

# XI CONGRESO LATINOAMERICANO DE BIOLOGÍA MATEMÁTICA SoLaBiMa 2019

Universidad Católica de Maule

22-25 de Octubre de 2019, Talca, Chile



## Análisis inmunoinformático y evolutivo de proteínas de *Trypanosoma cruzi* para el diseño de una vacuna de subunidades multiepítope

**Maria Virginia Ramírez\***

Universidad de Los Andes, Mérida, Venezuela

### Resumen

La enfermedad de Chagas afecta a millones de personas y representa una de las infecciones parasitaria más importante de América Latina. Desde su descubrimiento, numerosas investigaciones se han enfocado en el entendimiento de la biología de su agente causal, *Trypanosoma cruzi*, así como también de la respuesta inmunológica que se desarrolla en el huésped mamífero producto de la infección. Los avances obtenidos han permitido identificar proteínas con potencial antigénico que podrían incorporarse en el diseño de una vacuna preventiva.

A partir de este conocimiento, esta investigación se plantea analizar éstas y otras proteínas de *T. cruzi* menos estudiadas, a fin de localizar regiones clasificadas como epítopes mediante herramientas inmunoinformáticas. Los epítopes a su vez estarán sujetos a pruebas evolutivas para determinar el tipo de presiones selectivas operando en estas regiones.

Finalmente, el diseño de un péptido quimérico permitirá la incorporación de estos epítopes, que en conjunto con un adyuvante adecuado, representaría un prototipo inmunestimulante con el potencial de ser utilizado como una vacuna de subunidades preventiva contra la enfermedad de Chagas.

Equipo de trabajo:

**Ascanio Rojas**<sup>1</sup>, Centro Nacional de Cálculo Científico de la Universidad de Los Andes, CeCalCULA, Mérida, Venezuela.

**Héctor Acosta**<sup>2</sup>, Laboratorio de Enzimología de Parásitos, LEP, Universidad de Los Andes, Mérida, Venezuela.



\*e-mail: mvirginiarm9@gmail.com

<sup>1</sup>e-mail: ascanio@ula.ve

<sup>2</sup>e-mail: hectoracosta@ula.ve

## Planteamiento

La enfermedad de Chagas, o tripanosomiasis americana, es una zoonosis causada por el protozooario flagelado *Trypanosoma cruzi*, reconocida por la Organización Mundial de la Salud como una de las 16 enfermedades tropicales olvidadas que afecta de 8 a 10 millones de personas en mundo, principalmente en América Latina donde la enfermedad es endémica.<sup>1,2</sup> Se estima que ocurren de 10.000 a 14.000 muertes anuales, y que entre 65 y 100 millones de personas están en riesgo de contraerla. A pesar del tiempo transcurrido desde su descubrimiento (1909)<sup>3</sup>, aún no se dispone de una terapia eficaz para todas las fases clínicas de la enfermedad de Chagas y tampoco existen vacunas para prevenirla.<sup>4</sup>

Su agente causal, *T. cruzi*, es un parásito hemoflagelado perteneciente al filo Euglenozoa, orden Kinetoplastida, y familia Trypanosomatidae.<sup>5</sup> Éste presenta un ciclo de vida digenético, en el que desarrolla distintos estadios (epimastigote, tripomastigote metacíclico, amastigote y tripomastigote sanguíneo) que varían de acuerdo a los hábitats que ocupa dentro del vector, triatomíneos pertenecientes a la familia Reduviidae y su huésped mamífero.<sup>6</sup> Además de la vía vectorial, la enfermedad de Chagas puede transmitirse de manera congénita, por transfusión sanguínea, mediante el trasplante de órganos de personas infectadas y por el consumo de alimentos contaminados con heces de un triatomíneo infectado.<sup>7</sup> Una vez ocurrida la exposición, la enfermedad presenta una fase aguda y una fase crónica.<sup>8</sup> La mayoría de personas infectadas en fase crónica (60-70% de los casos) son asintomáticas, no obstante, en el resto de los casos, los individuos manifiestan un cuadro clínico severo caracterizado por afecciones cardíacas, digestivas, o ambas simultáneamente.<sup>6</sup>

