

Caracterización bioinformática de blancos terapéuticos putativos identificados en cepas de *Staphylococcus aureus* resistentes a meticilina: Evolución molecular aplicada en la selección racional de candidatos

Bioinformatic characterization of putative therapeutic targets identified in strains of methicillin-resistant *Staphylococcus aureus*: molecular evolution applied to the rational selection of candidates

Andrés Julián Gutiérrez Escobar¹

Recibido: marzo/2011; Aceptado: mayo/2011

Resumen

La mortalidad causada por cepas de *Staphylococcus aureus* multiresistentes supera al virus del SIDA y al cáncer en países desarrollados y en países latinoamericanos ha presentado un aumento considerable en los últimos años. Aunque los esfuerzos realizados para la identificación de candidatos terapéuticos son considerables, aun se observa una carencia en la caracterización evolutiva molecular consistente de los mismos lo cual se refleja en el aumento de vacunas y fármacos imperfectos. En el presente artículo se propone una metodología bioinformática para caracterizar a nivel evolutivo proteínas que sean consideradas como blancos terapéuticos y se utiliza la enzima UDP-N-acetilglucosamina 1-carboxiviniltransferasa de *S. aureus*, como caso de estudio donde se identificó un modo de selección natural operante sobre la población de genes codificantes, indicando que es una enzima altamente adaptable bajo presión selectiva para explorar nuevos nichos biológicos.

Palabras clave: evolución molecular, blanco terapéutico, selección positiva.

Abstract

Mortality caused by multiresistant strains of *Staphylococcus aureus* is higher than human immunodeficiency virus/AIDS in the United States with the same tendency in Latin-American countries. Despite the various treatment options for *S. aureus* infections, it remains a major hospital- and community-acquired opportunistic pathogen. With the emergence of multidrug-resistant *S. aureus* strains, there is an urgent need for the discovery of new antimicrobial drug targets in the organism. In this paper we propose a molecular evolutionary approach to characterize proteins that are considered as therapeutic targets for *S. aureus*. Here, the UDP-N-acetylglucosamine 1-carboxivinyltransferase evolutionary history has been described and an exhaustive analysis made for detecting the natural selection type was made. UDP Phylogenetic tree topology was seen to be statistically robust and the data indicated strong positive selection operating on the UDP proteins meaning the capacity to explore new niches for the species.

Key words: molecular evolution, therapeutic target, positive selection.

¹ M.Sc. en Ciencias Biomédicas de la Universidad del Quindío. Centro de Investigación y Desarrollo de la Fundación Universitaria del Área Andina, Bogotá, Colombia, andresgutierrez@colombia.com

INTRODUCCIÓN

Staphylococcus aureus es una bacteria Gram-positiva causante de infección en humanos y en animales. Se estima que cerca del 20% de la población humana saludable es persistentemente colonizada por esta bacteria causando infecciones nasales y como patógeno oportunista causa el mayor número de infecciones asociadas al cuidado de la salud e infecciones comunitarias a nivel mundial (Kluytmans, Belkum, & Verbrugh, 1997). Su potencial virulento le permite desarrollar bacteriemias, neumonías, endocarditis, meningitis, síndromes de choque tóxico en adultos, lesiones en la piel, abscesos en pacientes pediátricos y mastitis en ganado vacuno (Cheung, Projan, & Gresham, 2002; Lowy, 1998; Projan & Novick, 1997).

La emergencia de cepas resistentes tanto a la meticilina (Methicillin-resistant *Staphylococcus aureus* [MRSA]) como a la vancomicina (Vancomycin-resistant *Staphylococcus aureus* [VRSA]) a nivel mundial es de gran preocupación, ya que el tratamiento convencional basado en antibióticos no funciona bien y hasta la fecha solo se ha desarrollado un fármaco denominado *la platensimicina*, que ha demostrado ser efectivo contra algunas de estas cepas multiresistentes pero que es económicamente costoso (Gould, 2005; Wang et al., 2007). En Colombia según los últimos tres reportes del equipo de vigilancia epidemiológica de la alcaldía mayor de Bogotá, se ha identificado que los niveles de morbo-mortalidad relacionados a esta bacteria han aumentado, así como el número de casos no reportados similar al panorama encontrado en México (Ávila et al., 1999; Secretaría Distrital de Salud, 2009).

