- 1 Preferential suppression of *Anopheles gambiae* host sequences allows detection
- 2 of the mosquito eukaryotic microbiome
- 3
- 4 Eugeni Belda^{1,2}, Boubacar Coulibaly³, Abdrahamane Fofana³,
- 5 Abdoul H. Beavogui⁴, Sekou F. Traore³, Daryl M. Gohl⁵, Kenneth D. Vernick^{1,2,*},
- 6 Michelle M. Riehle⁶
- 7
- 8 ¹Department of Parasites and Insect Vectors, Unit of Genetics and Genomics of
- 9 Insect Vectors, Institut Pasteur, Paris, France.
- 10 ²CNRS Unit of Hosts, Vectors and Pathogens (URA3012), Paris, France
- ³Malaria Research and Training Centre (MRTC), Faculty of Medicine and
- 12 Dentistry, University of Mali, Bamako, Mali.
- 13 ⁴Centre de Formation et de Recherche en Santé Rurale de Mafèrinyah, Conakry,
- 14 Republic of Guinea.
- 15 ⁵University of Minnesota Genomics Center, Minneapolis, Minnesota, USA
- 16 ⁶Department of Microbiology and Immunology, University of Minnesota,
- 17 Minneapolis, MN, USA.
- 18
- 19
- 20 * Correspondence:
- 21 email: kvernick@pasteur.fr
- 22 address: Institut Pasteur, GGIV, 28 rue du Dr Roux, 75724 Paris Cedex 15,
- 23 FRANCE

894 Supplementary Figure Legends

895

Figure S1: Rarefaction plots based on phylogenetic diversity. Samples were
grouped according to their amplification with and without PNA (for PNA
samples, data from all 4 PNA concentrations are pooled in this analysis).

899 Phylogenetic diversity was computed from rarified OTU tables at sampling

- 900 depths from 10 to 1000 sequences. At each sampling depth, the average and
- 901 standard error of 10 rarified OTU tables is represented.
- 902

903 Figure S2: Taxonomic profile of larval sample from 18s rRNA gene amplicons of 904 V4 and V9 hypervariable regions amplified at different concentrations of PNA 905 blocker after exclusion of mosquito and mammalian sequences. Chi-square tests 906 were carried out to test whether the presence or concentration of PNA blocker 907 affects the relative fractions of the eukaryotic lineages detected. There was no 908 significant difference in the composition of eukaryotic microbiota (chi-square 909 p=0.92 for V4 and p=1 for V9) across all samples, with and without PNA blocker, 910 showing that the presence of PNA blockers does not influence differential 911 amplification of any particular eukaryotic microbe, and instead functions solely 912 to block the amplification of the mosquito host template. Note that this Figure, as 913 compared to Figures 3 and 4, provides the full taxonomic list of the larval sample 914 only, with the Other category uncollapsed to show all lineages. The larval sample 915 yielded most of the microbial diversity, and for simplicity the other two samples 916 represented in Figures 3 and 4, Animal Bloodmeal and Human Bloodmeal, are 917 not included here, so some classes in Figures 3 and 4 are not represented in 918 Figure S4.

919

Figure S3: Multiple sequence alignment of representative 18s rRNA gene
sequences from mosquito and mammal (Silva 119 database) with *Mammalia* and *Anopheles* anneal-inhibiting blocking primers targeting the V4 hypervariable
region used in the present study. The alignment region corresponds to the 5' end
of the V4 hypervariable region.

925

31

- 926 **Figure S4**: Multiple sequence alignment of representative 18s rRNA gene
- 927 sequences from mosquito and mammal (Silva 119 database) with *Mammalia* and
- 928 *Anopheles* anneal-inhibiting blocking primers targeting the V9 hypervariable
- region used in the present study. The alignment region corresponds to the 3' end
- 930 of the V9 hypervariable region.
- 931
- 932 **Figure S5:** Multiple sequence alignment of the 18s rRNA gene V9 hypervariable
- 933 region of wild *A. coluzzii* (Acol) and *A. gambiae* (Agam) from Burkina Faso.
- Alignments represent the reversed and complemented genome region between
- 935 nucleotide coordinates X:23669218-23669319 of the Vectorbase AgamP4
- annotation. Sequence of the V9 PNA blocker (as shown in Table S1) is
- highlighted in purple.
- 938
- 939 **Figure S6:** Multiple sequence alignments as in Figure S5, but for the V4
- 940 hypervariable region. Alignments represent the genome region between
- 941 nucleotide coordinates X:24339450-24339510 of the Vectorbase AgamP4
- 942 annotation. Sequence of the V4 PNA blocker (as shown in Table S1) is
- 943 highlighted in purple.











