

To B or not to B: *Arsenophonus* as a source of B-vitamins in whiteflies.

Diego Santos-Garcia^{1,†,*}, Ksenia Juravel¹, Shiri Freilich², Einat Zchori-Fein³, Amparo Latorre^{4,5}, Andrés Moya^{4,5}, Shai Morin¹ and Francisco J. Silva^{4,5,†,*}

¹ Department of Entomology, The Hebrew University of Jerusalem, Rehovot, Israel

² Institute of Plant Sciences, Newe-Ya'ar Research Center, ARO, Ramat-Yishai, Israel

³ Department of Entomology, Newe-Ya'ar Research Center, ARO, Ramat-Yishai, Israel

⁴ Institute for Integrative Systems Biology (I2SysBio), Universitat de València-CSIC, Spain

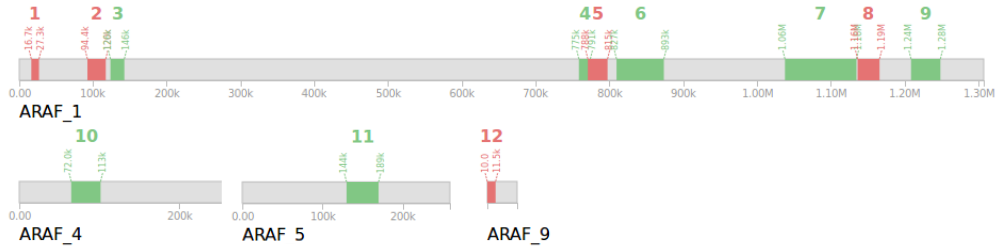
⁵ Unidad Mixta de Investigación en Genómica y Salud (FISABIO/Universitat de València-I2SysBio), València, Spain

† Equal contribution.

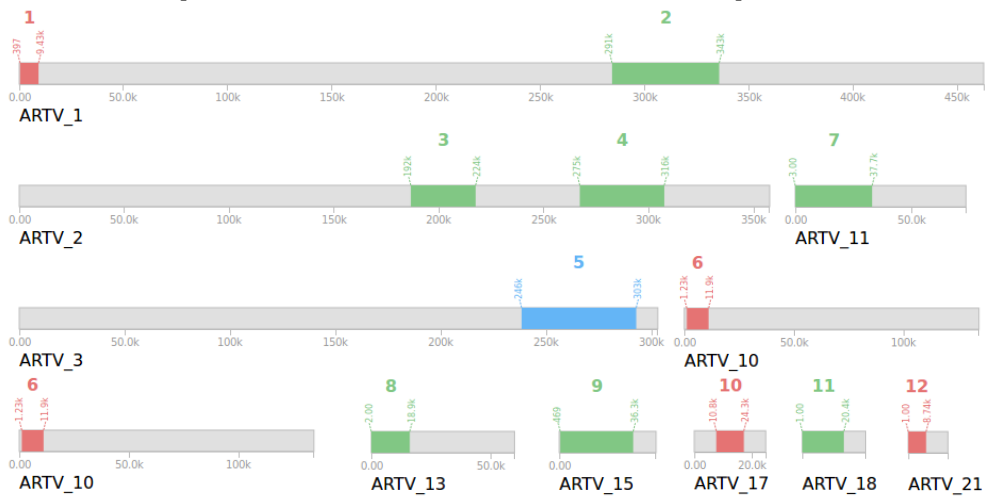
* For correspondence: diego.santos@mail.huji.ac.il and francisco.silva@uv.es