En cuanto a los mecanismos de prevención de la enfermedad, es el control vectorial el único existente; que si bien ha conseguido una reducción importante de la incidencia, no ha sido suficiente para prevenir nuevas infecciones, por lo que la permanencia de la enfermedad se sigue considerando una amenaza para la salud pública.<sup>9,10</sup> Por esta razón, y debido a la ausencia de un tratamiento 100% eficaz e inocuo, es necesario el desarrollo de una estrategia preventiva que

conlleve a un mejor control. Por lo que la vacunación representa un mecanismo clave para suplir esta necesidad, y en este contexto la inmunoinformática representa una importante herramienta.

La inmunoinformática es una rama de la biología computacional dedicada a la resolución de problemas inmunológicos, que se ha enfocado principalmente en el desarrollo de predictores de epítopes utilizados en el diseño de vacunas, el inmunodiagnóstico, y en la inmunoterapia personalizada contra enfermedades como el cáncer.<sup>11</sup> El avance de estas técnicas ha sido catalizado en gran medida por la necesidad de manejar la creciente cantidad de datos de interés inmunológico, lo que ha incentivado la creación de técnicas cada vez más sofisticadas y precisas que al mismo tiempo mejoran con la disponibilidad de más datos.<sup>11,12</sup> La aplicación de estas metodologías representa una ventaja importante ya que permite reducir el tiempo, los recursos y costos requeridos para la identificación experimental de epítopes mediante la disminución del repertorio de péptidos a evaluar. En vacunología, la inmunoinformática permite la predicción de epítopes a partir de datos genómicos, lo que propicia el diseño de vacunas basadas en epítopes, objetivo principal de esta investigación. Este enfoque sostiene que la creación de una proteína quimérica que contenga las regiones antigénicas representa un diseño racional capaz de desencadenar respuestas inmunitarias humorales y celulares adecuadas contra determinado patógeno.<sup>11-13</sup>

Actualmente existen diversos métodos computacionales enfocados en la identificación de regiones proteicas con el potencial de actuar como epítopes lineales y conformacionales reconocidos por anticuerpos y receptores de células B.<sup>11</sup> Desarrollo que comenzó en el siglo pasado con el surgimiento de predictores basados en la categorización de aminoácidos de acuerdo a sus características fisicoquímicas, otorgándoles valores de propensión para actuar como epítopes. Métodos que obtuvieron resultados modestos de predicción y que llevaron a la creación de nuevos predictores con mejor desempeño.<sup>11,14</sup> Así surgieron los métodos modernos basados en aprendizaje automatizado (*machine learning*), que varían de acuerdo a los datos utilizados para entrenar los algoritmos, y que ha sido utilizadas en varias investigaciones de enfermedades infecciosas, enfocadas tanto en la inmunización como el serodiagnóstico.<sup>11,14,15</sup> Por otro lado, basado en el

conocimiento de las interacciones fisicoquímicas que se presentan entre péptidos y moléculas MHC de clase I y II se han desarrollado programas enfocados en la predicción de péptidos de unión a estas.<sup>11</sup> Como en el caso anterior, el desarrollo de estos predictores comenzó el siglo pasado a partir de datos que surgieron experimentalmente sobre péptidos conocidos de unión a HLAs. Así, las predicciones se basaron en un primer momento en la detección de secuencias motivo encontradas en los péptidos, que se caracterizaban por contener residuos de anclaje.<sup>14,16</sup> Posteriormente, para considerar la contribución de los residuos totales, se desarrollaron nuevos métodos denominados matrices de motivos, que fueron superadas por las matrices de afinidad cuantitativas, que consideraban valores de afinidad de enlace. Seguidamente se desarrollaron nuevas metodologías que originaron los modelos de relaciones de actividad estructural cuantitativa, que tomaban en cuenta las interacciones entre cadenas laterales vecinas en la contribución individual de cada residuo.<sup>14,16</sup> Estos avances son actualmente superados por los modelos basados en aprendizaje automatizado, los cuales presentan algoritmos fundamentados en redes neuronales artificiales, máquinas de soporte vectorial (*support vector machines*), árboles de decisión o cadenas ocultas de Markov; que han aumentado en gran medida la capacidad de detección con precisiones que alcanzan el 95%.<sup>11,14,16</sup>