		10		20	30		40	50	60) .	70	80	
Dryctolagus_cuniculus_rabbit/349-431	GAGGT	GAAAT	тст	GGACCGGCGC	AAGACGG-A	Ċ	CAGAGCGA	AAG-CA-T	TTGCC-	AAGAATO	TTTTCAT	TAATCAAGA	AC <mark>G</mark> AAA <mark>GT</mark>
lomo_sapiens_human/340-421	GAGGT	GAAAT	тст	TGGACCGGCGC	AAGACGG-A	C	CAGAGCGA	AAG-CA-T	TTGCC-	AAGATTO	TTT-CAT	TATTCAAGA	CGAAAAGT
/icugna_pacos_alpaca/352-434	GAGGT	GAAA	тст	TGGACCGGCGC/	AAGACGG-A	C	CAGAGCGA	AAG-CA-T	TTGCC-	ΑΑ <mark>G</mark> ΑΑ <mark>Τ</mark>	TTTTCAT	ΤΑΑ <mark>Τ</mark> C ΑΑ G Α.	ACGAAAGT
0tolemur_garnettii_small-eared_galago/336-416	GAGGT	- <u>A</u> A A A <mark>T</mark>	тст	TCA <mark>ACCAGTG</mark> C/	AAGACGA	C	TATAGCAA	ACCA-T	TTCCC-	AAGAATO	TTTTCGT	TAAACAGGA	ΑΤΓΑΑΑΓ
lomo_sapiens_human/343-426	GAGGT	GAAAT	тст	TGGACCGGCGC	A A <mark>G A C G G - A</mark>	C	C A <mark>G A G C G</mark> A	AAG-CA-T	TTGCCC.	A A <mark>G</mark> A A <mark>T</mark> (TTTTCAT	TAA <mark>TC</mark> AA <mark>G</mark> A	A C <mark>G</mark> A A A <mark>G T</mark>
lomo_sapiens_human/354-437	AATTC	GAAAT	тст	TGGACCGGCGC/	AAGACGG-A	C	CAGAGCGA	AA <mark>G-C</mark> A-T	TTGCC-	A A <mark>G</mark> A A <mark>T</mark> (TTTTCAT	TAA <mark>TC</mark> AA <mark>G</mark> A	AC <mark>G</mark> AAA <mark>GT</mark>
ctidomys_tridecemlineatus_thirteen-lined_ground_squirrel/345-4	42 <mark>G A G G T</mark> -	GAAAT	тст	TGGACCGGCGC	AAGACGG-A	C	CAGAGCGA	AA <mark>G - C</mark> A - T	TTGCC-	AAGAATO	TTTTCAT	TAATCAAGA.	ACGAAAGT
<i>lomo_sapiens_human/341-423</i>	GAGGT	GAAAT	тст	TGGACCGGCGC	AGTTTGG-A	C	CAGAGTGA	AAG-CA-T	TTGCC-		TTTTCAT	TAATCAAGA	A T G A A A G T
lomo_sapiens_human/330-410	GAGGT	GAAAT	тст	GGACTGGCAG	TTAG-A	C	CAGAGCGA	AAG-CA-T	TTGCC-	AAGAATO	TTTTCAT	TAATCAAGA	ACGAAAGT
an_troglodytes_chimpanzee/351-433	GAGGT	GAAAT	тст	IGGACCGGCGC/	AGTTCGG-A	C	CAGAGTGA	AAG-CA-T	TTGCC-	AAGAATO	TTTTCAT	TAATCAAGA.	ACGAAAGT
ongo_abelii_Sumatran_orangutan/330-412	GAGGI	GAAAI	TCT	GGACCGGCGC	AGTICGG-A	C	CAGAGCGA	AAG-CA-I	TTGCC-	AAGAATO	TTTTCAT	TAACCAAGA.	ATGAAAGT
upaia_belangeri_northern_tree_shrew/343-425	GAGG	GAAAI		GGACCGGCGC	AAGACGG-A	C	CAGAGCGA	AAG-CA-I	TIGCC-	AAGAAIC	TTTTCAL	TAATCAAGA.	ACGAAAGI
lomo_sapiens_human/330-412	GAGG	GAAA	I CI	LCCACCGGAGC/	ATTICGG-A	C	CAGAGIGA	AAG-CA-I	TIGCC-	AAGAAIC	TTTTCAL	AACCAAGA.	ACGAAAG
Ayotis_lucifugus_little_brown_bat/338-420	CACCT		T CT			C	LAGAGCAA	AGG-CA-I	TTCCC-		TTTTCAT		
an_troglodytes_chimpanzee/324-406	CACCT	GAAAI	T CT			C	CACAGCIA	AAG-CA-I	TTCCC-		TTTTCAT		ACCAAAGT
rinaceus_europaeus_western_european_neugenog/339-421	CACCT				AACACCC A	C	CACACCCA		TTCCC		TTTTCAT		ACCAAACT
Alcrocebus_murinus_gray_mouse_lemur/348-430	CACCT				AAGACGG-A	C	CACACCCA		TTCCC		TTTTCAT		ACCAAAGT
nyotis_iuciiugus_iittie_brown_bat/340-422	CACCT		+ 2+		ACTTCCC A	C	CACACCCA	AAG CA-T	TTCCC		ATTTCCT		ATCAAAGT
Dipodomus ordii Ordis kongoroo rat/242_425	CACCT	CAAAT	+ 6+		ACACCC-A	C	CACACCCA		TTCCC		TTTTCAT		ACCAAACT
lomo saniens human/330-412	CACCT	CAAAT	TCT.		ACTTCCC-A	C	CAGAGCGA		TTCCC		TTTTCAT		ACCAAACT
Chotona princeps American pika/344-426	CACCT		TCT.		ACACCC-A	c	CAGAGCGA		TTCCC-		TTTTCAT		ACCAAACT
lomo saniens human/324_406	CACCT	CAAAT	TCT.		ACTTCCC-A	c	CAGAGCTA		TTCCC-		TTTTCAT		ACCAAACT