1 Supplementary Figures

A) *Arsenophonus* sp. of *Aleurodicus floccissimus*



B) *Arsenophonus* sp. of *Trialeurodes vaporariorum*



C) *Wolbachia* sp. of *Aleurodicus dispersus*

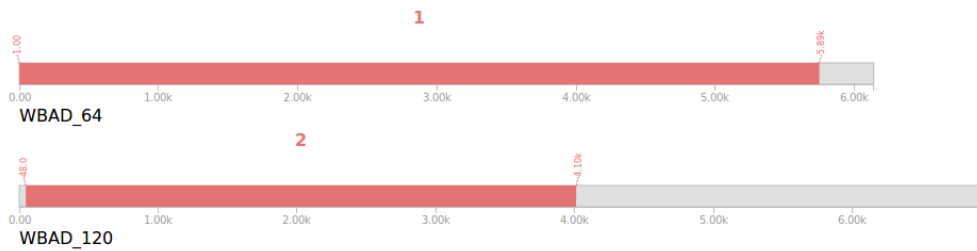


Figure 1: Prophages regions detected in *Arsenophonus* A) ARAF, B) ARTV and C) *Wolbachia* WBAD. Red bars represents incomplete prophages, green bars complete prophages and blue bars questionable complete prophages.

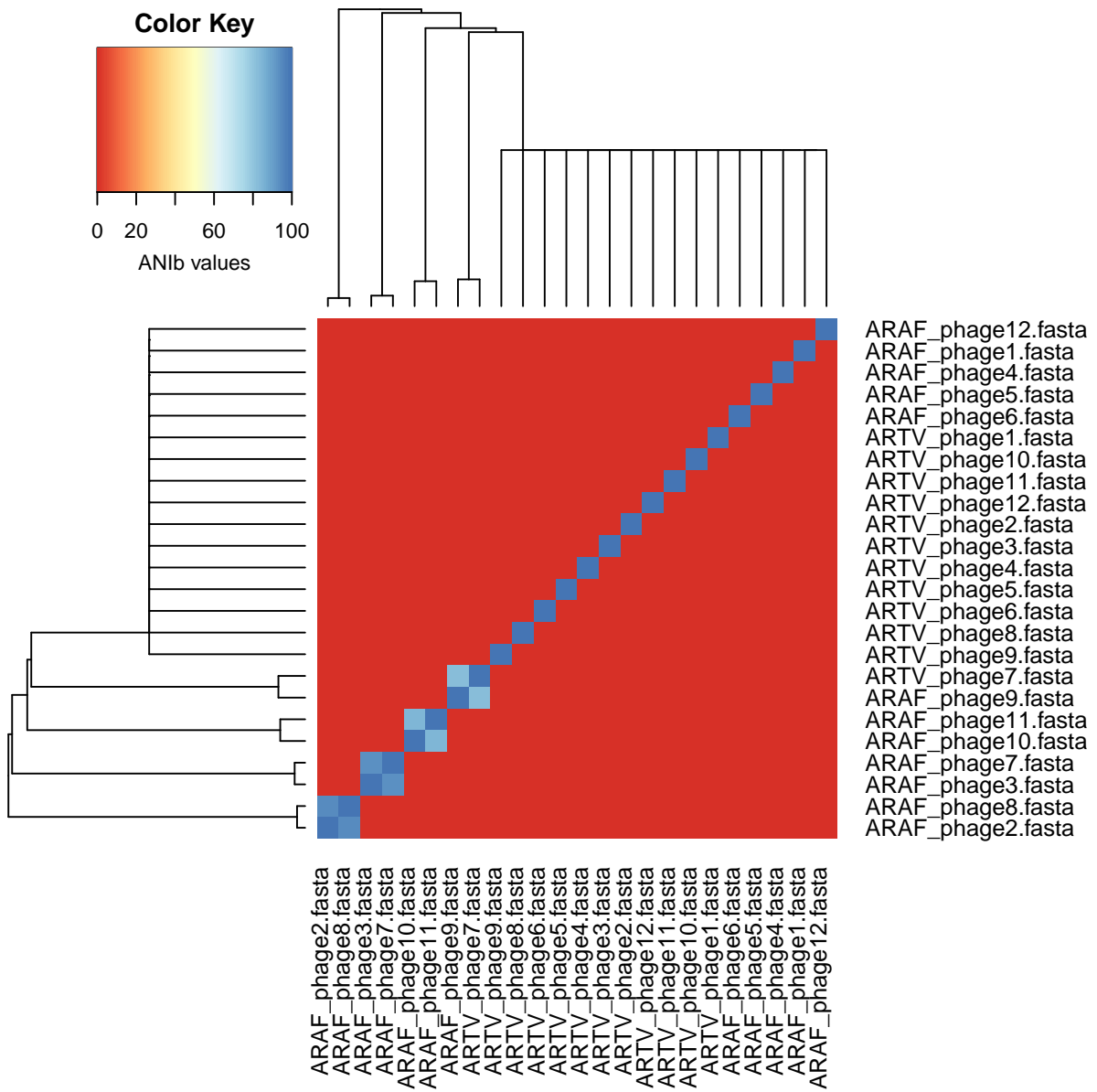


Figure 2: Hierarchical clustering of pairwise Average Nucleotide Identity (ANI) between the *Arsenophonus* ARAF and ARTV detected prophages.