Ahora bien, en adición al análisis inmunoinformático, esta investigación contemplará un enfoque evolutivo destinado a la determinación de las presiones selectivas actuando en las zonas de reconocimiento inmune de *T. cruzi*. En relación, es sabido que la coevolución huésped-patógeno se caracteriza por los cambios adaptativos recíprocos que se derivan de la interacción de las especies, donde la presión inmune del huésped y la evasión inmune del parásito representan los mecanismos claves de un proceso conocido como “carrera armamentista evolutiva”.<sup>17</sup> Dependiendo de los organismos patógenos, se ha encontrado que los genes que codifican antígenos tienden a ser muy variables como consecuencia de una selección diversificadora que conlleva a la evasión inmune. No obstante, existen organismos como *Mycobacterium tuberculosis* en donde se ha encontrado evidencias de una “hiper-conservación” de estos genes.<sup>17</sup> En el diseño racional de vacunas, el

estudio de la variabilidad de las proteínas portadoras de epítopes es fundamental, a fin de evitar la escogencia de aquellos cuya variabilidad les permita escapar del reconocimiento inmunitario.<sup>18</sup> Razón por la cual las proteínas seleccionadas en el parásito serán evaluadas mediante la búsqueda del tipo de presiones selectivas presente en los epítopes.

Por lo anterior expuesto, tomando en cuenta las ventajas de utilizar vacunas de subunidad con adyuvantes de nueva generación y dada la utilidad de los análisis inmunoinformáticos preliminares y de la integración del enfoque evolutivo expuesto, la presente investigación tiene como finalidad la escogencia de diferentes proteínas de *T. cruzi* adecuadas para el diseño de una vacuna de subunidades multiepítope.

### **Diseño metodológico**

La presente es una investigación en curso, por lo que la selección de los métodos más adecuados está actualmente en proceso, no obstante el planteamiento es el siguiente:

#### *Evaluación de predictores de epítopes de células B*

Se utilizarán epítopes de células B de *T. cruzi* definidos experimentalmente disponibles en la base de datos del IEDB (*Immune Epitope Database*)<sup>19</sup> y se utilizará la secuencia de los antígenos para evaluar los siguientes predictores: el índice de hidrofilia de Parker<sup>20</sup>, el método de flexibilidad de Karplus<sup>21</sup>, el predictor de accesibilidad de Emini<sup>22</sup>, el método de incidencia de giros  $\beta$  de Pellequer<sup>23</sup>, el predictor de propensión antigénica de Kolaskar<sup>24</sup>, el método de exposición a la superficie de Janin<sup>25</sup> y el predictor de polaridad de Ponnuswamy<sup>26</sup>, BepiPred 2.0<sup>27</sup> y ABCPred<sup>28</sup>. Se determinará un porcentaje de cobertura por antígeno por predictor de acuerdo a la cantidad de epítopes de la base de datos predichos.

#### *Evaluación de predictores de epítopes de unión a HLAs*

Como en el caso anterior se emplearán epítopes de unión a HLAs determinados experimentalmente presentes en IEDB y se utilizará su secuencia para evaluar el desempeño de los siguientes predictores: NetMHCpan-4.0<sup>29</sup>, MHCflurry<sup>30</sup>, NetMHC-4.0<sup>31</sup>, SMM<sup>32</sup>, IEDB Consensus<sup>33</sup>, Rankpep<sup>34</sup>, NetMHCpan-3.0<sup>35</sup>, SMMPMBEC<sup>36</sup>, SYFPEITHI<sup>37</sup>, BIMAS<sup>38</sup>, PREDEP<sup>39</sup>

y PAComplex<sup>40</sup>. Igualmente, el desempeño se medirá comparando la capacidad de predicción de los epítopes de la base de datos en los antígenos seleccionados.

