68
53	16.0	(9.9, 21.6)	< 0.001	39.8	(-1.5, 64.3)	0.06	0.21	1.00	0.68
54	18.0	(12.0, 23.5)	< 0.001	-38.3	(-121.5, 13.7)	0.18	0.04	0.65	0.65
55	41.5	(9.1, 62.3)	0.02	15.6	(9.5, 21.4)	< 0.001	0.11	1.00	0.68
63	16.9	(11.1, 22.4)	< 0.001	NA	(NA, NA)	NA	NA	NA	NA
66	16.5	(9.5, 23.0)	< 0.001	17.3	(2.9, 29.6)	0.02	0.92	1.00	0.92
85	16.9	(11.0, 22.4)	< 0.001	-409.3	(NA, NA)	NA	0.33	1.00	0.68
86	17.5	(11.3, 23.2)	< 0.001	4.4	(-33.4, 31.5)	0.79	0.42	1.00	0.68
88	13.0	(-0.2, 24.4)	0.05	18.2	(10.9, 24.9)	< 0.001	0.49	1.00	0.68
89	17.6	(10.5, 24.2)	< 0.001	13.7	(-1.4, 26.6)	0.07	0.64	1.00	0.72
90	16.9	(11.0, 22.4)	< 0.001	0.9	(-97.9, 50.4)	0.98	0.64	1.00	0.72
91	24.2	(7.3, 38.0)	0.01	15.2	(8.5, 21.5)	< 0.001	0.33	1.00	0.68
92	16.9	(11.0, 22.4)	< 0.001	-109.4	(NA, NA)	NA	0.46	1.00	0.68
93	16.3	(10.3, 21.9)	< 0.001	29.5	(-14.5, 56.6)	0.16	0.49	1.00	0.68

CI denotes confidence interval, FWER p-value adjusted p-value controlling the familywise error rate, and NA not available result due to insufficient data.

[†] For each haplotype locus (SERA-2, amino acid position), haplotype-matched (mismatched) VE was computed only including clinical malaria endpoint events with 3D7 matched (mismatched) malaria in terms of the haplotype locus. For SERA-2, haplotype-matched VE was based on 3 full amplicon 3D7 matched events and 3137.3 person-years in RTS,S vaccine recipients with event rate of 0.0010, and 3 events and 1468.3 person-years in control vaccine recipients with event rate of 0.0022. Haplotype-mismatched VE was based on 982 full amplicon 3D7 mismatched events and 3137.3 person-years in RTS,S vaccine recipients with event rate of 0.31, and 598 events and 1468.3 person-years in control vaccine recipients with event rate of 0.31.

[‡] Only amino acid positions with sufficiently high minor allele frequency were included in the analysis.

Table S15. Hazard Ratio Vaccine Efficacy (VE) Against 3D7 Matched and Mismatched Primary Clinical Malaria Through 12 Months Post Vaccination Follow-Up by SERA2 Haplotype and Position in Infants Aged 6-12 Weeks.

Estimates were stratified by study site.

Haplotype	Haplotype-Matched Efficacy[†]			Haplotype-Mismatched Efficacy[†]			Differential Efficacy		
Locus[‡]	VE (%)	95% CI	P-value	VE (%)	95% CI	P-value	P-value	FWER P-value	Q-value
SERA-2	57.3	(-66.2, 89.0)	0.26	30.7	(23.2, 37.4)	< 0.001	0.55	–	–
46	30.9	(23.5, 37.6)	< 0.001	NA	(NA, NA)	NA	NA	NA	NA
49	31.1	(23.6, 37.8)	< 0.001	23.8	(-16, 50)	0.2	0.90	1.00	0.91
50	30.1	(22.5, 36.9)	< 0.001	47.9	(17.9, 66.9)	0.003	0.47	1.00	0.91
51	31.5	(23.9, 38.3)	< 0.001	22.0	(-6.2, 42.6)	0.1	0.90	1.00	0.91
52	32.9	(24.2, 40.6)	< 0.001	27.4	(15.9, 37.3)	< 0.001	0.78	1.00	0.91
53	30.2	(22.6, 37.0)	< 0.001	50.7	(16.4, 70.9)	0.01	0.50	1.00	0.91
54	31.8	(24.4, 38.5)	< 0.001	-12.5	(-81.5, 30.3)	0.77	0.79	1.00	0.91
55	49.6	(20.4, 68.1)	0.003	30.0	(22.4, 36.9)	< 0.001	0.13	1.00	0.91
63	31.0	(23.6, 37.7)	< 0.001	NA	(NA, NA)	NA	NA	NA	NA
66	30.9	(22.8, 38.2)	< 0.001	30.3	(16.6, 41.7)	< 0.001	0.81	1.00	0.91
85	30.9	(23.5, 37.6)	< 0.001	NA	(NA, NA)	NA	NA	NA	NA
86	31.4	(23.9, 38.2)	< 0.001	21.1	(-11.7, 44.4)	0.17	0.91	1.00	0.91
88	27.9	(15.6, 38.3)	< 0.001	32.0	(23.8, 39.3)	< 0.001	0.63	1.00	0.91
89	31.6	(23.5, 38.9)	< 0.001	28.1	(14, 39.9)	< 0.001	0.85	1.00	0.91
90	31.0	(23.6, 37.7)	< 0.001	14.6	(-73.9, 58.1)	0.77	0.91	1.00	0.91
91	36.7	(21.3, 49.1)	< 0.001	29.6	(21.6, 36.8)	< 0.001	0.60	1.00	0.91
92	31.0	(23.5, 37.6)	< 0.001	NA	(NA, NA)	NA	NA	NA	NA
93	30.4	(22.9, 37.2)	< 0.001	43.3	(6, 65.8)	0.02	0.65	1.00	0.91

CI denotes confidence interval, FWER p-value adjusted p-value controlling the familywise error rate, and NA not available result due to insufficient data.

[†] For each haplotype locus (SERA-2, amino acid position), haplotype-matched (mismatched) VE was computed only including clinical malaria endpoint events with 3D7 matched (mismatched) malaria in terms of the haplotype locus. For SERA-2, haplotype-matched VE was based on 3 full amplicon 3D7 matched events and 3137.3 person-years in RTS,S vaccine recipients with event rate of 0.0010, and 3 events and 1468.3 person-years in control vaccine recipients with event rate of 0.0022. Haplotype-mismatched VE was based on 982 full amplicon 3D7 mismatched events and 3137.3 person-years in RTS,S vaccine recipients with event rate of 0.31, and 598 events and 1468.3 person-years in control vaccine recipients with event rate of 0.31.

[‡] Only amino acid positions with sufficiently high minor allele frequency were included in the analysis.

Table S16. Vaccine Efficacy (VE) Against 3D7 Matched and Mismatched Parasite Positivity at 18 Months Post Vaccine Dose 3 by CS C-terminus Haplotype and Position in Children Aged 5-17 Months.

Estimates were stratified by study site.

Haplotype	Haplotype-Matched Efficacy[†]			Haplotype-Mismatched Efficacy[†]			Differential Efficacy		
Locus[‡]	VE (%)	95% CI	P-value	VE (%)	95% CI	P-value	P-value	FWER P-value	Q-value
CS C-Terminus	53.2	(13.8, 74.6)	0.01	30.1	(17.5, 40.8)	< 0.001	0.19	–	–
DV10	23.2	(-24.5, 52.7)	0.28	31.5	(19.0, 42.0)	< 0.001	0.70	1.00	0.86
294	31.6	(19.2, 42.1)	< 0.001	6.2	(-126.9, 61.2)	0.89	0.55	1.00	0.75
295	30.8	(18.4, 41.3)	< 0.001	67.4	(10.4, 88.2)	0.03	0.48	1.00	0.70
298	28.5	(15.1, 39.8)	< 0.001	45.6	(22.7, 61.7)	< 0.001	0.13	1.00	0.59
299	30.6	(18.0, 41.2)	< 0.001	50.5	(-16.6, 79.0)	0.11	0.45	1.00	0.69
301	24.2	(-22.4, 53)	0.26	31.4	(19.0, 42.0)	< 0.001	0.74	1.00	0.86
302	31.2	(18.9, 41.7)	< 0.001	22.8	(-151.7, 76.3)	0.67	0.89	1.00	0.89
303	31.5	(19.2, 41.9)	< 0.001	-235.4	(NA, NA)	NA	0.42	1.00	0.69
LD[§]	46.9	(9.7, 68.8)	0.02	30.2	(17.5, 40.9)	< 0.001	0.29	1.00	0.66
Th2R	56.2	(20.6, 75.8)	0.01	29.9	(17.2, 40.6)	< 0.001	0.12	1.00	0.59
314^{LD}	30.5	(15.5, 42.9)	< 0.001	31.9	(15.0, 45.5)	< 0.001	0.87	1.00	0.89
317^{LD}	9.5	(-35.8, 39.6)	0.63	32.9	(20.5, 43.3)	< 0.001	0.17	1.00	0.59
318	22.5	(-6.8, 43.8)	0.12	32.7	(19.9, 43.4)	< 0.001	0.42	1.00	0.69
320	31.0	(18.6, 41.5)	< 0.001	50.0	(-44.1, 82.6)	0.2	0.75	1.00	0.86
321	16.5	(-16.6, 40.3)	0.29	33.4	(20.9, 44.0)	< 0.001	0.20	1.00	0.59
322	24.5	(3.7, 40.8)	0.02	34.2	(20.7, 45.4)	< 0.001	0.32	1.00	0.66
Th3R	46.7	(14.9, 66.6)	0.01	29.8	(17.0, 40.7)	< 0.001	0.24	1.00	0.62
352^{LD}	34.1	(21.5, 44.6)	< 0.001	12.9	(-23.2, 38.4)	0.43	0.13	1.00	0.59
356^{LD}	32.2	(19.8, 42.7)	< 0.001	16.8	(-29.7, 46.7)	0.42	0.40	1.00	0.69
357^{LD}	32.0	(12.8, 47.0)	< 0.001	30.7	(16.5, 42.5)	< 0.001	0.89	1.00	0.89
359	33.3	(20.8, 43.8)	< 0.001	13.3	(-26.4, 40.5)	0.46	0.19	1.00	0.59
361	36.9	(22.3, 48.8)	< 0.001	24.8	(7.3, 39.0)	0.01	0.18	1.00	0.59
376[¶]	32.7	(20.5, 43.0)	< 0.001	-24.7	(-130.7, 32.6)	0.48	0.05	1.00	0.59

CI denotes confidence interval, FWER p-value adjusted p-value controlling the familywise error rate, and NA not available result due to insufficient data.