S. aureus también coloniza una amplia gama de mamíferos de compañía como son perros y gatos y también animales de granja como cerdos, vacas y caballos, de hecho tiene la capacidad de colonizar gallinas y pavos. En vacas puede causar casos severos de mastitis con considerable impacto económico pero más alarmante es la presencia de cepas MRSA en estos animales, por ejemplo la cepa (ST) 398 que puede colonizar cerdos (Baptiste et al., 2005). Es necesario un entendimiento profundo de las características ecológicas, epidemiológicas y evolutivas de este patógeno, además de caracterizar la maquinaria molecular responsable de su interacción con diferentes hospederos. Sin embargo, todos los intentos de vacunas han sido infructuosos.

A nivel poblacional se ha determinado que *S. aureus* presenta linajes dominantes de los cuales se despren-

den linajes menores y estudios moleculares han logrado identificar que estos son altamente distintos y cada uno posee su propia maquinaria enzimática de superficie y de secreción con una regulación metabólica diferencial (Loeffler et al., 2005; Weese et al., 2005). En humanos se ha observado que los mayores linajes están conformados por los siguientes complejos clonales: CC1, CC5, CC8, CC9, CC12, CC15, CC22, CC25, CC30, CC45 y CC51 (Feil et al., 2003). Aquellos linajes que han adquirido el profago *mecA* adquieren resistencia conformándose como cepas MRSA asociadas a infección hospitalaria y están representadas por los clones CC5, CC8, CC22, CC30, CC45 y una cepa híbrida CC239 y las cepas MRSA asociadas a infección comunitaria son CC1, CC8, CC30, CC59 y CC80 (Robinson & Enright, 2003). Las vacas se infectan por sus propias líneas clonales y la infección cruzada con humanos es rara (Robinson & Enright, 2003; Robinson & Enright, 2004; Tristan et al., 2007).

Una de las técnicas de punta para la identificación de blancos terapéuticos son los análisis por balance de flujo a partir de reconstrucciones metabólicas empleando las secuencias genómicas de los microorganismos como molde. Varios estudios de simulación metabólica han sido desarrollados para esta especie logrando resultados importantes (Becker & Palsson, 2005; Heinemann, Kummel, & Panke, 2005). Sin embargo, solo un estudio ha demostrado la suficiente robustez como para proponer blancos terapéuticos que deben ser estudiados con mayor profundidad (Lee et al., 2009). Sin embargo, algo que también tienen en común estos estudios es su falta de descripción de los blancos terapéuticos allí identificados.

Entonces, en el presente estudio se propuso una caracterización evolutiva molecular profunda para los blancos metabólicos identificados por Leen et al. (2009), donde se evaluaron las relaciones por homología, su filogenia molecular y los cambios sinónimos y no sinónimos de las secuencias en cuanto a las implicaciones de estos análisis para la identificación de blancos terapéuticos.

METODOLOGÍA

Población de estudio

La población inicial de 44 enzimas fue filtrada bajo los siguientes parámetros: i) que fueran enzimas esenciales para el crecimiento bacteriano, ii) que es-

tuvieran presentes en todos los genomas anotados para *S. aureus* y iii) que presentaran evidencia experimental publicada. Después del proceso de filtrado, la población final de estudio quedó conformada por las siguientes seis enzimas: Transquetolasa (EC 2.2.1.1), Hidroximetilbilano sintetasa (EC 2.5.1.61), Metionina adenosiltransferasa (EC 2.5.1.6), UDP-N-acetilglucosamina 1-carboxiviniltransferasa (EC 2.5.1.7) (UDP), Fosfohistidina fosfotransferasa (2.7.1.69) y Acetil-Coa carboxilasa (EC 6.4.1.2).

Posteriormente se utilizó la herramienta Blast para hacer la búsqueda de secuencias homólogas contra los genomas (Altschul et al., 1997). Las cepas analizadas fueron las siguientes: *S. aureus* COL (número de acceso NC_002951), *S. aureus* JH1 (NC_009632), *S. aureus* JH9 (NC_009487), *S. aureus* MRSA252 (NC_002952), *S. aureus* MSSA476 (NC_002953), *S. aureus* Mu3 (NC_009782), *S. aureus* Mu50 (NC_002758), *S. aureus* MW2 (NC_003923), *S. aureus* N315 (NC_002745), *S. aureus* NCTC 8325 (NC_007795), *S. aureus* Newman (NC_009641), *S. aureus* RF122 (NC_007622), *S. aureus* USA300 (NC_007793) y *S. aureus* USA300_TCH1516 (NC_010079), usando como secuencia pregunta las enzimas pertenecientes a *S. aureus* N135. La población final fue de 28 secuencias. Los datos arrojados por Blast fueron corroborados utilizando la herramienta SMART (Letunic, Doerks, & Bork, 2006). Todas las secuencias fueron sometidas a un control empleando la herramienta Blast contra el genoma humano y el de otros mamíferos.