processia canensis cane rock hyrax/310=396	CACCC		TAT		ACACCCCC	CACA	CAGAGCGA	AAG-CA-T	TTCCC-	AAGAATO	TTTTCAT	TATCAACA	ACCAAACT
los taurus cattle/348=430	GAGG	GAAAT	TCT	TGGACCGGCGC	AAGACGG-A	C	CAGAGCGA	AAG-CA-T	TTGCC-	AAGAATO	TTTTCAT	TAATCAAGA	ACGAAAGT
teropus vampyrus large flying fox/343-425	GAGGT	GAAAT	TCT	GGACCGGCGC	AAGACGG-A	c	CAGAGCGA	AAG-CA-T	TTGCC-	AAGAATO	TTTTCAT	TAATCAAGA	ACGAAAGT
an troplodytes chimpanzee/337-419	GAGGT	GAAAT	тст	GGACCGGCGC	AAGACGG-A	c	CAGAGCGA	AAG-CA-T	TTGCC-	AAGAATO	TTTTCAT	TAATCAAGA	ACGAAAGT
lomo sapiens human/331-413	GAGGT	GAAAT	тст	TGGACTGGCAC	AGTTTGG-A	C	CAGAGCGA	AAG-CA-T	TTGCC-	AAGAATO	TTTTCAT	TAATCAAGA	ACGAAAGT
lomo sapiens human/345-427	GAGGT	GAAAT	тст	TGGACCGGCGC	AAGACGG-A	C	CAGAGCGA	AAG-CA-T	TTGCC-	AAGAATO	TTTTCAT	TAATCAAGA	ACGAAAGT
rinaceus_europaeus_western_European_hedgehog/370-452	GAGGT	GAAAT	тст	TGGACCGGCGC/	AAGACGG-A	C	CAGAGCGA	AAG-CA-T	TTGCC-	AAGAATO	TTTTCAT	TAATCAAGA	ACGAAAGT
rocavia_capensis_cape_rock_hyrax/340-422	GAGGT	GAAAT	тст	TGGACCGGCGC	AAGACGG-A	C	CAGAGCGA	AAG-CA-T	TTGCC-	ΑΑ <mark>G</mark> ΑΑ <mark>Τ</mark>	TTTTCAT	ΤΑΑΤΟΑΑΘΑ	ACGAAAGT
arsius_syrichta_Philippine_tarsier/313-394	GAGGT	GAAAT	TAT	TGGACTGGCGC	AAGACGG-A	C	CAGAGCGA	AAGGCA-T	TTGCC-	ΑΑ <mark>G</mark> ΑΑ <mark>Τ</mark>	TTT A A	ΤΑΑ <mark>Τ </mark> ΑΑ <mark>Γ</mark> ΑΑ Α	A C G A A A G T
quus_caballus_horse/343-425	GAGGT	GAAAT	тст	TGGACCGGCGC	AAGACGG-A	C	CAGAACAA	AAG-CA-T	TTGCC-	ΑΑ <mark>G</mark> ΑΑ <mark>Τ</mark>	TTTTCAT	ΤΑΑ <mark>Τ C</mark> ΑΑ <mark>G</mark> Α.	A C G A A A G T
"ursiops_truncatus_bottlenosed_dolphin/348-430	GAGGT	GAAAT	тст	TGGACCGG <u>C</u> GC/	AAGACGG-A	C	CAGAGCGA	AAG-CA-T	TTGCC-	ΑΑ <mark>G</mark> ΑΑ <mark>Τ</mark>	TTTTCAT	ΤΑΑ <mark>Τ</mark> CΑΑ GΑ	ACGAAAGT
ongo_abelii_Sumatran_orangutan/330-412	GAGGT	- GAAAT	тст	TGGACCGGTGC/	AGTTCGG-A	C	CGGAGCAA	AA <mark>G - C</mark> A - T	TTGCC-	AAGAATO	TTTTCAT	ΤΑΑ <mark>ς </mark> ΑΑ <mark>ς</mark> Α.	Α <mark>Τ Γ</mark> ΑΑΑ <mark>ΓΤ</mark>
Dipodomys_ordii_Ord's_kangaroo_rat/327-409	GAGGT	- <mark>G</mark> A A A <mark>T</mark>	тст	TGGACCGGCGC	AAGACGG-A	C	CAGAGCGA	AA <mark>G-C</mark> A-T	TTGCC-	A A <mark>G</mark> A A <mark>T</mark> (TTTTCAT	TAA <mark>TC</mark> AA <mark>G</mark> A	AC <mark>G</mark> AAA <mark>GT</mark>
Chotona_princeps_American_pika/340-422	GAGGT	GAAAT	тст	TGGACCGGCGC/	AAGACGG-A	C	CAGAGCGA	AA <mark>G-C</mark> A-T	TTGCC-	A A <mark>G</mark> A A <mark>T</mark> (TTTTCAT	TAA <mark>TC</mark> AA <mark>G</mark> A	AC <mark>G</mark> AAA <mark>GT</mark>
arsius_syrichta_Philippine_tarsier/360-442	GAGGT	GAAAT	тст	TGGACCGGCGC	AAGACGG-A	C	CAGAGCGA	AA <mark>G - C</mark> A - T	TTGCC-	AAGAATO	TTTTCAT	TAATCAAGA.	ACGAAAGT
Dipodomys_ordii_Ord's_kangaroo_rat/306-388	GAGGT	GAAAT	тст	TGGACCGGCGC/	AAGACGG-A	C	CAGAGCGA	AAG-CA-T	TTGCC-		TTTTCAT	TAATCAAGA	ACGAAAGT
Dtolemur_garnettii_small-eared_galago/315-397	GAGGT	GAAAT	тст	ICCACCGGCGC/	AAGACGG-A	C	CAGAGCTA	AAG-CA-T	TTGCC-	AAGAATO	TTTTCAT	TAATCAACA	ACGAAAGT
holoepus_hoffmanni_Hoffmann's_two-fingered_sloth/344-426	GAGGT	GAAAT	ТСТ	IGGACCGGCGC/	AAGACGG-A	C	CAGAGCGA	AAG-CA-T	TTGCC-	AAGAATO	TTTTCAT	TAATCAAGA.	ACGAAAGT
