**Borrador introducción**

La bioinformática es una disciplina científica que surgió por la necesidad de estudiar datos biológicos, que se pueden definir, de manera sofisticada, como información de secuencias moleculares, contenido genómico y análisis del producto génico, fenotípico (1). Así, esta disciplina integra interdisciplinariamente las ciencias exactas y naturales con la computación, en orden de solucionar cualquier problema relacionado a la biología o entender el funcionamiento de los fenómenos que componen a los seres vivos (1). Los métodos generados dentro de esta área han sido utilizados en una gran cantidad de áreas, incluyendo biología molecular, genómica, biología de sistemas, biotecnología, agricultura, biomédica e incluso medicina (1). Se puede entender, entonces, cómo la biología computacional ha revolucionado la forma en la cual se estudian los organismos vivos, de la mano de la biología molecular, resaltando así su gran importancia y contribución al estado del arte del conocimiento. Por ejemplo, la base de datos BLAST de ncbi es una herramienta computacional de análisis de secuencias, disponible para cualquiera que posea una conexión a internet (2). Sus algoritmos permiten buscar cualquier secuencia de ADN, ARN u proteínas en sus bases de datos, en un tiempo computacional muy rápido y, si se ha secuenciado previamente, los resultados proveerán el organismo al cual pertenece, en qué región de su genoma se encuentra, qué función posee e, incluso, el porcentaje de identidad con la secuencia buscada (2). Esto ha permitido soportar innumerables estudios en el campo de la biología.

Concomitante a la aparición de los medios moleculares para estudiar la composición más básica de los organismos, como la secuenciación de nucleótidos para el estudio de genes y genomas, por ejemplo, sucede el surgimiento de las “ómicas”, ramas de la biología molecular dedicadas al estudio del comportamiento de diversos sistemas moleculares que componen organismos y ecosistemas (3). En este caso de estudio es importante extender la discusión en el sentido de una de estas ramas, la interactómica (4). En específico, ésta integra la proteómica y a genómica, estudiando las redes de interacción proteica, las cuales, no sólo dependen de la expresión génica, sino, de muchos otros factores celulares. Estas redes de interacción buscan dilucidar, también, las interacciones proteína-proteína (PPI) (4). En el estudio de esta materia se han estandarizado tres formas de aproximación principales, métodos *in silico*, *in vitro* e *in vivo*.

En cuanto a los métodos *in silico* existen varios enfoques específicos (4). En general, se suelen hacer búsquedas en bases de datos por medio de minería de datos, esto, con la finalidad de hallar posibles interactores, término utilizado para referirse a proteínas que presentan PPI, como un indicio para el estudio posterior con otras herramientas (4). Usualmente, estos se pueden realizar como un complemento analítico a los métodos experimentales, puesto que aún se requiere validación en orden de disminuir la cantidad de falsos positivos (4).

Existen tres enfoques principales en la metodología *in silico*. Primero, la búsqueda de ortólogos, la cual se refiere a encontrar proteínas “homólogas”, es decir, que compartan origen evolutivo común con aquella de estudio (citar), en organismos ya caracterizados o “modelo”, en orden de encontrar candidatos para PPI en el organismo de interés o, si ya se han validado estas interacciones obteniendo datos pertinentes, relacionar la funcionalidad de las proteínas halladas con ortólogos que ya han sido estudiados (4). Este enfoque halla su fundamento en el hecho que muchas proteínas ortólogas conservan su función en el proceso evolutivo, por lo que no deben usarse parálogos. Estos se refieren a homólogos que han surgido por medio de duplicaciones génicas (5). Esto implica que, en muchos casos, adquieren funcionalidades diferentes entre sí, por lo que no son ideales para este método. Por otro lado, se encuentra el acercamiento basado en estructura. Este se refiere a la utilización de modelos tridimensionales computarizados de las proteínas, los cuales, poseen información muy precisa de los dominios proteicos su superficie y, en algunos casos, su afinidad de unión (4). Por este medio, se pueden predecir PPI, pues se ha observado que, si dos proteínas poseen dominios similares, es probable que éstas interactúen entre sí (4). Claramente, en muchos casos no existen modelos proteicos para aquella proteína que se esté estudiando, por lo que se pueden buscar indicios por ortólogos o, incluso, realizar un modelaje experimental *de novo*. Finalmente, el acercamiento basado en coevolución se encuentra basado en el hecho que, si dos proteínas poseen un proceso de coevolución, dado por presión selectiva, en función de una relación funcional estas, se puede sospechar, van a encontrarse juntas en el proceso de heredabilidad (4). Así, al estudiar la presencia o ausencia de pares proteicos, de manera recurrente, se pueden predecir PPI´s entre ellas (4). Aun así, este método posee varias limitaciones relacionadas con eventos aleatorios que suceden en el proceso evolutivo (4). En esta parte debo encontrar más fuentes que el artículo 3, para no “sesgar” la información.