Alineamiento de secuencias y reconstrucción filogenética

Se utilizó la herramienta Muscle para realizar los alineamientos de secuencias de proteínas y de nucleótidos correspondientes a cada gen empleando parámetros por defecto (Edgar, 2004). Los cambios dS (sinónimos) y dN (no sinónimos) fueron computados usando la prueba Z y la prueba Tajima usando MEGA 4.0 y el servidor SNAP (Korber, 2000; Tamura, Dudley, Nei, & Kumar, 2007). La reconstrucción filogenética se desarrolló en MEGA 4.0 bajo 1000 repeticiones bootstrap empleando el modelo evolutivo Neighbor Joining. La detección de los modos de selección operantes sobre las secuencias proteicas de la enzima UDP se desarrollaron usando SELECTON (Stern et al., 2007).

RESULTADOS

El siguiente paso con las seis enzimas fue compararlas contra el proteoma humano, resultando sorprendentemente en que solo la enzima UDP cumplió los requerimientos para poder continuar con el análisis al no presentar homólogos estructurales y funcionales con el humano.

Análisis filogenético y evolutivo

El árbol filogenético muestra un evento de duplicación basal entre los diferentes alelos UDP y una ramificación que termina con los representantes actuales de la proteína, todos los nodos presentaron valores bootstrap superiores a 50% indicativo de su robustez (figura 1).

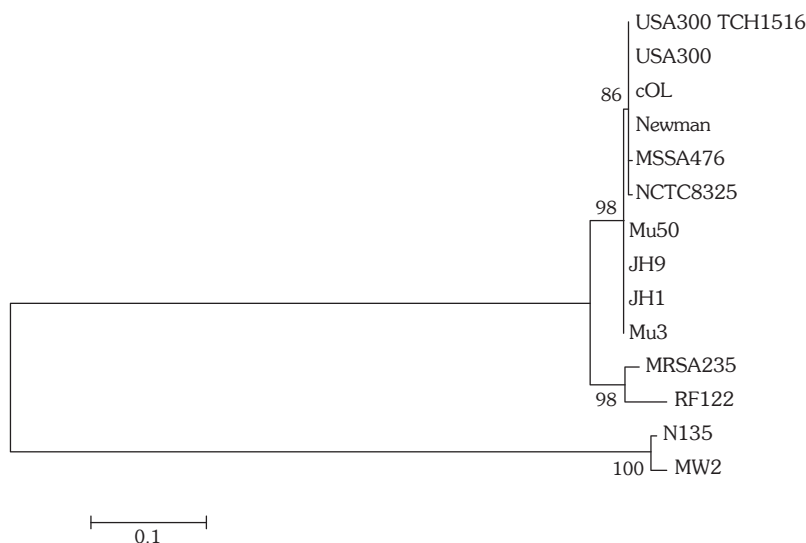
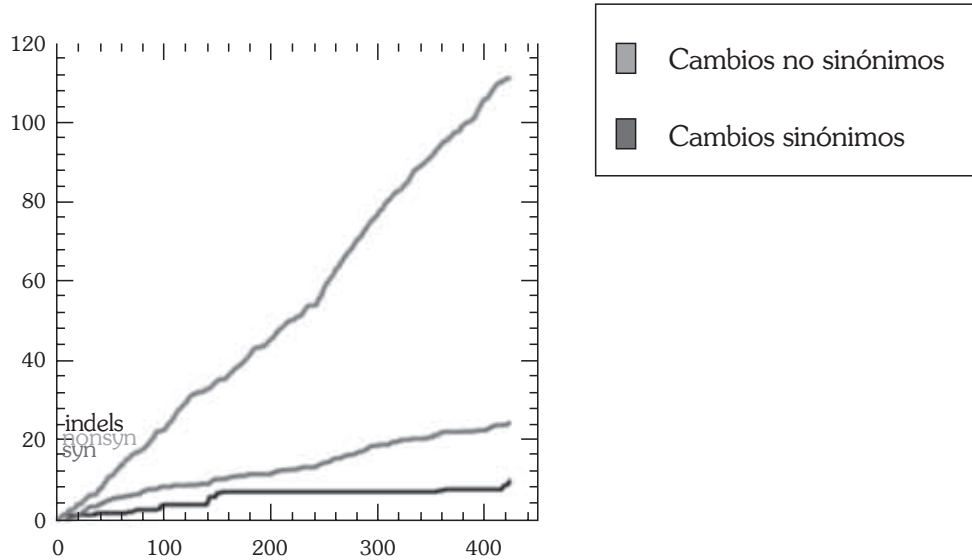