rinaceus_europaeus_western_European_hedgehog/346-428	GAGG	GAAA	I CI	GGACCGGCGC	AAGACGG-A	C	CAGAGCGA	AAG-CA-I	TIGCC-	AAGAAIC	TTTTCAL	TAATCAAGA.	ACGAAAGI
avia_porcellus_Domestic_guinea_pig/342-423	GAG-T-	GAAAI	L C	GGACCCG GC/	AAGACGG-A	C	CAAAGCGA	AAG-CA-I	TTCCC-	AAGAAIC	TTTTCAL		AGAAAGI
ongo_abelii_Sumatran_orangutan/329-411	CACCT	GAAAI			ACTTCCC A	C	CAGAGCIA	AAG-CA-I	TTCCC-		TTTTCAL		ACGAAAGI
orilla_gorilla_gorilla_western_lowland_gorilla/341-423	CACCT				ACT TTCC	C		AAG-CA-T	TTCCC		TTTTCAA		ACGAAAGT
an_troglodytes_cnimpanzee/342=424	CACCT		TCC		ACCACCA A	C		AAG C CT	TTCCC		CTTTCAT		
nopheles_maculatus/438-520	CACCT	CAAAT	TCC		ACCACCC-A	C	ACAACCCA		TTCCC	AACCATC	CTTTCAT		ACCAAACT
nopheles_gamblae_strrEs1/453-541	CACCT	CAAAT	TCC			c	TCAACCCA	AAC-C-CT	TTCCC	ATCCATC	CTTTCAT		ACCAAACT
nopheles_taradit_1/500-502	CACCT	CAAAT	TCC			c	TCAACCCA	AAG-C-CT	TTCCC-	ATCCATC	CTTTCAT	CATCAACA	ACCAAACT
nopheles_corresensis/440-522	GAGGT	GAAAT	TCG	AGACCGTCGC	AGGACCG-A	C	TGAAGCGA	AAG-C-GT	TTGCC-	ATGGATO	CTTTCAT	TAATCAAGA	ACGAAAGT
nopheles stephensi Asian malaria mosquito/429-511	GAGGT	GAAAT	TCG	TAGACCGTCGT	AGGACCA-A	c	TGAAGCGA	AAG-C-GT	TTGCC-	ATGGATO	CTTTCAT	TAATCAAGA	ACGAAAGT
nopheles gambiae str. PEST/398-480	GAGGT	GAAAT	TCG	TAGACCGTCGT	AGGACCC-A	c	AGAAGCGA	AAG-C-GT	TTGCC-	AAGGATO	CTTTCAT	TAATCAAGA	ACGAAAGT
nopheles gambiae African malaria mosquito/458-540	GAGGT	GAAAT	TCG	TAGACCGTCGT	AGGACCC-A	C	AGAAGCGA	AAG-C-GT	TTGCC-	AAGGATO	CTTTCAT	TAATCAAGA	ACGAAAGT
nopheles clowi/540-622	GAGGT	GAAAT	TCG	TAGACCGTCGC	AGGACCG-A	C	TGAAGCGA	AAG-C-GT	TTGCC-	ATGGATO	CTTTCAT	TAATCAAGA	ACGAAAGT
nopheles_albimanus/408-490	GAGGT	GAAAT	TCG	TAGACCGTCGT	AAGACTA-A	C	TAAAGCGA	AAG-C-AT	TTGCC-	AAGGATO	CTTTCAT	TAATCAAGA.	ACGAAAGT
nopheles_farauti_4/516-598	GAGGT	- GAAAT	TCG	TAGACCGTCGC	AGGACCG-A	C	TGAAGCGA	AAG-C-GT	TTGCC-	ATGGATO	CTTTCAT	ΤΑΑ <mark>Τ</mark> CΑΑ GΑ.	ACGAAAGT
nopheles_funestus_African_malaria_mosquito/443-525	GAGGT	GAAA	TCG	TAGACCGTCGT/	AGGACCA-A	C	TGAAGCGA	AAG-C-GT	TTGCC-	ATGGATO	сттстт	ΤΑΑ <mark>Τ</mark> C ΑΑ G Α.	AC <mark>G</mark> AAA <mark>G</mark> T
nopheles_gambiae_strPEST/459-541	GAGGT	GAAAT	TCG	TAGACCGTCGT	AGGACCC-A	C	AGAAGCGA	AAG-C-GT	TTGCC-	AAGGATO	CTTTCAT	ΤΑΑ <mark>Τ</mark> ΟΑΑ <mark>Γ</mark> Α.	AC <mark>G</mark> AAA <mark>G</mark> T
nopheles_punctipennis/404–486	GAGGT	GAAA	TCG	TAGACCGTCGT	AAGACTA-A	C	TAAAGCGA	AAG-C-AT	TTGCC-	ATGGACO	CTTTCAT	TAA <mark>TC</mark> AA <mark>G</mark> A	AC <mark>G</mark> AAA <mark>G</mark> T
nopheles_plumbeus/413-495	GAGGT	GAAAT	TCG	TAGACCGCCGT	AAGACTA-A	C	CGAAGCGA	AAG-C-AT	CTGCC-	ATGGATO	CTTTCAT	TAA <mark>TC</mark> AA <mark>G</mark> A	AC <mark>G</mark> AAA <mark>GT</mark>
nopheles_gambiae_strPEST/458-540	GAGGT	GAAAT	TCG	AGACCGTCGT	AGGACCC-A	C	AGAAGCGA	AAG-C-GT	TTGCC-	AAGGATO	CTTTCAT	TAATCAAGA	AC <mark>G</mark> AAA <mark>G</mark> T
nopheles_walkeri/404-486	GAGGT	GAAA	TCG	AGACCGTCGT	AAGACTA-A	C	TGAAGCGA	AAG-C-AT	TTGCC-	ATGGACO	CTTTCAT	TAATCAAGA.	ΑС <mark>G</mark> AAA <mark>G</mark> T
noph_V4_block/1-60			G	AGACCGTCGT	IIIICC-A	C	AGAAGCGA	AAG-C-GT	TTGCC-	AAGGATO	CTTTCAT	AATCA	
/ammal_V4_block/1-60			T	GGACCGGCGC	IIIIGG-A	C	CAGAGCGA	AAG-CA-T	ITGCC-	AAGAATO	TTTCAT	AATCA	

