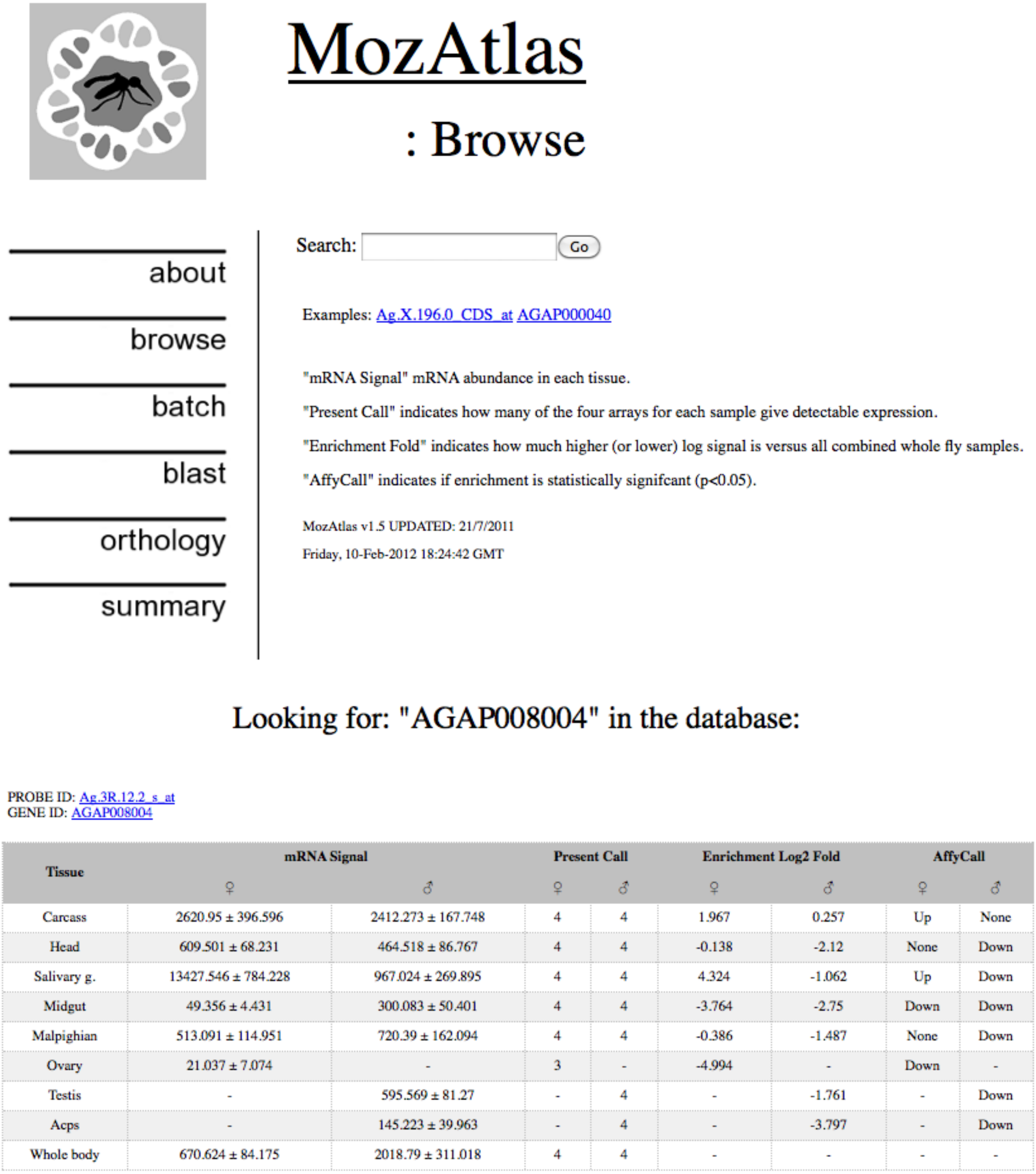
Table S1. Oligonucleotide primers used for expression analysis

