

Caracterización *in silico* y análisis filogenético de la malato sintasa de *Paracoccidioides brasiliensis* asociada a su proceso de infección

In silico characterization and phylogenetic analysis of the malate synthase from *Paracoccidioides brasiliensis* related to its infection process

Julio K. Huayhua¹, Luis A. Rodríguez¹ y Gustavo A. Sandoval¹

RESUMEN

Objetivo: Realizar la caracterización *in silico* de la proteína malato sintasa de *Paracoccidioides brasiliensis* relacionada con su proceso de infección. **Materiales y métodos:** Se obtuvo la secuencia aminoacídica de malato sintasa de *P. brasiliensis* (código de acceso: AQ75800.3) del GenBank. Dicha secuencia fue utilizada para determinar sus principales parámetros bioquímicos, dominios y regiones conservadas, predicción de estructuras secundarias y modelamiento tridimensional, empleando las herramientas bioinformáticas ProtParam, Clustal Omega, PHD Secondary Structure Prediction, SWISSMODEL, Phyre2 y Robetta, respectivamente. La visualización y comparación de los modelos tridimensionales se realizó empleando PyMol. Finalmente, se buscó el modelo de sustitución aminoacídica con modelgenerator_v_85 y se usó para generar una filogenia de máxima verosimilitud en PHYML, la cual se visualizó en FigTree. **Resultados:** La malato sintasa de *P. brasiliensis* es una proteína básica de 60.9 kDa, la cual presenta 53.1% de hélices α , 6.3% de láminas β y 40.6% de loops internos con plegamiento de tipo TIM beta/alpha-barrel. Se observó una estructura conservada entre diversos organismos. **Conclusión:** La malato sintasa de *P. brasiliensis* es una transferasa básica de mediano peso molecular que presenta estructura conservada y que estaría relacionada con su proceso de infección y el de otros organismos relacionados.

Palabras clave: malato sintasa, *Paracoccidioides brasiliensis*, adhesina, transferasa.

ABSTRACT

Objective: To perform an *in silico* characterization of the malate synthase protein from *Paracoccidioides brasiliensis* related to its infectious process. **Materials and methods:** The amino acid sequence of the malate synthase from *P. brasiliensis* was obtained from GenBank (accession code: AQ75800.3). This sequence was used to determine its main biochemical parameters, conserved domains and regions, prediction of secondary structures and tridimensional modelling, using several bioinformatics tools such as ProtParam, Clustal Omega, PHD Secondary Structure Prediction, SWISSMODEL, Phyre2 and Robetta, respectively. Visualization and model comparison was performed in PyMol. Finally, a model of amino acid substitution was obtained using modelgenerator_v_85, and using to generate a phylogeny of maximum likelihood in PHYML and visualized in FigTree. **Results:** The malate synthase from *P. brasiliensis* is a basic protein of 60.9 kDa, and is composed by 53.1% of α -helix, 6.3% of β -sheet and 40.6% of internal loops, with a folding type of TIM beta/alpha-barrel. It was also observed a conserved structure between several related organisms. **Conclusion:** The malate synthase from *P. brasiliensis* is a basic transferase with medium molecular weight with a conserved structure which is related to its infectious process and of the other related organisms.

Key words: malate synthase, *Paracoccidioides brasiliensis*, adhesin, transferase.

¹ Grupo de Estudio en Bioinformática Estructural. Facultad de Ciencias Biológicas. Universidad Nacional Mayor de San Marcos. Lima 01. Perú.

INTRODUCCIÓN

La paracoccidiomicosis es una micosis sistémica, la cual es una infección con gran prevalencia en diferentes países de América Latina. Dicha enfermedad se extiende desde el sur de México, hasta Argentina y Uruguay. Esta condición es causada por el hongo dimórfico *Paracoccidioides brasiliensis*, y se adquiere por vía inhalatoria. Su hábitat en la naturaleza aún no está definido, sin embargo, fue descrita por primera vez en 1908, en São Paulo-Brasil por Adolfo Lutz [24]. La infección se da por la inhalación de esporas en la fase filamentosa, y dentro de los pulmones se transforma a la fase levaduriforme (forma parasitaria), siendo la temperatura el principal factor de esta transición morfológica (25-37 °C). La mayoría de las personas infectadas sólo desarrollan una paracoccidioidomicosis (PCM) asintomática o subclínica, que puede progresar hacia una enfermedad con múltiples formas clínicas, que dependen de factores en el huésped o ambientales y de la virulencia fúngica. En pacientes con inmunidad deprimida esta infección suele ser mortal. La infección es más frecuente en varones que en mujeres, lo que sugiere una relación con las condiciones hormonales [16].