Oryctolagus_cuniculus_rabbit/1-81 Homo sapiens human/1-82 Vicugna_pacos_alpaca/1-82 Otolemur garnettii small-eared galago/1-67 Homo sapiens human/1-83 Homo sapiens human/1-83 Ictidomys tridecemlineatus thirteen-lined ground squirrei Homo_sapiens_human/1-82 Homo_sapiens_human/1-80 Pan_troglodytes_chimpanzee/1-82 Pongo abelii Sumatran orangutan/1-82 Tupaia belangeri northern tree shrew/1-83 Homo_sapiens_human/1-82 Myotis_lucifugus_little_brown_bat/1-83 Pan_troglodytes_chimpanzee/1-82 Erinaceus europaeus western European hedgehog/1-83 Sequences Microcebus murinus gray mouse lemur/1-83 Myotis_lucifugus_little_brown_bat/1-83 Pongo abelii Sumatran orangutan/1-82 Dipodomys ordii Ord's kangaroo rat/1-82 Homo sapiens human/1-82 Ochotona_princeps_American_pika/1-86 Homo_sapiens_human/1-82 Procavia capensis cape rock hyrax/1-83 Bos_taurus_cattle/1-83 Pteropus vampyrus large flying fox/1-83 Pan_troglodytes_chimpanzee/1-83 Homo sapiens human/1-82 Homo_sapiens_human/1-83 Erinaceus_europaeus_western_European_hedgehog/1-83 Procavia capensis cape rock hyrax/1-83 Tarsius_syrichta_Philippine_tarsier/1-83 Equus caballus horse/1-83 Tursiops_truncatus_bottlenosed_dolphin/1-83 Pongo_abelii_Sumatran_orangutan/1-82 Dipodomys ordii Ord's kangaroo rat/1-83 Ochotona_princeps_American_pika/1-83 Tarsius syrichta Philippine tarsier/1-83 Dipodomys_ordii_Ord's_kangaroo_rat/1-83 Otolemur_garnettii_small-eared_galago/1-78 Choloepus_hoffmanni_Hoffmann's_two-fingered_sloth/1-8. Erinaceus_europaeus_western_European_hedgehog/1-83 Cavia porcellus Domestic guinea pig/1-83 Pongo_abelii_Sumatran_orangutan/1-82 Gorilla_gorilla_gorilla_western_lowland_gorilla/1-82 Pan_troglodytes_chimpanzee/1-84 Anopheles_maculatus/1-89 Anopheles_farauti_1/1-85 Sequences Anopheles_torresiensis/1-85 Anopheles annulipes A/1-85 Anopheles_stephensi_Asian_malaria_mosquito/1-85 Anopheles_gambiae_str._PEST/1-85 Anopheles gambiae African malaria mosquito/1-85 Anopheles_clowi/1-85 Anopheles albimanus/1-85 Mosquito Anopheles_farauti_4/1-85 Anopheles_funestus_African_malaria_mosquito/1-85 Anopheles_punctipennis/1-85 Anopheles_plumbeus/1-84 Anopheles_gambiae_str._PEST/1-85 Anopheles walkeri/1-83 Mammal_block_I-short_1391f/1-40 Anoph_V9_block/1-46