Primer	Sequence	Experiment[a]
cE5-F1	5′-CAGAACGCGAGAGCAACAGC-3′	1, 2
cE5-R1	5′-GGCGTGTTCCTCGAGAAGGC-3′	1, 2
cE5-R2	5′-AGATGAAGGAAGGTGCGGAG-3′	1, 2
cE5-Fm	5′-TTTGATCCAAGCCTTCTCGAGGAAC-3′	3
cE5-F3	5′-CTGGTACAAACATCTATGCC-3′	3
AP	5′-GGCCACGCGTCGACTAGTAC$(T)_{17}$-3′	3
RpS7-F	5′-GGCGATCATCATCTACGTGC-3′	1, 2
RpS7-R	5′-GTAGCTGCTGCAAACTTCGG-3′	1, 2

a. (1) RT-PCR on first strand cDNA. (2) One-step RT-PCR. (3) 3′-RACE

Figure S1. Sex- and tissue-specific expression profiles of the *An. gambiae* genes encoding cE5 (AGAP008004), apyrase (AGAP011971), salivary peroxidase (AGAP010735) and D7r1 (AGAP008280) as determined by the microarray analysis of Baker et al. (2011). Data reported below were retrieved searching the web server http://www.tissue-atlas.org/.



MozAtlas

: Browse

Search:

Examples: Ag.X.196.0_CDS_at AGAP000040

"mRNA Signal" mRNA abundance in each tissue.

"Present Call" indicates how many of the four arrays for each sample give detectable expression.

"Enrichment Fold" indicates how much higher (or lower) log signal is versus all combined whole fly samples.

"AffyCall" indicates if enrichment is statistically signifcant (p<0.05).

MozAtlas v1.5 UPDATED: 21/7/2011
Friday, 10-Feb-2012 18:24:42 GMT

Looking for: "AGAP008004" in the database:

PROBE ID: Ag.3R.12.2_s_at
GENE ID: AGAP008004

Tissue	mRNA Signal		Present Call		Enrichment Log2 Fold		AffyCall	
	♀	♂	♀	♂	♀	♂	♀	♂
Carcass	2620.95 ± 396.596	2412.273 ± 167.748	4	4	1.967	0.257	Up	None
Head	609.501 ± 68.231	464.518 ± 86.767	4	4	-0.138	-2.12	None	Down
Salivary g.	13427.546 ± 784.228	967.024 ± 269.895	4	4	4.324	-1.062	Up	Down
Midgut	49.356 ± 4.431	300.083 ± 50.401	4	4	-3.764	-2.75	Down	Down
Malpighian	513.091 ± 114.951	720.39 ± 162.094	4	4	-0.386	-1.487	None	Down
Ovary	21.037 ± 7.074	-	3	-	-4.994	-	Down	-
Testis	-	595.569 ± 81.27	-	4	-	-1.761	-	Down
Acps	-	145.223 ± 39.963	-	4	-	-3.797	-	Down
Whole body	670.624 ± 84.175	2018.79 ± 311.018	4	4	-	-	-	-

Looking for: "AGAP011971" in the database:

PROBE ID: Ag.3L.17.0_CDS_at
GENE ID: AGAP011971

Tissue	mRNA Signal		Present Call		Enrichment Log2 Fold		AffyCall	
	♀	♂	♀	♂	♀	♂	♀	♂
Carcass	9.908 ± 4.037	5.418 ± 0.042	1	0	0	0	NA	NA
Head	19.221 ± 11.275	5.52 ± 0.042	2	0	0	0	NA	NA
Salivary g.	17097.232 ± 1584.821	7.364 ± 0.694	4	0	8.599	0	Up	NA
Midgut	5.689 ± 0.112	5.577 ± 0	0	0	0	0	NA	NA
Malpighian	5.573 ± 0.345	5.228 ± 0	0	0	0	0	NA	NA
Ovary	6.518 ± 0.013	-	0	-	0	-	NA	-
Testis	-	5.128 ± 0.016	-	0	-	0	-	NA
Acps	-	7.003 ± 0.089	-	0	-	0	-	NA
Whole body	44.105 ± 9.502	6.32 ± 0.165	4	0	-	-	-	-