Respecto a la proteína malato sintasa, esta enzima cataliza la condensación aldólica de glioxilato con acetil-CoA para formar L-malato, lo cual corresponde al segundo paso del ciclo del glioxilato, una ruta alternativa al ciclo del ácido tricarbóxico presente en bacterias, hongos y plantas [7]. Da Silva Neto et al. [5] han reportado que en este organismo, dicha enzima se comporta como una adhesina, que participa en la unión de las células fúngicas en el tejido del huésped, promoviendo así la infección.

Entre las terapias actuales, se observa la importancia de atacar la pared celular para prevenir o inhibir una infección; sin embargo, lo demostrado por Da Silva Neto et al. [5] hace que esta enzima se vuelva un interesante candidato para el diseño de drogas que puedan neutralizarla y resultaría pertinente hacer un estudio del estado de conservación de la misma en otros organismos.

Por tal motivo, el presente trabajo se propone realizar una caracterización *in silico* de la proteína malato sintasa de *P. brasiliensis* y profundizar en el conocimiento de su proceso de infección.

MATERIALES Y METODOS

Obtención de las secuencias aminoacídicas

La secuencia primaria correspondiente a la proteína malato sintasa de *Paracoccidioides brasiliensis* fue obtenida a través de la base de datos del GenBank (código de acceso AQ75800.3). Dicha secuencia fue utilizada para las posteriores predicciones *in silico* y obtención de la información estructural de dicha proteína.

Determinación de parámetros bioquímicos y dominios conservados

Para determinar los principales parámetros bioquímicos de la proteína en estudio se utilizó la herramienta ProtParam del ExPASy (<http://web.expasy.org/protparam/>) [8]. La búsqueda de dominios conservados presentes en dicha proteína fue realizada mediante la herramienta Prosite del ExPASy (<http://prosite.expasy.org/>) [6].

Alineamiento múltiple de secuencias aminoacídicas

Para realizar el alineamiento múltiple, se obtuvieron las secuencias aminoacídicas de *Bacillus anthracis* (código de acceso PDB: 3CUX_A) y *Escherichia coli* (código de acceso PDB: 3CUZ_A). Dichas secuencias fueron alineadas y editadas utilizando la herramienta Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) [22].

Predicción de estructuras secundarias y alineamiento estructural

El análisis de predicción de estructuras secundarias presentes en la proteína malato sintasa de *P. brasiliensis* fue realizada empleando la herramienta PHD Secondary Structure Prediction Method (<http://npsa-prabi.ibcp.fr>) [20]. Por otro lado, el alineamiento estructural se realizó en la plataforma Cn3D (versión 4.3.1) utilizando como molde a la proteína malato sintasa de *B. anthracis* [14].

Modelamiento tridimensional de la malato sintasa de P. brasiliensis

El modelo tridimensional por homología de la enzima en estudio, se obtuvo empleando la herramienta SWISSMODEL (<http://swissmodel.expasy.org/>) [15], a partir de la información depositada en la base de datos del Protein Data Bank (PDB). Por otro lado, se empleó la herramienta Phyre2 Protein

Homology/analogy Recognition Engine (<http://www.sbg.bio.ic.ac.uk/~phyre2/>) en modo intensivo [11], así como el servidor Robetta Full Chain Protein Structure Prediction (<http://rosetta.bakerlab.org/>) [13], para la predicción de estructuras mediante modelamiento por threading y ab initio, respectivamente. La visualización de los modelos obtenidos se realizó empleando el programa PyMol Molecular Graphics System (versión 1.3). La evaluación de los modelos obtenidos se hizo a través del plot de Ramachandran generado por el servidor RAMPAGE (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>) [21].

Análisis filogenético de proteínas seleccionadas

Para hacer un análisis filogenético en base a las secuencias primarias, se realizó un BLASTp [1] en el NCBI y se seleccionaron 28 secuencias con distinto grado de identidad. Se alinearon con Clustal Omega [22] y se editaron utilizando el programa Bioedit (versión 7.2.5). Se encontró el modelo de sustitución aminoacídica con modelgenerator_v_85 y se usó para generar una filogenia de máxima verosimilitud en PHYML [9], la cual se visualizó en el programa FigTree (versión 1.4.0).