Mammalian

	10	20	30		40	50	60	70	80	90	
	CTACACACCCCCCC	CCTACT	ACCCATTCCA	TCCTT			CCCATCCC-	C-CCCCCCCC	TCCCCCC	ACCC	- CCCT
	un ACACACCUCCCUTC	UCTACT	ACCUATTOUA	luuli	TA-U		COUNTCOU	c-ccuccuudu-		A C 00	
	GIACACACCGCCCGIC	GCTACT	- ACCGAIIGGA		- A - G	- GAGGCCCT	CGGAICGG-			A C GG	C - C I
	GTACACACCGCCCGTC	GCTACT	- ACCGATTGGA	TGGTT-	TA-G	- GAGGCCCT	CGGATCGG-	CCCCCCCCGGGG	TCGGCCC	A C G G	CCG-
	CTACACACTCCCCATT	CCTACT	- ACCAATTCCA	TCATT-			CCCATCAC-	TCCTACTCC			T
	CTACACACCCCCCCC	CCTACT	ACCONTROCA	TCCTT		CACCCCCT	CCCATCCC	CECECECECE	TCC CCCC		CCCT
	GTACACACCUCCUTC	GUTACI	- ACCUATIGGA	16611-		- GAGGEEEE	CGGALCGG-	CCCCCCCCCCC		ACGG	
	GTACACACCGCCCGTC	GCTACT	- ACCGATTGGA	TGGTT-	TA-G	- GAGGCCCT	CGGATCGG-	CCCCCCCCCCCC	TCGGCCC	A C G G	CCCT
irrel/1-86	GTACACACCGCCCGTC	GCTACT	- ACCGATTGGA	TGGTT-	TA-G	- GAGGCCCT	CGGATCGG-	CCCCCCCCCCCC	TCGCCCC	A CAAAGG	GCCT
	CTACACACTCCCCCC	TCTACT	ACTCATACCA	TCCTT	TA CT	CACCCCT	CTCATCAC	CCCCCCCCCC	TCALLCCCC	A C T C	CCCT
	GTACACACI GCCCGCC	COTICT	CCT CLTTCCL	TCCTT		CACCCCC	CCCATCAG-				CCCT
	GTACACACIGCA-GIC	GCTACT	- CCIGAIIGGA			- GAGGCCCT	CGGATCAG-	C-CCG-CIGGG		ACIG	
	GTACACACTGCCCGTC	ACTACT	- ACTGATAGGA	TGATT-	T A - G	- GAGGCCCT	CGGATCGG-	CCCCCCCCGGG	TCAGCCC	A C T G	CCCT
	GTACACACCGCCGGTA	GCTACT	- ACTGATTGGA	TGGTT-	TA-G	- GAGGCCCT	CAGATCGG-	CCCCCCCCCCC	TCAGCCC	A C T G	CCCT
	CTACACACCCCCCCC	CCTACT	ACCONTICOA	TCCTT	TA CT	CACCCCT	CCCATCCC	CCCCCCCCCCC		A	CCCT
	GTACACACCGCCCGTC	GTIGT	ACCUATIONA	TCCTT			COUNTCOU			A	CCCT
	GTACACATCGCCAGTC	ACTACT				- GAGGCCCT	CGGALCGG-	CCCCCCCCC-C	ICAGCCC	A C I G	
	GTACACTCCACCTGTC	GCTACT	- ACCGATTGGA	TGGTT-	TA-G	- GAGGCCCT	CGGGGTGGG-	CCCCCCCCGGGG	TGGGCCC	A C C G	CCC T
	GTACACACTGCCAGTC	ACTACT	- ACTGATTGGA	TGGTT-	CA-G	- GAGGCCCT	TGGATCGG-	CCCCCCCAGG	TCAGCCC	A C T G	CCCT
07	CTACACACCCCCCCC	CCTACT	ACCONTICOA	TCCTT		CACCCCT	CCCATCCC	CCCCCCCCCCC		A	CCCT
05	GTACACACCCCCCCCC	CCTACT	ACCUATIOUA				COUNTCOU			A C 00	
	GTACACACCGCCCGTC	GUTACI	- ACCGATIGGA		A - G	- GAGGCCCT	CGGALCGG-	CCCCCCCCCCCC		ACGG	
	GTACACACCGCCCGTC	GCTACT	- ACCGATTGGA	TGGTT-	• T A – A T	- TAGGCCCT	CGGATCGG-	CCCCTCCGGGG	TCGGCCT	A C G G	CCC T
	GTACACACCGCCGGTA	GCTACT	- ACTGATTGGA	TGGTT-	TA-G	- GAGGCCCT	TGGATCGG-	CCCCCCAGG	TCAGCCC	A C T G	CCCT
	CTACACACCCCCCCC	CCTACT	ACCONTICOA	TCCTT	TACT	CACCCCT	CCCATCCC	CCCCCCCCCCC		A C CC	CCC
	GTACACACCCCCCCCC	GCTACT	ACCUATIOUA			- UAUUUUUU	CUUATCUU-			A C UU	
	GIACACACCGCCAGII	GCTACT	- ACTGATIGGA			- GAGGCCCT	CAGAGCGG-	CCCCCCCCCCC	ICAGCCC	ACIG	CCGT
	GTACACACCGCCCGTC	GCTACT	- ACCGATTGGA	TGGTTT	TA-G	GGAGGCCCT	CGGATCCG-	CCCCCCCCGGGG	TCGGCCC	A C G G	- GCCCT
	GTACACACTCCCACTC	CCTACT	- ACTGATTGGA	TCCTT-		- GAGGCCCT	TCCATCCC-	CCCCCCCCCC		A C T C	CCCT
	CTACACACCCCCCCCC	CCTACT	ACCONTICON	TCCTT	TACT	CACCCCCT	CCCATCCC	CCCCCCCCCC			CCCT
	GTACACACCCCCCCC	GUTACI	ACCUATTUGA		TA-U	- GAGGCCCT	CUGATCUG-	CCCCCCCCCCCA		A	
	GTACACACCGCCCGTC	GCTACT	- ACCGATIGGA		• I A - G	- GAGGCCCT	CGGATCGG-	CCCCCCCCCCCC	• TCGGCCC	A C GG	CCCT
	GTACACACCGCCCGTC	GCTACT	- ACCGATTGGA	TGGTT-	TA-G	- GAGGCCCT	CGGATCGG-	CCCCCCCCGGGG	TCGGCCC	A C G G	CCCT
	GTACACACCCCCCC	CCTACT		TCCTT-		- CACCCCT	CCCATCCC-	CCCCCCCCCCCC	TCGCCCC	A C C C	CCCT
	CTACACACTCCACCTT	CCTACT	CETCATTCCA	TCCTT	TACT	CACCCCCT	CCCATCAC	C CCC CCCCC			CCCT
	GTACACACIGCAGGTT	UCTACT	- CCTGATTGGA		TA-U	- UAGUCCCT	CUUATCAU-		ICAUCCC	A	
	GTACACACCGCCCGTC	GCTACT	- ACCGATTGGA	TGGTT-	• T A - G	- GAGGCCCT	CGGATCGG-	CCCCCCCCCCCC	- <mark>T</mark> CG GCCC	A C GG	CCCT
83	GTACACACCGCCCGTC	GCTACT	- ACCGATTGGA	TGGTT-	TA-G	- GAGGCCCT	CGGATCGG-	CCCCCCCCGGGG	TCGGCCC	A C G G	CCCT
	GTACACACCCCCC	GCTACT	- ACCGATTGGA	TGGTA-	TA-G	- GAGGCCCT	CGGATCGG-	CCCCCCCCCCA	TCGGTCC	A C C G	CCCT
	CTACACACCCCCCCC	CCTACT	ACCONTICOA	TCCTT	TACT	CACCCCT	CCCATCCC	CCCCCCCCCCC		A C CC	CCAT
	GTACACACCCCCCCCC	UCTACT	ACCUATTOUA		TA-U	- GAGGCCCT	COUATCOU-			A C GG	CCAT
