A NEW WEB-BASED GENOMICS RESOURCE FOR BIOINFORMATICS ANALYSIS OF *RHIPICEPHALUS (BOOPHILUS) MICROPLUS*: CATTLETICKBASE.

Corresponding Author: Ala Lew-Tabor, Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Queensland, Brisbane, Australia. E-Mail: a.lewtabor@uq.edu.au

Authors: Bellgard, Matthew; Moolhuijzen, Paula; Guerrero, Felix; Appels, Rudi; Schibeci, David; Rodriguez Valle, Manuel; Barrero, Roberto; Hunter, Adam; Lew-Tabor, Ala

Keywords

Genomics, tick, bioinformatics

Introduction

A new web-based genomics resource for bioinformatics analysis of *Rhipicephalus (Boophilus) microplus*: CattleTickBase

Bellgard, M.I.[1,2], Moolhuijzen, P.M.[1,2], Guerrero, F.D.[1,3], Appels, R.[2], Schibeci, D.[2], Rodriguez Valle, M.[1,4], Barrero, R.A.[2], Hunter, A.[2], *Lew-Tabor, A.E.[1,2,4]

[1]Cooperative Research Centre for Beef Genetic Technologies, Armidale, NSW, Australia; [2]Centre for Comparative Genomics, Murdoch University, Perth, Western Australia 6150, Australia; [3]USDA-ARS Knipling Bushland US Livestock Insect Research Laboratory, 2700 Fredericksburg Rd., Kerrville, TX 78028, USA; [4]Queensland Alliance for Agriculture & Food Innovation, The University of Queensland, Qld, 4072, Australia.

The *Rhipicephalus (Boophilus) microplus* genome is large and complex in structure (~7.1Gb consisting of over 70% repetitive DNA), making a genome sequence difficult to assemble and costly to resource the required bioinformatics. In light of this, a consortium of international collaborators was formed to pool resources to begin sequencing this genome. We have sequenced, assembled and annotated over 2 Gb of sequence comprised from Cot-filtered genomic DNA, ~10,000 BAC-ends, 15 targeted whole BACs ~1.5 Mb (Guerrero et al. 2010, Bellgard, unpublished), ~21 Mb from eight whole transcriptome library assemblies representing over 75,000 transcripts (Wang et al. 2007; Barrero et al., unpublished) and microRNAs (Barrero et al. 2011). The Cot-filtered genomic DNA encodes 144 Mbp of assembled contigs derived from three Cot re-association experiments that utilised methodologies to select sequences that are depleted in repetitive elements, from randomly sheared genomic DNA fractions, to preferentially enrich putative gene coding regions (Guerrero et al 2010; Bellgard et al. unpublished). Our consortium has acquired genome and transcriptome sequence data at approximately 0.9x coverage of the gene-coding regions of the *R. microplus* genome. We also have integrated several gene expression datasets derived from selected tick life stages and key organs (Rodriguez Valle et al. 2010, Barrero et al. 2011). A web-based resource was developed to enable the tick research scientific community to access our datasets and conduct analysis through a single web-based bioinformatics environment called YABI. The collective bioinformatics resource is termed 'CattleTickBase' and also includes access to public databases in order to undertake *R. microplus* sequence analysis against all available resources within a single integrated resource. The YABI tool at CattleTickBase will facilitate access and manipulation of cattle tick genome sequence data as the genome sequencing of *R. microplus* progresses.