

A data-driven approach to construct a molecular map of *Trypanosoma cruzi* to identify drug and vaccine targets

Swarsat Kaushik Nath^{1#}, Preeti P.¹, Trapti Sharma¹, Priya Kumari¹, Sweety Shinde¹, Nikita Garg¹, Kartavya Mathur¹, Nevidita Arambam¹, Divyank Harjani¹, Manpriya Raj, Garwit Kwatra¹, Sayantan Venkatesh¹, Alakto Choudhoury¹, Saima Bano¹, Prashansa Tayal¹, Mahek Sharan¹, Ruchika Arora¹, Ulrich Strych², Peter J. Hotez^{2,3}, Maria Elena Bottazzi^{2,3} and Kamal Rawal^{1*#}

1 Amity Institute of Biotechnology, Amity University Uttar Pradesh, India.

2 Texas Children's Hospital Center for Vaccine Development, Departments of Pediatrics and Molecular Virology and Microbiology, National School of Tropical Medicine, Baylor College of Medicine, Houston, TX, USA.

3 Department of Biology, Baylor University, Waco, Texas, USA.

*Corresponding Author:

Dr. Kamal Rawal

Centre for Computational Biology and Bioinformatics, AIB
Amity University, Noida, Uttar Pradesh, India.

Email address: kamal.rawal@gmail.com

#These authors have contributed equally.

Website: <https://tinyurl.com/Tcruzipathwaymapx>

S. No.	Supplementary Figure no.	Description	Link
1.	Supplementary Figure 1	Supplementary Figure 1: The schematic diagram describing the flowchart to achieve the methodology.	https://tinyurl.com/supfig1
2.	Supplementary Figure 2	Supplementary Figure 2: The standard notations for the graphical representation of species and edges in cell designer v.4.4.2.	https://tinyurl.com/supfig2
3.	Supplementary Figure 3	Supplementary Figure 3: Docking of Tc24 protein with Benznidazole and Nifurtimox using AutoDock Vina. (a. Benznidazole, b. Nifurtimox). The top scoring pose of docking affinity against Tc24 for Benznidazole is -6.9 kcal/mol while the docking affinity for Nifurtimox is -6.7 kcal/mol.	https://tinyurl.com/supfig3
4.	Supplementary Figure 4	<p>Supplementary Figure 4A: <i>Trypanosoma cruzi</i> CL Brener Esmeraldo-like. Enriched GO terms of genes (a. Biological process, b. Cellular component, c. Molecular function). The size of the dots and their color gradient (according to the legend) reflects the log₁₀ value of the p value of each GO term, larger dots represent higher significance (all p values <0.05). X and Y axes represent semantic spaces which have no intrinsic meaning. REVIGO uses multi-dimensional scaling to reduce the dimensionality of a matrix of the GO terms' pairwise semantic similarities. This results in semantically similar GO terms remaining close together in the plot.</p> <p>Supplementary Figure 4B: <i>Trypanosoma cruzi</i> CL Brener Non-Esmeraldo-like. Enriched GO terms of genes (a. Biological process, b. Cellular component, c. Molecular function). The size of the dots and their color gradient (according to the legend) reflects the log₁₀ value of the p value of each GO term, larger dots represent higher significance (all p values <0.05). X and Y axes represent semantic spaces which have no intrinsic meaning. REVIGO uses multi-dimensional scaling to reduce the dimensionality of a matrix of the GO terms' pairwise semantic similarities. This results in semantically similar GO terms remaining close together in the plot.</p>	https://tinyurl.com/supfig4
5.	Supplementary Figure 5	Supplementary Figure 5A: <i>Trypanosoma cruzi</i> CL Brener Non-Esmeraldo-like. "Interactive graph" view of REVIGO (a. Biological process, b. Cellular component, c. Molecular function). Bubble color indicates the user-provided p-value. Highly similar GO terms are linked by	https://tinyurl.com/supfig5

		edges in the graph, where the line width indicates the degree of similarity.	
		Supplementary Figure 5B: <i>Trypanosoma cruzi</i> CL Brener Esmeraldo-like. "Interactive graph" view of REVIGO (a. Biological process, b. Cellular component, c. Molecular function). Bubble color indicates the user-provided p-value. Highly similar GO terms are linked by edges in the graph, where the line width indicates the degree of similarity.	
6.	Supplementary Figure 6	Supplementary Figure 6: Gene ontology (GO) classification, KEGG/MetaCyc analysis of the <i>T. cruzi</i> genes (a.) <i>T. cruzi</i> CL Brener Esmeraldo like (b.) <i>T. cruzi</i> CL Brener Non-Esmeraldo like). The percentage regulated genes based on GO annotations and KEGG/MetaCyc pathways. (c) word cloud graph depicting enriched Metabolic Pathway (1) <i>Trypanosoma cruzi</i> CL Brener Esmeraldo-like (2) <i>Trypanosoma cruzi</i> CL Brener Non-Esmeraldo-like.	https://tinyurl.com/supfig6

* % regulated genes are the percentage of the background or GO/pathway-related genes.