



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

**ANÁLISIS EVOLUTIVO DE LOS COMPONENTES DE LA CASCADA DE
FOSFORRELEVO DE LA ESPORULACIÓN EN PROCARIOTAS**

TESIS

**QUE PARA OPTAR POR EL GRADO DE:
Maestro en Ciencias**

PRESENTA:

PAOLA ALICIA MARTÍNEZ AMADOR

TUTOR PRINCIPAL

ROSA MARÍA GUTIÉRREZ RÍOS

Instituto de Biotecnología, UNAM campus Morelos

MIEMBROS DEL COMITÉ TUTOR

MIGUEL ÁNGEL CEVALLOS

**Centro de Ciencias Genómicas,
UNAM campus Morelos**

DIEGO CLAUDIO CORTEZ QUEZADA

**Centro de Ciencias Genómicas,
UNAM campus Morelos**

Cuernavaca, Morelos

Enero 2017



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

El presente trabajo se realizó bajo la asesoría de la Dra. Rosa María Gutiérrez Ríos y la supervisión técnica del Dr. Carlos Daniel Velázquez Hernández en el laboratorio adscrito al Departamento de Microbiología Molecular del Instituto de Biotecnología de la Universidad Nacional Autónoma de México campus Morelos.

Este proyecto se realizó con recursos del programa de apoyo de la dirección general de asuntos del personal académico (DGAPA) bajo el programa de apoyo a proyectos de investigación e innovación tecnológica (PAPIIT) con número de proyecto IN204515 y durante mis estudios de maestría fui becada por el Consejo Nacional de Ciencia y Tecnología (CONACyT) bajo el programa de maestría y doctorado en ciencias bioquímicas.

DEDICATORIA

Para mis abuelas

Felisa Valencia Arista

Irma Yolanda Zaragoza López

Sus enseñanzas las llevo en el corazón.

AGRADECIMIENTOS

Al Dr. Enrique Merino por darme la oportunidad de integrarme a la unidad de investigación a su cargo, por sus valiosos consejos académicos, por la calidad de su trato y por ser un ejemplo a seguir.

A la Dra. Rosa María Gutiérrez Ríos por todo el apoyo brindado, por la confianza depositada en mí desde el inicio de este trabajo, por toda su comprensión y guía y por ser una mujer profesional admirable.

Al Dr. Carlos Daniel Velázquez Hernández por enseñarme con infinita paciencia, por estar presente en cada tropiezo y por su invaluable amistad.

Al Dr. Miguel Ángel Cevallos y Dr. Claudio Diego Cortez quienes enriquecieron mi trabajo con sus valiosas opiniones y correcciones.

Mtro. Alejandro Abdalá y Dr. Esteban Peguero por toda la disposición prestada y por inspirarme a buscar siempre la excelencia en todo.

A todos mis compañeros del laboratorio por los momentos tan gratos que he compartido a su lado, por sus enseñanzas y por hacer de mi estancia una experiencia inolvidable rodeada siempre de un ambiente de trabajo estupendo.

A mi familia porque me han apoyado en todas las actividades y metas que me he propuesto y me han ayudado con mucho amor todos los días.

A mis amigos del IBT porque lo único que me llevo son los bellos momentos que viví a su lado y porque me enseñaron a disfrutar la vida entre científicos.

ÍNDICE GENERAL

DEDICATORIA.....	II
AGRADECIMIENTOS.....	III
INDICE DE TABLAS Y FIGURAS.....	V
LISTA DE ABREVIATURAS.....	VII
RESUMEN.....	1
ABSTRACT.....	2
INTRODUCCIÓN.....	3
Sistemas de transducción de señales.....	3
Proteínas sensoras histidina-cinasas.....	5
Dominios sensores de las histidina-cinasas.....	6
Reguladores de respuesta.....	7
<i>Bacillus subtilis</i> como Modelo de estudio.....	8
Los componentes de la cascada de fosforrelevo en <i>Bacillus subtilis</i>... 8	8
Regulación de la cascada de fosforrelevo en <i>Bacillus subtilis</i>.....	10
Destinos celulares en <i>Bacillus subtilis</i>.....	11
Evolución de proteínas multidominio.....	13
Dinámica evolutiva de los sistemas de dos componentes.....	14
El modelo Clostridia.....	19
PLANTEAMIENTO DEL PROBLEMA.....	21
JUSTIFICACIÓN	22
PREGUNTA DE INVESTIGACIÓN.....	23
OBJETIVO GENERAL	23
OBJETIVOS PARTICULARES	23
HIPÓTESIS	23
MÉTODOS	24
ESQUEMA DE LA ESTRATEGIA EXPERIMENTAL.....	26
RESULTADOS	27
DISCUSIÓN	42
CONCLUSIONES	50
PERSPECTIVAS	50
BIBLIOGRAFÍA	51

ÍNDICE DE TABLAS Y FIGURAS

Figura 1. A) Esquema de los sistemas de transducción de señales de dos componentes (rojo) y de fosforrelevo (azul) **B)** Diversidad de arquitecturas de proteínas que conforman los sistemas de transducción de señales. p. 4

Figura 2. Esquema general de la organización de dominios en las proteínas histidina-cinasas y en reguladores de respuesta. p. 7

Figura 3. A) Componentes de la cascada del fosforrelevo en *B. subtilis* B) Arquitecturas de las cinco cinasas de *B. subtilis*, las barras negras representan dominios transmembranales. p. 9

Figura 4. Representación esquemática de los distintos tipos celulares en las comunidades de *Bacillus subtilis*. p. 12

Figura 5. Evolución de los sistemas de dos componentes. a) sistema canónico de dos componentes: se expande y diversifica a través de la duplicación de genes y la subsecuente divergencia b) dos posibles modelos para el origen del fosforrelevo de la esporulación en *B. subtilis*. p. 17

Figura 6. Arquitecturas de los modelos de sistemas de transducción de señales probados experimentalmente en Bacilli y Clostridia. p. 27