[†] For each haplotype locus (CS C-terminus, epitope, amino acid position), haplotype-matched (mismatched) VE was computed only including any malaria endpoint events with 3D7 matched (mismatched) malaria in terms of the haplotype locus. For CS C-terminus, haplotype-matched VE was based on 8 full amplicon 3D7 matched events and 4580 assessed participants in the RTS,S vaccine group with event rate of 0.0017, and 9 events and 2337 assessed participants in the control vaccine group with event rate of 0.0039. Haplotype-mismatched VE was based on 277 full amplicon 3D7 mismatched events and 4580 assessed participants in the RTS,S vaccine group with event rate of 0.060, and 201 events and 2337 assessed participants in the control vaccine group with event rate of 0.086.

[‡] Only amino acid positions with sufficiently high minor allele frequency were included in the analysis.

[§] Linkage disequilibrium haplotype includes Th2R and Th3R amino acid positions marked as ^{LD}.

[¶] Amino acid position 376 is included in CST.

Table S17. Vaccine Efficacy (VE) Against 3D7 Matched and Mismatched Parasite Positivity at 18 Months Post Vaccine Dose 3 by CS C-terminus Haplotype and Position in Infants Aged 6-12 Weeks.

Estimates were stratified by study site.

Haplotype	Haplotype-Matched Efficacy[†]			Haplotype-Mismatched Efficacy[†]			Differential Efficacy		
Locus[‡]	VE (%)	95% CI	P-value	VE (%)	95% CI	P-value	P-value	FWER P-value	Q-value
CS C-Terminus	15.6	(-106.9, 65.6)	0.71	12.0	(-8.0, 28.3)	0.22	0.89	–	–
DV10	24.0	(-56.9, 63.2)	0.46	11.4	(-8.8, 27.9)	0.25	0.65	1.00	0.97
294	11.2	(-9.0, 27.7)	0.26	48.4	(-98.1, 86.5)	0.34	0.44	1.00	0.97
295	11.7	(-8.2, 28.0)	0.23	41.6	(-211.8, 89.1)	0.53	0.62	1.00	0.97
298	10.4	(-11.1, 27.8)	0.32	21.1	(-22.7, 49.3)	0.29	0.59	1.00	0.97
299	12.3	(-7.6, 28.5)	0.21	4.2	(-271.8, 75.3)	0.95	0.93	1.00	0.97
301	26.6	(-49.1, 63.9)	0.39	11.3	(-9.0, 27.8)	0.26	0.57	1.00	0.97
302	12.2	(-7.6, 28.3)	0.21	10.6	(NA, NA)	NA	1.00	1.00	1.00
LD[§]	7.2	(-104.4, 57.9)	0.85	12.3	(-7.7, 28.6)	0.21	0.93	1.00	0.97
Th2R	16.1	(-102.2, 65.2)	0.70	12.0	(-8.1, 28.3)	0.22	0.88	1.00	0.97
314^{LD}	15.1	(-8.7, 33.8)	0.19	8.1	(-20.7, 30.1)	0.54	0.63	1.00	0.97
317^{LD}	17.8	(-43.2, 52.8)	0.49	11.5	(-9.3, 28.3)	0.26	0.78	1.00	0.97
318	8.3	(-40.1, 39.9)	0.69	12.9	(-8.1, 29.7)	0.21	0.83	1.00	0.97
320	12.9	(-6.7, 28.9)	0.18	-582.6	(NA, NA)	NA	0.36	1.00	0.97
321	19.9	(-20.9, 46.8)	0.29	10.5	(-10.9, 27.8)	0.31	0.60	1.00	0.97
322	17.2	(-11.5, 38.5)	0.21	9.4	(-14.2, 28.1)	0.40	0.59	1.00	0.97
324	9.7	(-14.9, 29.1)	0.41	15.6	(-11.9, 36.4)	0.24	0.68	1.00	0.97
327	11.1	(-10.1, 28.3)	0.28	17.2	(-26.7, 45.9)	0.38	0.74	1.00	0.97
Th3R	23.1	(-56.7, 62.2)	0.47	11.4	(-9.0, 28.0)	0.25	0.69	1.00	0.97
352^{LD}	13.8	(-6.9, 30.4)	0.18	2.2	(-54.0, 38.0)	0.92	0.61	1.00	0.97
354^{LD}	15.1	(-4.2, 30.9)	0.12	-85.2	(NA, NA)	NA	0.03	0.69	0.69
357^{LD}	14.0	(-19.7, 38.2)	0.37	11.3	(-11.7, 29.6)	0.31	0.87	1.00	0.97
359	11.6	(-9.2, 28.4)	0.25	15.7	(-35.5, 47.6)	0.48	0.83	1.00	0.97
361	11.4	(-14.6, 31.5)	0.36	12.9	(-13.2, 32.9)	0.30	0.92	1.00	0.97
376[¶]	11.9	(-8.0, 28.2)	0.22	18.5	(NA, NA)	NA	0.82	1.00	0.97

CI denotes confidence interval, FWER p-value adjusted p-value controlling the familywise error rate, and NA not available result due to insufficient data.

[†] For each haplotype locus (CS C-terminus, epitope, amino acid position), haplotype-matched (mismatched) VE was computed only including any malaria endpoint events with 3D7 matched (mismatched) malaria in terms of the haplotype locus. For CS C-terminus, haplotype-matched VE was based on 7 full amplicon 3D7 matched events and 3991 assessed participants in the RTS,S vaccine group with event rate of 0.0017, and 4 events and 2010 assessed participants in the control vaccine group with event rate of 0.0020. Haplotype-mismatched VE was based on 215 full amplicon 3D7 mismatched events and 3991 assessed participants in the RTS,S vaccine group with event rate of 0.054, and 125 events and 2010 assessed participants in the control vaccine group with event rate of 0.062.

[‡] Only amino acid positions with sufficiently high minor allele frequency were included in the analysis.

[§] Linkage disequilibrium haplotype includes Th2R and Th3R amino acid positions marked as ^{LD}.

[¶] Amino acid position 376 is included in CST.

Table S18. Vaccine Efficacy (VE) Against 3D7 Matched and Mismatched Parasite Positivity at 18 Months Post Vaccine Dose 3 by SERA2 Haplotype and Position in Children Aged 5-17 Months.

Estimates were stratified by study site.

Haplotype	Haplotype-Matched Efficacy[†]			Haplotype-Mismatched Efficacy[†]			Differential Efficacy		
Locus[‡]	VE (%)	95% CI	P-value	VE (%)	95% CI	P-value	P-value	FWER P-value	Q-value
SERA-2	-40.6	(NA, NA)	NA	31.5	(19.3, 41.9)	< 0.001	0.63	–	–
49	29.1	(16.1, 40.0)	< 0.001	58.1	(26.3, 76.2)	0.003	0.07	0.79	0.62
50	31.4	(19.1, 41.9)	< 0.001	25.2	(-54.9, 63.9)	0.43	0.87	1.00	0.96
53	30.6	(18.2, 41.2)	< 0.001	63.3	(NA, NA)	NA	0.21	1.00	0.71
54	31.7	(19.3, 42.1)	< 0.001	22.0	(-49.9, 59.5)	0.46	0.72	1.00	0.89
66	27.0	(12.6, 39.1)	< 0.001	42.6	(24.3, 56.5)	< 0.001	0.11	1.00	0.62
88	35.3	(17.0, 49.5)	< 0.001	29.4	(15.0, 41.3)	< 0.001	0.53	1.00	0.73
89	29.4	(15.5, 40.9)	< 0.001	36.9	(16.7, 52.2)	0.001	0.45	1.00	0.71
90	30.7	(18.3, 41.3)	< 0.001	59.0	(-18.6, 85.8)	0.1	0.34	1.00	0.71
91	30.6	(2.1, 50.9)	0.04	31.4	(18.5, 42.2)	< 0.001	0.96	1.00	0.96
92	31.1	(18.8, 41.5)	< 0.001	89.1	(NA, NA)	NA	0.43	1.00	0.71
93	32.0	(19.8, 42.3)	< 0.001	-6.3	(-193.4, 61.5)	0.91	0.39	1.00	0.71

CI confidence interval, FWER p-value adjusted p-value controlling the familywise error rate, and NA not available result due to insufficient data.

[†] For each haplotype locus (SERA-2, amino acid position), haplotype-matched (mismatched) VE was computed only including any malaria endpoint events with 3D7 matched (mismatched) malaria in terms of the haplotype locus. For SERA-2, haplotype-matched VE was based on 2 full amplicon 3D7 matched events and 4580 assessed participants in the RTS,S vaccine group with event rate of 0.00044, and 1 event and 2337 assessed participants in the control vaccine group with event rate of 0.00043. Haplotype-mismatched VE was based on 285 full amplicon 3D7 mismatched events and 4580 assessed participants in the RTS,S vaccine group with event rate of 0.062, and 211 events and 2337 assessed participants in the control vaccine group with event rate of 0.090.

[‡] Only amino acid positions with sufficiently high minor allele frequency were included in the analysis.

Table S19. Vaccine Efficacy (VE) Against 3D7 Matched and Mismatched Parasite Positivity at 18 Months Post Vaccine Dose 3 by SERA2 Haplotype and Position in Infants Aged 6-12 Weeks.

Estimates were stratified by study site.