#### *Pruebas de selección*

Se utilizará varios algoritmos para evaluar presiones selectivas positivas y negativas en los epítopes de células B y T de las bases de datos empleadas en la evaluación de predictores. Para ello se usarán el método de máxima verosimilitud de probabilidad de efectos fijos FEL (*Fixed Effects Likelihood*)<sup>41</sup> para la detección de selección negativa, y el modelo de efectos mixtos de evolución, MEME (*Mixed Effects Model of Evolution*)<sup>42</sup> para la determinación de selección positiva.

#### *Elección de proteínas*

Se escogerá una serie de proteínas evaluadas en investigaciones experimentales tanto *in vivo* como *in silico* catalogadas por su potencial para generar respuesta inmunitaria. Sus respectivas secuencias nucleotídicas y aminoacídicas serán localizadas en bases de datos especializadas como la del Centro Nacional para la Información Biotecnológica NCBI (*National Center for Biotechnology Information*)<sup>43</sup> y TriTrypDB<sup>44</sup>. En cada caso se acotará la unidad de tipificación y cepa correspondiente.

#### *Análisis inmunoinformático y evolutivo de las proteínas de interés*

Los predictores seleccionados en las etapas de evaluación serán utilizados para elegir los epítopes a incluir en la vacuna de subunidades, los que a su vez serán analizados mediante FEL<sup>41</sup> y MEME<sup>42</sup> para la búsqueda de presiones selectivas, y corroborados por pruebas bioquímicas y genómicas *in silico*.

#### *Diseño del prototipo inmunoestimulante*

Una vez determinados los péptidos que actuarían como blancos inmunológicos que cumplan con las premisas impuestas en nuestro estudio, y validadas por datos bioquímicos, estructurales y genómicos, se procederá a evaluar los posibles adyuvantes a los cuales asociarlos. Se utilizarán enlazadores adecuados entre los epítopes y el adyuvante, y mediante modelado por dinámica molecular se obtendrá la estructura tridimensional resultante del péptido quimérico.

## Referencias

1. Rassi, A. Jr, Rassi, A. & Marin-Neto, J. Chagas disease. *Lancet*. 375, 1388-1402, [https://doi.org/10.1016/S0140-6736\(10\)60061-X](https://doi.org/10.1016/S0140-6736(10)60061-X) (2010).
2. WHO, "Chagas disease (American trypanosomiasis)," Fact Sheet 340 (2013).
3. Arriagada, K. Tipificación de linajes de *Trypanosoma cruzi* en individuos con enfermedad de Chagas cardiópatas y no cardiópatas. Trabajo especial de grado, Universidad de Chile, Chile (2014).
4. Pérez-Aguilar, M. *et al.* Neuroimmunoendocrine System During Infection by *Trypanosoma cruzi*: Mechanisms of Immunoregulation. *EC Microbiology*. 14.9, 581-596 (2018).
5. Domingo, M. Ubicación subcelular, purificación, caracterización y estudios de interacción proteína-proteína de la enolasa de epimastigotes de *Trypanosoma cruzi*. Trabajo especial de grado, Universidad de Los Andes, Venezuela (2004).
6. Bontempi, I. Desarrollo de una estrategia inmunomoduladora para el control de la infección generada por *Trypanosoma cruzi*. Trabajo especial de grado, Universidad Nacional del Litoral, Argentina (2015).
7. Khatoon, N., Ojha, R., Mishra, A., & Prajapati, V. Examination of antigenic proteins of *Trypanosoma cruzi* to fabricate an epitope-based subunit vaccine by exploiting epitope mapping mechanism. *Vaccine*. 36, 6290-6300, <https://doi.org/10.1016/j.vaccine.2018.09.004> (2018).
8. Elisei, R. *et al.* Immunogenomic screening approach to identify new antigens for the serological diagnosis of chronic Chagas' disease. *Applied microbiology and biotechnology*. 102, 6069-6080, <https://doi.org/10.1007/s00253-018-8992-7> (2018).
9. Avendaño-Rangel, F. & Rey, K. Avances y desafíos en el control de la enfermedad de Chagas en Venezuela y un estudio de caso en el estado Mérida. *Consciencia y Diálogo*. 6, 155-164, (2015).
10. Quijano-Hernandez, I., & Dumonteil, E. Advances and challenges towards a vaccine against Chagas disease. *Human vaccines*. 7, 1184-1191, <https://doi.org/10.4161/hv.7.11.17016> (2011).
11. Backert, L., & Kohlbacher, O. Immunoinformatics and epitope prediction in the age of genomic medicine. *Genome medicine*. 7(1), 119, <https://doi.org/10.1186/s13073-015-0245-0> (2015).
12. Kazi, A. *et al.* Current progress of immunoinformatics approach harnessed for cellular- and antibody-dependent vaccine design. *Pathogens and Global Health*. 112(3), 123-131, <https://doi.org/10.1080/20477724.2018.1446773> (2018).

