

Unique organization and unexpected expression patterns of *Anopheles gambiae* Vitellogenin genes.

Patricia A Romans, Department of Zoology, University of Toronto

Most insects have multiple *Vitellogenin* (*Vg*) genes of varying identity to each other dispersed in their genomes. *A. gambiae* has a variable number of *Vg* genes, all highly identical to each other and to *Aedes aegypti* *VgA1*. These consist of several tandem genes, whose numbers vary both within and between mosquito strains, and a single “dispersed” gene located about 20 kb 3’ to the tandem array on chromosome 2R, in division 18A of the polytene chromosome map. Intergene regions between the tandem genes are also highly identical, as would be expected when tandem repeat units are 9.3 kb in length. Although the dispersed gene coding region and its immediately flanking sequences are only distinguishable from tandem genes by a few nucleotide changes, sequences begin to differ between them immediately 5’ to the likely transcriptional start sites and within the 3’ untranslated regions. They are completely divergent 10 bp 5’ to the TATA boxes and 60 bp 3’ to the termination codons. Further, sequences 5’ and 3’ to the tandem array diverge completely from sequences between tandem genes within 250 bp of protein coding sequence.

Identification of 3’ end differences in *Vg* EST sequences permitted production of tandem and dispersed gene type-specific probes as well as PCR amplification of tandem genes and their transcripts separately from amplification of the dispersed gene and its transcripts. Southern blot and qPCR results yielded estimates of tandem gene numbers ranging from a low of 3 in the G3 strain to a high of 7 in PEST, the strain used for genome sequencing. qRT-PCR experiments showed that the dispersed *Vg* gene is transcribed at low levels even in non-blood fed females, as well as earlier and more abundantly than any tandem gene after blood feeding.

2.5 kb of sequence from the promoter region of a gene internal to the tandem array could not drive expression of a reporter gene in *Drosophila*, although an equivalent amount of the *Ae. aegypti* *VgA1* promoter does so quite successfully. Apparently, at least some of the tandem gene regulatory information is located peripheral to the cluster. The promoter region of the dispersed gene is being tested for its ability to drive expression of a transgene in *Ae. aegypti* in response to blood feeding and stimulation by 20-OH ecdysone.

Mosquito *Vg* gene transcription is induced by 20-OH ecdysone binding to a nuclear Ecdysone Receptor, and thereby to specific sequences in the *Vg* gene promoter region, at least in the mRNA accumulation phase during the first 24 hours post blood meal. After a peak in mRNA levels about this time and modest decreases until about 30 hr PBM, *Vg* mRNA abundance then decreases precipitously until it is barely detectable at 36 hr PBM. Virtually nothing is known about whether posttranscriptional processes assist in regulating *Vg* mRNA abundance and stability late in the gonotrophic cycle or, indeed, at any time PBM. The *Drosophila* EcR gene and many downstream target genes of the EcR are regulated in part by microRNAs which are highly conserved among species. Target sites of several of these miRNAs, including the EcR target of miR-14 are also conserved between *Drosophila* and *A. gambiae*. Because of results suggesting that regulation of *Vg* mRNA stability may be important both normally and in malaria-infected females, see below, we are investigating the potential of miRNA to regulate *Vg* transcript abundance.

Vg mRNAs show slightly reduced abundance by 24 hr post infection with the rodent malaria, *Plasmodium yoelii nigeriensis*. This reduction is more obvious at 36 hr PI, and is statistically significant throughout a second gonotrophic cycle, $P < 0.05$. Previously, parasite-induced fecundity reduction had been shown to affect reproductive physiology in the ovary, but not in the fat body. Both tissues begin to show pathology at the same time, a time coincident with the appearance of malaria ookinetes at the basal lamina of the midgut and contact with mosquito hemolymph. We are now examining whether the effects on fat body are direct effects

of parasite infection, or whether they are indirect, due to reduced 20-OH ecdysone production by ovaries. Specifically, we are looking to determine whether the reduction in *Vg* transcript abundance is due to interference with transcriptional regulation and/or mRNA stability.

Data from a recent digital Northern experiment suggest that *Vg* transcripts may be degraded from their 3' ends in females infected with late ookinetes and/or early oocysts of *P. berghei*. ESTs from the 5' ends of *Vg* mRNAs significantly outnumbered ESTs from the 3' ends, $P < 0.001$. This disparity between 5' and 3' ESTs was also seen in uninfected females at the same time post blood meal, but to a much less significant extent, $0.05 > P > 0.025$. While these results might be an artifact, it is difficult to envision how cDNA libraries constructed using oligo-dT to prime reverse transcription would over-represent anything but 3' ends. We plan to perform qRT-PCR on selected 5', mid-transcript and 3' regions in *Vg* mRNA in order to determine whether and to what extent post-transcriptional regulation controls *Vg* mRNA abundance at various times in the gonotrophic cycle in both normal and malaria-infected females.

These results were obtained with grant support from Insect Biotech Canada, the Natural Sciences and Engineering Research Council of Canada and the Rockefeller Foundation to PAR, from NIH to Frank Collins (Notre Dame), and from a studentship awarded by the government of Egypt to Ashraf Ahmed in Hilary Hurd's lab (Keele).