Figura 7.1 Contexto genómico de las cinco H-cinasas huérfanas de *Bacillus subtilis*. p. 28

Figura 7.2 Contexto genómico del regulador Spo0F en Bacilli y Clostridia. p. 29

Figura 7.3 Contexto genómico de la fosfotransferasa Spo0B en Bacilli. p. 29

Figura 7.4 Contexto genómico del regulador de respuesta Spo0A en Bacilli y Clostridia. p. 30

Figura 7.5 Contexto genómico de la fosfatasa Sda en Bacilli. p. 30

Figura 7.6 Contexto genómico de la fosfatasa Kipl en Bacilli y Clostridia. p. 31

Figura 7.7 Contexto genómico de la fosfatasa Spo0E en Bacilli. p. 31

Figura 7.8 Contexto genómico de las fosfatasas Rap en Bacilli y Clostridia. p. 32

Figura 8. Distribución en procariotas de dominios homólogos a los sistemas de transducción de señales descritos en Bacilli y Clostridia. p. 34

Figura 9. H-cinasas huérfanas por phylum. Total de H-cinasas huérfanas por bacteria en cada phylum ordenados por distancia evolutiva 16s. p. 36

Figura 10. H-cinasas huérfanas por Clase en Firmicutes. Total de H-cinasas huérfanas por bacteria en cada Clase. Clasificadas como esporulantes o no esporulantes, no se sabe y bacterias que tienen cinasas huérfanas pero no tienen ningún elemento de la cascada de fosforrelevo. p. 37

Figura 11. Distribución en Bacilli y Clostridia de las H-cinasas homólogas a los modelos propuestos y clasificadas de acuerdo a su capacidad de esporular. p. 38

Figura 12. Conservación de los elementos del fosforrelevo en Firmicutes. Diversidad de arquitecturas de la vía completa en las clases Bacilli, Clostridia, Erisipelotrichia y Negativicutes. p. 39

Figura 13. Filogenia 16s ribosomal de la conservación de los elementos de la cascada de la esporulación en las Clases Bacilli (rojo), Clostridia (azul), Erisipelotrichia (rosa) y Negativicutes (verde). p. 40

Figura14.- Formación de la cascada del fosforrelevo en Firmicutes. p. 48

LISTA DE ABREVIATURAS

ADN	Ácido desoxirribonucleico	TM	Dominios transmembranales
Asp	Residuo aspartato		
ATP	Adenosin trifosfato		
<i>B. subtilis</i>	<i>Bacillus subtilis</i>		
CA	Dominio catalítico y de unión a ATP	Acidobacteria	Ac
CE	Combinatorial Extension	Actinobacteria	At
COG	Clusters of Orthologous Groups of proteins	Alphaproteobacteria	A
DHp	Dominio de dimerización y fosfotransferasa	Aquificae	Aq
FSF's	Superfamilias	Armatimonadetes	Ar
GAF	cGMP-fosfodiesterasas, adenilato ciclasas y FhIA	Bacteroidetes	Ba
GeneBank	Nucleotide database	Betaproteobacteria	B
GO	Ontología genética	Caldiserica	Cd
H-cinasa	Histidina-cinasa	Chlamydiae	Cm
His	Residuo histidina	Chlorobi	Cb
HK	Histidina-cinasa	Chloroflexi	Cf
HPt	Proteína fosfotransferasa	Chrysiogenetes	Cf
KEGG	Kyoto Encyclopedia of Genes and Genomes	Cyanobacteria	Cy
MIPS	Munich Information Center for Protein Sequences	Deferribacteres	Df
OCSs	Sistemas de un componente	Deinococcus-Thermus	Dt
PAS	Dominios Per-ARNT-Sim	Deltaproteobacteria	D
Phrs	Pentapéptidos	Dictyoglomi	Dc
Rap	Fosfatasa aspartil-fosfato	Elusimicrobia	El
REC	Dominio receptor del regulador de respuesta	Epsilonproteobacteria	E
RR	Regulador de respuesta	Fibrobacteres	Fb
SCOP	Clasificación Estructural de las proteínas	Firmicutes	Fi
Spo0A~P	Spo0A en su forma fosforilada	Fusobacteria	Fu
TCS	Sistemas de dos componentes	Gammaproteobacteria	G
		Gemmatimonadetes	Ge
		Nitrospirae	Nt
		Planctomycetes	Pl
		Spirochaetes	S
		Synergistetes	Sy
		Tenericutes	T
		Thermodesulfobacteria	Th
		Thermotogae	Tt
		Verrucomicrobia	V

(Salvado et al. 2015)

RESUMEN

En el presente trabajo, se analizaron proteínas con “dominios homólogos” a los presentes en cada componente de la cascada de fosforrelevo para esporular de *Bacillus subtilis*. Para ello, se utilizó la base de datos de procariotas de nuestro grupo, denominada Gene Context Tool III, que almacena los ~3000 organismos totalmente secuenciados disponibles en la base de datos KEGG y GeneBank. Se clasificaron las arquitecturas modulares de cada proteína en base al tipo de dominios que las conforman y se evaluó su variabilidad y prevalencia dentro de cada phylum bacteriano.

Se encontraron 12 H-cinasas huérfanas con arquitecturas distintas involucradas en el proceso de esporulación tanto en Bacilli como en Clostridia, algunas de ellas están muy distribuidas a lo largo de casi todos los phyla bacterianos, y otras se conservan únicamente en Firmicutes. Encontramos una marcada preferencia por las arquitecturas pequeñas y poco complejas, y algunos de los dominios más conservados son el supradominio catalítico, el dominio sensor PAS y en pocos casos los dominios son variables. En contraste, los dominios de las transferasas y fosfatasas no presentan dominios variables, con la excepción del dominio Response_reg en Spo0F que resultó ser muy versátil.