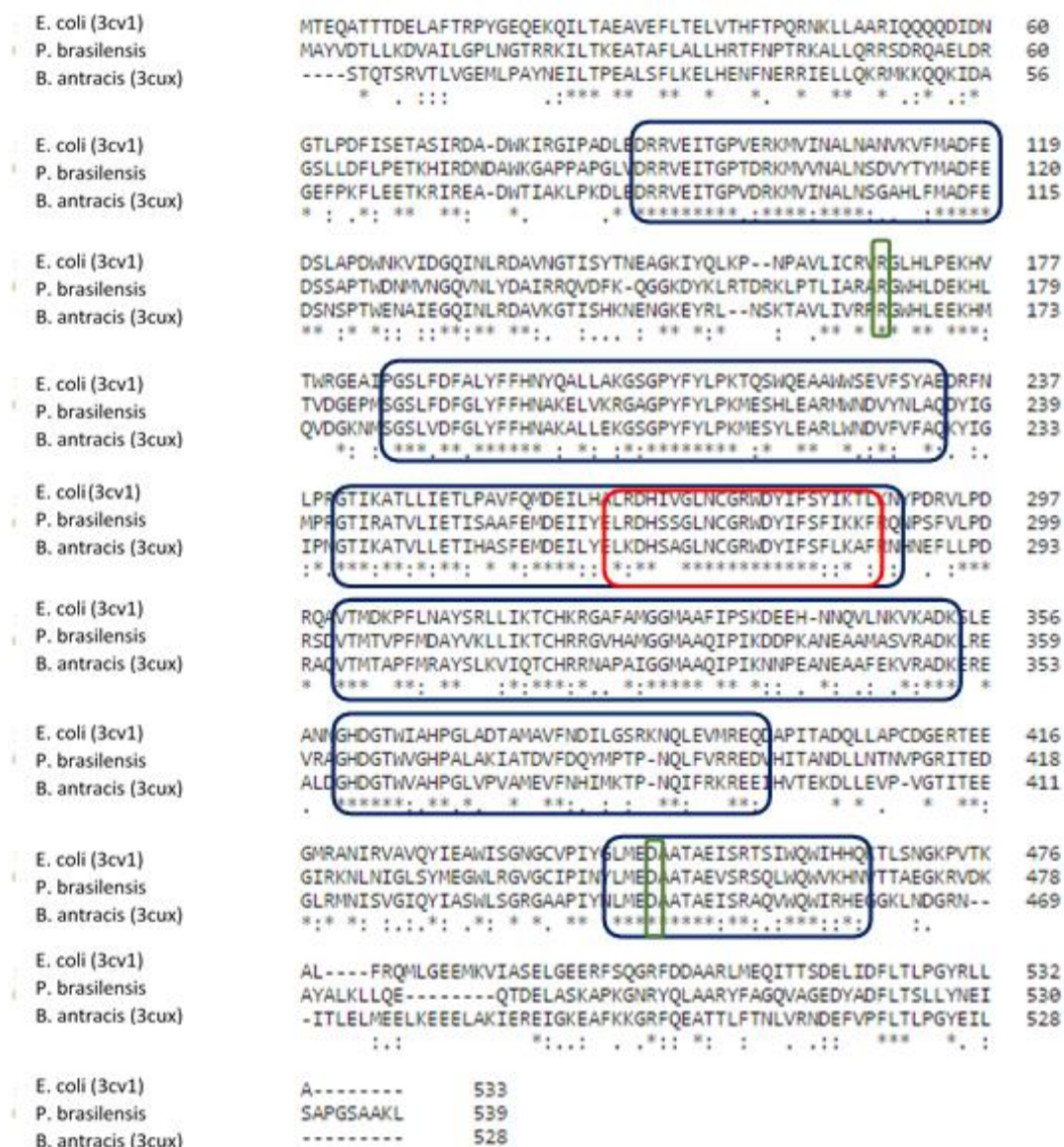


Fig 1. Alineamiento de secuencias aminoacídicas entre las proteínas malato sintasa de *E. coli*, *P. brasiliensis* y *B. anthracis*. Identidad completa (*) y residuos con sustituciones conservativas. Nótese las regiones con mayor identidad (azul), el dominio conservado de la familia proteinca (rojo) y los residuos del sitio activo (verde)

