

## **AnoEST: *A. gambiae* component of VectorBase.**

Evgenia V Kriventseva, Evgeny M Zdobnov,  
George K. Christophides and Fotis C Kafatos

*EMBL, Meyerhofstr.1, 69117 Heidelberg, Germany,*

krivents@embl.de

I will present our analysis of 215,634 EST and cDNA sequences of a major vector of human malaria *Anopheles gambiae* structured into the AnoEST database. The expressed sequences are grouped into clusters using genomic sequence as template and associated with inferred functional annotation, including the following: corresponding Ensembl gene prediction, putative orthologous genes in other species, homology to known proteins, protein domains, associated Gene Ontology terms, and corresponding classification into broad GO-slim functional groups. AnoEST is a vital resource for interpretation of expression profiles derived using recently developed *A. gambiae* cDNA microarrays. Using these cDNA microarrays, we have experimentally confirmed the expression of 7961 clusters during mosquito development. Of these, 3100 are not associated with currently predicted genes. Moreover, we found that clusters with confirmed expression are nonbiased with respect to the current gene annotation or homology to known proteins. Consequently, we expect that many as yet unconfirmed clusters are likely to be actual *A. gambiae* genes. AnoEST is publicly available at <http://komar.embl.de>, and is also accessible as a Distributed Annotation Service (DAS). AnoEST constitutes one of the functional genomics modules of a new integrated genomic data resource for multiple vectors of disease, VectorBase (<http://www.vectorbase.org>).

1. AnoEST: Toward *A. gambiae* functional genomics.  
Genome Res. 2005 May 17  
<http://www.genome.org/cgi/doi/10.1101/gr.3756405>