Haplotype	Haplotype-Matched Efficacy[†]			Haplotype-Mismatched Efficacy[†]			Differential Efficacy		
Locus[‡]	VE (%)	95% CI	P-value	VE (%)	95% CI	P-value	P-value	FWER P-value	Q-value
SERA-2	59.0	(NA, NA)	NA	13.5	(-5.9, 29.4)	0.16	0.61	–	–
49	16.0	(-3.1, 31.6)	0.09	-36.0	(-179.8, 33.9)	0.40	0.21	1.00	0.87
50	14.5	(-4.8, 30.2)	0.13	-20.3	(-284.7, 62.4)	0.76	0.60	1.00	0.87
51	12.1	(-8.0, 28.5)	0.22	34.2	(-21.7, 64.5)	0.18	0.34	1.00	0.87
52	13.7	(-10.0, 32.4)	0.23	14.0	(-14.3, 35.2)	0.30	0.99	1.00	0.99
53	14.5	(-4.7, 30.2)	0.13	-30.3	(NA, NA)	NA	0.60	1.00	0.87
54	13.9	(-5.6, 29.8)	0.15	15.0	(-140.3, 69.9)	0.76	0.99	1.00	0.99
55	-24.2	(-308.6, 62.2)	0.72	14.5	(-4.8, 30.2)	0.13	0.58	1.00	0.87
66	17.1	(-3.8, 33.8)	0.10	4.6	(-33.6, 31.9)	0.78	0.46	1.00	0.87
86	15.3	(-4.1, 31.1)	0.11	-9.7	(-110.2, 42.8)	0.78	0.46	1.00	0.87
88	4.6	(-29.9, 29.9)	0.77	18.1	(-3.2, 35.0)	0.09	0.39	1.00	0.87
89	14.1	(-7.2, 31.2)	0.18	13.0	(-23.3, 38.6)	0.43	0.95	1.00	0.99
90	14.0	(-5.4, 29.8)	0.15	13.5	(-259.7, 79.2)	0.84	0.94	1.00	0.99
91	-3.8	(-58.7, 32.1)	0.86	16.7	(-3.0, 32.6)	0.09	0.34	1.00	0.87

CI denotes confidence interval, FWER p-value adjusted p-value controlling the familywise error rate, and NA not available result due to insufficient data.

[†] For each haplotype locus (SERA-2, amino acid position), haplotype-matched (mismatched) VE was computed only including any malaria endpoint events with 3D7 matched (mismatched) malaria in terms of the haplotype locus. For SERA-2, haplotype-matched VE was based on 1 full amplicon 3D7 matched event and 3991 assessed participants in the RTS,S vaccine group with event rate of 0.00025, and 1 event and 2010 assessed participants in the control vaccine group with event rate of 0.00050. Haplotype-mismatched VE was based on 220 full amplicon 3D7 mismatched events and 3991 assessed participants in the RTS,S vaccine group with event rate of 0.055, and 130 events and 2010 assessed participants in the control vaccine group with event rate of 0.065.

[‡] Only amino acid positions with sufficiently high minor allele frequency were included in the analysis.

REFERENCES

1. Picard Tools - By Broad Institute. Available from: http://broadinstitute.github.io/
2. Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improv assemblies. Bioinforma Oxf Engl 2011;27(21):2957–63.
3. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler Bioinforma Oxf Engl 2009;25(14):1754–60.
4. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment se Mol Biol 1990;215(3):403–10.
5. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nuc Res 1999;27(2):573–80.

Statistical Analysis Plan for the First Tier Sieve Analysis of Breakthrough Sequences in the GlaxoSmithKline Phase III RTS,S/AS01$_{\mathrm{E}}$ Malaria Vaccine Trial

1 Tier 1 Sieve Analysis Objectives

The objectives are to evaluate whether and how the efficacy of the vaccine to prevent defined malaria endpoints depends on the following characteristics of the infecting malaria parasite(s):

(a) full match vs. mismatch to the vaccine sequence 3D7 haplotype of the C-terminus of the CSP protein,

(b) full match vs. mismatch to the 3D7 haplotype of four epitope sub-regions of the C-terminus of the CSP protein: TH2R, TH3R, Unnamed epitope region, Combined set of residues in TH2R and TH3R in linkage disequilibrium,

(c) full match vs. mismatch to the 3D7 haplotype of amino acid residues at individual sites in the C-terminus of CSP,

(d) genetic distances to the vaccine (in the C-terminus of CSP and in each of the epitope sub-regions in (b)),

(e) the number of repeats of the NANP region of the CSP protein, and

(f) the number of distinct genomic variants of infecting malaria parasites (multiplicity of infection, MOI).

In addition, the sieve analyses for (a), (c), (d) and (f) will be repeated for the SERA-2 protein. This analysis serves as a control for the analysis of the CSP protein, given that SERA-2 is not in the vaccine and hence we assume that there should not be sieve effects in SERA-2. Sieve analyses (a)–(c) are primary whereas sieve analyses (d)–(f) are secondary. For analysis (b) the CST3 epitope region was also of interest; however there was no variability in the haplotype in malaria endpoint cases such that a sieve analysis is not possible.

The exact definitions of the malaria genetic variables analyzed in (a)–(f) are finalized based on treatment-blinded descriptive data analysis. A 'sieve effect' is defined as a statistically significant result of differential vaccine efficacy to prevent a defined malaria endpoint with a given founder malaria infection according to a defined genetic variable/feature of the infecting parasite; we refer to such a feature as a genotype or genetic distance. We also assess 'post-infection genomic effects', defined as a statistically significant result of a different mean value of the genetic variable/feature of the infecting parasite in vaccine recipient malaria endpoint cases versus placebo recipient malaria endpoint cases. The former sieve effect analyses include the entire attention to protocol (ATP) cohort in the analysis and assess genotype-specific vaccine efficacy target parameters that parallel overall vaccine efficacy target parameters that ignore malaria genomics, whereas the latter analyses include malaria endpoint cases only. Thus the post-infection analyses do not directly assess differential vaccine efficacy; nevertheless if a post-infection genomic effect is detected it constitutes evidence for vaccine pressure on the malaria genome that may suggest differential vaccine efficacy

and motivate future experiments. Moreover, the post-infection analyses may have greater power for some scenarios than the vaccine efficacy sieve analyses.

The sieve analyses are based on the following defined malaria endpoints and cohorts, which link back to the primary study publications in 2011 and 2012:

1. Primary case definition of clinical malaria in the ATP cohort (first or only episode of clinical malaria with >5000 parasites per cubic millimeter or a severe case, occurring at least 14 days post third immunization and within 385 days later, representing a right edge of the Month 14 study visit window),
2. Primary case definition of clinical malaria in the ATP as defined in 1. that also satisfies the protocol-definition of severe clinical malaria, and
3. Malaria infection at the Month 20 post first vaccination visit with >0 parasites per cubic millimeter (any malaria infection).

Primary analyses are performed only including the study sites with larger number of cases (Agogo, Kintampo, Kombewa, Nanoro, and Siaya). The primary analyses pool over these five sites and secondary analyses provide results separately for each of these sites. In addition, secondary analyses are done for all study sites pooled together. All analyses are performed separately for children (enrolled at 5 to 17 months of age) and infants (enrolled at 6 to 12 weeks of age).

Two types of sieve analyses are performed. 'Global' sieve analyses assess differential vaccine efficacy against parasite genotypes with genotype defined by a single number representing a feature of the C-terminus of the CSP protein (or of SERA-2), typically a low-dimensional categorical feature or an ordered categorical distance to the corresponding protein region of the 3D7 vaccine sequence. In contrast, 'local' sieve analyses scan individual amino acid sites to assess differential vaccine efficacy against parasite genotypes with genotype defined by the amino acid character relative to the corresponding amino acid character at the same site in the vaccine immunogen sequence. The sieve analyses are conducted on a multiply aligned data set of malaria sequences.

2 Pre-Screening of C-terminus CSP Amino Acid Sites and Regions for 3D7 Match Versus 3D7 Mismatch Sieve Analysis

To maximize statistical power to detect sieve effects, a pre-specified process is used to select a subset of the amino acid sites in the C-terminus of CSP to include in the global and local sieve analyses. The filtering procedures are based on treatment blinded data (and overall malaria endpoint case counts for the vaccine and placebo groups) and are summarized below.

- **Variability filter for local site-scanning sieve analysis**: For site-scanning sieve analysis, we remove all sites that have minimal sequence diversity measured across both treatment groups combined. Specifically, for each of the three malaria endpoints and 14 analyses separately, we select a variability threshold based on the number of malaria endpoint cases with evaluable sequence data that have a minority amino acid variant in at least one of their founder sequences. The threshold will be selected such that if all subjects with a minority

Table 1-shell. Required number of cases with a minority variant in at least one founder sequence to qualify an amino acid site for site-scanning sieve analysis

		Required Number of Cases (Number of Sites Qualifying)		
Site	Age Cohort	Malaria Endpoint 1	Malaria Endpoint 2	Malaria Endpoint 3
5 Major Pooled	6-12 Weeks			
Agogo	6-12 Weeks			
Kintampo	6-12 Weeks			
Kombewa	6-12 Weeks			
Nanoro	6-12 Weeks			
Siaya	6-12 Weeks			
11 Pooled	6-12 Weeks			
5 Major Pooled	5-17 Months			
Agogo	5-17 Months			
Kintampo	5-17 Months			
Kombewa	5-17 Months			
Nanoro	5-17 Months			
Siaya	5-17 Months			
11 Pooled	5-17 Months			

variant (in match to 3D7 vs. mismatch) at a given site in at least one of their founder sequences were in one of the treatment groups, then a significant sieve effect using Fisher's exact test would narrowly miss rejecting the null hypothesis of no sieve effect with 2-sided $p < 0.05$; i.e., adding one more subject with a minority variant would yield 2-sided $p < 0.05$.

The thresholds (defined in terms of requiring x or more cases with a match and x or more cases with a mismatch for a given x) cannot be computed until treatment-unblinding because they depend on the total numbers of infected vaccine recipients and infected placebo recipients. Table 1-shell provides a shell table of the thresholds for site-scanning analysis that will be provided in the statistical analysis report of the sieve reports.

- **Variability filter for other 3D7 match versus mismatch sieve analysis**:

  The same type of filter is used for the other primary binary mark analyses for the full CSP protein and for each of the epitope region analyses. Thus similar tables as Table 1-shell will be constructed for all of the other primary analyses of full 3D7 match vs. mismatch.