13. Soria-Guerra, R. *et al.* An overview of bioinformatics tools for epitope prediction: implications on vaccine development. *Journal of biomedical informatics*. 53, 405-414, <https://doi.org/10.1016/j.jbi.2014.11.003> (2015).
14. Sanchez-Trincado, J. *et al.* Fundamentals and methods for T-and B-cell epitope prediction. *Journal of immunology research*. 2017, <https://doi.org/10.1155/2017/2680160> (2017).
15. Potocnakova, L., Bhide, M., & Pulzova, L. An introduction to B-cell epitope mapping and in silico epitope prediction. *Journal of immunology research*. 2016, <http://dx.doi.org/10.1155/2016/6760830> (2016).
16. Konstantinou, G. T-Cell Epitope Prediction. In *Food Allergens* (pp. 211-222). Humana Press, New York, NY (2017).
17. Comas, I. *et al.* Human T cell epitopes of Mycobacterium tuberculosis are evolutionarily hyperconserved. *Nature genetics*. 42(6), 498, <https://doi.org/10.1038/ng.590> (2010).
18. Raymond, D. *et al.* Conserved epitope on influenza-virus hemagglutinin head defined by a vaccine-induced antibody. *Proceedings of the National Academy of Sciences*. 115(1), 168-173, <https://doi.org/10.1073/pnas.1715471115> (2018).
19. Vita, R. *et al.* The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res*. 8, D339-D343, <https://doi.org/10.1093/nar/gky1006> (2019).
20. Parker, J., Guo, D., & Hodges, R. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry*. 25(19), 5425-5432, <https://doi.org/10.1021/bi00367a013> (1986).
21. Karplus, P., & Schulz G. Prediction of chain flexibility in proteins. *Naturwissenschaften*. 72:212-213 (1985).
22. Emini, E. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *Journal of Virology*. 55:836-839, PMC255070 (1985).
23. Pellequer, J., Westhof, E., & Van Regenmortel, M. Correlation between the location of antigenic sites and the prediction of turns in proteins. *Immunology letters*. 36(1), 83-99, [https://doi.org/10.1016/0165-2478\(93\)90072-A](https://doi.org/10.1016/0165-2478(93)90072-A) (1993).
24. Kolaskar, A., & Tongaonkar, P. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS letters*. 276(1-2), 172-174, [https://doi.org/10.1016/0014-5793\(90\)80535-Q](https://doi.org/10.1016/0014-5793(90)80535-Q) (1990).
25. Janin, J. *et al.* Conformation of amino acid side-chains in proteins. *Journal of molecular biology*. 125(3), 357-386, [https://doi.org/10.1016/0022-2836\(78\)90408-4](https://doi.org/10.1016/0022-2836(78)90408-4) (1978).
26. Ponnuswamy, P., Prabhakaran, M., & Manavalan, P. Hydrophobic packing and spatial arrangement of amino acid residues in globular proteins. *Biochimica et Biophysica Acta*