En orden de permitir que estos métodos computacionales se nutran de información real y de validar las predicciones de posibles PPI´s, se deben llevar a cabo experimentos de interactómica. Uno de los diseños experimentales más utilizados es el ensayo de doble híbrido. Consiste en utilizar el sistema de regulación de la expresión génica de levaduras GAL4, específicamente, el dominio DBD de unión a DNA y el dominio AD de activación (7). El primero se encarga de reclutar ciertos factores de transcripción a la región génica específica y, el segundo, permite el reclutamiento de factores de activación (6). La base yace en el hecho que estos dominios pueden estar presentes en diferentes proteínas (6). Así, se preparan dos plásmidos, como mínimo, pero dependiendo de la naturaleza del objetivo experimental y la cantidad de proteínas que se desean evaluar, esto puede cambiar. En uno se ubica el cDNA de la proteína “cebo” fusionada con el dominio BD (6, 7). En el otro, se ubica el cDNA de la proteína   
“presa” con el dominio AD. En caso de existir interacción entre ambas proteínas, el acercamiento físico de los dominios causará el evento de iniciación de la transcripción, en este caso, de un gen reportero previamente clonado en la célula de levadura (6, 7). Así, el gen reportero permitirá medir los niveles de interacción por medio del grado de expresión de este y su reporte (7). Una de las principales limitaciones de este método es la presencia de falsos positivos, debido a que suelen presentarse interacciones no específicas que pueden activar la cascada transcripcional generando señales que no necesariamente corresponden a interacciones entre las proteínas reportadas (7). No obstante, se pueden realizar correcciones estadísticas sobre los niveles de expresión en orden de corregir este tipo de error. Este método sigue siendo muy útil para el estudio de la interactómica.

Con estas técnicas, se puede abordar diversos sistemas biológicos, estudiándolos desde la perspectiva de redes de interacción. Uno de los sistemas más estudiados por estos mecanismos es la interacción plata patógeno, dada su importancia académica y práctica, sobre todo en el campo de la agricultura, tema que se tratará más adelante. Dado esto, es imperativo entender el contexto de esta relación entre organismos. Esta relación es muy compleja debido a la gran cantidad de eventos físicos y químicos que la componen. Se puede abordar como una compleja red de interacciones entre moléculas producidas por la planta y aquellas secretadas por la bacteria. El primer contacto inicia con la llegada de la bacteria al tejido de la planta, donde debe superar las barreras físicas de ésta, compuesta por la cutícula, la pared celular, la membrana celular y los compuestos antimicrobianos de expresión constitutiva (8, 9). La principal forma en la cual superan las defensas físicas es ingresando por medio de las estomas, heridas, lenticelas u otros puntos débiles (10). La planta posee PRR´s, receptores de membrana de reconocimiento de patrones moleculares que reconocen sustancias específicas, usualmente endógenas, de las bacterias patógenas (8). Estas moléculas se conocen como PAMP´s, definidos como patrones moleculares asociados a patógenos (8). Así, después de este proceso de reconocimiento, la planta activa una respuesta conocida como PTI (Inmunidad desencadenada por PAMP) de carácter general y bastante efectiva en la mayoría de los casos, pues es suficiente para defender a la planta de la posible infección (8, 9). Adicionalmente, las células vegetales pueden generar más respuestas parecidas reconociendo DAMP´s o marcadores de daño celular (10).

La respuesta PTI está asociada a la vía de señalización MAPK (10). Posterior a la transducción de señales intracelulares se activan respuestas celulares como el cierre estomático, limitando la carga bacteriana, producción de ROS, producción de compuestos antimicrobianos como las fitoalexinas y disminución del paso de nutrientes del citoplasma al apoplasto, disminuyendo así la tasa de crecimiento bacteriana, entre otros (10). En muchos casos, los patógenos producen factores de avirulencia o proteínas efectoras (8). Su mecanismo de acción es la inhabilitación de la respuesta PTI, suprimiendo los componentes moleculares involucrados en la formación de la respuesta (10). La respuesta del hospedero ante un suceso de este tipo es específica a la proteína efectora bacteriana, expresando un factor de resistencia (proteína R) activando así una segunda fase de inmunidad conocida como ETI (Inmunidad desencadenada por efector) (8, 9). En una gran cantidad de los casos, esta respuesta es acompañada de un efecto colateral conocido como muerte celular programada o respuesta hipersensible, donde el tejido vegetal muere de forma localizada (8). Una ETI puede ser tan simple como la interacción de sólo un efector patógeno y una proteína R específica. Así, se hace énfasis en la importancia de considerar la respuesta ETI como una red de interacción de proteínas. Estas interacciones y respuestas de un lado al otro se conocen como el modelo de zig-zag (10).