  Even without using the vaccine and placebo case counts, we can determine that the pooled-arm case counts are too low to support the sieve analyses of the severe malaria endpoint 2 for the full CSP protein. Therefore the sieve analyses will only be conducted for malaria endpoints 1 and 3. Moreover, based on treatment-blinded analysis we expect that several of the individual site analyses for malaria endpoints 1 and/or 3 will not be possible. In addition to Table 1-shell above, the statistical report implementing this SAP will include parallel tables for the full CSP protein and each of the epitope regions. These tables will be constructed immediately after unblinding to determine which of the sieve analyses can be conducted.

- **Variability filter for global distance-based sieve analysis:**

  Global sieve analyses are conducted using the same variability filter above based on full match vs. mismatch. The rationale for not using additional filtering is that distances (at the epitope or C-terminus CSP level) aggregate sequence differences over many sites and rare minor variants may contribute to a vaccine sieve effect. With Hamming distances, this approach affords a simple interpretation of a distance as the percentage (or equivalently number) of mismatches among all sites in the protein region under consideration.

3 Structure of the Sieve Analysis Statistical Methods

We consider sieve analysis statistical methods that are either in a competing risks failure time analysis framework or in a relative risk framework that considers final outcome data but ignores failure times. Analyses of the first two malaria endpoints use the former framework because subjects are followed prospectively from study entry until occurrence of the clinical malaria endpoint, whereas analyses of the third malaria endpoint (any malaria infection) use the latter framework because of the cross-sectional sampling of malaria sequences. The failure time analysis framework is advantageous because it can provide more interpretable results if the vaccine efficacy changes over time and it more fully accounts for the partial follow-up of some study subjects.

The competing risks failure time methods define the failure time T as the time from 14 days after the third immunization until the malaria endpoint under consideration out to a maximum of 385 days later (399 days after the third immunization date), which is one way to define the right edge of the Month 14 visit window. Subjects who remain endpoint-free at the Month 14 visit are right-censored at the earliest date of the Month 14 visit date or 385 days. Subjects who are lost to follow-up prior to the Month 14 visit or 385 days are right-censored on the date of last contact. At the time of diagnosis of the malaria endpoint, the malaria genotype is denoted by J. The primary analyses consider J as a dichotomous genomic feature, the indicator of whether a founder parasite mismatches the vaccine sequence at the given unit level of analysis. The secondary analyses consider J as an ordered categorical genetic distance to the vaccine with 3 or more categories (e.g., a Hamming distance based on the C-terminus of CSP).

A scaled genetic distance $\widetilde{V}$ between an infecting malaria amino acid sequence and the vaccine sequence may be computed as

$$\widetilde{V} = 1 - \frac{2\sum_i S(x_i, y_i)}{\sum_i S(x_i, x_i) + \sum_i S(y_i, y_i)},$$

where $S(x, y)$ is the from-insert-AA-to-subject-AA entry in a given substitution matrix, x_i is the AA at position i in the vaccine immunogen sequence, y_i is the AA at position i in the subject's founder sequence, the summations are over screened-in sites in the region of CSP under consideration (i.e., a particular epitope region or the C-terminus of the CSP protein).

The matrix $S(x, y)$ is chosen to yield a Hamming distance for the primary analysis, for which $\widetilde{V}$ is the proportion of AA sites that differ between an infected subject's sequence and the vaccine sequence. Based on the treatment-blinded descriptive data analysis, for any given founder sequence

all of the reads are identical, such that the Hamming distance for a founder is equivalent to a count variable that tallies the number of amino acid mismatches to the vaccine. We report analyses using this count scale. Future exploratory analyses not implemented from this SAP may use a matrix that accounts for the *AT* bias of malaria sequences, for example using a matrix estimated in Paila et al. (2008).

3.1 Failure Time Analysis Vaccine Efficacy Parameters for Malaria Endpoints 1. and 2.

The sieve analysis conducts estimation and testing of two measures of genotype-specific vaccine efficacy (VE) in the competing risks failure time analysis framework, defined in terms of:

- Genotype-specific cumulative incidence functions [cumulative VE], and
- Genotype-specific hazard functions [instantaneous VE].

With Z the indicator of assignment to the vaccine group, the two measures of vaccine efficacy are defined as follows.

- Discrete genotype-specific cumulative vaccine efficacy

$$VE^{\text{cum}}(t,j) = \left[1 - \frac{P(T \leq t, J = j | Z = 1)}{P(T \leq t, J = j | Z = 0)}\right] \times 100\%,$$

- Discrete genotype-specific instantaneous vaccine efficacy

$$VE^{\text{haz}}(t,j) = \left[1 - \frac{h(t,j|Z = 1)}{h(t,j|Z = 0)}\right] \times 100\%,$$

where j takes values 0 (full match to 3D7) and 1 (mismatch) for the primary analyses and takes counting number values up to 10 for the secondary analyses of genetic distances.

3.2 Binary Outcome Analysis Vaccine Efficacy Parameters for Malaria Endpoint 3. (Cross-Sectional Any Malaria Infection at the Month 20 Visit)

For assessing sieve effects on the any malaria infection endpoint, the binary outcome Y is the indicator of whether a subject has parasitemia at the Month 20 visit. The analysis is based on the set of subjects from whom malaria sequence data are available at the Month 20 visit. Complete-case methods will be used to analyze the data, which for validity rely on the assumption that the status of attending the Month 20 visit and of obtaining malaria genomic sequence data are independent of subject characteristics. If diagnostics show that these assumptions are clearly violated then missing data methods may also be applied that allow the status of missing data to depend on subject characteristics.

The target vaccine efficacy parameter is an analog of the cumulative incidence vaccine efficacy parameter defined above for the failure time analyses:

- Discrete genotype-specific cumulative vaccine efficacy

$$VE^{\text{bin}}(j) = \left[1 - \frac{P(Y=1, J=j|Z=1)}{P(Y=1, J=j|Z=0)}\right] \times 100\%,$$

where again j takes values 0 (full match to 3D7) and 1 (mismatch) for the primary analyses and takes counting number values up to 10 for the secondary analyses of genetic distances.

3.3 Post-Infection Sieve Analysis for Malaria Endpoints 1., 2., 3.

For the post-infection sieve analyses that condition on being a malaria case under any of the three endpoint definitions, the parameter of interest is the mean difference of the genomic feature between vaccine recipient cases vs. placebo recipient cases:

- Mean difference in the genomic feature

$$\text{MeanDiff} = E[J|Y=1, Z=1] - E[J|Y=1, Z=0].$$

For dichotomous J the mean difference parameter is the additive difference in the mismatch frequency for the vaccine vs. placebo group, and for ordered categorical J it is the additive difference in the mean of J for the vaccine vs. placebo group. The conditional post-infection analyses only assess a causal effect of assignment to vaccine vs. placebo if the analysis includes all baseline covariates predictive of both occurrence of the malaria endpoint under consideration and the malaria parasite genomic feature J (Shepherd and others, 2006). For this reason, all of the post-infection analyses adjust for baseline covariates in an effort to assess a causal effect of assignment to vaccine.

4 Multiplicity of Infection (MOI) and Multiple Outputation

Given that different founder infections are usually from separate transmission events from different mosquitos, we specify the mark J in the above vaccine efficacy parameters to represent the genotype/genetic distance of a randomly sampled founder infection of a malaria endpoint case. This gives the vaccine efficacy parameters interpretations in terms of differential vaccine efficacy against the malaria endpoint with a given founder virus. For estimation and testing of each of the vaccine efficacy parameters, the multiple outputation method of Follmann et al. (2003) is used. In particular, a data set is contructed randomly sampling a single founder infection from each malaria endpoint case, and a statistical method designed for a single mark per case is applied, which yields point estimates and variance estimates for parameters that are parts of the vaccine efficacy parameters of interest. This process is repeated M times, and the procedures of Follmann et al. (2003) applied to obtain the overall multiple outputation (MO) point estimate and 95% CI about the vaccine efficacy parameter of interest, and a 2-sided p-value for testing for a sieve effect. The 2-sided p-value is computed based on a Wald statistic and a normal approximation where the numerators and denominators of the Wald statistic are separate averages of point estimates and variance estimates, respectively, over the set of multiply outputated data sets.

The number M is chosen following formula (3) in Follmann et al. (2003), which ensures that the results do not depend on the randomness of the multiple outputation sampling; i.e., the results are very similar to what would be obtained by exhaustive multiple outputation covering all possible data sets taking one founder from each malaria endpoint case. For subjects with multiple founder infections, each founder infection is sampled with equal probability, regardless of the read counts supporting each founder infection. The reason for this is to give the vaccine efficacy parameters an interpretation on the per founder unit instead of on the per parasite unit that would be achieved by weighting the sampling of founder infections by read counts. The choice of the per founder interpretation is based on the fact that the vaccine operates by blocking malaria infection, but is thought to not impact the number of read counts that emerge after a founder infection.

5 Implementation of targeted maximum likelihood estimation (tMLE)

The targeted maximum likelihood estimation (tMLE) method of Benkeser and others (2014) is used for estimation and testing of the cumulative genotype-specific vaccine efficacy parameters for malaria endpoints 1 and 2. tMLE is a general method that optimizes unbiasedness, validity, and robustness, and includes model selection for covariates predictive of the malaria endpoint and for censoring (van der Laan and Rubin, 2006; Moore and van der Laan, 2009; Stitelman and van der Laan, 2010; Van der Laan and Rose, 2011). tMLE is also used for estimation and testing of the MeanDiff parameters for malaria endpoint 3, implemented with the R package tmle available at the Comprehensive R Archive Network (CRAN). Objective implementation of tMLE requires specification of the set of covariates to adjust for and the set of learners to use in the Super Learner cross-validation procedure (Van der Laan and Rose, 2011). The following covariates are adjusted for: Study site, distance from the nearest outpatient facility, distance from the nearest inpatient facility, pneumococcal vaccination status, the anthropometric measures height, weight, and head circumference, hemoglobin, calendar time of vaccination, the estimated cumulative rate of malaria in the placebo group of the study site of the subject, and the receipt of IPTi vaccinations for the 6-12 week old cohort.