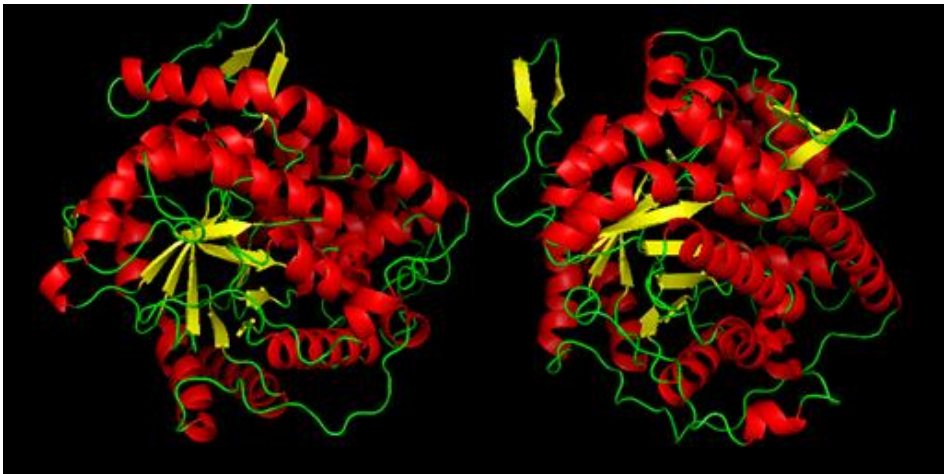


Fig. 3. Modelo tridimensional de la proteína malato sintasa de *Paracoccidioides brasiliensis* empleando modelamiento por homología (izquierda) y *threading* (derecha). Las α hélices y láminas β se muestran de color rojo y amarillo respectivamente, mientras que los loops internos se muestran de color verde.

Tabla 1. Distribución de aminoácidos en el plot de Ramachandran del servidor RAMPAGE.

Modelo*	Aminoácidos Favorecidos (%)	Aminoácidos permitidos (%)	Aminoácidos en region atipica (%)
H1	94	4	1.9
H2	94.6	3.6	1.7
T1	96.6	2.3	1.1
T2	95.8	3	1.2
T3	89.8	6.3	3.9
A1	90.5	6.3	3.2
A2	95.3	3	1.7
A3	90.5	6.3	3.2
A4	95	3.5	1.5
A5	95	2.8	2.2

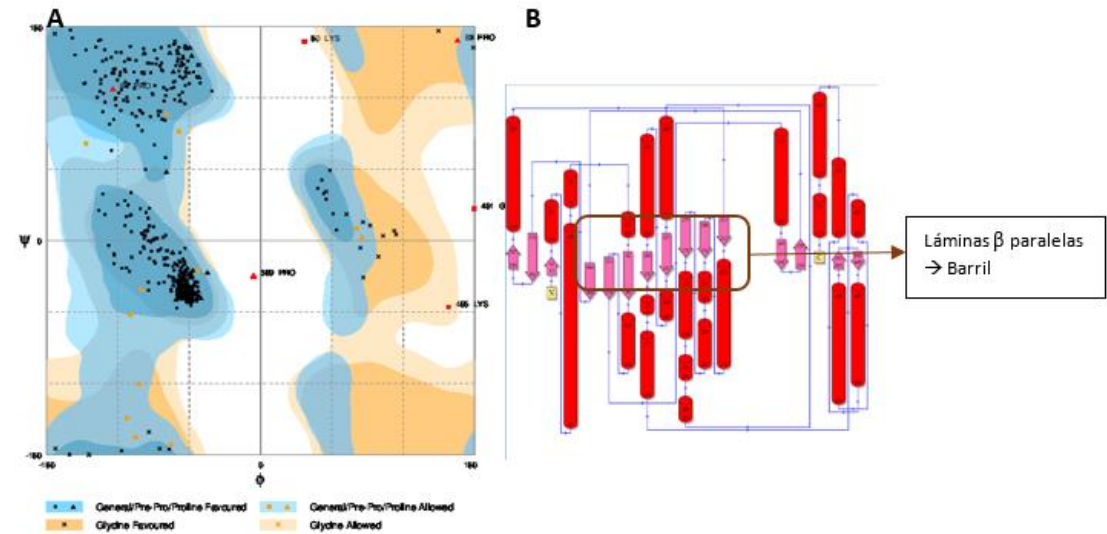


Fig. 4. A) Plot de Ramachandran del modelo T1, obtenido con el servidor RAMPAGE. B) Topología del modelo. Nótese la distribución de las láminas β paralelas que conforman el “barril” del modelo tridimensional.