Se observó una amplia distribución de la proteína Spo0F, Kipl y Rap, contrario a la fosfotransferasa Spo0B, el regulador Spo0A y las fosfatasas Sda y Spo0E que resultaron ser exclusivas de Firmicutes. Adicionalmente, en este mismo Phylum se descubrieron las formas reducidas de la proteína Spo0B denominada SPOB_a, y de Spo0A denominada Spo0A_C, como ancestros de las proteínas completas.

Finalmente, evaluamos la evolución de los componentes de la vía en bacterias clasificadas en formadoras de esporas y no formadoras de esporas. La filogenia explica claramente como la trayectoria evolutiva de la vía de señalización inicia desde un sistema de dos componentes en Clostridia hacia una cascada de fosforrelevo en Bacilli hasta llegar a organismos que cuentan con todos los complejos niveles de control.

ABSTRACT

In this study, we analyzed proteins with "homologous domains" to those present in each component of the sporulation phosphorelay of *Bacillus subtilis*. To this end, from the prokaryotic database Gene Context Tool III developed in our group, which stores the ~ 3000 fully sequenced organisms available in the KEGG and GeneBank database. We extracted the modular architectures of each protein. The extracted proteins were classified based on the type of domains that comprise them and their variability and prevalence were evaluated within each bacterial phylum.

Twelve orphan H-kinases with different architectures involved in the sporulation initiation were found in both Bacilli and Clostridia, some of which are very distributed throughout almost all bacterial phylum, and others are only conserved in Firmicutes. We found a marked preference for small and simple architectures, some of the most conserved domains are the catalytic supradomain, the PAS domain and in few cases domains are variable. In contrast, domains of phosphotransferases and phosphatases do not have variable domains, except Spo0F which it's such a versatile domain.

A broad distribution of the Spo0F, KipI and Rap proteins was observed, contrary to the phosphotransferase Spo0B, the response regulator Spo0A and Sda and Spo0E phosphatases that turned out to be exclusive to Firmicutes. Additionally, in same Phylum the reduced proteins SPOB_a and Spo0A_C were discovered as ancestors of the complete proteins.

Finally, we evaluated the evolution of pathway components in bacteria classified as spore forming and non-spore forming. The phylogeny clearly explains how the evolutionary trajectory of the signaling pathway starts from two-component system in Clostridia to a phosphorelay in Bacilli to reach organisms that have all the complex levels of control.

INTRODUCCIÓN

Sistemas de transducción de señales

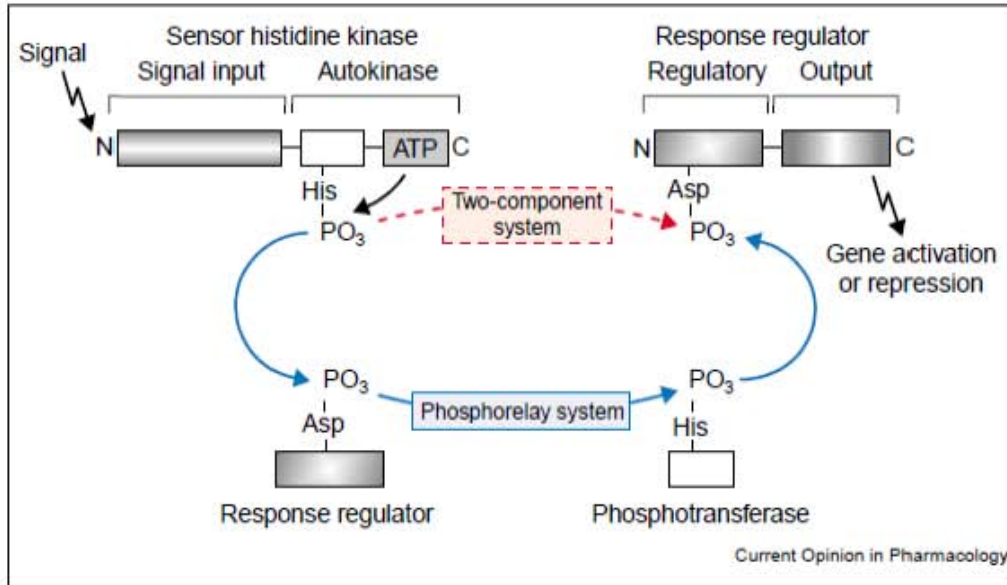
El mecanismo para la transducción de señales consiste en reconocer señales exteriores y convertir esta información a una entidad química (un grupo fosforilo) que modifica la actividad funcional de las proteínas; de esta manera las bacterias pueden responder a una enorme gama de señales y factores de estrés (Capra and Laub 2012; Fabret, Feher, and Hoch 1999).

En conjunto, las arquitecturas de las histidina-cinasas (HKs), reguladores de respuesta (RR) e histidina-cinasas híbridas (HKRR) representan el 94% de todos los sistemas de dos componentes (TCS) y fosforrelevo. Existen ~50 formas alternativas en las que los dominios HK, HPt (fosfotransferasa) y RR se organizan en proteínas individuales para generar diversos diseños de transducción de señales. La siguiente clasificación es la más representativa:

- 1) **Sistemas de dos componentes:** se compone de una proteína receptora HK que detecta una señal de entrada específica y la traduce en una respuesta mediante la autofosforilación en un residuo de histidina (His) conservado. Este grupo fosforilo se transfiere entonces a un aspartato conservado (Asp) de una proteína RR afín, que una vez fosforilada, conduce a un cambio en la expresión génica (Figura 1A). (Capra and Laub 2012)
- 2) **Cascada de fosforrelevo:** El sistema más largo hasta ahora descrito es de cuatro componentes, este tipo de cascadas conforma redes reguladoras complejas que permiten mayores niveles de control y facilitan la interpretación de múltiples señales positivas y negativas. En esta vía, la detección de señal también está mediada por una HK, que fosforila en un residuo Asp a un RR que carece de un dominio de unión a ADN (REC); éste a su vez, transfiere el grupo fosfato a un residuo His de una fosfotransferasa (HPt), que finalmente, dona el grupo fosfato a un RR terminal que actúa como factor de transcripción (Figura 1A) (Capra and Laub 2012).
- 3) **Formas híbridas:** Alrededor del 25% de las cascadas de fosforrelevo utilizan proteínas híbridas donde las proteínas HK, RR y, a veces, el dominio HPt son contiguas dentro de la misma cadena polipeptídica para generar un fosforrelevo más

sofisticado. La abundancia de formas híbridas oscila entre menos del 1% en Firmicutes y más del 10% en Cyanobacterias de todas las proteínas TCS y de fosforrelevo (Fabret, Feher, and Hoch 1999; Salvado et al. 2015)(Figura 1B).

