

PREDICCIÓN IN SILICO DE LA FUNCIÓN DE LOS NSSNP DEL GEN MBL2 VINCULADO A LA MANIFESTACIÓN DE LA TUBERCULOSIS PULMONAR

PREDICTION IN SILICO OF THE FUNCTION OF THE NSSNP OF THE MBL2 GENE LINKED TO THE MANIFESTATION OF PULMONARY TUBERCULOSIS

Luis Máximo Villanueva Zúñiga¹, María Gracia Alvarez Valdivia¹, Víctor Alexander Solís Salazar¹, Diego E. Valencia¹

(1) Universidad Católica de Santa María, Arequipa - Perú

RESUMEN: *Introducción:* El objetivo del presente trabajo fue evaluar mediante técnicas computacionales el potencial daño de los polimorfismos de nucleótido único no sinónimo del gen MBL2. *Material y Métodos:* Se realizó un screening de todas las mutaciones presentes en el gen MBL2, encontrándose 176 polimorfismos en la región codificante. Ellos fueron sometidos al análisis de los servidores Panther, i-mutant 3.0, PROVEAN, MutPred2 y PolyPhen2. *Resultados:* Se encontraron quince (15) polimorfismos con respuesta de datos dañina o deletérea en todos los servidores; de ellos, dos (2) están ubicados en los sitios activos, en comparación con los bolsillos enzimáticos que tiene la proteína. *Conclusiones:* Los polimorfismos A173D y G131D tiene potencial efecto deletéreo sobre el sitio activo de la proteína, por lo que se recomienda su evaluación posterior en ensayos in vivo para corroborar la información obtenida.

Palabras clave: Polimorfismos, screening, mutación, MBL2.

ABSTRACT: *Introduction:* The objective of the present work was to evaluate by means of computational techniques the potential damage of single nucleotide polymorphisms not synonymous with the MBL2 gene. *Material and Methods:* All mutations present in the MBL2 gene were screened and 176 polymorphisms were found in the coding region. They were submitted to the analysis of the servers Panther, i-mutant 3.0, PROVEAN, MutPred2 and PolyPhen2. *Results:* Fifteen (15) polymorphisms with harmful or deleterious data response were found in all servers; of these, two (2) are located in active sites, in comparison with the enzymatic pockets that the protein has. *Conclusions:* The polymorphisms A173D and G131D have a potential deleterious effect on the active site of the protein, so it is recommended to evaluate them later in vivo to corroborate the information obtained.

Keywords: Polymorphisms, screening, mutation, MBL2.

INTRODUCCIÓN

Mundialmente la tuberculosis es la primera causa de muerte por agente infeccioso, en Perú, la reciente Ley de Prevención y Control de la TB en el Perú (Ley 30287) declara un tema de interés nacional por la alta tasa de casos y mortalidad, existe una lucha constante para combatir esta enfermedad. Se ha demostrado que el gen MBL2 codifica la lectina de unión a manosa que se encuentra en el suero (1), esta proteína codificada pertenece a la familia de la colectina, que es un elemento importante del sistema inmunitario innato. Se ha evidenciado controversialmente la deficiencia de este gen como un factor protector o inductor a enfermedades infecciosas como lo es el mycobacterium (2), ya que se conoce la función del gen de opsonización del Mycobacterium (3), para una posterior fagocitosis, y la resistencia a mecanismos de protección intracelular como la inhibición de la maduración del fagosoma. se podría decir que está destinado a tener una vida intracelular, y la autofagia para el control de la infección, es incierta(4).

La lectina fijadora de manosa es una macromolécula grande que tiene una estructura similar a un ramo.

La subunidad estructural básica de MBL es un homotrímico de polipéptidos de MBL, entrelazados en una triple hélice(5).