Looking for: "AGAP008280" in the database:

PROBE ID: Ag.3R.59.0_CDS_at
GENE ID: AGAP008280

Tissue	mRNA Signal		Present Call		Enrichment Log2 Fold		AffyCall	
	♀	♂	♀	♂	♀	♂	♀	♂
Carcass	5.877 ± 0	6.381 ± 0.465	0	0	0	0	NA	NA
Head	11.457 ± 0.991	9.59 ± 0.544	1	0	0	0	NA	NA
Salivary g.	2563.578 ± 70.29	12.646 ± 1.359	4	1	8.459	0	Up	NA
Midgut	5.324 ± 0.046	5.278 ± 0	0	0	0	0	NA	NA
Malpighian	5.729 ± 0.485	5.236 ± 0.004	0	0	0	0	NA	NA
Ovary	5.187 ± 0	-	0	-	0	-	NA	-
Testis	-	7.853 ± 0	-	0	-	0	-	NA
Acps	-	6.428 ± 0.336	-	0	-	0	-	NA
Whole body	7.286 ± 0.276	6.811 ± 0.152	0	0	-	-	-	-

Looking for: "AGAP010735" in the database:

PROBE ID: Ag.3L.638.0_CDS_at
GENE ID: AGAP010735

Tissue	mRNA Signal		Present Call		Enrichment Log2 Fold		AffyCall	
	♀	♂	♀	♂	♀	♂	♀	♂
Carcass	8.146 ± 1.864	5.355 ± 0.064	1	0	0	0	NA	NA
Head	30.237 ± 12.817	6.711 ± 0.041	3	0	-0.762	0	None	NA
Salivary g.	16887.781 ± 1445.617	6.541 ± 0.917	4	0	8.363	0	Up	NA
Midgut	5.309 ± 0.002	5.306 ± 0	1	0	0	0	NA	NA
Malpighian	12.573 ± 7.343	5.236 ± 0	1	0	0	0	NA	NA
Ovary	5.336 ± 0.129	-	0	-	0	-	NA	-
Testis	-	5.097 ± 0.002	-	0	-	0	-	NA
Acps	-	5.307 ± 0.011	-	0	-	0	-	NA
Whole body	51.277 ± 8.627	6.922 ± 0.032	4	0	-	-	-	-

PROBE ID: Ag.3L.1742.1_s_at
GENE ID: AGAP010735

Tissue	mRNA Signal		Present Call		Enrichment Log2 Fold		AffyCall	
	♀	♂	♀	♂	♀	♂	♀	♂
Carcass	10.158 ± 4.319	5.31 ± 0	1	0	0	0	NA	NA
Head	20.577 ± 6.296	5.548 ± 0.173	3	0	-1.644	0	Down	NA
Salivary g.	14481.531 ± 1399.842	6.337 ± 0.894	4	0	7.815	0	Up	NA
Midgut	5.281 ± 0.006	5.276 ± 0	0	0	0	0	NA	NA
Malpighian	18.483 ± 13.265	5.218 ± 0	1	0	0	0	NA	NA
Ovary	5.187 ± 0	-	0	-	0	-	NA	-
Testis	-	5.095 ± 0	-	0	-	0	-	NA
Acps	-	5.289 ± 0	-	0	-	0	-	NA
Whole body	64.289 ± 7.617	5.245 ± 0.097	4	0	-	-	-	-