Esta interacción también se debe abordar desde la naturaleza del patógeno. Existen tres categorías principales. Los biótrofos se alimentan de tejido vegetal vivo (11). Los necrótrofos, por otra parte, requieren que las células mueran induciendo este evento por medio de toxinas producidas por la bacteria invasora (11). Finalmente, los hemibiótrofos combinan ambas estrategias, donde, en la primera fase biótrofa, secretan sustancias que inhiben la muerte celular programada de la planta y, posterior a esto, una segunda fase necrótrofa, donde secretan toxinas para generar la muerte de la célula y alimentarse de los compuestos degradados (11). En el caso de las bacterias es difícil clasificar de forma rígida a determinada especie debido a que muchas presentan resultados experimentales variable (12). Es importante mencionar que *Xathomonas.spp*, bacteria objeto en este estudio, suele ser considerada como un patógeno biótrofo, a tal punto, que produce efectores, como AvrBs3, para producir hipertrofia celular en su hospedero (12, 13). La importancia de esta clasificación para patógenos yace en la respuesta que la planta dispone en orden de combatir a cada tipo de agente infeccioso (12). La respuesta sistémica adquirida (SAR) está determinada por la señalización por parte del ácido salicílico y es activada por un agente biótrofo o hemibiótrofo, relacionado, por ende, con la respuesta hipersensible localizada (16). Por otro lado, la respuesta de resistencia sistémica inducida (ISR) es activada por patógenos biótrofos por medio de sustancias como el metil-jasmonato y el etileno, por lo que suele tener como objetivo la manutención de sus células vivas y la producción de agentes antimicrobianos os sustancias inhibitorias (14, 16) (15). Usualmente, algunos organismos de la rizosfera generan un tipo de predisposición activando la vía del metil-jasmonato/etileno cuyo efecto es una posterior respuesta ISR mucho más eficiente y rápida al entrar la planta en contacto con un patógeno (14, 15). Las interacciones planta patógeno son muy complejas dada su larga historia de coevolución, así que es en este punto donde se recalca la importancia de herramientas operacionales eficientes de bioinformática que permiten estudiar estos fenómenos de forma mucho más fácil y completa.

Los efectores bacterianos tienen blancos proteicos en la planta muy específicos, así, como la funcionalidad que modifican en base a esta interacción. Aún así, se ha descubierto que efectores con historias evolutivas independientes muchas veces convergen en cuanto a sus blancos, de manera directa e indirecta, afectando sistemas celulares inmunes de la planta en conjunto, concepto reconocido como “hubs celulares” (17). En cuanto a efectores, se debe considerar el hecho que el proceso de coevolución planta patógeno ha sucedido por millones de años, por lo que, por medio de adaptaciones y exaptaciones, la variedad de estas proteínas que ha surgido es muy alta. Por lo tanto, su mecanismo de acción varía mucho en función de la fisiología del hospedero. Muchos patógenos poseen efectores cuya función es superar la barrera física descrita previamente, estos, entonces, son proteínas de degradación, por ejemplo, de la pared celular vegetal, que permitirán la entrada del parásito (17). Como respuesta a mecanismos PTI descritos, algunos efectores evitan el cierre de estomas alterando la vía de señalización del metil-jasmonato (16). Relacionado con esta primera línea de defensa, algunos otros bloquean la vía de señalización MAPK, disminuyen la efectividad del reconocimiento de PAMP´s actuando sobre los PRR, ebmascaran PAMP´s específicos o inhiben enzimas proteolíticas del hospedero disminuyendo así la efectividad de la respuesta PTI (16). Dependiendo del tipo de patógeno y su clasificación según el estado celular de la planta, los efectores pueden ser producidos con el objetivo de generar la muerte celular o, incluso, mantener a la célula viva inhibiendo la respuesta de hipersensibilidad (16). Estos efectos se consiguen por medio de la disrupción de las vías de metil-jasmonato/etileno y ácido salicílico ya descritas, dependiendo de tipo de patógeno y sus requerimientos nutricionales (16). Por último, con una gran cantidad de fines, muchos efectores han evolucionado con la capacidad de alterar la expresión génica en su hospedero de diversas maneras, por ejemplo, alterando el silenciamiento por medio de RNA, inhibiendo de manera directa algunos factores de transcripción de la planta o, incluso, por medio de TALE´s, efectores que actúan directamente como factores de transcripción específicos para diversos promotores en el genoma vegetal. Dada la complejidad de estos mecanismos, se requiere un gran poder de análisis para poder estudiar estos fenómenos.