Cada cadena polipeptídica tiene cuatro dominios:

- una región rica en cisteína N-terminal de 21 aminoácidos involucrada en la oligomerización,
- un dominio tipo colágeno de 59 aminoácidos,
- un α de 30 aminoácidos α - helicoidal, dominio hidrofóbico en espiral del cuello en espiral, que es crucial para iniciar la oligomerización, y
- un dominio de reconocimiento de carbohidratos C-terminal de 188 aminoácidos.

En el presente estudio se toma como hipótesis el factor protector dado por las mutaciones que presentan deficiencia en la codificación del gen MBL2(6-8). Se sabe también que cada gen posee millones de mutaciones. El objetivo de este trabajo fue crear una herramienta para reducir el número total de mutaciones a las más representativas, para que sean protagonistas de próximos estudios más específicos, se logra simplificar la búsqueda con esta herramienta, todo esto usando mecánica molecular.

Correspondencia:

Luis Máximo Villanueva Zúñiga

Celular: 920147780

E-mail: villanuevamaximo1@gmail.com

MATERIALES Y MÉTODOS

Diseño del estudio: Análisis de mutaciones en servidores.

Espacio y Tiempo: se realizaron los análisis del 25 de febrero al 08 de marzo del presente año; todo el trabajo fue realizado en el Centro de Investigación de Ingeniería Molecular (CIIM-UCSM).

Universo, población y muestra: Se utilizó como muestra todas las mutaciones (2507) del gen MBL2.

Criterios de selección: se tomaron como criterios de inclusión aquellas mutaciones codificantes Missense y Nonsense (184)

Recolección, gestión y análisis de datos: se analizaron los datos en 5 servidores diferentes:

1. I-Mutant (<http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi>)
2. PROVEAN (http://provean.jcvi.org/seq_submit.php)
3. MutPred2 (<http://mutpred.mutdb.org/#qform>)
4. PolyPhen-2 (<http://genetics.bwh.harvard.edu/pph2/>)
5. Panther (<http://www.pantherdb.org/tools/csnpscoreForm.jsp>).

Consideraciones éticas y diseminación: el presente trabajo no requiere aprobación de un comité de ética.

METODOLOGIA PROCEDIMENTAL

1.- Obtención de la secuencia

La secuencia de nucleótidos del gen humano MBL2 se obtuvo a partir de la página web del National Centre for Biotechnology Information (NCBI). La secuencia de aminoácidos de la proteína MBL2 fue obtenida en formato FASTA a partir de la base de datos UniProt (<http://www.uniprot.org/>).

2.- Mapeo de los polimorfismos

Todos los SNPs reportados del gen MBL2 se obtuvieron de la base de datos NCBI dbSNP (<http://www.ncbi.nlm.nih.gov/snp>). Se mapeó un total de 2507 SNPs activos en humanos en la secuencia del gen MBL2. El número de SNPs de diferentes clases funcionales es reportado en la Fig. 1. En la región no codificante se mapearon 84 SNPs, 676 en la región 3'UTR, 52 en la región 5'UTR, 956 en la región intrónica y 176 fueron en la región codificante y otros 563 SNPs activos en humanos.

3.- Análisis por medio del servidor Panther

El servidor PANTHER (Análisis de proteínas a través de relaciones evolutivas) (<http://pantherdb.org/tools/cSNPscoreForm.jsp>) calculó la duración de un aminoácido dado que se ha conservado evolutivamente entre varias especies y predice el efecto de ese cambio específico de aminoácidos sobre el efecto estructural y funcional sobre la proteína.

4.- Análisis utilizando el i-mutant 3.0

Este servidor (<http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi>) se basa en el algoritmo de la Máquina de vectores de soporte (SVM) para predecir la estabilidad de la proteína debido a las variaciones de aminoácidos individuales. Puede predecir cambios en la estabilidad de las proteínas mediante el uso de secuencias o estructuras de proteínas. Tiene una precisión general del 77% cuando la predicción se basa en la secuencia de proteínas.