PROBE ID: Ag.3L.1742.0_CDS_at
GENE ID: AGAP010735

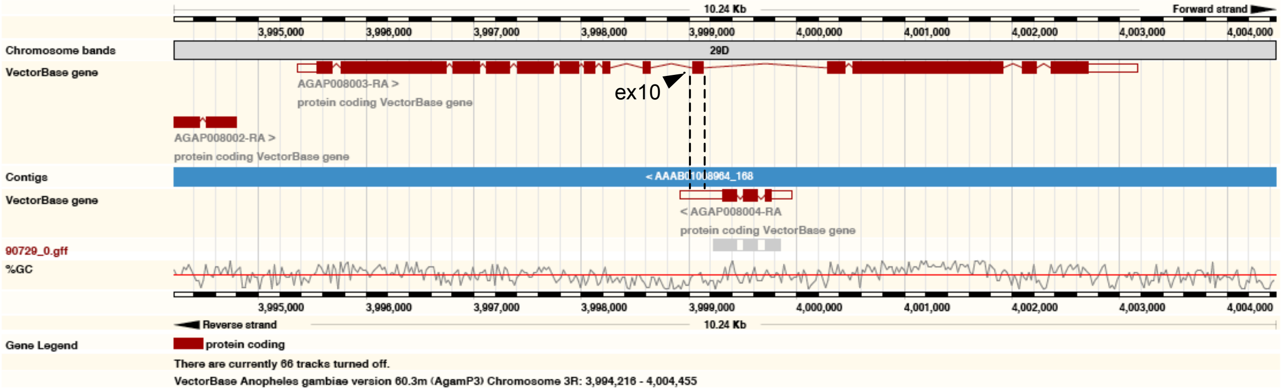
Tissue	mRNA Signal		Present Call		Enrichment Log2 Fold		AffyCall	
	♀	♂	♀	♂	♀	♂	♀	♂
Carcass	6.188 ± 0.949	5.247 ± 0.008	0	0	0	0	NA	NA
Head	7.462 ± 1.365	5.456 ± 0.031	1	0	0	0	NA	NA
Salivary g.	3627.104 ± 516.332	5.367 ± 0.089	4	0	8.389	0	Up	NA
Midgut	5.524 ± 0.248	5.276 ± 0	0	0	0	0	NA	NA
Malpighian	6.097 ± 0.879	5.218 ± 0	1	0	0	0	NA	NA
Ovary	5.187 ± 0	-	0	-	0	-	NA	-
Testis	-	5.095 ± 0	-	0	-	0	-	NA
Acps	-	5.275 ± 0	-	0	-	0	-	NA
Whole body	10.818 ± 1.497	5.678 ± 0.108	0	0	-	-	-	-

Figure S2. The 3’UTR of the longest *cE5* mRNA isoform ($cE5L_{950}$) with the potential binding sites for aga-miR-989 and aga-miR-981 (in blue) are shown. Polyadenilation signals AAUAAA are highlighted in red. The 99 nt region corresponding to a small exon of gene AGAP008003 encoded on the other DNA strand is highlighted in yellow.