- (BBA)-Protein Structure. 623(2), 301-316, [https://doi.org/10.1016/0005-2795\(80\)90258-5](https://doi.org/10.1016/0005-2795(80)90258-5) (1980).
27. Jespersen, M. *et al.* BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res.* 45(W1), W24-W29, <https://doi.org/10.1093/nar/gkx346> (2017).
  28. Saha, S., & Raghava, G. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins.* 65 (1): 40-48, <https://doi.org/10.1002/prot.21078> (2006).
  29. Jurtz, V. *et al.* NetMHCpan-4.0: Improved Peptide–MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. *Journey of Immunology.* 199: 3360–3368, <https://doi.org/10.4049/jimmunol.1700893> (2017).
  30. O'Donnell, T. *et al.* MHCflurry: Open-Source Class I MHC Binding Affinity Prediction. *Cell Syst.* 7: 129-132.e4, <https://doi:10.1016/j.cels.2018.05.014> (2018).
  31. Andreatta, M. *et al.* Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics.* 32: 511-517, <https://doi.org/10.1093/bioinformatics/btv639> (2015).
  32. Peters, B., & Sette, A. Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC Bioinformatics.* 6:132, <https://doi:10.1186/1471-2105-6-132> (2005).
  33. Moutaftsi, M. *et al.* A consensus epitope prediction approach identifies the breadth of murine TCD8-cell responses to vaccinia virus. *Nat Biotechnol.* 24: 817–819, <https://doi:10.1038/nbt1215> (2006).
  34. Reche, P., Glutting, J-P., & Reinherz, E. Prediction of MHC class I binding peptides using profile motifs. *Human Immunology.* 63: 701-709, [https://doi:10.1016/S0198-8859\(02\)00432-9](https://doi:10.1016/S0198-8859(02)00432-9) (2002).
  35. Nielsen, M., & Andreatta, M. NetMHCpan-3.0; improved prediction of binding to MHC class I molecules integrating information from multiple receptor and peptide length datasets. *Genome Medicine.* 8: 33, <https://doi:10.1186/s13073-016-0288> (2016).
  36. Kim, Y. *et al.* Derivation of an amino acid similarity matrix for peptide: MHC binding and its application as a Bayesian prior. *BMC Bioinformatics.* 10: 394, <https://www.biomedcentral.com/1471-2105/10/39> (2009).
  37. Rammensee, H-G. *et al.* SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics.* 50: 213–219, <https://doi:10.1007/s002510050595> (1999).
  38. Parker, K., Bednarek, M., & Coligan, J. Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. *The Journal of Immunol.* 152: 163–175 (1993).

39. Altuvia, Y., Schueler, O., & Margalit, H. Ranking potential binding peptides to MHC molecules by a computational threading approach. *The Journal of Molecular Biology*. 249: 244-250, <https://doi:10.1006/jmbi.1995.0293> (1995).
40. Liu, I-H., Lo, Y-S., & Yang, J-M. PAComplex: a web server to infer peptide antigen families and binding models from TCR-pMHC complexes. *Nucleic Acids Research*. 39: W254-W260, <https://doi:10.1093/nar/gkr434> (2011).
41. Kosakovsky, P. & Frost, S. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution*. 22, 1208-1222, <https://doi.org/10.1093/molbev/msi105> (2005).
42. Murrell, B. *et al.* Detecting individual sites subject to episodic diversifying selection. *PLoS Genet*. 8, e1002764, <https://doi.org/10.1371/journal.pgen.1002764> (2012).
43. National Center for Biotechnology Information. [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)
44. TriTryp, D. B. Kinetoplastid genomic resources Database. <https://tritrypdb.org/tritrypdb>