Let Z denote the vaccine variable, S the study site, D both distances from nearest clinic, A anthropometric measures, C month calendar time, R the estimated cumulative probability of the primary malaria endpoint by 385 days in the placebo group at each site, and V the receipt of IPTi vaccinations for the 6-12 week old cohort. The following table shows the models used in the ensemble of learners. We list type of model and the adjustment variables used in each model, where *glm* denotes a binomial generalized linear model with logit link and *gam* denotes a generalized additive model (R package gam by Trevor Hastie available at CRAN). We will use the notation $f(Y)$ to denote that variable Y was fit using a series of indicator variables, $s(Y, x)$ to denote that variable Y was fit using a polynomial spline of degree x, and Y^*X to denote that an interaction term was fit between variables X and Y. The stepwise procedures listed will be performed using both Alkaike Information Criteria and Bayesian Information Criteria and will be performed using both forward selection and backwards elimination. For the library of models for the censoring mechanism, each model listed is fit with month of follow-up included as a series of dummy variables, a linear term, a log-linear term, and a polynomial spline of degree 3.

Model	Covariates
Malaria endpoints	
glm	Z
glm	$Z^* f(S)$
glm	$Z^* R$
gam	$Z, s(R,3)$
glm	$Z, f(S), R, A, D, f(C), V$
stepwise glm	$Z^* f(S), Z^* R, f(A), D, f(C), V$
Censoring	
glm	$\emptyset$
glm	Z
glm	D
gam	$s(D,3)$
stepwise glm	$f(S), R, f(A), D, f(C), V$

6 Summary of Planned Sieve Analyses

Table 1 summarizes the planned analyses for malaria endpoints 1. and 2., citing the sieve analysis methods that are used. Similarly, Table 2 summarizes the planned analyses and methods for malaria endpoint 3. In all cases the analysis entails estimation of the target genotype-specific vaccine efficacy parameter for each specified genotype of interest with a 95% confidence interval as well as the conduct of a test for the null hypothesis of no sieve effect. For the genotype-specific vaccine efficacy parameters that vary with time, the analysis will include plotting the estimates over the entire follow-up period (14 days post third immunization to 385 days later) and statistical tests for whether the genotype-specific vaccine efficacy varies over time. Most of the analyses include two analyses, one not adjusted for covariates and one adjusted for covariates, as noted by 'Unadj' and 'Adj' in the tables. Adjusted analyses control for the covariates listed above in Section 5. The calendar time of vaccination together with the study time of the clinical malaria endpoint provides some control for heterogeneity in malaria intensity across seasons. For covariate-adjusted analyses, covariates to adjust for are omitted if they prevent the complete data set from having at least 95% of subjects with the full set of covariates to adjust for. This justifies the analysis approach that restricts the analysis to the set of subjects with complete baseline covariate data.

The unadjusted analyses use a separate placebo risk for each site or otherwise restrict vaccine vs. placebo comparisons to be within individual sites. This means that the analyses that include the five major sites or that include all sites assess unconditional vaccine efficacy parameters (not site-specific) but that allow placebo malaria endpoint rates to differ across the sites.

Where there is evidence of a significant sieve effect, interaction/effect modification analyses may be conducted to assess if the sieve effect is restricted to or stronger in certain subgroups. These subgroups will include groupings of sites such as West versus East Africa (which differ in 3D7 haplotype prevalence) and malaria intensity, where malaria intensity may be defined for each site by the estimated incidence of the malaria endpoint under consideration in the placebo group. In addition, if in the analysis of overall vaccine efficacy there is evidence that the vaccine efficacy is

restricted to certain subgroups, then the sieve analysis may be repeated excluding the unprotected subgroups.

In addition to the vaccine efficacy sieve analyses that are conducted in the entire ATP cohort, the post-infection "case-conditional" analyses compare malaria parasite genomic features and MOI features between vaccine recipient cases and placebo recipient cases. The vaccine efficacy analyses allow the results to be presented in terms of genotype-specific vaccine efficacy that measure relative incidences of the malaria outcome based on prospective follow-up, and are parallel to the original primary analyses of vaccine efficacy. The case-conditional analyses condition on infection and consequently require stronger modeling assumptions to be interpreted in terms of prospective genotype-specific vaccine efficacy (Gilbert and others, 1998; Gilbert, 2000), but have advantages of clear graphical descriptives and potentially providing increased statistical power.

For the sieve analyses of malaria genotype variables for which at least 90% of malaria cases have data on the genotype of interest, complete-case analysis methods will be used that exclude the subjects with missing data from the analysis. This choice is justified by the fact that if the missing data rate is low then inferences from the complete-case method are expected to be very similar to those obtained via "missing data" versions of the sieve analysis methods that can provide valid inferences allowing the status of missing sequence data to depend on subject characteristics. Based on the descriptive analysis we expect the bar of 90% to be met for all of the genotype variables except for the number of NANP repeats. With moderate to high rates of missing data use of complete-case methods can have a substantial negative impact on the validity and precision of inferences (Little and others, 2012). Accordingly, for the number of NANP repeats the sieve analyses will be conducted both with complete-case methods and with missing data methods. The missing data methods use augmented inverse probability weighting and are valid under a missing at random assumption [Sun and Gilbert (2014) and Juraska and Gilbert (2014)]. These methods allow the probability that a subject is missing sequence data to depend on the subject's data, for example this occurs because malaria sequences are more likely to be missing in subjects with very low parasitemia (which is most relevant for the any malaria infection endpoint).

Tables 1 and 2 include a series of footnotes to very brief summarize how the specified methods would be used. The reader is referred to the publications for additional details, and the appendix provides a small amount of additional detail.

6.1 Binary Full 3D7 Match Sieve Analysis vs. Hamming Distance Sieve Analysis

The primary analyses assess sieve effects in terms of differential vaccine efficacy to prevent a malaria case with a fully vaccine-matched malaria parasite genome of a founder infection, versus a founder infection with some mismatch. The secondary analyses assess sieve effects in terms of differential vaccine efficacy to prevent a malaria case with the number of mismatches (equivalent to Hamming distance) of a founder infection. Each analysis is conducted separately for the full C-terminus CSP haplotype sequence or from an individual epitope region sequence, namely for TH2R, TH3R, Unnamed, the combined set of TH2R and TH3R amino acid positions identified with treatment-blinded linkage disequilibrium analysis, and the full SERA-2 protein.

6.2 Treatment Blinded Descriptive/Graphical Analysis of Malaria Genetic Features

Descriptive analyses of the genotypes are conducted based on treatment-blinded data, to finalize genotype definitions for the sieve analysis. The genotype variables are defined solely based on treatment-blinded data. The graphical descriptive analyses include barplots and boxplots of the distributions of the genotypes of infected subjects on the per founder virus level. The descriptive analysis will also explore information about the distribution of the multiplicity of infection. Graphics for this purpose will include histograms across subjects showing how often a pair of MO samples have different Hamming distances or 3D7 match/mismatch values, and summary statistics of the proportions of subjects in which MO creates variability.

Additional sieve analysis research beyond the scope of this SAP will include covariability/linkage disequilibrium analysis of the CSP protein, to assess whether and how to specify genotype variables for sieve analysis defined as grouped sets of covarying sites with putative structural or functional significance.

6.3 Graphical Display of Estimation Results

We summarize core graphical output of results that would be generated.

- Genotype-specific cumulative malaria endpoint rates over time by treatment group with numbers at risk and cumulative numbers of endpoints by treatment group displayed along the time axis (endpoints 1., 2.)
- Differences in genotype-specific cumulative malaria endpoint rates over time (placebo − vaccine) with 95% pointwise and simultaneous confidence bands (endpoints 1., 2.)
- Genotype-specific cumulative vaccine efficacy over time with 95% pointwise and simultaneous confidence bands (endpoints 1., 2.)
- Distance-specific cumulative malaria endpoint rates by 14 months *vs.* the distance by treatment group with 95% pointwise confidence bands (endpoints 1., 2.)
- Distance-specific cumulative vaccine efficacy by 14 months (endpoints 1., 2.) *vs.* the distance with 95% pointwise confidence bands
- Distance-specific vaccine efficacy at Month 20 (endpoint 3.) *vs.* the distance with 95% pointwise confidence bands
- Genotype-specific hazard rates of malaria endpoint over time by treatment group with 95% pointwise and simultaneous confidence bands (endpoints 1., 2.)
- Genotype-specific instantaneous vaccine efficacy over time with 95% pointwise and simultaneous confidence bands (endpoints 1., 2.)
- Distance-specific instantaneous vaccine efficacy over time *vs.* the distance with 95% pointwise confidence bands (endpoint 1., 2.)

- Distance-specific binary endpoint vaccine efficacy *vs.* the distance with 95% pointwise confidence bands (endpoint 3.)

6.4 Notation in Tables 1 and 2

In Tables 1 and 2 below,

- "full-match" indicates a malaria case founder infection with 100% of reads matching the 3D7 vaccine sequence whereas "else" indicates the complementary event of at least one mismatch in at least one read (it turns out that for every founder infection, all of the reads are identical);
- "# mismatches" denotes the number of amino acid positions mismatched to 3D7.

The tables divide the sieve analyses into four types: (Cum. VE) Based on cumulative vaccine efficacy through 385 days beyond 14 days after the third immunization including the whole ATP cohort; (Inst. VE) Based on instantaneous vaccine efficacy through 385 days beyond 14 days after the third immunization including the whole ATP cohort; (X-Sec. VE) Based on vaccine efficacy from the cross-sectional measurement of malaria endpoint 3 at Month 20 including all subjects with samples at Month 20; and (Conditional) Conditional on malaria endpoint status only including endpoint cases in the analysis. The different analysis types are included because they are geared to detect different types of sieve effects and provide a relatively thorough assessment of potential vaccine pressure on malaria.