Figura 1. Reconstrucción filogenética a partir de 14 genes codificantes para la enzima UDP-N-acetilglucosamina 1-carboxiviniltransferasa de 14 cepas diferentes de *S. aureus*. Árbol filogenético desarrollado en MEGA 4.0 bajo 1000 repeticiones bootstrap y el modelo de evolución Neighbor Joining.

Los análisis estadísticos indican una fuerte tendencia hacia un modelo de selección positiva operando sobre estos genes (prueba Z 0.001, Stat 6.467), lo cual explica porqué estos genes están acumulando más cambios de tipo no sinónimo que sinónimo. De forma interesante, los datos del test Tajima (-0,594803) sugieren que esta población está saliendo o se encuentra sufriendo un efecto cuello de botella (figura

2AB y tabla 1). Además, los procesos de selección positiva ocurren específicamente en ciertos residuos lo cual es determinante a la hora de seleccionar un blanco terapéutico ya que estas posiciones muestran que residuos pueden ser fácilmente adaptables y cuáles no ante nuevas condiciones fisicoquímicas del medio externo.

A.



B.



Figura 2. Cambios sinónimos y no sinónimos y detección de tipo de selección puntual en 14 secuencias proteicas de la enzima UDP-N-acetilglucosamina 1-carboxiviniltransferasa. A) Tendencia poblacional de cambios dS y dN en 14 secuencias proteicas usando el servidor SNAP. B) Tipo de selección operante por residuo en la proteína UDP de la cepa N135 a partir de SELECTON.

Tabla 1. Parámetros evolutivos de la enzima UDP-N-acetilglucosamina 1-carboxiviniltransferasa en 14 cepas de *S. aureus*.

Prueba	Parámetros					
Tajima	M	S	Ps	θ	II	D
	14	6	0.014252	0.004482	0.002610	-1.498927
Z	Neutral/Stat	Positiva/Stat	Purificadora/Stat			
	0.165/1.396	0.077/1.436	1.000/1.409			
SNAP	dS	dN	dS/dN	pS/pN		
	0.1960	0.2004	1.0126	1.0123		

La prueba Tajima y la prueba Z fueron estimadas en MEGA 4.0 y para ambas pruebas todas las posiciones con gaps o datos faltantes fueron removidas del grupo de datos. Las abreviaturas usadas corresponden a: M = número de sitios, S = número de sitios segregantes, Ps = S/M, $\theta = Ps/a1$ y π = diversidad nucleotídica, D es el estadístico Tajima. El servidor SNAP fue usado con uso de codones para procarionotas.

DISCUSIÓN

Lee et al. (2009) propusieron seis candidatos a blancos terapéuticos basándose en análisis de reconstrucción metabólica a partir de la secuencia genómica de varias cepas de *S. aureus* resistentes a meticilina. En dicho estudio el autor plantea la necesidad de encontrar genes esenciales para el crecimiento del microorganismo fincados en el genoma de cada cepa. Cinco de los seis genes reportados en este estudio presentaron homólogos contra el proteoma humano lo cual los descalifica para seguir avanzando en análisis más detallados.

El candidato putativo restante es la enzima UDP-N-acetilglucosamina 1-carboxiviniltransferasa (UDP). Esta cataliza el sexto paso de la biosíntesis de los aminoácidos aromáticos en plantas, hongos y bacterias y pertenece a la ruta metabólica del chiquimato que no se encuentra en taxones vertebrados. Esta enzima presenta dos regiones altamente conservadas. La primera, es parte del sitio activo y se encuentra asociada a la resistencia a glifosato y la segunda localizada en la región C terminal que contiene el residuo catalítico (lisina) importante para su actividad. Esta enzima se considera como blanco terapéutico para el desarrollo de fármacos y vacunas.