A)



B)

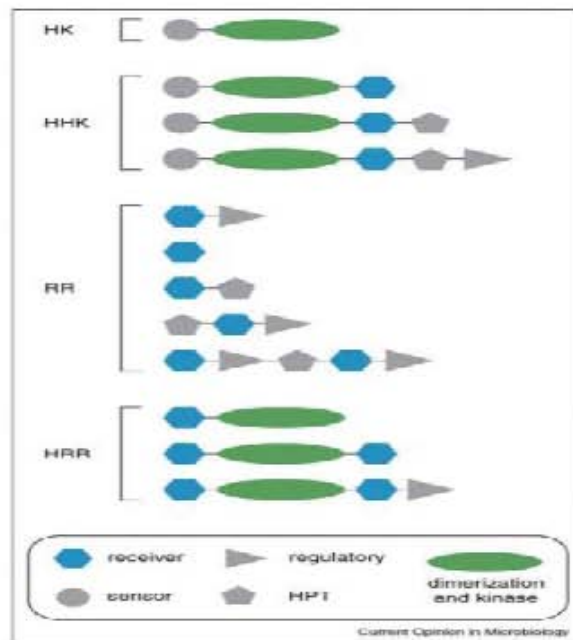


Figura 1. A) Esquema de los sistemas de transducción de señales de dos componentes (rojo) y de fosforrelevo (azul) (Keith Stephenson and Hoch 2002a) **B)** Diversidad de arquitecturas de proteínas que conforman los sistemas de transducción de señales (Capra and Laub 2012).

Cabe mencionar que en eucariotas la transducción de señales es del tipo fosforrelvos serina/treonina. Sin embargo, los sistemas His/Asp también han sido identificados en hongos, plantas y mohos mucilaginosos (Barba-Ostria 2014), inclusive en algunas arqueobacterias (Ashby 2006).

Proteínas sensoras histidina-cinasas

Las histidina-cinasas (HKs) sensan una señal exterior y la transmiten a su RR afín, se extienden a lo largo de la membrana citoplásmica donde transducen señales tanto extracitoplásmicas como citoplásmicas, y pueden tener actividad de autocinasa o fosfatasa (Stewart 2010).

Las HKs se componen de dos grandes dominios 1) El dominio autocinasa que alberga un residuo His altamente conservado y se compone por el dominio de dimerización y fosfotransferasa (DHP) y el dominio catalítico y de unión a ATP (CA); 2) El dominio sensor N-terminal que es sumamente variable y se clasifican topológicamente como extracitoplásmico, transmembranal (TM) y citoplásmico (Figura 2) (Capra and Laub 2012; Fonslow et al. 2013; Jung et al. 2012).

En procariontes, del 60% - 90% de los genes TCS tienden a agruparse en operones como pares de HK-RR, lo que permite la duplicación o transferencia lateral de una vía de señalización completa. Por otro lado, ~20% de los genes son de tipo huérfanos, es decir, HK y RR no se encuentran juntos en el genoma, esto debido a que sólo un gen de un operón se duplica o ambos se duplican y uno se pierde (Salvado et al. 2015). Sin embargo, un gen huérfano de una HK puede conservar la capacidad de fosforilar el RR del operón de la que se derivó. Tales eventos de duplicación, junto con un cambio en el dominio de entrada de la cinasa, son un mecanismo esencial para la generación de sistemas cruzados de regulación en las que múltiples señales independientes pueden desencadenar la misma respuesta. Del mismo modo, la duplicación de sólo el RR de un par HK-RR puede conducir a un escenario en el que un solo sensor HK puede generar múltiples respuestas. Un ejemplo clásico lo encontramos en la cascada de fosforrelvo para la esporulación en *B. subtilis*. (Capra and Laub 2012; Wu et al. 2013)

Dominios sensores de las histidina-cinasas

A pesar de los avances y décadas de estudios, las identidades exactas de las moléculas activadoras de las HK todavía siguen siendo desconocidas (Wu et al. 2013). En general, se ha observado que las arquitecturas modulares son de gran importancia, pues los dominios sensores están separados estructuralmente de los dominios catalíticos por una región de engarce que evita el contacto directo entre ellos. Esto es evolutivamente relevante porque las mutaciones en los dominios sensores no tendrán efectos perjudiciales en los dominios catalíticos, generando así una variedad de especificidades de ligandos que conservan los módulos reguladores corriente abajo en el genoma (Schultz and Natarajan 2013).

Algunos dominios sensores que han sido descritos funcionalmente son:

Dominios sensores globulares: Algunos ejemplos de dominios citoplásmicos incluyen los dominios GAF (~9% de HK) que reconocen ficocianobilina, formato y iones metálicos. Los dominios PAS (33% de HK) que unen flavina y utilizan sus propiedades redox y fotosensibles para sensor oxígeno y luz. La ortología y arquitectura tridimensional de los dominios PAS se conserva ampliamente desde bacterias hasta plantas y animales superiores, la repetición en tándem de este dominio pudiera ser un poderoso mecanismo para filtrar respuestas, amplificar la señal o detectar múltiples señales, inclusive péptidos (Henry and Crosson 2011; Narikawa et al. 2004; Taylor and Zhulin 1999; Vreede et al. 2003). Otros como HNOB y HNOBA usan grupos prostéticos o iones para sensor óxido nítrico, y el dominio PocR une pequeños hidrocarburos. Los dominios de detección extracitoplásmicos incluyen el dominio PBPb de unión a solutos, el dominio CACHE que une glicerol, citrato, azúcares y iones metálicos; y una serie de dominios CHASE que son miembros divergentes de la superfamilia CACHE. La mayoría de estas superfamilias están relacionadas evolutivamente a nivel de "fold" (Stephen Spiro and Ray Dixon 2010).