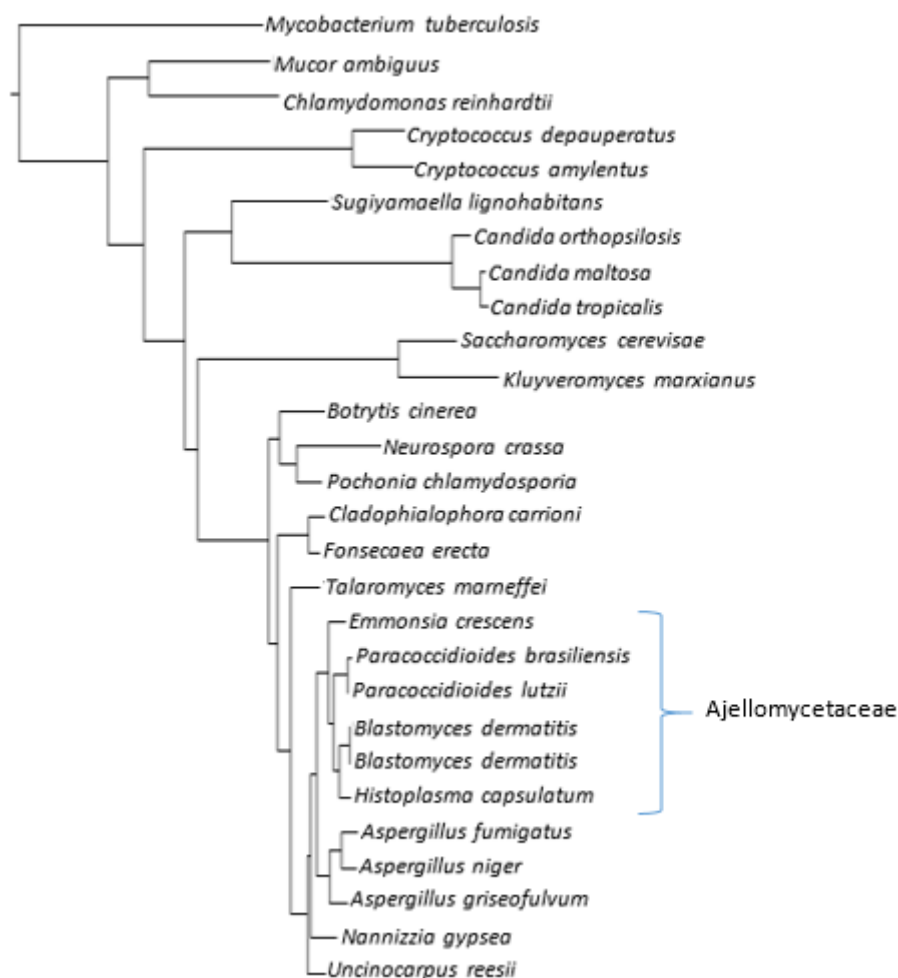


Fig. 5. Árbol filogenético construido de secuencias aminoácidas de la proteína malato sintasa de diferentes organismos. Se señala la familia Ajellomycetaceae, a la cual pertenece *P. brasiliensis*. Visualizado por FigTree (1.4.0).

DISCUSIÓN

La paracoccidiomicosis es una micosis sistémica, infección con gran prevalencia en países de América Latina. El principal factor de esta transición morfológica es la temperatura. (25°C a 37°C). Su cuadro clínico es tan complejo, que incluso se ha reportado una confusión con sarcoidosis, llevando a un primer tratamiento poco exitoso [19].

El presente informe presenta un estudio *in silico* de la proteína malato sintasa de *Paracoccidioides brasiliensis*. Esta enzima cataliza la condensación aldólica de glioxilato con acetil-CoA para formar L-malato, lo cual corresponde al segundo paso del ciclo del glioxilato, una ruta alternativa al ciclo del ácido tricarboxílico presente en bacterias, hongos y plantas [7].