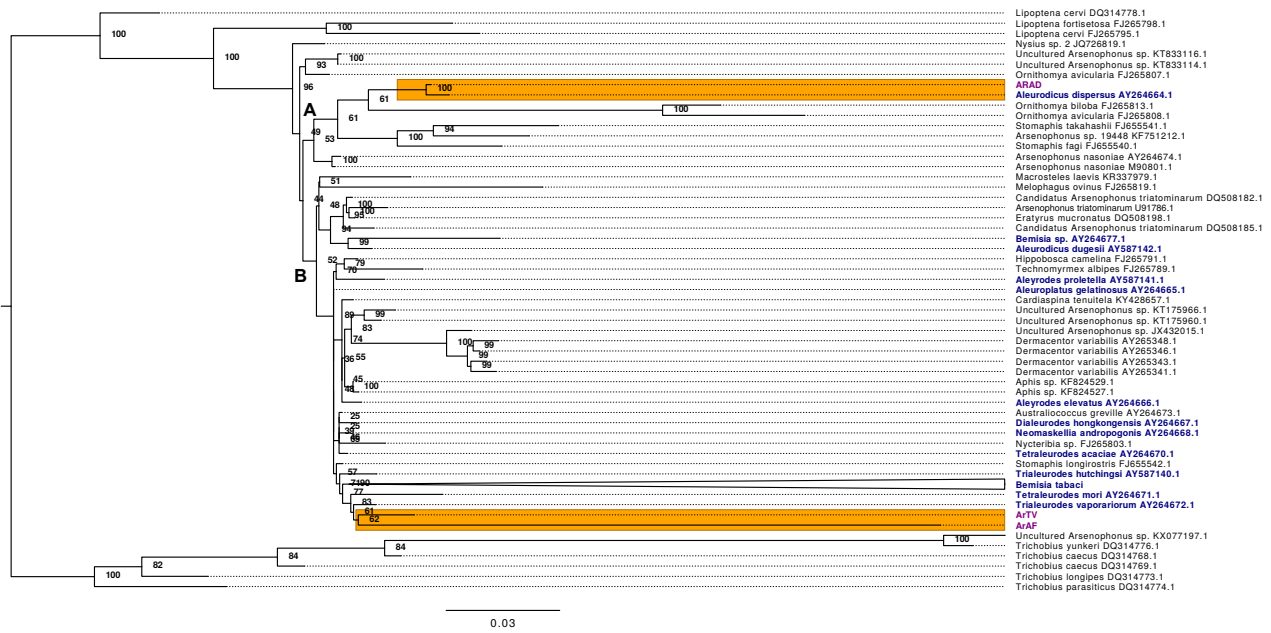


Figure 3: Midpoint rooted maximum likelihood tree obtained for several *16S rRNA* genes from *Arsenophonus* infecting different hosts. The tree was inferred using a TVM+F+R3 substitution model and 1000 ultrafast bootstraps (node labels). *Arsenophonus* from *Aleurodicus dispersus* (ARAD), *A. floccissimus* (ARAF) and *Trialeurodes vaporariorum* (ARTV) are highlighted in purple. *Arsenophonus* from different whiteflies are highlighted in blue. Two potential clades are denoted with letters A and B