En la presente investigación se determinó que la proteína UDP sigue un modo de selección positiva, ya que tanto los datos arrojados por la prueba Z como los de la prueba Tajima, sugieren que los genes codificantes de la enzima están acumulando gradualmente cambios no sinónimos favorables para la adaptación rápida a nuevos nichos (figura 2A), mejorando así el fitness de la especie; lo cual bajo criterio del autor es

desventajoso porque el sistema inmune o el fármaco desarrollado podrían ser factores selectivos que estimulen la virulencia del patógeno. Es aquí donde es importante determinar el modo de selección operante y sus implicaciones sobre un gen antes de proponerlo como candidato vacunal o farmacológico, ya que una vacuna derivada de la proteína UDP sería considerada como una vacuna imperfecta (Gandon, Mackinnon, Nee & Read, 2001).

Cada residuo de la enzima UDP (figura 2B) tiene un peso selectivo específico, lo cual indica que no todos tienen la misma probabilidad de cambio. Una vacuna basada en péptidos debe contemplar todas las posibilidades de cambio permitidas por la termodinámica implícita de la estructura proteica y el modo de selección de los residuos que la compongan, es por este motivo, que la estrategia vacunal actual busca generar constructos multiantígenos tomando información de diferentes regiones de las proteínas. En conclusión, la enzima UDP probablemente no sea un buen candidato debido a su volatilidad evolutiva.

El hecho de identificar estos genes a partir del análisis del metaboloma predicho, es una enorme ventaja pero este tipo de análisis deben ser más robustos en términos evolutivos. Actualmente, existen varias referencias donde se ha identificado la relación entre los sistemas de vacunación y la estimulación de la virulencia en las poblaciones patógenas. El sistema inmune puede ser un factor selectivo conducente a la producción de cepas más virulentas en forma similar al modelo predador y presa. Los genes esenciales y aquellos implícitos en la transmisión hospedero-hospedero son buenos candidatos terapéuticos ya que al ser bloqueados impiden el crecimiento o la disper-

sión del patógeno, pero también cuando un candidato vacunal o terapéutico no produce la extinción total del patógeno solo aumentará la virulencia del mismo contra hospederos sin previa exposición a la infección.

Finalmente, se proponen tres características que todo blanco terapéutico o vacunal debe tener desde el punto de vista bioinformático para ser considerado en estudios experimentales: i) determinar los modos de evolución y tipos selectivos operantes sobre los genes, empleando el mayor número de genomas secuenciados que sea posible, ii) desarrollar los estudios poblacionales locales con el alelo específico para falsar la hipótesis evolutiva identificada en (i) y iii) identificar el tipo de selección operante sobre residuos puntuales de las proteínas, con el fin de identificar las posibles mutaciones de acuerdo a parámetros termodinámicos.