Dominios receptores transmembranales: Estos incluyen una variedad de receptores con múltiples dominios TM (7TM, 8TM y 5TM) que pueden sensor estímulos como la luz vía el grupo prostético rodopsina. **Dominios transmisores de señal:** Esta categoría incluye hélices que transmiten la señal en forma de cambio conformacional de un dominio a otro. Algunos ejemplos son el dominio HAMP (~ 31% de HK) y el módulo S-helix (Stephen Spiro and Ray Dixon 2010).

Reguladores de respuesta

La mayoría de los RR (63%) consisten de dos dominios: El dominio N-terminal REC que recibe el grupo fosfato de la HK y el dominio efector C-terminal con actividad de unión al ADN. Algunos ejemplos de dominios efectores son el winged-helix, helix-turn-helix y LytTR, otros dominios enzimáticos no bien caracterizados incluyen AAA+ (Actividad ATPasa), GGDEF (nucleótido ciclasa), EAL (guanilato fosfodiesterasa) y HD-GYP (fosfodiesterasa). Por otro lado, el 17% de los RR existen como dominios REC independientes y tienen las actividades enzimáticas de fosfotransferencia y autodefósforilación, pero no regulan directamente las respuestas efectoras y son funcionalmente análogos a los dominios REC situados dentro de las HKs híbridas. El resto de los RR son conservados en proteínas multidominio (Fonslow et al. 2013; Galperin, Nikolskaya, and Koonin 2001).

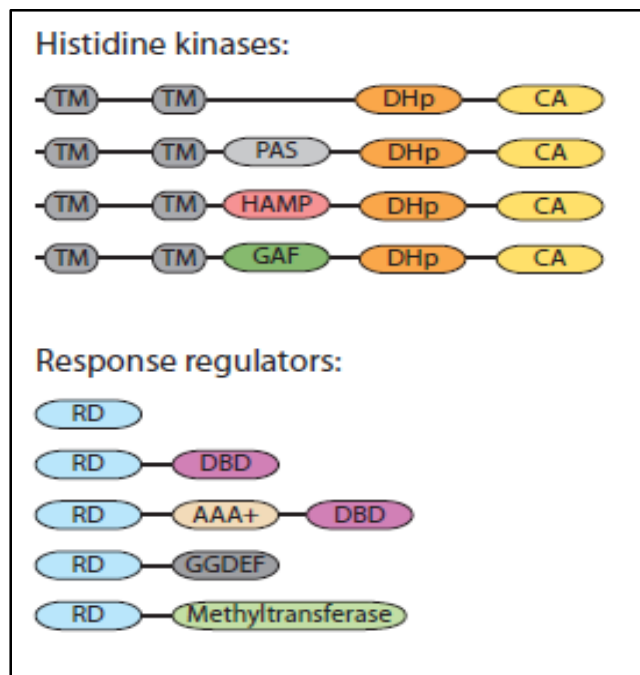


Figura 2. Esquema general de la organización de dominios en las proteínas histidina-cinasas, los dominios transmembranales (TM) en gris, los dominios catalíticos en naranja y los dominios sensores en diversos colores. En los reguladores de respuesta el dominio receptor se representa en azul (RD) y los dominios efectores en diversos colores (Capra and Laub 2012).

***Bacillus subtilis* como Modelo de estudio**

Bacillus subtilis (*B. subtilis*) es un bacilo Gram positivo, móvil y aerobio, que pertenece al género Bacilli del *Phylum* Firmicutes, capaz de crecer en ambientes terrestres y acuáticos incluyendo suelos, raíces de plantas y el tracto gastrointestinal de animales.

Una de las características más importantes de *B. subtilis* es que tiene propiedades únicas de diferenciación celular, tales como la formación de biopelículas, canibalismo (González-Pastor 2011) y, especialmente, es formador de estructuras de resistencia denominadas endosporas.

El inicio de la esporulación en *B. subtilis* es regulado por una cascada de fosforrelevo ampliamente estudiada (Figura 3), esta cascada tiene la ventaja de tener múltiples pasos de fosfotransferencia que aumentan la precisión de la respuesta, actuando como un circuito de integración de señales positivas y negativas para la toma de decisiones hacia los distintos tipos de diferenciación celular (Burbulys, Trach, and Hoch 1991; Fabret, Feher, and Hoch 1999).

A pesar de que *B. subtilis* ha sido un poderoso modelo para estudiar los determinantes morfológicos, bioquímicos y genéticos de la esporulación todavía se desconoce cómo se regula la entrada a esta y cómo ha evolucionado a nuevos esquemas de producción de esporas. (Earl et al. 2010; de Hoon, Eichenberger, and Vitkup 2010; Hutchison, Miller, and Angert 2014)

Los componentes de la cascada de fosforrelevo en *Bacillus subtilis*

En *B. subtilis* existen cinco HK que alimentan el flujo de fosfatos a través de la cascada de fosforrelevo denominadas KinA, KinB, KinC, KinD y KinE (miembros de la clase IIIB), cada una de las cinco HK puede fosforilar al REC **Spo0F** en su Asp54. Esta proteína consta de un solo dominio y su función es donar el grupo fosfato a la fosfotransferasa dimérica **Spo0B** en su His30, la cual nuevamente transfiere el grupo fosfato al factor de transcripción **Spo0A** en un residuo Asp situado en la porción N-terminal (Figura 3). La fosforilación de ésta última proteína (**Spo0A~P**) promueve un cambio conformacional para pasar de su estado dimérico (no fosforilado) al estado monomérico (fosforilado) y estimular, a través de su dominio C-

terminal, el reconocimiento del sitio de unión a ADN conocido como caja 0A (Figura 3A) (Driks 2002; Fabret, Feher, and Hoch 1999; Hoch and Varughese 2001; Lewis et al. 2000; Keith Stephenson and Hoch 2002a; Varughese 2002).