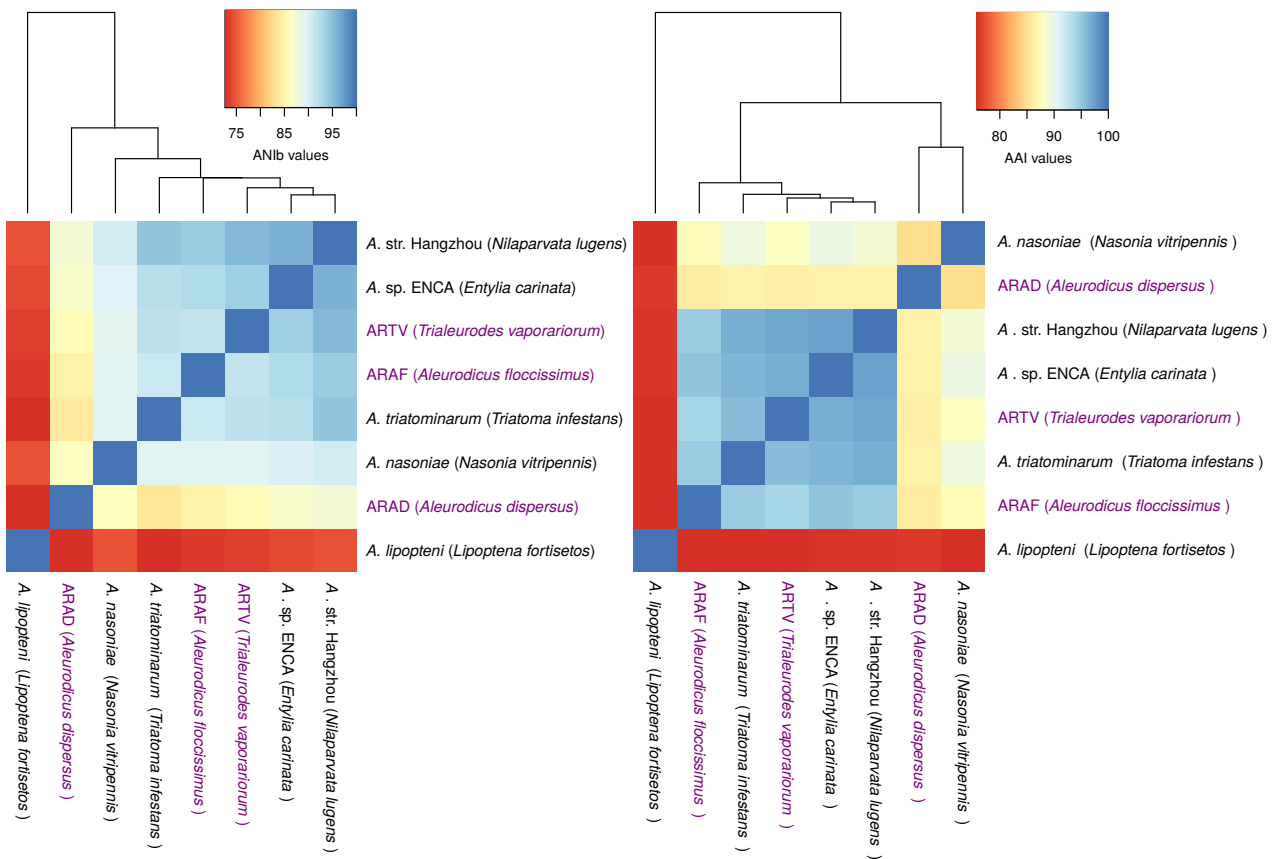


Figure 4: Hierarchical clustering of pairwise Average Nucleotide Identity (ANI, left) and Average Amino Acid Identity (AAI, right) of several *Arsenophonus* genomes. *Arsenophonus* from *Aleurodicus dispersus* (ARAD), *A. floccissimus* (ARAF) and *Trialeurodes vaporariorum* (ARTV) are highlighted in purple. Names for the eukaryotic hosts are shown after the strain names inside parentheses.