5.- Predicción de daño basado en la homología de la secuencia usando PROVEAN

PROVEAN (Protein Variation Effect Analyzer) (<http://provean.jcvi.org>) es una herramienta de predicción basada en las secuencias que estima el efecto de una variación en la secuencia proteica con respecto a la función de la proteína. Las coordenadas en el cromosoma para cada uno de los SNP fueron obtenidas a partir de dbSNP para el análisis PROVEAN. El formato de entrada fue un sistema de coordenadas basado en los residuos separados por coma (número del cromosoma, coordenadas del SNP, orientación, sustitución de nucleótido). Se predice que un nsSNP presente en la región codificante de un gen es "perjudicial" si el puntaje de predicción está por debajo del valor -2.5, y "neutral" si el puntaje es superior a dicho valor.

6.- Predicción de daño utilizando MutPred 2

Este servidor es una herramienta (<http://mutpred.mutdb.org/>) que detecta el efecto de las mutaciones sin sentido en la estructura de la proteína al proporcionarle la secuencia de la proteína en un formato FASTA, posición de mutación, tipos silvestres y mutantes. El resultado contiene el puntaje Mutpred2, que es la probabilidad de que la sustitución del aminoácido sea patógena (> 0.50 se considera patógena), también nos muestra los valores de PROSITE y ELM afectados, y una tabla con los cambios estructurales que se produjeron en la pérdida/ganancia de una hélice, una hebra o un bucle. Además, proporciona información sobre otros cambios estructurales, en su caso, y un valor P.

7.- Predicción de daño basado en la homología de la estructura usando PolyPhen-2

PolyPhen-2 (Polymorphism Phenotyping-2) (<http://genetics.bwh.harvard.edu/pph2/>) es una herramienta web para la anotación de SNPs no sinónimos. La herramienta emplea una regla empírica específica la cual abarca tanto las consideraciones físicas como las comparativas para predecir las posibles consecuencias funcionales de la sustitución de un aminoácido en la estructura y función de una proteína humana. PolyPhen-2 utiliza la secuencia de la proteína en formato FASTA como archivo de entrada y estima la influencia de un SNP en particular o la variación de un aminoácido en una posición dada en la secuencia. La herramienta calcula el puntaje del conteo independiente de posición específica (PSIC) para cada una de las variantes y calcula el puntaje de la diferencia entre las variables.

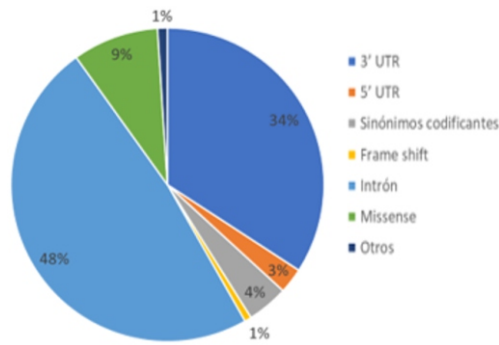
RESULTADOS

Se evaluaron 176 mutaciones correspondientes nsSNP, quienes cumplieran con los criterios de inclusión para el estudio.

En la figura 1 se aprecia la distribución de los tipos de polimorfismos ubicados en todo el gen MBL2; pudiendo observar que la mayoría de los polimorfismos se encuentra en la zona intrónica, correspondiente a una frecuencia cercana a la mitad del gen; seguido de esto se encontró la zona 3' UTR.

Cada uno de los servidores utiliza diferentes criterios, los cuales se ven reflejados en la variabilidad y estrategia computacional que utilizan; por ello se pueden obtener diferentes resultados en cada uno de ellos.