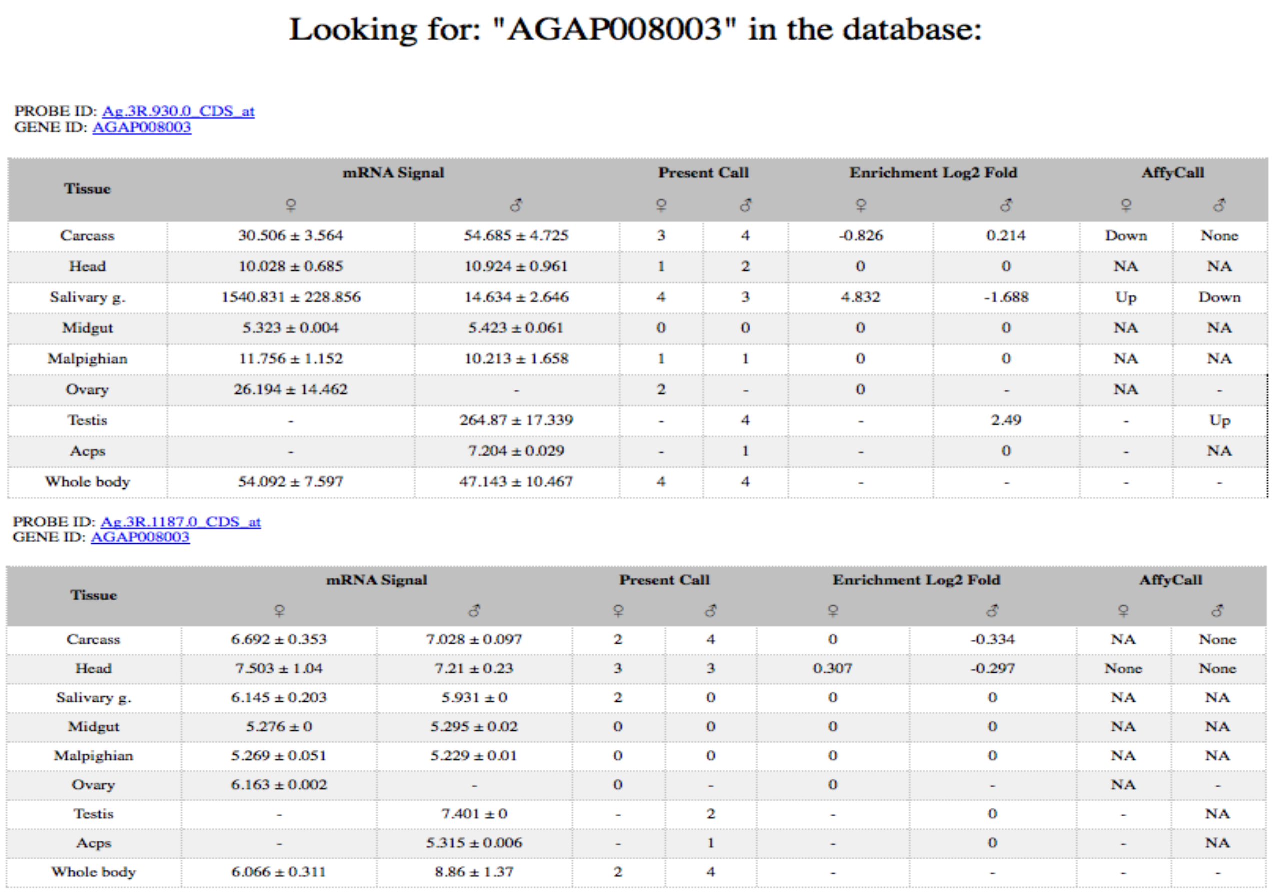
```
UUGCAGGAGCAUCAAAUAUGUUAUAUCACGGAAUAUGUAGAACGCAUUUAAGCAUCAAUAAAAUUACAGAAACGGUAUUA   80
        | ||||  ||||||||||||||
        cauggug-augcaguguagugu (aga-miR-989)

CACCUAGUUCCAUCAGUUCGAUGAGCUCUUUGUACAACAAAAAUUGGCUUGAAGGAAUAACAACACAACAGACACUGGUA  160

CAAACAUCUAUGCCCUAACUUACCCUUACCGGUUUCGCCGACAUUUUUCACGACGCUUUCCAGCUUUGCCUGCACCUUGG  240
                            | ||||||| |||||      ||||
                            guccaaagcagcugu------ugcuu (aga-miR-981)

UUUGCACAUUCUCCGCACCUUCCUUCAUCUUAGCAAGGAUACCUUUACGCCGCAAAGAAAUGUACCAAACAAACAUGAGU  320

AUACUGCGCAAACAAUCAUUUUGCUGAUUGCUAAGAAAUUAGGGAAUUAAUAAAUUAUUAUAUAGGGACAACAAAAUGCA  400

ACAUUCUAACUAGCUAUACAAUGAAAAGCAUAACAUUAAAUACACAUUCAAAUAAACCACAGCUUCUAGUUUCCCAAUAU  480

GAUUGCAAUACG 492
```

Figure S3. (A) Map of the genomic region encoding AGAP008004 (*cE5*) and AGAP008003 genes. Coding regions are shown as red boxes, 5'UTRs and 3'UTR as empty boxes, red lines mark introns. The arrowhead points to exon 10 of AGAP008003, which corresponds to the 3'UTR of the *cE5L* mRNAs. The map was obtained by VectorBase (http://www.vectorbase.org). **(B)** Sex- and tissue-specific expression profile of the *An. gambiae* AGAP008003 gene as determined by the microarray analysis of Baker et al. (2011). Data retrieved searching the web server http://www.tissue-atlas.org/.