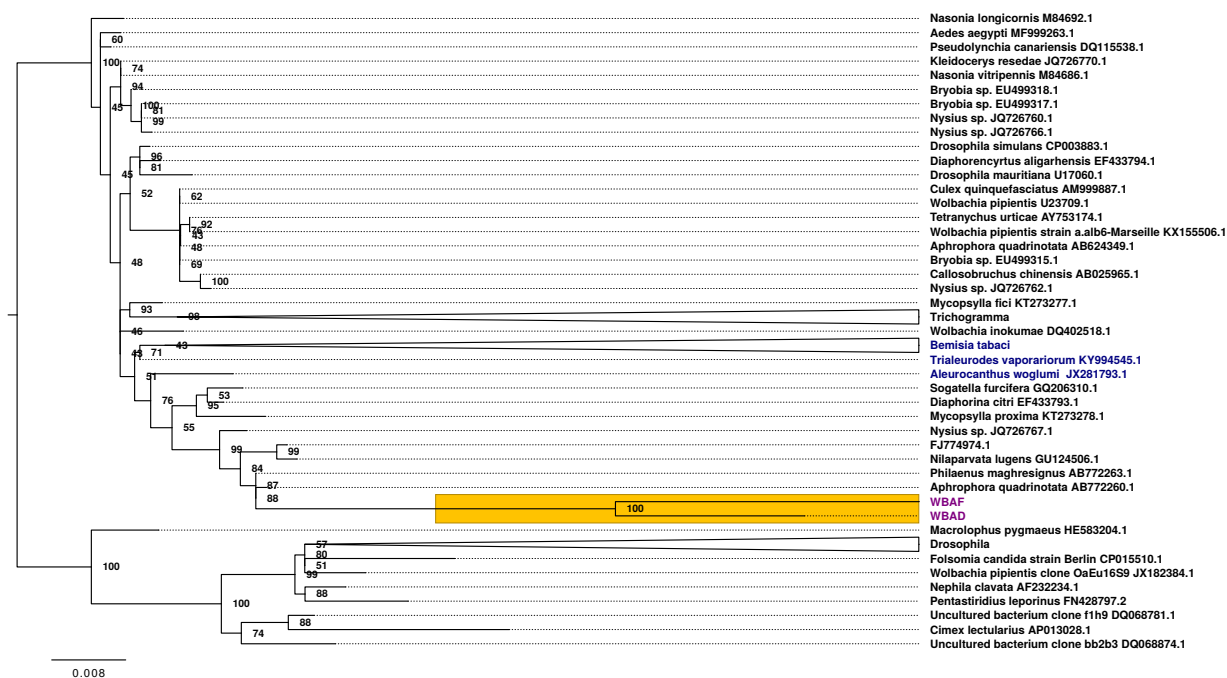


Figure 5: Midpoint rooted maximum likelihood tree obtained for several *16S rRNA* genes from *Wolbachia* infecting different hosts. The tree was inferred using a TN+F+R2 substitution model and 1000 ultrafast bootstraps (node labels). *Wolbachia* from *Aleurodicus dispersus* (WBAD) and *A. floccissicums* (WBAF) are highlighted in purple. *Wolbachia* from different whiteflies are highlighted in blue.