Fig. 1 Distribución de los polimorfismos en los diferentes segmentos del MBL2



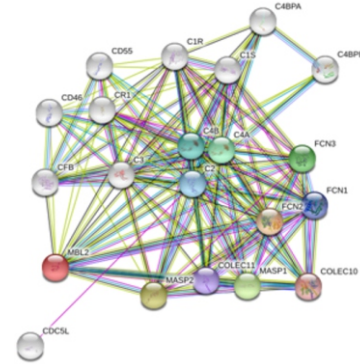
En la tabla 1 se presentan los valores diagnósticos obtenidos según cada servidor.

Tabla 1. Frecuencia de los valores diagnóstico de cada uno de los servidores

| Servidor | Frecuencia | Porcentaje |
|------------------------------------|------------|------------|
| PANTHER | | |
| Posiblemente dañinas | 45 | 23.6% |
| Probablemente dañinas | 59 | 30.9% |
| Probablemente benignas | 87 | 45.5% |
| I-Mutant (SMV3) | | |
| Gran disminución de la estabilidad | 83 | 43.5% |
| Gran incremento de la estabilidad | 2 | 1.0% |
| Estabilidad neutral | 106 | 55.5% |
| PROVEAN | | |
| Deletéreo | 112 | 58.6% |
| Neutral | 79 | 41.4% |
| MutPred2 | | |
| Patogénico | 56 | 29.3% |
| Normal | 135 | 70.7% |
| PolyPhen2 | | |
| Probablemente dañino | 85 | 44.5% |
| Posiblemente dañino | 26 | 13.6% |
| Benigno | 80 | 41.9% |

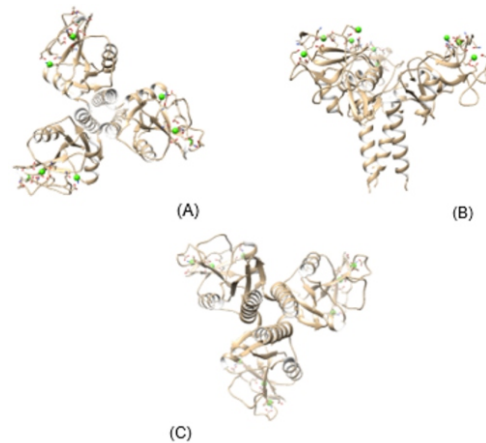
En la figura 2 se muestran las redes de las isoformas de empalme o las modificaciones posteriores a la traducción están colapsadas, es decir, cada nodo representa todas las proteínas producidas por un único locus de genes codificantes de proteínas. Los nodos coloridos son los que forma la primera capa de interacciones, los blancos son los que componen la segunda capa de interacciones. Las mutaciones que fueron reconocidas como dañinas o que se encontraban alteradas su función fueron: F246L, D221A, W208G, F195L, G185S, F183C, A173D, E170G, T164P, V151G, G131D, R52C, G45S, L12R y L10P. Se hizo la búsqueda del sitio activo de la proteína, así mismo se buscó la coincidencia de aminoácidos que se encuentren en este sitio. Se encontró que dos mutaciones: A173D y G131D han sido ubicadas en este lugar. Por lo que estas mutaciones son recomendadas para su empleo en estudios poblacionales para corroborar la información obtenida en el trabajo.

Fig. 2 Nodos de interacción de la proteína MBL2



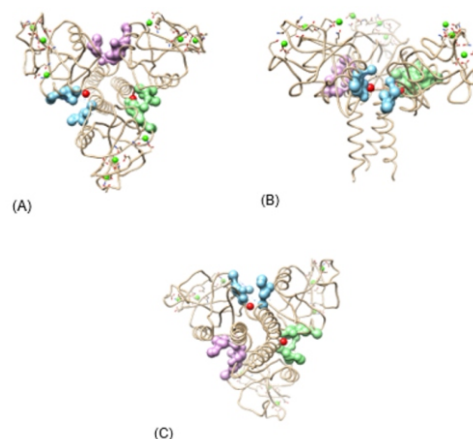
La figura 3 corresponde a la proteína nativa en sus dos vistas: superior lateral e inferior.