REFERENCIAS

- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller W. & Lipman D. J. (1997). «Gapped BLAST & PSI-BLAST: a new generation of protein database search programs». *Nucleic Acid Research* 25(17), 3389-33402.
- Ávila, C., Cashat, M., Aranda, E., León, A. R., Justiniani, N., Pérez, L., ... Herrera, E. L. (1999). Prevalencia de infecciones nosocomiales en niños: encuesta de 21 hospitales en México. *Salud pública de México*, 41(1), 18-25.
- Baptiste, K. E., Williams, K., Willams, N. J., Watret, A., Clegg, P. D., Dawson, S., ... Hart, C. A. (2005) Methicillin-resistant staphylococci in companion animals. *Emerging Infectious Diseases*, 11(12), 1942-1944.
- Becker, S. A., & Palsson, B. O. (2005). Genome-scale reconstruction of the metabolic network in *Staphylococcus aureus* N315: an initial draft to the two dimensional annotation. *BMC Microbiol.* 5(1), 8-20.
- Cheung, A. L., Projan, S. J. & Gresham. H. (2002). The genomic aspect of virulence, sepsis and resistance to killing mechanisms in *Staphylococcus aureus*. *Current Infectious Disease Reports*, 4(5), 400-410.
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy & high throughput. *Nucleic Acids Res.* 32(5):1792-1797.
- Feil, E. J., Cooper, J. E., Grundmann, H., Robinson, D. A., Enright, M. C., Berendt, T., ... Day, N. P. (2003). How clonal is *Staphylococcus aureus*? *Journal of Bacteriology*, 185(11), 3307-3316.
- Gandon, S., Mackinnon, M. J., Nee, S. & Read, A. F. (2001). Imperfect vaccines and the evolution of pathogen virulence. *Nature*, 414(6865), 751-756.
- Gould, I. M. (2005). The clinical significance of methicillin-resistant *Staphylococcus aureus*. *Journal of Hospital Infection*, 61(4), 277-82.
- Heinemann, M., Kummel, R., & Panke, S. (2005). In silico genome-scale reconstruction and validation of the *Staphylococcus aureus* metabolic network. *Biotechnology and Bioengineering*. 92(7), 850-864.
- Kluytmans, J., Belkum, A. V., & Verbrugh, H. (1997). Nasal carriage of *Staphylococcus aureus*: epidemiology, underlying mechanisms, and associated risks. *Clinical Microbiology Reviews*, 10(3), 505-520.
- Korber, B. (2000). Computational Analysis of HIV Molecular Sequences. En G. R. Allen & G. H. Learn (Eds.), *HIV Signature & Sequence Variation Analysis* (pp. 55-72). Dordrecht, Netherlands: Kluwer Academic Publishers.
- Lee, D. S., Burd, H., Liu, J., Almaas, E., Wiest, O., Barabási, A., ... Kapatral V. (2009). Comparative Genome-Scale Metabolic Reconstruction and Flux Balance Analysis of Multiple *Staphylococcus aureus* Genomes Identify Novel Antimicrobial Drug Targets. *Journal of Bacteriology* 191(12), 4015-4024.
- Letunic, I., Doerks, T., & Bork, P. (2006). SMART 6: recent updates and new developments. *Nucleic Acids Research*, 37(Database issue), D229-D232.
- Loeffler, A., Boag, A. K., Sung, J., Lindsay, J. A., Guardabassi, L., Dalsgaard, A., ... Lloyd, D. H. (2005) Prevalence of methicillin-resistant *Staphylococcus aureus* among staff and pets in a small animal referral hospital in the UK. *Journal of Antimicrobial Chemotherapy*, 56(4), 692-697.
- Lowy, F. (1998). *Staphylococcus aureus* infections. *New England Journal of Medicine*, 339(8), 520-532.

- Projan, S. I., & Novick, R. P. (1997). Molecular basis of pathogenicity. En K. B. Crosslye & G. L. Archer (Eds.), *The staphylococci in human diseases* (pp.65-109). New York, United States: Churchill Livingstone.
- Robinson, D. A., & Enright, M. C. (2003). Evolutionary models of the emergence of methicillin-resistant *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy*, 47(12), 3926-3934.
- Robinson, D. A., & Enright, M. C. (2004). Evolution of *Staphylococcus aureus* by large chromosomal replacements. *Journal of Bacteriology*, 186(4), 1060-1064.
- Secretaria Distrital de Salud. (2009). Boletín epidemiológico de infecciones asociadas al cuidado de salud. Bogotá, Colombia.
- Stern, A., Doron-Faigenboim, A., Erez, E., Martz, E., Bacharach, E., & Pupko, T. (2007). Selection 2007: advanced models for detecting positive and purifying selection using a Bayesian inference approach. *Nucleic Acids Research*, 35(2), 506-511.
- Tamura, K., Dudley, J., Nei, M., & Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* 24(8), 1596-1599.
- Tristan, A., Bes, M., Meugnier, H., Lina, G., Bozdogan, B., Courvalin, P., ... Etienne, J. (2007). Global distribution of Panton- Valentine leukocidin--positive methicillin-resistant *Staphylococcus aureus*, 2006. *Emerging Infectious Disease*, 13(4), 594-600.
- Wang, J., Kodali, S., Lee, S. H., Galgoci, A., Painter, R., Dorso, K., ... Singh, S. B. (2007). Discovery of platencin, a dual FabF and FabH inhibitor with in vivo antibiotic properties. *Proceedings of the National Academy of Sciences of the United States of America*, 104(18), 7612-7616.
- Weese, J. S., Rousseau, J., Traub-Dargatz, J. L., Willey, B. M., McGeer, A. J., & Low, D. E. (2005). Community-associated methicillin-resistant *Staphylococcus aureus* in horses and humans who work with horses. *Journal of the American Veterinary Medical Association*, 226(4), 580-5833.