KinA y KinE son HKs citoplasmáticas que contienen tres dominios PAS y responden a la inducción de la esporulación (Figura 3B). Aunque el papel exacto de cada dominio PAS se desconoce, es probable que puedan ser sinérgicos o antagonistas o inducir la formación de tetrámeros (P. Eswaramoorthy and Fujita 2010; Prahathees Eswaramoorthy et al. 2011; K Stephenson and Hoch 2001; L. Wang et al. 2001). KinC contiene un único dominio PAS que se activa en respuesta a la pérdida de potasio; y KinD es la única cinasa que cuenta con un segmento extracitoplasmático que le confiere más capacidad de responder a ligandos como piruvato, propionato, butirato y malato que se acumulan o agotan en el exterior (Wu et al. 2013). Se sabe que sólo KinA y KinB son necesarios para iniciar la esporulación al generar altos niveles de Spo0A~P. En contraste, KinC, KinD y KinE juegan un papel en la transición del crecimiento exponencial a la fase estacionaria y en la producción de biopelículas mediante la acumulación de bajos niveles de Spo0A~P (Devi et al. 2015; McLoon et al. 2011). Sin embargo, cuando la vía con KinA y KinB está inhibida, KinC es capaz de iniciar la esporulación. Curiosamente, solo KinC y KinD pueden ser capaces de fosforilar directamente a Spo0A y así evitar los intermediarios Spo0F-Spo0B (Phillips and Strauch 2002). Adicionalmente, se ha demostrado que KinD actúa como fosfatasa en ausencia de piruvato, y se cree que puede mantener bajos niveles de Spo0A~P para retrasar la esporulación (Jiang et al. 2000; Wu et al. 2013).

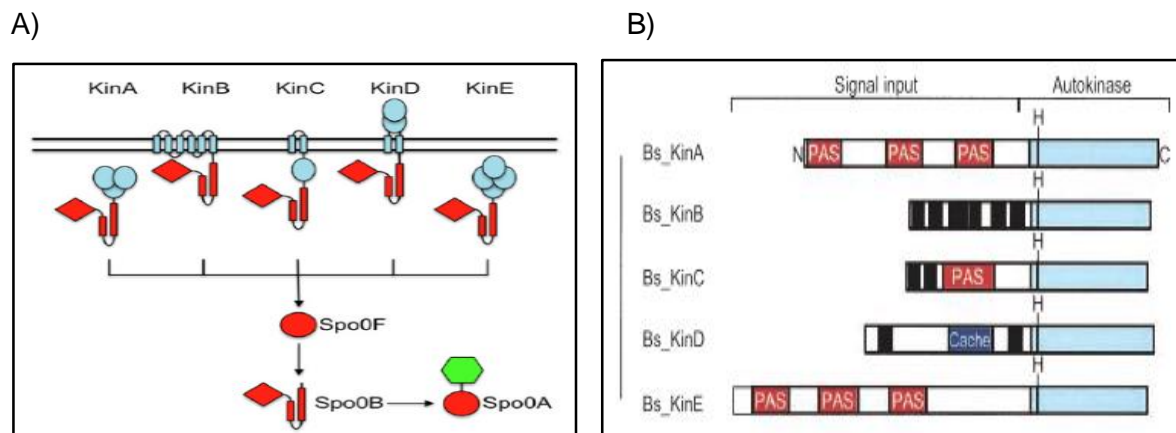


Figura 3. A) Componentes de la cascada del fosforrelievo en *B. subtilis* (Wu et al. 2013) **B)** Arquitecturas de las cinco cinasas de *B. subtilis*, las barras negras representan dominios transmembranales. (Keith Stephenson and Hoch 2002b)

Regulación de la cascada de fosforrelevo en *Bacillus subtilis*

Las fosfatasas de los sistemas de transducción de señales son un grupo de proteínas que modulan la actividad de las HK y de los RR afectando el estado de fosforilación y determinan cuándo y por cuánto tiempo estarán activos. Las fosfatasas están presentes tanto en bacterias Gram-positivas como Gram-negativas y son críticas para una variedad de funciones fisiológicas incluyendo la esporulación, la competencia, resistencia a los antibióticos y la transición a la fase estacionaria. Además, generan circuitos de retroalimentación y actúan como puntos de control.

La regulación por fosfatasas es crítica para la heterogeneidad de poblaciones genéticamente idénticas en *B. subtilis* (Mitrophanov and Groisman 2008). El flujo de fosfatos a través de la cascada de fosforrelevo al inicio de la esporulación es regulado por la acción de varias fosfatasas como las proteínas Rap, Spo0E, KipI y Sda que son capaces de drenar los grupos fosforilo de la cascada manteniendo a la célula en un estado de suspensión (fosforilación y defosforilación a tasas casi iguales) hasta alcanzar gradualmente una concentración umbral de Spo0A~P, que sólo se alcanza cuando las fosfatasas son antagonizadas, permitiendo así la integración de señales para evitar un inicio innecesario de la esporulación (Fujita and Losick 2005; Piggot 1996; Sonenshein 2000).