(A)



(B)



Looking for: "AGAP008003" in the database:

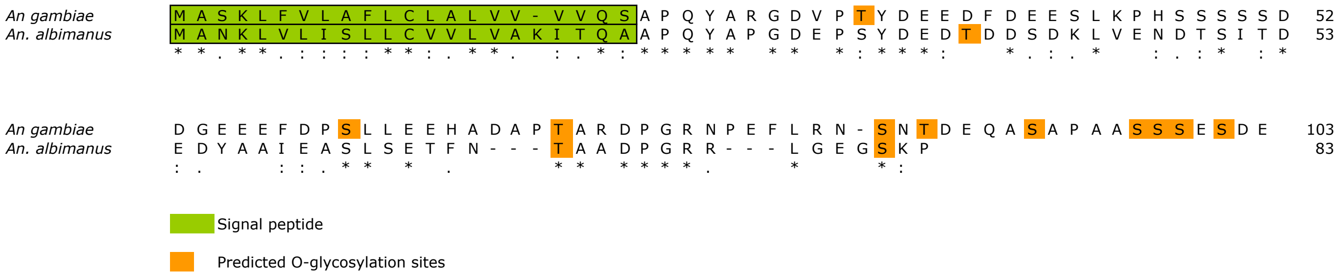
PROBE ID: Ag.3R.930.0_CDS_at
GENE ID: AGAP008003

Tissue	mRNA Signal		Present Call		Enrichment Log2 Fold		AffyCall	
	♀	♂	♀	♂	♀	♂	♀	♂
Carcass	30.506 ± 3.564	54.685 ± 4.725	3	4	-0.826	0.214	Down	None
Head	10.028 ± 0.685	10.924 ± 0.961	1	2	0	0	NA	NA
Salivary g.	1540.831 ± 228.856	14.634 ± 2.646	4	3	4.832	-1.688	Up	Down
Midgut	5.323 ± 0.004	5.423 ± 0.061	0	0	0	0	NA	NA
Malpighian	11.756 ± 1.152	10.213 ± 1.658	1	1	0	0	NA	NA
Ovary	26.194 ± 14.462	-	2	-	0	-	NA	-
Testis	-	264.87 ± 17.339	-	4	-	2.49	-	Up
Acps	-	7.204 ± 0.029	-	1	-	0	-	NA
Whole body	54.092 ± 7.597	47.143 ± 10.467	4	4	-	-	-	-

PROBE ID: Ag.3R.1187.0_CDS_at
GENE ID: AGAP008003

Tissue	mRNA Signal		Present Call		Enrichment Log2 Fold		AffyCall	
	♀	♂	♀	♂	♀	♂	♀	♂
Carcass	6.692 ± 0.353	7.028 ± 0.097	2	4	0	-0.334	NA	None
Head	7.503 ± 1.04	7.21 ± 0.23	3	3	0.307	-0.297	None	None
Salivary g.	6.145 ± 0.203	5.931 ± 0	2	0	0	0	NA	NA
Midgut	5.276 ± 0	5.295 ± 0.02	0	0	0	0	NA	NA
Malpighian	5.269 ± 0.051	5.229 ± 0.01	0	0	0	0	NA	NA
Ovary	6.163 ± 0.002	-	0	-	0	-	NA	-
Testis	-	7.401 ± 0	-	2	-	0	-	NA
Acps	-	5.315 ± 0.006	-	1	-	0	-	NA
Whole body	6.066 ± 0.311	8.86 ± 1.37	2	4	-	-	-	-