The tables indicate that for the Cum. VE, Inst. VE, and X-Sect. VE analyses, the number of NANP repeats is also analyzed both as a binary mark (the same number of repeats as the vaccine sequence vs. else) and as the number of repeats (irrespective of the number in the vaccine sequence).

Table 1. Planned sieve analyses for malaria endpoints 1. and 2. (first or only clinical malaria)

Genotype Feature	Analysis Type	Target Parameter	Sieve Null Hypothesis	Method
(a) Full 3D7 Haplotype CSP	Cum. VE	$VE^{\text{cum}}(t, j)$ j = full-match, else	$H_0 : VE^{\text{cum}}(t = 14, j = fm)$ $= VE^{\text{cum}}(t = 14, j = else)$	NPMLE Wald test[1] (Unadj) tMLE Wald test[2] (Adj)
	Conditional	MeanDiff	$H_0 : \text{MeanDiff} = 0$	tMLE Wald test[3] (Adj)
	Inst. VE	Prob. else $VE^{\text{haz}}(j)$ j =full-match, else	$H_0 : VE^{\text{haz}}(j = fm)$ $= VE^{\text{haz}}(j = else)$	(Lunn and McNeil, 1995)/ (Gilbert, 2000) (Unadj)
(b) Full 3D7 Haplotype TH2R, TH3R, Unnamed, TH2R+TH3R Epitopes	Cum. VE	Same as (a)	Same as (a)	Same as (a)
		Same as (a)	Same as (a)	Same as (a)
	Conditional	Same as (a)	Same as (a)	Same as (a)
		Same as (a)	Same as (a)	Same as (a)
	Inst. VE	Same as (a)	Same as (a)	Same as (a)
		Same as (a)	Same as (a)	Same as (a)
(c) AA Site-Scanning CSP	Cum. VE	Same as (a)	Same as (a)	Same as (a)
		Same as (a)	Same as (a)	Same as (a)
	Conditional	Same as (a)	Same as (a)	Same as (a)
		Same as (a)	Same as (a)	Same as (a)
	Inst. VE	Same as (a)	Same as (a)	Same as (a)
		Same as (a)	Same as (a)	Same as (a)

[1]NPMLE of Aalen and Johansen (1978) with influence-curve based 95% confidence intervals and Wald-based tests for a sieve effect.

[2]tMLE of Benkeser and others (2014); tMLE is targeted maximum likelihood, a general method that optimizes unbiasedness, validity, and robustness, and includes model selection for covariates predictive of the malaria endpoint and for censoring (van der Laan and Rubin, 2006; Moore and van der Laan, 2009; Stitelman and van der Laan, 2010).

[3]tMLE for comparing the additive difference of frequency of amino acid mismatch to the vaccine sequence for vaccine recipient cases versus placebo recipient cases (van der Laan and Rubin, 2006), implemented with the R package tmle available at the Comprehensive R Archive Network (CRAN).

Table 1 Continued. Planned sieve analyses for malaria endpoints 1. and 2. (first or only clinical malaria)

Genotype Feature	Analysis Type	Target Parameter	Sieve Null Hypothesis	Method
(d) Whole C-terminus CSP and Epitope Distances	Cum. VE	$VE^{\text{cum}}(t,j)$ j =# mismatches	$H_0 : VE^{\text{cum}}(t=14,j)$ constant in j	NPMLE Wald test[1] (Unadj)
	Conditional	MeanDiff Mean # mismatches	H_0 : MeanDiff $= 0$	tMLE Wald test[4] (Adj)
	Inst. VE	$VE^{\text{haz}}(j)$ j =# mismatches	$H_0 : VE^{\text{haz}}(j)$ constant in j	(Juraska and Gilbert, 2013, 2014)/ (Sun and Gilbert, 2012; Gilbert and Sun, 2014)[5] (Unadj)
(e) Number of NANP repeats	Cum. VE	$VE^{\text{cum}}(t,j)$ j = same #, differ.	$H_0 : VE^{\text{cum}}(t=14, j=same)$ $= VE^{\text{cum}}(t=14, j=differ.)$	NPMLE Wald test[1] (Unadj) tMLE Wald test[2] (Adj)
	Conditional[6]	MeanDiff Mean # repeats	H_0 : MeanDiff $= 0$	tMLE Wald test[4] (Adj)
	Inst. VE	$VE^{\text{haz}}(j)$ j = same #, differ.	$H_0 : VE^{\text{haz}}(j=same)$ $= VE^{\text{haz}}(j=differ.)$	(Lunn and McNeil, 1995)/ (Gilbert, 2000) (Unadj)
(f) Number of Distinct Genomic Variants	Cum. VE	$VE^{\text{cum}}(t,j)$ j =1 vs. $j>1$	$H_0 : VE^{\text{cum}}(t=14,j=1)$ $= VE^{\text{cum}}(t=14,j>1)$	NPMLE Wald test[1] (Unadj) tMLE Wald test[2] (Adj)
	Conditional[6]	MeanDiff Mean # variants	H_0 : MeanDiff $= 0$	tMLE Wald test[4] (Adj)
	Inst. VE	$VE^{\text{haz}}(j)$ j =1 vs. $j>1$	$H_0 : VE^{\text{haz}}(j=1)$ $= VE^{\text{haz}}(j>1)$	(Lunn and McNeil, 1995)/ (Gilbert, 2000) (Unadj)
	Inst. VE			(Follmann, 2014)[7] (Unadj, Adj)

[4]tMLE for comparing the mean of the number of amino acid mismatches to the vaccine for vaccine recipient cases versus placebo recipient cases (van der Laan and Rubin, 2006), implemented with the R package tmle.

[5]While the Juraska and Gilbert (2013, 2014) and Sun and Gilbert (2012); Gilbert and Sun (2014) methods assess the same target VE parameter, they make different assumptions. The goodness-of-fit tests and graphical procedures of Juraska and Gilbert (2013) and Sun and others (ress), respectively, will be used to decide the most appropriate method. The 2013 and 2012 papers describe complete-case methods and the two 2014 papers describe missing data methods that use augmented inverse probability weighting.

[6]This method does not use the vaccine sequence as a reference sequence.

[7]Proportional hazards analysis for estimation and testing of vaccine efficacy defined as the multiplicative reduction (vaccine versus placebo) in the mean number of distinct genomes of transmitted founder malaria parasites.

Table 2. Planned sieve analyses for malaria endpoint 3. (any infection)

Genotype Feature	Analysis Type	Target Parameter	Sieve Null Hypothesis	Method
(a) Full 3D7	X-Sec. VE	$VE^{\text{bin}}(j)$	$H_0 : VE^{\text{bin}}(j = fm)$	NPMLE Wald test[1] (Unadj)
Haplotype		j = full-match, else	$= VE^{\text{bin}}(j = else)$	tMLE Wald test[2] (Adj)
CSP	Conditional	MeanDiff	H_0: MeanDiff=0	tMLE Wald test[3] (Adj)
		Prob. else		
	Conditional	MeanDiff	H_0: MeanDiff = 0	tMLE Wald test[3] (Adj)
		Mean # mismatches		
(b) Full 3D7	X-Sec. VE	Same as (a)	Same as (a)	Same as (a)
Haplotype		Same as (a)	Same as (a)	Same as (a)
TH2R, TH3R,	Conditional	Same as (a)	Same as (a)	Same as (a)
Unnamed, TH2R+TH3R		Same as (a)	Same as (a)	Same as (a)
Epitopes	Conditional	Same as (a)	Same as (a)	Same as (a)
		Same as (a)	Same as (a)	Same as (a)
(c) AA Site-	X-Sec. VE	Same as (a)	Same as (a)	Same as (a)
Scanning		Same as (a)	Same as (a)	Same as (a)
CSP	Conditional	Same as (a)	Same as (a)	Same as (a)
		Same as (a)	Same as (a)	Same as (a)

[1]NPMLE of Aalen and Johansen (1978) with influence-curve based 95% confidence intervals and Wald-based tests for a sieve effect.

[2]tMLE of Benkeser and others (2014); tMLE is targeted maximum likelihood, a general method that optimizes unbiasedness, validity, and robustness, and includes model selection for covariates predictive of the malaria endpoint and for censoring (van der Laan and Rubin, 2006; Moore and van der Laan, 2009; Stitelman and van der Laan, 2010).

[3]tMLE for comparing the additive difference of frequency of amino acid mismatch to the vaccine sequence for vaccine recipient cases versus placebo recipient cases (van der Laan and Rubin, 2006), implemented with the R package tmle available at the Comprehensive R Archive Network (CRAN).

Table 2 Continued. Planned sieve analyses for malaria endpoint 3. (any infection)

Genotype Feature	Analysis Type	Target Parameter	Sieve Null Hypothesis	Method
(d) Whole C-terminus CSP and Epitope Distances	X-Sec. VE	$VE^{\text{bin}}(j)$ $j =$# mismatches	$H_0 : VE^{\text{bin}}(j)$ constant in j	NPMLE Wald test[1] (Unadj)
	Conditional	MeanDiff Mean # mismatches	$H_0 :$ MeanDiff $= 0$	tMLE Wald test[4] (Adj)
(e) Number of NANP repeats	X-Sec. VE	$VE^{\text{bin}}(j)$ $j =$ same #, differ.	$H_0 : VE^{\text{bin}}(j = same)$ $= VE^{\text{bin}}(j = differ.)$	NPMLE Wald test[1] (Unadj) tMLE Wald test[2] (Adj)
	X-Sec. VE	$VE^{\text{bin}}(j)$ $j =$ # addnl repeats	$H_0 : VE^{\text{bin}}(j)$ constant in j	NPMLE Wald test[1] (Unadj)
	Conditional[6]	MeanDiff Mean # repeats	$H_0 :$ MeanDiff $= 0$	tMLE Wald test[3] (Adj)
(f) Number of Distinct Genomic Variants	X-Sec. VE	$VE^{\text{bin}}(j)$ $j =$ same #, differ.	$H_0 : VE^{\text{bin}}(j = same)$ $= VE^{\text{bin}}(j = differ.)$	NPMLE Wald test[1] (Unadj) tMLE Wald test[2] (Adj)
	Conditional[6]	MeanDiff Mean # variants	$H_0 :$ MeanDiff $= 0$	tMLE Wald test[4] (Adj)
	X-Sec. VE			(Follmann, 2014)[7] (Unadj, Adj)

[4]tMLE for comparing the mean of the number of amino acid mismatches to the vaccine for vaccine recipient cases versus placebo recipient cases (van der Laan and Rubin, 2006), implemented with the R package tmle.