La fosfatasa Sda provoca un retraso en el inicio de la esporulación evitando la autofosforilación de KinA y KinB mediante la interacción entre los dominios catalíticos y de fosfotransferencia, es regulada por DnaA cuando el inicio de la replicación es perturbada (Burkholder, Kurtser, and Grossman 2016; Rowland et al. 2004). A diferencia de Sda, KipI es inhibidor de la reacción de autofosforilación de KinA y no tiene ningún efecto sobre KinB, es antagonizado por KipA y es regulado por la disponibilidad de nitrógeno y de ciertos azúcares, probablemente bloquea la esporulación en ambientes ricos en nutrientes (Bick et al. 2009; Cunningham and Burkholder 2009; Ling Wang et al. 1997).

La familia de las proteínas Rap (fosfatasas aspartil-fosfato) está integrada por 11 proteínas Rap que regulan la señalización en los RR afines mediante la defosforilación o bloqueando la actividad de unión al ADN. Se sabe que RapA, RapB, RapE y RapH promueven la defosforilación del regulador de respuesta Spo0F~P. Estas proteínas están sujetas al control de σ^H y sus antagonistas son los pentapéptidos (Phrs) PhrA, PhrE y PhrH

respectivamente codificados en operón. Las proteínas Rap previenen la ocurrencia simultánea de dos alternativas fisiológicas como competencia y esporulación, y se ha sugerido que el transporte de Phr's representa los puntos de control y sincronización de la tasa de acumulación de Spo0A~P (Gallego del Sol and Marina 2013; López and Kolter 2010; Lopez, Vlamakis, and Kolter 2009; Tojo, Hirooka, and Fujita 2013; Ling Wang et al. 1997). La regulación del RR Spo0A~P es realizada por los miembros de la familia Spo0E, YisI y YnzD. A bajas concentraciones, Spo0A es un regulador negativo que reprime la transcripción de AbrB que a su vez evita la expresión de los promotores *spo0H* (σ^H) y *spo0E*, también se ha demostrado su regulación por el inicio de la replicación (Burbulys, Trach, and Hoch 1991; Ireton and Grossman 1994).

Destinos celulares en *Bacillus subtilis*

Como las personas, por lo general las células “deliberan” con mucho cuidado antes de comprometerse a cambios drásticos en su estilo de vida (Stephens 1998). Algunos organismos son capaces de diferenciarse en subpoblaciones de células fenotípicamente distintas, pero genéticamente idénticas. La diferenciación de distintos tipos celulares es un mecanismo eficiente en *B. subtilis* para adaptar su desarrollo a un nicho ecológico particular en el que coexiste con otros microorganismos, además, la división del trabajo permite la alta optimización de los recursos (De Jong, Veening, and Kuipers 2010). *B. subtilis* posee al menos tres reguladores maestros de respuesta: Spo0A, DegU y ComA que en su forma fosforilada (DegU~P, ComA~P, Spo0A~P) activan la expresión de un subconjunto de genes necesarios para la diferenciación a un linaje celular específico. (Lopez, Vlamakis, and Kolter 2009)

El RR DegU~P provoca la producción y secreción de proteasas, que conduce a la diferenciación de una subpoblación de células denominadas mineros. Del mismo modo, ComA~P induce la vía hacia la producción de surfactina, una fracción de los productores de surfactina se somete a un proceso secundario de diferenciación para convertirse en células competentes. Por último, cuando los niveles de Spo0A~P son bajos se activan las cascadas hacia la producción de polisacáridos y proteínas que generan la matriz extracelular (Vlamakis et al. 2013) y el canibalismo mediante la segregación de toxinas que matan a sus hermanos sensibles para ser utilizados como alimento; y cuando hay altos

niveles de Spo0A~P se inducen más de 125 genes involucrados en esporulación (López and Kolter 2010; Stragier and Losick 1996).

Algunas de las señales implicadas en la inducción de la esporulación son: el aumento de población celular, en consecuencia la falta de nutrientes desencadena la acumulación de Spo0A~P y la neutralización de las fosfatasa del fosforrelevo (Sonenshein 2000). Spo0A~P actúa como regulador positivo en la transcripción de genes esenciales de esporulación (operones *spolIA*, *spolIG*, *spolIE*) que inician la cascada de activación de seis diferentes factores σ (subunidades de ARN polimerasa): σ^A , σ^H , σ^F , σ^E , σ^G y σ^K que culminan con la formación de la endospora (Brown et al. 1994; Fujita and Losick 2005; Phillips and Strauch 2002).

Las endosporas se desarrollan en condiciones de estrés o desfavorables, son capaces de sobrevivir durante años a la acción de sustancias químicas, calor, desecación, disturbios mecánicos, radiación UV y enzimas, se dispersan fácilmente por el viento y pueden migrar grandes distancias antes de germinar en células vegetativas cuando el ambiente es más favorable (Mongkoltharuk 2012). Son de gran importancia biológica, pues han sido utilizadas como biosímetros, fungicidas, bactericidas, fertilizantes, probióticos, simbiontes y confieren efectos benéficos para la salud en alimentos debido a que el 30% de la microbiota intestinal total son bacterias formadoras de esporas (Browne et al. 2016; Garrity et al. 2004; Nicholson 2002).

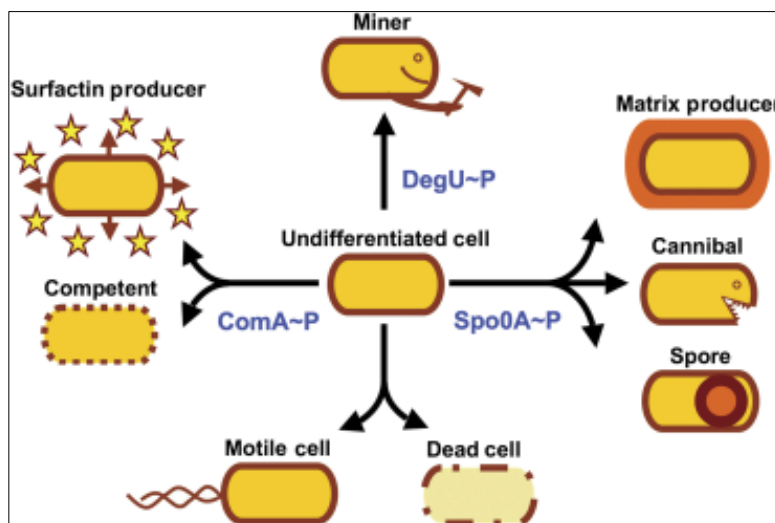


Figura 4. Representación esquemática de los distintos tipos celulares en las comunidades de *Bacillus subtilis*. (López and Kolter 2010)

Evolución de proteínas multidominio

Las proteínas se componen de una o varias unidades evolutivas llamadas “dominios”, estos determinan la función y las relaciones evolutivas de las proteínas, por lo tanto, es importante entender los principios de las combinaciones de dominios y las interacciones entre ellos.