La relevancia del estudio de interacciones planta-patógeno radica en muchos aspectos. En cuanto al impacto que las enfermedades en plantas producen en nuestra sociedad, las implicaciones son inmensas. Este hecho se evidencia con la estimación que, debido sólo a las enfermedades en cultivos de consumo humano, existen alrededor de 800 millones de personas a nivel mundial que no poseen una alimentación adecuada (18). Sólo en relación con la enfermedad causada por el patógeno en estudio, la “Mancha foliar”, se puede aseverar que comprende una alta proporción de los casos de enfermedad en cultivos y afecta tanto a monocotiledóneas como dicotiledóneas (19). En orden de solucionar este problema, los esfuerzos se han decantado a la producción de organismos genéticamente modificados por medio del uso de DNA recombinante, produciendo cepas con mejor resistencia a enfermedades (20). Teniendo esto en cuenta, es claro que, si se logran identificar las interacciones específicas entre los efectores de patógenos y sus blancos en las plantas, estas modificaciones genéticas para conferir resistencia a los cultivos se harían inmensamente más eficientes. Es aquí donde un estudio como este adquiere una gran relevancia a nivel práctico.

El objetivo de este estudio radica en dilucidar redes de interacción en cuanto a la interacción planta patógeno dados los datos de un ensayo doble híbrido entre *Manihot esculenta* y *Xanthomonas*, utilizando una gran cantidad de herramientas bioinformáticas y computacionales, encontrar los blancos específicos en términos de rutas metabólicas y enzimas específicas de los 16 efectores bacterianos estudiados, además, de la producción de un modelo predictivo de redes de interacción (no definitivo). En base a los resultados de estos análisis, se estudiarán en detalle algunos puntos específicos de esta interactómica estudiada, los cuales, posean importancia biológica y funcional relevante y hubs de interacción que resulten de interés académico o práctico. Se espera generar una metodología por medio de la cual se puedan estudiar relaciones de este tipo en otros casos de interacciones planta-patógeno con un valor predictivo y eficiencia altos, facilitando así el estudio en esta área. En este caso de estudio específico, se postula que existen relaciones claras en cuanto a los blancos metabólicos y redes enzimáticas en función de los efectores estudiados y que estos interactúan indirectamente entre sí, por lo cual, se prevé que, por medio del análisis computacional, se encontrarán diferentes blancos en común para grupos de efectores en particular. Paralelo a esto, se postula que existen dominios proteicos en los efectores que determinan sus blancos en el hospedero y, en gran parte, las relaciones propuestas anteriormente. Con base a este hecho, se encontrarán redes de interacción con un alto grado de importancia a nivel del desarrollo de la enfermedad infecciosa por *Xanthomonas* en plantas de este tipo. Por último, se hipotetiza que con las herramientas computacionales que se tienen a disposición podrá generarse un modelo predictivo que, como mínimo y basado en datos experimentales de estudio interactómico pueda predecir posibles interacciones entre efectores patogénicos y genes de resistencia o blancos metabólicos en su hospedero. Este modelo podrá ser enriquecido con información presente en bases de datos, en orden de mejorar su capacidad predictiva.