Fig. 3 Estructura tridimensional de la proteína MBL. A) Vista superior, B) Vista lateral, C) Vista inferior.



En la figura 4 se muestran imágenes de las estructuras proteínas con sus respectivos pockets, se han identificado tres bolsillos con puntajes iguales, cada uno correspondiente a las interacciones de las proteínas para conformar el complejo. Vistas superior, lateral e inferior.

Fig. 4 Ubicación de los bolsillos en la estructura tridimensional de la proteína MBL. A) Vista superior, B) Vista lateral, C) Vista inferior.



DISCUSIÓN

Exitosamente se encontraron 15 polimorfismos que mostraban una función alterada en cuanto al análisis en los distintos servidores (F246L, D221A, W208G, F195L, G185S, F183C, A173D, E170G, T164P, V151G, G131D, R52C, G45S, L12R y L10P), pero descartando las mutaciones en comparación del sitio activo obtenemos 2 mutaciones **A173D y G131D**, que demuestran una correlación de aminoácidos con la secuencia original.

Los polimorfismos en este gen han sido vinculados a la aparición y variabilidad de diversas enfermedades. El trabajo de Rantala(9) ha demostrado una asociación entre los niveles de la proteína MBL y las infecciones del tracto respiratorio en jóvenes (10); dando pie a la interpretación de posibles polimorfismos en el gen MBL2 que estén determinando esto(9).

También se han vinculado polimorfismos del gen MBL2 como un factor de riesgo para la transmisión vertical del VIH-1(11).

En lo que respecta a tuberculosis, estudios demuestran que el codón 57, podrían considerarse como factor de riesgo para el desarrollo de tuberculosis(12).

Desde hace algún tiempo se ha venido manteniendo la idea de la relación entre esta proteína MBL y los polimorfismos de su gen MBL2 en la susceptibilidad para la tuberculosis. Un meta-análisis en el 2010 ha sugerido que existe una pobre caracterización de los genotipos de MBL2 reportados, y que no están directamente asociados con la susceptibilidad a la infección pulmonar de tuberculosis; sin embargo, en estos pacientes hay una presencia elevada de MBL(13). Un estudio realizado en la India encontró que el alelo "B" del MBL está asociado a la protección contra la infección de MT, sin embargo, no hubo influencia sobre la tasa de conversión en el esputo(14).

Lo que obtiene a partir del trabajo, es la especificidad en futuras mutaciones que pueden ser favorables o desfavorables a la infección por mycobacterium, el presente artículo fomenta a futuras investigaciones en medir ambas mutaciones y determinar su acción agonista y antagonista, en una población real.

CONCLUSION

Se han encontrado 2 mutaciones clave A173D y G131D, las cuales se recomienda utilizar un estudio experimental in vivo para corroborar su daño potencial.

CONFLICTOS DE INTERES

Los autores no manifiestan ningún tipo de conflicto de interés.

AGRADECIMIENTOS

Agradecer al Centro de Investigación en Ingeniería Molecular (CIIM) por habernos brindado sus instalaciones para la realización de este trabajo, al Centro de Investigación y Estudios Médicos (CIEM) por las oportunidades que siempre nos brindan durante nuestra formación en postgrado.

