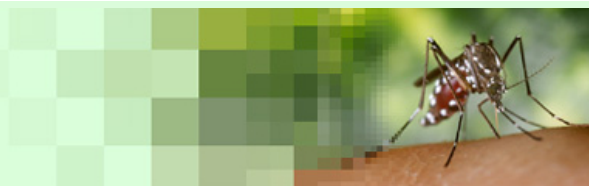


The VectorBase Newsletter

Issue 11 - Spring 2011



News

Funding Opportunity

The NIAID-funded VectorBase invites research investigations called Driving Biological Projects (DBPs) from organizations housing experimental laboratories having expertise and knowledge in working with arthropod vectors of human pathogens. We anticipate making two awards of up to \$600K each by April 1, 2012. A two page Letter of Intent (LOI) is due August 1, 2011. Full proposals will be solicited by invitation only, with a deadline of November 1, 2011. See the full solicitation from the VectorBase site¹ and send questions to vector@nd.edu.

Anopheles Genomes Cluster

Samples are now in sequencing at the Broad Institute for the NHGRI/NIAID-supported project, "Genome Analysis of Vectorial Capacity in Major Anopheles Vectors of Malaria Parasites". VectorBase looks forward to working together with the Broad and members of the vector community to process and host these new genomes and their meta-data. Please visit VectorBase often for project information, updates, and data as they become available.

Genomes and Data

Rhodnius prolixus

An early pre-release of the *Rhodnius prolixus* genome is available. This annotation includes: repeat masking, ESTs and RNA-seq transcriptome alignments, UniProt peptide alignments, GenScan based gene models and an transcriptome derived geneset.



Kissing bug

Further data sets will be added in the coming months with the expected pre-release of the first genebuild for the June release.

¹<http://www.vectorbase.org/Other/News/?id=140>

Glossina morsitans

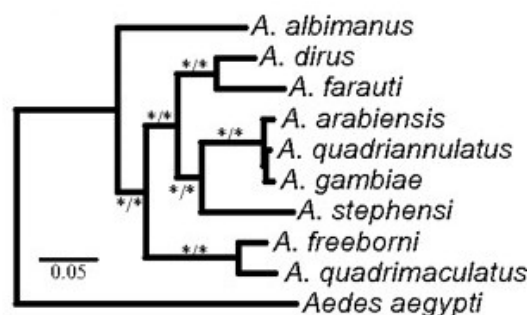
The genome of the Tsetse fly, *Glossina morsitans morsitans*, has been sequenced and assembled by the Wellcome Trust Sanger Institute. A pre-release of the assembly will be available in the June release. Like the *Rhodnius* pre-release this will contain alignments of transcriptome and peptide data and *ab initio* gene predictions.



Tsetse fly

Next-generation RNA-seq

VectorBase now contains next-generation transcriptome sequence data sets (RNA-seq). Over the past few releases we have added data from 9 Anopheline species available for querying via the BLAST server and as data tracks served to the *Anopheles gambiae* browser via DAS. Further, the genome browser is now capable of displaying community members' own indexed BAM files for all species.



As more data sets become available VectorBase will continue to make these available to the community. A major update for *Aedes aegypti* is due for the June release including 1.3 M Roche 454 reads and 41 M Illumina reads from the Short Read Archive (SRA). We encourage our communities to contact us early to discuss visualization of newer generation sequencing data sets through VectorBase including resequencing data.

Upcoming Outreach

Kansas City, USA, June 9 - 12, 2011
<http://www.k-state.edu/agc/symp2011/>

VectorBase staff will be attending a number of meetings in the coming months giving workshops, talks or posters. If you are attending any of the following look out for us and come chat about vector genomics.

Journées Ouvertes en Biologie, Informatique et Mathématiques (JOBIM 2011)
Paris, France, June 28 - July 1, 2011
<http://www.pasteur.fr/JOBIM2011>

7th Annual BioMalPar Conference, Biology and Pathology of the Malaria Parasite
EMBL Heidelberg, Germany 16-18th May 2011
<http://embl.de/training/events/2011/BMP11-01>

Fifth International Meeting on "Molecular and Population Biology of Mosquitoes and Other Disease Vectors"
Kolymbari, Crete, Greece, 24-30 July 2011
http://www.anobase.org/conferences/home_2011.html

5th Annual Arthropod Genomics Symposium Arthropod Genomics 2011: Exploring Diversity, Relating Similarity

Social Networks

We're still on Twitter² and Facebook³!

The VectorBase Editors.
newsletter@vectorbase.org

VectorBase Top Tip – Variant Effect Predictor

Did you know you can predict the effect of variants within the vector genomes? A variant effect predictor is available through the 'Manage your data' button found in the left hand navigation menu of the genome browser. Users can upload a text file containing the variation data in a tab-separated format (chromosome, start, end, allele, strand) and the predictor will return the consequence of this variation with information about the affected gene and transcript, any amino acid changes and whether this variation is currently in the database. Full documentation relating to both input and output file formats can be found at <http://www.vectorbase.org/info/website/upload/var.html>.

The screenshot shows the 'Variant Effect Predictor' tool interface. On the left is a navigation menu with 'Variant Effect Predictor' selected. The main content area has a title 'Variant Effect Predictor:' followed by a description: '(Formerly SNP Effect Predictor). This tool takes a list of variant positions and alleles, and predicts the effects of each of these on any overlapping features (transcripts, regulatory features) annotated in Ensembl. The tool will accept substitutions, insertions and deletions as input, uploaded as a list of tab separated values - [more about required format](#). There is a limit of 750 variations that can be processed at any one time. You can upload a file that contains more entries, however anything after the 750 line will be ignored. If your file contains more than 750 variations you can split your file into smaller chunks and process them one at a time, or you may wish to use the [variation API](#) or a standalone [perl script](#) which you can run on your own machine to generate the same results as this web tool.'

The interface includes the following fields and controls:

- Upload file:** A section with a 'Species:' dropdown menu set to 'Anopheles gambiae (Anopheles gambiae)', a 'Name for this upload (optional):' text input, and a 'Paste file:' text area.
- Upload file:** A 'Choose File' button next to the text 'No file chosen'.
- or provide file URL:** A text input field.
- Next >** A green button at the bottom.

²<http://twitter.com/vectorbase>

³<http://facebook.com/vectorbase>