**Referencias:**

1. Ibrokhim Y. Abdurakhmonov (July 27th 2016). Bioinformatics: Basics, Development, and Future, Bioinformatics - Updated Features and Applications, Ibrokhim Y. Abdurakhmonov, IntechOpen, DOI: 10.5772/63817. Available from: <https://www.intechopen.com/books/bioinformatics-updated-features-and-applications/bioinformatics-basics-development-and-future>
2. Ye, J., McGinnis, S., & Madden, T. L. (2006). BLAST: improvements for better sequence analysis. Nucleic acids research, 34(Web Server issue), W6–W9. https://doi.org/10.1093/nar/gkl164
3. Vailati-Riboni M., Palombo V., Loor J.J. (2017) What Are Omics Sciences? In: Ametaj B. (eds) Periparturient Diseases of Dairy Cows. Springer, Cham. <https://doi.org/10.1007/978-3-319-43033-1_1>
4. Shachuan Feng, Li Zhou, Canhua Huang, Ke Xie & Edouard C Nice (2015) Interactomics: toward protein function and regulation, Expert Review of Proteomics, 12:1, 37-60, DOI: 10.1586/14789450.2015.1000870
5. Jensen R. A. (2001). Orthologs and paralogs - we need to get it right. Genome biology, 2(8), INTERACTIONS1002. <https://doi.org/10.1186/gb-2001-2-8-interactions1002>
6. Paiano, A., Margiotta, A., De Luca, M., & Bucci, C. ( 2019). Yeast two‐hybrid assay to identify interacting proteins. Current Protocols in Protein Science, 95, e70. doi: 10.1002/cpps.70
7. Brückner, A., Polge, C., Lentze, N., Auerbach, D., & Schlattner, U. (2009). Yeast two-hybrid, a powerful tool for systems biology. International journal of molecular sciences, 10(6), 2763–2788. <https://doi.org/10.3390/ijms10062763>
8. Gupta, R., Lee, S. E., Agrawal, G. K., Rakwal, R., Park, S., Wang, Y., & Kim, S. T. (2015). Understanding the plant-pathogen interactions in the context of proteomics-generated apoplastic proteins inventory. Frontiers in plant science, 6, 352. <https://doi.org/10.3389/fpls.2015.00352>
9. Peyraud, R., Dubiella, U., Barbacci, A., Genin, S., Raffaele, S., & Roby, D. (2017). Advances on plant-pathogen interactions from molecular toward systems biology perspectives. The Plant journal : for cell and molecular biology, 90(4), 720–737. <https://doi.org/10.1111/tpj.13429>
10. Jean Bigeard, Jean Colcombet, Heribert Hirt, Signaling Mechanisms in Pattern-Triggered Immunity (PTI), Molecular Plant,Volume 8, Issue 4, 2015, Pages 521-539. ISSN 1674-2052. <https://doi.org/10.1016/j.molp.2014.12.022>.
11. Koeck, M., Hardham, A. R., & Dodds, P. N. (2011). The role of effectors of biotrophic and hemibiotrophic fungi in infection. Cellular microbiology, 13(12), 1849–1857. <https://doi.org/10.1111/j.1462-5822.2011.01665.x>
12. Kraepiel, Y. and Barny, M.‐A. (2016), Gram‐negative phytopathogenic bacteria, all hemibiotrophs after all?. Molecular Plant Pathology, 17: 313-316. doi:10.1111/mpp.12345
13. Marois, E, (2002). The Xanthomonas Type III Effector Protein AvrBs3 Modulates Plant Gene Expression and Induces Cell Hypertrophy in the Susceptible Host. Molecular Plant-Microbe Interactions® 2002 15:7, 637-646
14. Van Wees, S. C., de Swart, E. A., van Pelt, J. A., van Loon, L. C., & Pieterse, C. M. (2000). Enhancement of induced disease resistance by simultaneous activation of salicylate- and jasmonate-dependent defense pathways in Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United States of America, 97(15), 8711–8716. <https://doi.org/10.1073/pnas.130425197>
15. Choudhary, D. K., Prakash, A., & Johri, B. N. (2007). Induced systemic resistance (ISR) in plants: mechanism of action. Indian journal of microbiology, 47(4), 289–297. <https://doi.org/10.1007/s12088-007-0054-2>
16. Toruño, T. Y., Stergiopoulos, I., & Coaker, G. (2016). Plant-Pathogen Effectors: Cellular Probes Interfering with Plant Defenses in Spatial and Temporal Manners. Annual review of phytopathology, 54, 419–441. <https://doi.org/10.1146/annurev-phyto-080615-100204>
17. Mukhtar, M. S., Carvunis, A. R., Dreze, M., Epple, P., Steinbrenner, J., Moore, J., … Payne, T. (2011). Independently evolved virulence effectors converge onto hubs in a plant immune system network. *Science*, *333*(6042), 596–601. <https://doi.org/10.1126/science.1203659>
18. Taylor, W. I., & Achanzar, D. (1972). Catalase test as an aid to the identification of Enterobacteriaceae. *Applied Microbiology*, *24*(1), 58–61. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=380547&tool=pmcentrez&rendertype=abstract>
19. Hayes, R. J., Trent, M. A., Truco, M. J., Antonise, R., Michelmore, R. W., & Bull, C. T. (2014). The inheritance of resistance to bacterial leaf spot of lettuce caused by Xanthomonas campestris pv. vitians in three lettuce cultivars. Horticulture research, 1, 14066. <https://doi.org/10.1038/hortres.2014.66>
20. Buiatti, M., Christou, P., & Pastore, G. (2013). The application of GMOs in agriculture and in food production for a better nutrition: two different scientific points of view. *Genes & nutrition*, *8*(3), 255–270. https://doi.org/10.1007/s12263-012-0316-4