REFERENCIAS BIBLIOGRÁFICAS

- [1] Mandal RK, Khan MA, Hussain A, Dar SA, Aloufi S, Jawed A, et al. Association of MBL2 gene polymorphisms with pulmonary tuberculosis susceptibility: trial sequence meta-analysis as evidence. *Infect Drug Resist.* 2019;12:185.
- [2] Areeshi MY, Mandal RK, Akhter N, Dar SA, Jawed A, Wahid M, et al. A meta-analysis of MBL2 polymorphisms and tuberculosis risk. *Sci Rep.* 2016;6:35728.
- [3] Wu Y, Yang X, Ting-di Chen ZZ, You Y. Genetic polymorphisms of exon 1 of MBL2 contribute to tuberculosis risk especially in Asian populations: an updated meta-analysis of 26 studies. *Infect Drug Resist.* 2018;11:1237.
- [4] Zhang C, Li X, Zhang W, Wang Y, Fan G, Wang W, et al. Common genetic variant rs3802842 in 11q23 contributes to colorectal cancer risk in Chinese population. *Oncotarget.* 2017;8(42):72227.
- [5] Cao Y, Wang X, Cao Z, Wu C, Wu D, Cheng X. Genetic polymorphisms of MBL2 and tuberculosis susceptibility: a meta-analysis of 22 case-control studies. *Arch Med Sci AMS.* 2018;14(6):1212.
- [6] Shi J, Xie M, Wang J, Xu Y, Xiong W, Liu X. Mannose-binding lectin two gene polymorphisms and tuberculosis susceptibility in Chinese population: a meta-analysis. *J Huazhong Univ Sci Technolog Med Sci.* 2013;33(2):166–171.
- [7] Zhang N, Zhuang M, Ma A, Wang G, Cheng P, Yang Y, et al. Association of levels of mannose-binding lectin and the MBL2 gene with type 2 diabetes and diabetic nephropathy. *PLoS One.* 2013;8(12).
- [8] Capparelli R, Iannaccone M, Palumbo D, Medaglia C, Moscarello E, Russo A, et al. Role played by human mannose-binding lectin polymorphisms in pulmonary tuberculosis. *J Infect Dis.* 2009;199(5):666–672.
- [9] Rantala A, Lajunen T, Juvonen R, Bloigu A, Silvennoinen-Kassinen S, Peitso A, et al. Mannose-binding lectin concentrations, MBL2 polymorphisms, and susceptibility to respiratory tract infections in young men. *J Infect Dis.* 2008;198(8):1247–1253.
- [10] Eisen DP, Minchinton RM. Impact of mannose-binding lectin on susceptibility to infectious diseases. *Clin Infect Dis.* 2003;37(11):1496–1505.
- [11] Boniotto M, Braida L, Pirulli D, Arraes L, Amoroso A, Crovella S. MBL2 polymorphisms are involved in HIV-1 infection in Brazilian perinatally infected children. *Aids.* 2003;17(5):779–780.
- [12] Da Cruz HLA, Da Silva RC, Segat L, de Mendonça Gomes MSZ, Brandão LAC, Guimarães RL, et al. MBL2 gene polymorphisms and susceptibility to tuberculosis in a northeastern Brazilian population. *Infect Genet Evol.* 2013;19:323–329.
- [13] Denholm JT, McBryde ES, Eisen DP. Mannose-binding lectin and susceptibility to tuberculosis: a meta-analysis. *Clin Exp Immunol.* 2010;162(1):84–90.
- [14] Singla N, Gupta D, Joshi A, Batra N, Singh J, Birbian N. Association of mannose-binding lectin gene polymorphism with tuberculosis susceptibility and sputum conversion time. *Int J Immunogenet.* 2012;39(1):10–14.
- [15] Secuencia Fasta
>sp|P11226|MBL2_HUMAN Mannose-binding protein C
OS=Homo sapiens OX=9606 GN=MBL2 PE=1 SV=2
MSLFPSPPLLLSMVAASYSETVTTCEDAQKTCPAVIACSS
PGINGFPGKDGDRDGTGKGEKGEPPGQGLRGLQGPPGKLG
PPGNPGSPGSPGPKGQKGDGPKSPDGDSSLAASERKA
LQTEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKA
LCVKFQASVATPRNAENGAIQNLIKEAFLGITDEKTEG
QFVDLTGNRLTYTNWNEGEPNNAAGSDEDCVLLLNKGQ
WNDVPCSTSHLAVCEFFI

Recibido el 24 de octubre del 2019 y aceptado para su publicación el 12 de diciembre del 2019