Además, Da Silva Neto et al. [5] han reportado que esta enzima se comporta como una adhesina en este organismo, participando

en la unión de las células fúngicas al tejido del huésped, promoviendo así el proceso de infección.

Este grupo de proteínas presenta entre 530-570 aa y es altamente conservada entre las especies [4], lo cual puede ser confirmado en el alineamiento de secuencias. Esto se debe a una alta homología con las provenientes de otros organismos, especialmente en la región central y el sitio activo. Al analizar los patrones de plegamiento, también se observó grandes regiones conservadas, pudiendo predecir la conformación a partir del alineamiento estructural, y una riqueza en aminoácidos que favorecen la conformación de α -hélices.

A partir de los modelos obtenidos, se pudo observar que la proteína se incluye dentro de la clasificación a/b con plegamiento tipo TIM beta/alpha-barrel (SCOP) [18], a&b (CATH) [23], con una predominancia de α -hélices y una zona central de láminas β paralelas en

forma de barril, característica de la familia de proteínas malato sintasa. Estos resultados concuerdan con lo descrito por autores como Anstrom et al. [2] y Lohman et al. [14] con esta enzima en otras especies. Con respecto al sitio activo, los residuos R(168) y D(449) forman parte de láminas β . Esta enzima, además, tiene como ligandos a dos iones, ya sea Ca^{+2} o Mg^{+2} , tal como ocurre con esta enzima en especies como *E. coli* y *B. anthracis* [14]; sin embargo, los modelos obtenidos por homología solo conservan uno de los sitios de unión a dichos iones.

Del análisis filogenético mostrado en la Fig. 5, se observa que la formación de grupos parece dependiente solo de la cercanía filogenética de las especies analizadas y no guardaría relación con la capacidad de infección de las especies. Por ejemplo, *C. orthopsilosis* y *C. tropicalis* son agentes causales conocidos de candidiasis [3] pero no forman un grupo separado de *C. maltosa*, especie que no suele ser encontrada en aislados clínicos [3] y es considerada solo un patógeno potencial [25]. Los géneros *Paracoccidioides*, *Blastomyces*, *Histoplasma* y *Emmonia* pertenecen a la familia Ajellomycetaceae [17], forman un solo cluster en el árbol y son patógenos humanos importantes, a diferencia de las especies de *Aspergillus*, que suelen ser oportunistas. Es posible que *P. brasiliensis* no sea la única especie de la familia Ajellomycetaceae cuya malato sintasa posee función de adhesión [5] sino que regiones determinadas de esta proteína estén relacionadas con la mayor patogenicidad de la familia. En caso de *M. tuberculosis*, su enzima malato sintasa tiene importancia en la adhesión a fibronectina/laminina de líneas celulares de epitelio pulmonar [12]. Sin embargo, se observa en la filogenia que la malato sintasa de *M. tuberculosis* se encuentra bastante alejada de la de la familia Ajellomycetaceae por lo que la modificación que le otorga capacidad de adhesión a la malato sintasa de *P. brasiliensis* podría ser independiente y encontrarse en otra región de la proteína.

CONCLUSIONES

La enzima malato sintasa de *P. brasiliensis* es una transferasa básica de mediano peso molecular, con un predominio de hélices α en su conformación, y una región central con láminas β paralelas. Esta proteína presenta una secuencia y estructura muy conservada entre organismos relacionados y no relacionados.

Esta conservación resultaría conveniente para su rol en la adhesión del microorganismo a un hospedero en un proceso de infección. Esto la constituye como un posible blanco para el desarrollo de drogas inhibitoras.

AUTOR DE CORRESPONDENCIA

Dr. Gustavo A. Sandoval
Facultad de Ciencias Biológicas
Universidad Nacional Mayor de San Marcos
Av. Venezuela s/n - Ciudad Universitaria
Lima - Perú.
Teléfono: +51-1-6197000 Anexo 1558.
E-mail: gsandovalp@unmsm.edu.pe