[5]While the Juraska and Gilbert (2013, 2014) and Sun and Gilbert (2012); Gilbert and Sun (2014) methods assess the same target VE parameter, they make different assumptions. The goodness-of-fit tests and graphical procedures of Juraska and Gilbert (2013) and Sun and others (ress), respectively, will be used to decide the most appropriate method. The 2013 and 2012 papers describe complete-case methods and the two 2014 papers describe missing data methods that use augmented inverse probability weighting.

[6]This method does not use the vaccine sequence as a reference sequence.

[7]Proportional hazards analysis for estimation and testing of vaccine efficacy defined as the multiplicative reduction (vaccine versus placebo) in the mean number of distinct genomes of transmitted founder malaria parasites.

The tables specify analyses that separately address malaria genomics and malaria MOI. The method of Juraska and Gilbert (2013) may be applied for estimation and testing of the bivariate vaccine efficacy parameter $VE^{\text{haz}}(j_1, j_2)$, where j_1 is the genomic information (indicator of full match to 3D7 or the genetic distance) and j_2 is the MOI.

References

Aalen, OO and Johansen, S. (1978). An empirical transition matrix for non-homogeneous markov chains based on censored observations. Scandinavian Journal of Statistics **5**, 141–150.

Benkeser, D, Gilbert, PB and Carone, M. (2014). Efficient and robust estimation of cumulative incidence in competing risks settings. Technical report, Fred Hutchinson Cancer Research Center.

Follmann, D. (2014). Incorporating founder virus information in vaccine field trials. Biometrics **in press**.

Gilbert, PB. (2000). Comparison of competing risks failure time methods and time-independent methods for assessing strain variations in vaccine protection. Statistics in Medicine **19**, 3065–3086.

Gilbert, PB, Self, S and Ashby, M. (1998). Statistical methods for assessing differential vaccine protection against human immunodeficiency virus types. Biometrics **54**, 799–814.

Gilbert, PB and Sun, Y. (2014). Testing for vaccine efficacy against a spectrum of pathogen sequences in stratified mark-specific proportional hazards models with missing marks, with application to the RV144 HIV vaccine efficacy trial. Journal of the Royal Statistical Society Series C, in press.

Grambsch, P and Therneau, T. (1994). Proportional hazards tests and diagnostics based on weighted residuals. Biometrika **81**, 515–526.

Juraska, M and Gilbert, PB. (2013). Mark-specific hazard ratio model with multivariate continuous marks: An application to vaccine efficacy. Biometrics **69**, 328–337.

Juraska, M and Gilbert, PB. (2014). Mark-specific hazard ratio model with missing multivariate marks. under review.

Kalbfleisch, JD and Prentice, RL. (2002). The Statistical Analysis of Failure Time Data, Second Edition.. New York: John Wiley and Sons.

Little, Roderick J, D'Agostino, Ralph, Cohen, Michael L, Dickersin, Kay, Emerson, Scott S, Farrar, John T, Frangakis, Constantine, Hogan, Joseph W, Molenberghs, Geert, Murphy, Susan A and others. (2012). The prevention and treatment of missing data in clinical trials. New England Journal of Medicine **367**(14), 1355–1360.

Lunn, M and McNeil, D. (1995). Applying Cox regression to competing risks. Biometrics **51**, 524–532.

Moore, Kelly L and van der Laan, Mark J. (2009). Application of time-to-event methods in the assessment of safety in clinical trials. U.C. Berkeley Division of Biostatistics Working Paper Series **Working Paper 248**.

Paila, Umadevi, Kondam, Rohini and Ranjan, Akash. (2008). Genome bias influences amino acid choices: analysis of amino acid substitution and re-compilation of substitution matrices exclusive to an at-biased genome. Nucleic acids research **36**(21), 6664–6675.

Shepherd, B, Gilbert, PB, Jemiai, Y and Rotnitzky, A. (2006). Sensitivity analyses comparing outcomes only existing in a subset selected post-randomization, conditional on covariates, with application to HIV vaccine trials. Biometrics **62**, 332–342.

Stitelman, Ori M and van der Laan, Mark J. (2010). Collaborative targeted maximum likelihood for time to event data. The International Journal of Biostatistics **6, Issue 1**, Article 21.

Sun, Y and Gilbert, PB. (2012). Estimation of stratified mark-specific proportional hazards models with missing marks. Scandinavian Journal of Statistics **39**, 34–52. PMCID: PMC3601495.

Sun, Y, Li, M and Gilbert, PB. (in press). Goodness-of-fit test of the stratified mark-specific proportional hazards model with continuous mark. Computational Statistics and Data Analysis.

Van der Laan, Mark J and Rose, Sherri. (2011). Targeted learning: causal inference for observational and experimental data. Springer.

van der Laan, Mark J and Rubin, Daniel. (2006). Targeted maximum likelihood learning. U.C. Berkeley Division of Biostatistics Working Paper Series **Working Paper 213**.

7 Appendix: Sieve Analysis Methods

The primary sieve analysis assesses the cumulative vaccine efficacy parameters for time-point $t = \tau$ fixed at the end of follow-up for the primary case definition (Month 14). The cumulative parameters are prioritized to the instantaneous parameters because prevention of malaria until the end of follow-up is maximally clinically relevant and the core approach to analyzing the instantaneous parameters make proportional hazards assumptions that may be violated for the malaria trial given the evidence for waning vaccine efficacy wanes over time. To interpret the analyses of the instantaneous parameters, the Grambsch and Therneau (1994) test and graphical procedures may be used to assess the cause-specific proportional hazards assumption for the case of discrete genotypes. For ordered categorical or approximately continuous distance genotypes, the (Sun and others, ress) tests and graphical procedures will be used to assess the distance-specific proportional hazards assumption.

7.1 Models and Estimation Procedures Without Covariate Adjustment

For the analyses that do not adjust for covariates, the cumulative incidence functions $F(t, j|Z = z) := P(T \leq t, J = j|Z = z)$, $z = 0, 1$, will be estimated by the nonparametric maximum likelihood estimator (NPMLE)(Aalen and Johansen, 1978), implemented in the R `cmprsk` package, based on the Kaplan-Meier estimate of the overall survival function and the Nelson-Aalen estimates of the genotype-specific cumulative hazard functions (see, e.g., Kalbfleisch and Prentice, 2002). The adjusted analyses that estimate the same parameter use targeted maximum likelihood estimation (unpublished method of David Benkeser, Peter Gilbert, and Marco Carone).

To estimate $VE^{\text{haz}}(t, j)$ for each level j of a discrete genotype, we posit the competing risks Cox model

$$h(t, j|Z = z) = h_0(t, j) \exp\{\widetilde{\alpha}_j z\}.$$

The genotype-specific log hazard ratios $\widetilde{\alpha}_j$ are estimated by the maximum partial likelihood estimation method. The parameter $VE^{\text{haz}}(t, j)$ is estimated by the method of Juraska and Gilbert (2014b) if the assumptions of the method, tested using procedures described in the paper, hold; if the method's assumptions are violated, then we will use the alternative method of Sun and Gilbert (2012); Gilbert and Sun (2014) for estimation and inference about the same parameter.

7.2 Hypothesis Testing

Statistical tests for a "sieve effect" assess the following null hypothesis:

H_0: Constant vaccine efficacy against all exposing genotypes.

We use the following procedure to test this null hypothesis in terms of the cumulative vaccine efficacy parameter $VE^{\text{cum}}(\tau, j)$, where τ is 385 days after 14 days post-third immunization. Here j denotes the binary mark 3D7 match vs. 3D7 mismatch, or the Hamming distance represented as

a count variable (number of mismatches to the vaccine). The following procedure can be used for unadjusted or covariate-adjusted analysis, and is performed for each multiply outputated data set.

Denote $\theta_{zj} = F(t_0, j|Z = z)$, where $z = 0, 1$, $j = 1, \ldots, J$, $J \geq 2$, is an ordered categorical genotype, and t_0 is fixed. Let $\boldsymbol{\theta} = (\theta_{01}, \theta_{11}, \ldots, \theta_{0J}, \theta_{1J})^T$ and $\widehat{\boldsymbol{\theta}}$ be the Aalen-Johansen estimator for $\boldsymbol{\theta}$. Then

$$\sqrt{n}\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathsf{N}_{2J}(\mathbf{0}, \boldsymbol{V}).$$

Consider the influence curve-based variance estimator $\widehat{\boldsymbol{V}}$ for $\boldsymbol{V}$.

Next consider the transformation $\boldsymbol{g}(\boldsymbol{\theta}) = \left(\log\frac{\theta_{01}}{\theta_{11}}, \ldots, \log\frac{\theta_{0J}}{\theta_{1J}}\right)^T$ and denote

$$\dot{\boldsymbol{g}}(\boldsymbol{\theta}) = \left(\frac{\partial g_j(\boldsymbol{\theta})}{\partial \theta_i}\right)_{\substack{i=1,\ldots,2J \\ j=1,\ldots,J}}.$$

The delta method yields

$$\sqrt{n}\left(\boldsymbol{g}(\widehat{\boldsymbol{\theta}}) - \boldsymbol{g}(\boldsymbol{\theta})\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathsf{N}_J(\mathbf{0}, \boldsymbol{\Sigma}),$$

where $\boldsymbol{\Sigma} = \dot{\boldsymbol{g}}(\boldsymbol{\theta})^T \boldsymbol{V} \dot{\boldsymbol{g}}(\boldsymbol{\theta})$. Consider the estimator $\widehat{\boldsymbol{\Sigma}} = \dot{\boldsymbol{g}}(\widehat{\boldsymbol{\theta}})^T \widehat{\boldsymbol{V}} \dot{\boldsymbol{g}}(\widehat{\boldsymbol{\theta}})$ for $\boldsymbol{\Sigma}$.