	GTACACACCGCCGGTC		- ACCGATIGGA		- A - G	- GAGGCCCT	CGGALGGG-	CCCCCCCCCCCC		A C GG	CCCT
	GTACACACCGCCCGTC	GCTACT	- ACCGATTGGA	TGGTT-	TA-G	- GAGGCCCT	CGGATCGG-	CCCCCCCCGGGG	TCGGCCC	A C G G	CCCT
	GTACACACTGCCGGTA	GCTACT	- ACTGATTGGA	TGGTT-	TA-G	- GAGGCCCT	TGGATCGG-	C-CCCCCTCCC	TAAGCCC	A C T G	CCCT
	CTACACACCCCCCCC	CCTACT	ACCONTICON	TCCTT	TACT	CACCCCCT	CCCATCCC	CECCECCCCC	TCC CCCC		CCCT
	GTACACACCUCCUTC	GUTACI	- ACCUATTUUA	I G G I I	TA-U	- GAGGCCCT	COUATCOU-			A	
	GTACACACCGCCCGTC	GCTACT	- ACCGATTGGA	TGGTT-	• T A - G	- GAGGCCCT	CGGATCGG-	CCCCCCCCCCCC	• <mark>T</mark> CG GCCC	A C G G	CCCT
	GTACACACCGCCCGTC	GCTACT	- ACCGATTGGA	TGGTT-	TA-G	- GAGGCCCT	CGGATCGG-	CCCCCCCCGGGG	TCGGCCC	A C G G	CCCT
	GGACACTCCCCCC	GCTACT	- ACCGAATGGA	TGGTT-	TA-G	- GAGGCCCT	CTGATCGG-	CCCCCCCCCCCC	TCGGCCC	A C GG	CCCT
	CTACACACTCCCCCTT	CCTACT	ACCCATTCCA	TCCTT	TACT	CACCCCCT	CACATCCC	CCTCCCTCC			CT
	GTACACACITUCCCUTT	GUTACI	- ACCUATTUCA		TA-U	- UAUUUUUU	CAGATCOG-			AC	
/1-83	GTACACACCGCCCGTC	GCTACT	- ACCGATTGGA	IGGII-	• T A - G	- GAGGCCCT	CGGATCGG-	CCCCCCCCCCCC	- <mark>T</mark> CG GCCC	A C GG	CCCT
83	GTACACACCGCCCGTC	GCTACT	- ACCGATTGGA	TGGTT-	TA-G	- GAGGCCCT	CGGATCGG-	CCCCCCCCGGGG	TCGGCCC	A C G G	CCCT
	GTACACACCACCATT	GCTACT	- ACCTATTGGA	TGGTT-	TA-G	- GAGGCCCT	CAGATGGG-	CCCAGCGGGGG	TTGGCCC	A C G G	CCCT
	CTCCACACTCCTCCTA	CCTACT	ACTCATTCAA	TCCTT	TACT	CACCCCT	CCCATTCC	C CCCCCTCCA		A T T	CCCT
	GTUCACACIGCIGGIA	COTACT		TCCTT			TCCATCCC	C-CCCGCTGGA	TCA		CCCT
	GTACACACIGUUU	GUTACI	- ACTGATAGGA			- GAGGELLET	IGGAICGG-	C-CC16CC666	ILAGUUU	AUIU	
	GTACACACTGCCCGCC	TCTACT	- ACTGATAGGA	TGGTT-		- GAGGCCCT	CTGATCAG-	C-CCTGCCGGG	TCAGCCC	<u>ACG</u> T	GCCCCT
	GTACACACCGCCCGTC	GCTACT	AACCGAT-GGA	TAATO	TAAG	- GGGGTCTC	TGGAGGCAA	CGCCTTCCGCC	NCNNCCTT-	GT - GA GC	TGCA
	GTACACACCCCCCC	GCTACT	- ACCGAT-GGA	TTCTT-	TA-G	- GAGGTCTC	TCCACCCA-	CCCCTTCCCCC	ATTC-CACC	$\mathbf{GC} - \mathbf{GA} \mathbf{GT}$	
	CTACACACCCCCCCC	CCTACT	ACCONT COM	TTCTT	TACT	CACCTCTC	TCCACCCA	CCCCTTCCCCC	ATTC-CACC	CC-CA	
	GTACACACCCCCCCCC	GUTACI	- ACCUAT- GUA		TA-U	- GAGGTETC	TUUAUUCA-		ATTC-CACC	UC-UAUT	
	GTACACACCGCCCGTC	GATACT	- ACCGAT - GGA	TICTT-	IA-G	- GAGGTCTC	TGGAGGCA-	CCCCCCCCCCCCCC	ATIC-CACC	GI-GAGT	
	GTACACACCGCCCGTC	GCTACT	- ACCGAT - GGA	TTATT-	TA-G	- GAGGTCTC	TGGAGGCA-	TACCTTCCGCG	GTTC-CTTC	GT-GAGC	TGC A
	GTACACACCGCCCGTC	GCTACT	- ACCGAT - GGA	TTATT-	TA-G	- GAGGTCTC	TGGAGGCA-	CACCTTCCCCC	ATTC-CTTC	$\mathbf{GT} - \mathbf{GA} \mathbf{GT}$	TGCA
	CTACACACCCCCCC	CCTACT	ACCONT COA	TTATT	TACT	CACCTCTC	TCCACCCA-	CACCTTCCCCC		CT-CACT	TCCA
	GTACACACCGCCCGTC	COTACT	ACCOAT GOA	TTCTT			TCCACCCA	CACCT CCCCC		CT CA CT	
	GTACACCGCCCGTC	GUTACT	- ACCGAT-GGA	IICII-	TA-G	- GAGGICIC	IGGAGGCA-		ATTC-CACC	GI-GAGT	
	GTACACACCGCCCGTC	GCTACT	- ACCGAT - GGA	TATT-	TA-G	- GAGGTCTT	TGAAGACG-	AACCTATGCTGCT	GCTC-CTC-	GT - GG GC	– – <mark>C A</mark> – C
	GTACACACCGCCCGTC	GCTACT	- ACCGAT - GGA	TTCTT-	TA-G	- GAGGTCTC	TGGAGGCA-	CGCCTTCCCCC	ATTC-CACC	GT - GA GT	TCCA
	CTACACACCCCCCC	CCTACT	- ACCCAT-CCA	TTATT	TA-CT	- CACCTTTC	TCCACCCT-	TACCTTCCCCC	CTTC-CTTC	CT-CACC	
	CTACACACCCCCCCCC	CCTACT	ACCONTOUR		TAC	CACCTOT	TCCACCTC		COTOCOTO	CTCCA	
	GIACACACCGCCCGIC	GUTACI	- ACCUAT-GGA	TATT-	TA-G	- GAGGICIC	TUGAUGIG-	AACGI CGCAA	UCICCUTC-	UIGUAUI	A <mark>G</mark> C
	GTACACACCGCCCGTC	GCTACT	- ACCGAT - GGA	TTATT-	TA-G	- GAGGTCTC	TGGAGATG-	AACCTATCCTC-1	GCTC-CTC-	GC-GGGC	– – <mark>CA</mark> – C
	GTACACACCGCCCGTC	GCTACT	- ACCGAT - GGA	TTATT-	TA-G	- GAGGTCTC	TGGAGGCA-	CACCTTCCCCC	ATTC-CTTC	GT-GAGT	TCCA
	CTACACACCCCCCC	CCTACT	- ACCCAT-CCA	TTATT	TA-CT	- CACCTCTC	TCCACCTA-	ATCOTTCCCAC	CCC-CTC-	CT-CCCT	A C C
	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	CCTACT	ACCONTECT		TAC		TOUNDU IA-	CO COCAO-		<u>u - uu u</u>	A U C
		GUTACI	ACCUATIGGE		TA-G	- GAUGUUUI					
	GCCCGTC	GCTACT	- ACCGAT - GGA	TATT-	A - 1	- IIIGTCTC	GGAGGY				
								- 8			_
											عكيو
Consensus											