Figure S4. Alignment of the *An. gambiae* cE5 to the *An. albimanus* anophelin. The alignment was obtained using the ClustalW2 Multiple Sequence Alignment program at the EBI website (http://www.ebi.ac.uk/Tools/msa/clustalw2/). Signal peptides are boxed and highlighted in green. Potential O-glycosylation sites, as predicted by the NetOGlyc 3.1 server (Julenius et al., 2005; http://www.cbs.dtu.dk/services/NetOGlyc/) are highlighted in orange. Identities (*), conservative (:) and semiconservative (.) substitutions are marked.



```
An gambiae      M A S K L F V L A F L C L A L V V - V V Q S A P Q Y A R G D V P T Y D E E D F D E E S L K P H S S S S S D   52
An. albimanus   M A N K L V L I S L L C V V L V A K I T Q A A P Q Y A P G D E P S Y D E D T D D S D K L V E N D T S I T D   53
                * * . * * . : : : : * * : . * * .   : . * : * * * * *   * *   * : * * * :     * . : . *     : . : *   : *

An gambiae      D G E E E F D P S L L E E H A D A P T A R D P G R N P E F L R N - S N T D E Q A S A P A A S S S E S D E   103
An. albimanus   E D Y A A I E A S L S E T F N - - - T A A D P G R R - - - L G E G S K P                                    83
                : .       : : . * *   *   .         * *   * * * * .       *       * :
```

🟩 Signal peptide

🟧 Predicted O-glycosylation sites

Supplementary material and methods

1. Serine protease inhibition assays: enzymes, buffers and substrates.

The **source and concentration of the enzymes** used in the serine protease screen was as follows:

- Thrombin (0.03 nM), α-chymotrypsin (0.075 nM), plasmin (0.4 nM) and chymase (0.9 nM) were purchased from Sigma (St. Louis, MO);
- skin β-tryptase (0.01nM) was purchased from Promega (Madison, WI);
- activated coagulation factor X (FXa) (0.1nM) was purchased from EMD Biosciences (La Jolla, CA);
- FXIIa (0.1 nM) was purchased from Haematologic Technologies Inc. (Essex Junction, VT);
- kallikrein (0.04nM) was purchased from Fitzgerald Industries International (Concord, MA);
- elastase (0.18nM) was purchased from Elastin Products (Owensville, MO);
- FXIa (0,06 nM) and tissue plasminogen activator (t-PA) (0.06 nM) from Molecular Innovations (Southfield, MI);
- matriptase (0.1 nM) from R&D Systems (Minneapolis, MN);
- cathepsin G (6.7 nM) from Enzo Life Sciences (Plymouth Meeting, PA);
- sequencing grade trypsin (0.1 nM) was purchased from Roche (Chicago, IL).

The **assay buffers** were as follows:

1) elastase and chymase, 50 mM Hepes buffer pH 7.4, 100 mM NaCl, 0.01% Triton X-100;

2) trypsin, α-chymotrypsin, factor XIa, factor XIIa and thrombin, 50 mM Tris buffer pH 8.0, 150 mM NaCl, 20 mM $CaCl_2$, 0.01% Triton X-100;

3) β-tryptase, 50 mM Tris pH 8.0, 50 mM NaCl, 0.05% Triton X-100;

4) kallikrein and matriptase, 20 mM Tris buffer pH 8.5, 150 mM NaCl, 0.02% triton X-100;

5) factor Xa, 20 mM Tris buffer pH 8.0, 200 mM NaCl, 5 mM $CaCl_2$, 0.1%BSA;

6) t-PA, 20 mM Tris buffer pH 8.5, 0.05% Triton X-100;

7) cathepsin G, 50 mM Tris buffer pH 7.4, 150 mM NaCl, 0.01% Triton X-100;

8) plasmin, 50mM Tris pH 8, 150 mM NaCl, 0.01% Triton X-100.

The **substrates and sources** were as follows:

- Suc-Ala-Ala-Pro-Val-AMC for elastase; Boc-Asp-Pro-Arg-AMC for thrombin and plasmin; Boc-Gln-Ala-Arg-AMC for trypsin and factor XIa (Sigma, St. Louis, MO);
- Boc-Phe-Ser-Arg-AMC for β-tryptase; Suc-Leu-Leu-Val-Tyr-AMC for chymase (Bachem, King of Prussia, PA);
- Suc-Ala-Ala-Pro-Val-AMC for α-chymotrypsin and chymase (EMD Biosciences, La Jolla, CA);
- methylsulfonyl-D-cyclohexylalanyl-Gly-Arg-AMC acetate for factor Xa, factor XIIa, t-PA, matriptase and kallikrein (American Diagnostica Inc., Stamford, CT).