REFERENCIAS BIBLIOGRÁFICAS

1. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997;25(17):3389-402.
2. Anstrom, D. M., & Remington, S. J. (2006). The product complex of *M. tuberculosis* malate synthase revisited. *Protein science*, 15(8),2002-2007.
3. Blanco-Blanco, María Teresa, et al. *Candida orthopsilosis* fungemias in a Spanish tertiary care hospital: Incidence, epidemiology and antifungal susceptibility. *Revista Iberoamericana de Micología*, 2014, vol. 31, no 2, p. 145-148.
4. Bruinenberg P.G., Blaauw M., Kazemier B., Ab G., Cloning and sequencing of the malate synthase gene from *Hansenula polymorpha*. *Yeast* 6:245-254(1990).
5. Da Silva Neto, B. R., da Silva, J. D. F., Mendes-Giannini, M. J. S., Lenzi, H. L., de Almeida Soares, C. M., & Pereira, M. (2009). The malate synthase of *Paracoccidioides brasiliensis* is a linked surface protein that behaves as an anchorless adhesin. *BMC microbiology*, 9(1), 1.
6. de Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, et al. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res.* 2006;34(Web Server issue):W362- 5.
7. Dixon, G. H., Kornberg, H. L., & Lund, P. (1960). Purification and properties of

- malate synthetase. *Biochimica et biophysica acta*, 41(2), 217-233.
8. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the ExPASy Server. In: Walker JM, editor. *The Proteomics Protocols Handbook*. Humana Press; 2005. p. 571-607.
 9. Guindon, Stéphane. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic biology*, 2010, vol. 59, no 3, p. 307-321.
 10. Keane, T. M., Creevey, C. J., Pentony, M. M., Naughton, T. J., & McInerney, J. O. (2006). Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC evolutionary biology*, 6(1), 1.
 11. Kelley LA et al. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* 10, 845-858
 12. Kinhikar, A. G., Vargas, D., Li, H., Mahaffey, S. B., Hinds, L., Belisle, J. T., & Laal, S. (2006). Mycobacterium tuberculosis malate synthase is a laminin-binding adhesin. *Molecular microbiology*, 60(4), 999-1013.
 13. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* 2004;32(Web Server issue):W526-31.
 14. Lohman, J. R., Olson, A. C., & Remington, S. J. (2008). Atomic resolution structures of Escherichia coli and Bacillus anthracis malate synthase A: Comparison with isoform G and implications for structure-based drug discovery. *Protein Science*, 17(11), 1935-1945.
 15. Marco Biasini, Stefan Bienert, Andrew Waterhouse, Konstantin Arnold, Gabriel Studer, Tobias Schmidt, et al. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research*; (1 July 2014) 42 (W1): W252-W258.
 16. Marques, S. A. (2003). Paracoccidioidomycosis: atualização epidemiológica, clínica e terapêutica. *Anais brasileiros de dermatologia*, 135-146.
 17. Muñoz, J. F., Gauthier, G. M., Desjardins, C. A., Gallo, J. E., Holder, J., Sullivan, T. D. & Gujja, S. (2015). The dynamic genome and transcriptome of the human fungal pathogen Blastomyces and close relative Emmonsia. *PLoS Genet*, 11(10), e1005493.
 18. Murzin A. G., Brenner S. E., Hubbard T., Chothia C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247, 536-540
 19. Pujol-Riqué, M., Ruiz, S., Alonso-Tarrés, C., & Cañete, C. (2011). Micosis pulmonar por Paracoccidioides brasiliensis: confusión peligrosa con sarcoidosis. *Radiologia*, 53(6), 560-563.
 20. Rost B, Sander C. Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol.* 1993 Jul 20;232(2):584-99.
 21. S.C. Lovell, I.W. Davis, W.B. Arendall III, P.I.W. de Bakker, J.M. Word, M.G. Prisant, J.S. Richardson and D.C. Richardson (2002) Structure validation by C α geometry: phi, psi and C β deviation. *Proteins: Structure, Function & Genetics*. 50: 437-450.
 22. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol.* 2011 Oct 11;7:539.
 23. Sillitoe, I., Lewis, T. E., Cuff, A., Das, S., Ashford, P., Dawson, N. L., & Lehtinen, S. (2015). CATH: comprehensive structural and functional annotations for genome sequences. *Nucleic acids research*, 43(D1), D376-D381.
 24. Vargas, J., & Vargas, R. (2009). Paracoccidioidomycosis. *Revista de enfermedades Infecciosas y Tropicales*, 1, 49.
 25. Yoshida, Minoru; Hashimoto, Kazunori. Potential pathogenicity of Candida maltosa IAM 12248. *Agricultural and Biological Chemistry*, 1986, vol. 50, no 8, p. 2119-2120.

Recibido: 23/09/2016
Aceptado: 28/11/2016