Figure S4





Table S1. Primers used for Anneal Blocking and PNA Blocking

Anneal Blockers

Anneal Blockers		
Hypervariable Region	Name	Sequence (5'->3')
	Mammal_block_I-short_1391f	GCCCGTCGCTACTACCGATTGG/ideoxyl//ideoxyl//ideoxyl//ideoxyl//TTAGTGAGGCCCT/3SpC3/
٧J	Anoph_V9_block	GCCCGTCGCTACCGATGGATTATTTA/ideoxyl//ideoxyl
	18S_V9_1391_F_Nextera*	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTACACACCGCCCGTC
	18S_V9_EukBr_R_Nextera*	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG TGATCCTTCTGCAGGTTCACCTAC
VA	Mammal_V4_block	TGATTAATGAAAACATTCTTGGCAAATGCTTTCGCTCTGGTCC/ideoxyl//ideoxyl
**	Anoph_V4_block	TGATTAATGAAAGCATCCTTGGCAAACGCTTTCGCTTCTGTGG/ideoxyl//ideoxyl//ideoxyl//ideoxyl//ideoxyl/ACGACGGTCTAC/3SpC3/
	18S_V4F_Nextera*	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG CCAGCASCYGCGGTAATTCC
	18S_V4R_Nextera*	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGACTTTCGTTCTTGATYRA

*The 18S targetting portions of the primer are in bold, adaptor tails for sequencing are in plain text

PNA Blockers		
Hypervariable Region	Name	Sequence (5'->3')
V9	AgV9-PNA	CCGTGCCAACTGCAAC
V4	AgV4-PNA	ACGCCCAGGTACACC
Hypervariable region	1 Forward primer*	Reverse primer*
V9	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTACACACCGCCCGTC	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG TGATCCTTCTGCAGGTTCACCTAC
V4	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCAGCASCYGCGGTAATTC	CGTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGACTTTCGTTCTTGATYRA

*The 18S targetting portions of the primer are in bold, adaptor tails for sequencing are in plain text