It is of interest to test the null hypothesis

$$H_0 : VE(t_0, j) = VE(t_0) \text{ for all } j = 1, \ldots, J.$$

To this end, consider approximating $\log\frac{\theta_{0j}}{\theta_{1j}}$ as a function of j with a straight line fit to the points $\log\frac{\widehat{\theta}_{01}}{\widehat{\theta}_{11}}, \ldots, \log\frac{\widehat{\theta}_{0J}}{\widehat{\theta}_{1J}}$:

$$\log\frac{\widehat{\theta}_{0j}}{\widehat{\theta}_{1j}} \approx \beta_0 + \beta_1(j-1).$$

Then H_0 is equivalent to $H_0^* : \beta_1 = 0$. Denote $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$ and

$$\boldsymbol{X} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & J-1 \end{pmatrix}.$$

Using the generalized least squares method, we estimate $\boldsymbol{\beta}$ with $\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^T \widehat{\boldsymbol{\Sigma}}^{-1} \boldsymbol{X})^{-1} \boldsymbol{X}^T \widehat{\boldsymbol{\Sigma}}^{-1} \boldsymbol{g}(\widehat{\boldsymbol{\theta}})$. If $\widehat{\boldsymbol{\Sigma}}$ is the true asymptotic covariance matrix of $\boldsymbol{g}(\widehat{\boldsymbol{\theta}})$, then

$$\sqrt{n}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathsf{N}_2(\mathbf{0}, \boldsymbol{Q}),$$

where $\boldsymbol{Q} = (q_{ij})_{i,j=1,2} = (\boldsymbol{X}^T \widehat{\boldsymbol{\Sigma}}^{-1} \boldsymbol{X})^{-1}$. Finally, the Wald test of H_0^* is based on the test statistic $\widehat{\beta}_1 / \sqrt{q_{22}/n}$ that is approximately $\mathsf{N}(0, 1)$ under validity of H_0^*.

For testing for a sieve effect in terms of hazard ratios, i.e. for testing $H_0 : VE^{\text{haz}}(j)$ is constant in j, the Lunn and McNeil (1995) test will be used for each multiply outputated data set.

7.3 Local Sieve Analysis Methods

Each amino acid site will be tested with four site-scanning sieve analysis methods to identify those that discriminate the vaccine and placebo groups. A q-value multiplicity adjustment procedure will also be applied for each method separately to limit the false discovery rate to 20%.

The four methods are:

1. Nonparametric maximum likelihood estimation (NPMLE) Wald test of genotype-specific cumulative vaccine efficacy through 14 months that does not incorporate subject covariates (**Cum. VE NPMLE Wald test**)
2. Targeted maximum likelihood estimation (tMLE) Wald test of genotype-specific cumulative vaccine efficacy through 14 months that incorporates subject covariates (**Cum. VE tMLE Wald test**)
3. Proportional hazards model Wald test (Lunn and McNeil, 1995) analsis of genotype-specific instantaneous vaccine efficacy that does not incorporate subject covariates (**PH Wald test**)
4. Targeted maximum likelihood estimation (tMLE) Wald test comparing genotypes between vaccine recipient cases versus placebo recipient cases (**Cond. tMLE Wald test**)

The planned sieve analysis methods are implemented in R and are publicly available, several in R packages and others as R code on Peter Gilbert's web-site.

8 Statistical Implementation Notes

1. (Table 1 NPMLE Wald test (Unadj)): Implement the Aalen-Johansen estimator and tests stratifying by site, for the analyses that pool the 5 sites or that pool all 11 sites. This is implemented by a version of tMLE that only includes site as a covariate. To avoid problems with sparsity, for the analyses that pool together data from the 11 sites, the 6 non-major sites are pooled together and treated as a single unit. The stratification is done to accommodate the heterogeneity in background malaria risk across sites.
2. (Table 1 tMLE Wald test (Adj)): For estimating $VE^{cum}(t, j)$ over time t, only use the unadjusted AJ method as described in 1. above. The full tMLE is only used for the final time point of $t = 14$ months. For implementing full tMLE, use a relatively simple ensemble of learners, to help with computational speed given the multiple outputations.
3. (Table 1, Conditional MeanDiff tMLE Wald test): The mean difference is a difference in probabilities, so use a logit link and binary endpoint glm in the tMLE package.
4. (Table 1 Lunn and McNeil (1995)/Gilbert (2000) (Unadj)): Implement the method with a separate baseline hazard for each site, for the analyses that pool the 5 sites or that pool all 11 sites. This should be straightforward given that coxph() is called. If this turns out to be problematic to implement, then the stratification could be dropped.

5. (Table 1, continued distance analyses (d)): Most of these analyses use a count mark, which may be postponed until after the binary mark analyses are done.

6. (Table 1, continued distance analyses (d)): CST3 has no mark count variability, so it is not analyzed. The CSP, TH2R, TH3R, and Unnamed distance analyses have counts ranging from 0–10, 0–6, 0–4, and 0–2, respectively. The analyses are also conducted for a combined set of TH2R and TH3R amino acid positions identified with treatment-blinded linkage disequilibrium analysis, with range xx–xx (pending from Dan). The distance analysis is also done for the full SERA-2 protein.

7. (Table 1, continued NPMLE Wald test (Unadj)): Use an AJ estimator without any covariate adjustment/stratification for the site-pooled analyses. The reason is because we expect unstable estimation for some of the multiply outputated data sets given sparse mark counts at some levels.

8. (Table 1, continued Juraska/Gilbert/Sun SmoothMarks analysis): Try the Juraska and Gilbert (2013) method first; if the goodness of fit diagnostics are satisfactory, stop there. If they show violations of assumptions, implement the Sun and Gilbert (2012)/Gilbert and Sun (2014) method.

9. (Table 1, continued Juraska/Gilbert/Sun SmoothMarks analysis): If the Sun and Gilbert (2012)/ Gilbert and Sun (2014) method is implemented, stratify by site for the site-pooled analyses.

10. (Table 1, Conditional MeanDiff tMLE Wald test): The mean difference is a difference in means of a quantitative variable, and the analysis uses an identity link and quantitative endpoint glm in the tMLE package.

11. (Table 1, X-Sec. VE NPMLE Wald test (Unadj)): Use the same analysis as the Cum. VE analysis (Unadj) in Table 1(a).

12. (Table 1, X-Sec. VE tMLE Wald test (Adj)): Use the same analysis as the Cum. VE analysis (Adj) in Table 1(a).

13. (Table 1, Continued, (d) X-Sec. VE analysis (Unadj)): Use the same analysis as the Cum. VE analysis (d) (Unadj) in Table 1, Continued (a).

14. (Table 1, Continued, (e) X-Sec. VE analysis (Unadj)): Use the same analysis as the Cum. VE analysis (e) (Unadj) in Table 1, Continued (a).

15. For all full tMLE analyses, include as a baseline covariate the "intensity score" defined as the cumulative malaria endpoint rate in the placebo group, scaled to range between 0 and 1 and potentially effect modifier parameters in the Super Learner by this intensity score.

16. For the primary analyses of the binary mark 3D7 match vs. 3D7 mismatch within each of the 5 major sites, show the estimated cumulative incidence curves over time $t \in (0, 12]$ months after 14 days post-third immunization for each mark and for each of the vaccine and placebo groups, with 95% pointwise CIs, and show the estimated cumulative and additive vaccine efficacy over time for each mark (the additive VE is defined as $F(t_0, j|Z = 0) - F(t_0, j|Z = 1)$, $j = 1, 2$). For the analyses that pool the 5 sites or that pool all 11 sites, the above estimated cumulative

incidence and VE curves in continuous time will be replaced with discrete inference, stratified by site, at time points 1,...,12 months after 14 days post-third immunization. The reason for assessing the parameters of interest over time is so that we can directly check if the sieve effect changes with time. In addition, produce the ordinary time-varying VE plots for each of the 2 marks. This comment applies for endpoints 1. and 2.

17. For the "Smooth Marks' Hamming distance analyses, always use plotting of the estimated $VE^{cum}(t,j)$ with 95% confidence intervals, for both multiplicative and additive VE, as a core analysis (using the AJ and tMLE approaches). If the Hamming distance has 4 or fewer categories, this is the only estimation analysis; if 5 or more categories, then "smooth marks" analysis ala Juraska/Gilbert/Sun is warranted.

18. For the primary analysis, implementing the MO procedure includes estimating $F_1(t,v) \equiv P(T \leq t, V = v|Z = 1)$ and $F_2(t,v) \equiv P(T \leq t, V = v|Z = 0)$ for each of $v = 0$ and $v = 1$ and the estimation of the variances of these estimates, for each of the multiply outputed data sets. If the MO data set for estimating $F_z(t,v)$ has zero malaria endpoint events of type v, then the variance estimate is not defined. To handle this issue, the overall MO estimate of $\text{Var}(\widehat{F}_z(t,v))$ for each z and v averages over the subset of multiply outputed data sets for which the variance estimate is defined. In contrast the overall MO estimate of $F_z(t,v)$ averages the point estimates over all M multiply outputed data sets. In the presence of zero malaria endpoint events of type v, the same rules apply to estimation of the hazard ratios $\lambda(t,v|Z=1)/\lambda(t,v|Z=0)$, $v = 0, 1$, and the respective variances estimates.

19. Following from the last bullet, if the proportion of M outputed data-sets with zero malaria endpoint events of type v in at least one of the vaccine and placebo groups is $\leq 5\%$, the hypothesis test described in Section 7.2 will be used for evaluating the cumulative VE sieve effect and the Lunn and McNeil (1995) test will be used for evaluating the hazard-based VE sieve effect (the site-stratified versions of the tests will be used for analyses that pool the 5 major sites or that pool all 11 sites). If the proportion is $> 5\%$, for both the cumulative and hazard-based VE approach, sieve effect testing will be based on calculating the Fisher's exact test p-value for the 2-by-2 table in each of the M outputed data-sets, and then the p-values across MO runs will be combined using the procedure described in Follmann (2003, Section 4.1).