Las proteínas multidominio se forman por duplicación, divergencia y recombinación de dominios. Una vez que un dominio o proteína se ha duplicado, puede evolucionar a una función nueva, ya sea por la divergencia de secuencias o mediante la combinación con otros dominios. Los mecanismos que conducen a que los dominios sean utilizados y reutilizados como bloques de construcción de manera similar a un Lego son la recombinación homóloga denominada “domain shuffling” y la fusión/fisión de genes. La aparición de nuevas combinaciones está ligada a la especiación y a grupos filogenéticos específicos. Se ha observado que ~40 dominios individuales son reutilizados y se combinan en varios arreglos. Un ejemplo es la combinación de los dominios SH3 y SH2 que se repite en varias proteínas de transducción de señales, esta versatilidad de recombinación califica al par dominio SH3-SH2 como un "supradominio". Los supradominios son de dos a tres combinaciones de dominios que se encuentran en diferentes arquitecturas con vecinos N- y C-terminales distintos (Schultz and Natarajan 2013).

La serie ordenada N- a C-terminal de los dominios de una proteína es conocida como su “arquitectura”. Generalmente, existe una conservación del orden -N a C-terminal de los dominios, así como de su geometría tridimensional y sus relaciones funcionales. Debido a que las proteínas a menudo mantienen la geometría 3D de sus pliegues y por lo tanto de sus propiedades bioquímicas, pueden ser reconocidos sus homólogos a nivel de su estructura y función. Además, empíricamente, se ha demostrado que la estructura de las proteínas es de al menos 3 a 10 veces más conservadas que sus secuencias, de igual forma la geometría de los sitios de interacción entre dominios (Nasir and Caetano-anollés 2015).

Existen algunos esquemas de clasificación funcional de proteínas, como GenProtEC, GO, MIPS, basados en la clasificación COG (Cluster of orthologous genes) y CE (Combinatorial Extension) (Vogel et al. 2004). Otro esquema para la clasificación de los dominios de proteínas en función de sus relaciones estructurales, funcionales y evolutivas es la base de

datos SCOP, considerada el "estándar de oro" en la clasificación de los dominios de proteínas con estructuras 3D conocidas y proporciona información útil sobre la evolución de dominios agrupados en superfamilias (FSF's) que incluyen núcleos estructurales 3D y propiedades bioquímicas comunes con baja identidad de secuencia (podrían ser <15%), como las FSFs son más conservadas en la evolución, es más adecuado su uso para las comparaciones evolutivas profundas (Nasir and Caetano-anollés 2015).

Dinámica evolutiva de los sistemas de dos componentes

La mayoría de la transducción de señales en procariontes parece ser llevado a cabo por los sistemas de un solo componente (OCSs), estos son proteínas individuales que combinan propiedades tanto de un sensor como regulador. Se prevé que los OCSs son los progenitores de los TCSs debido a su baja complejidad, mayor abundancia y la más amplia variedad de dominios sensores y reguladores (Capra and Laub 2012). Sin embargo, los TCS también exhiben un grado sorprendente de cambio evolutivo y plasticidad, en el transcurso de la evolución los TCS han crecido dramáticamente en número dentro de muchos genomas bacterianos a través de la duplicación de genes y la posterior divergencia de los sistemas parálogos. No obstante, las mutaciones individuales responsables de establecer nuevas vías de post-duplicación y el orden en que deben ocurrir estas mutaciones permanecen indefinidas.

Un paso clave en la divergencia de los genes de dos componentes después de la duplicación es el establecimiento de rutas aisladas, como las HK que rara vez fosforilan RR no afines en vivo, incluso aquellos derivados de reciente duplicación. Por ejemplo, la mutación de los residuos de especificidad de la HK EnvZ para la unión con los de RstB, produce una EnvZ mutante que específicamente fosforila a RstA, mientras que la interacción con su RR habitual queda eliminada. La capacidad intrínseca de una HK de discriminar su RR afín de todos los no afines implica la existencia de residuos de especificidad en los sitios de interacción proteína-proteína que deben coevolucionar. La relación co-evolutiva entre HKs y RR afines es evidente en los árboles filogenéticos de estas proteínas. Además, proteínas afines HK y RR son a menudo codificados adyacentes entre sí en los genomas, apoyando aún más su relación de co-evolutivos. Esfuerzos a gran escala para deducir la interacción HKs y RR en varios genomas se han basado en el contexto del genoma (Capra and Laub 2012).

El mapeo de los residuos de especificidad en las vías de dos componentes ha permitido entender cómo la especificidad evoluciona después de la duplicación. Análisis de secuencia y filogenia han revelado que los residuos de especificidad que eran relativamente estáticos en ausencia de la duplicación de genes, exhiben una ráfaga de diversificación coincidente con un evento de duplicación de la vía. Por lo general, tanto una HK como el RR se diversifican y coevolucionan para conservar su interacción mientras que permanecen aislados de sus homólogos a partir de quienes se duplicaron. Una vez establecidas las vías aisladas, los residuos de especificidad se vuelven relativamente estáticos, preservando así las interacciones entre proteínas afines. Por ejemplo, en α -proteobacterias, la vía NtrB-NtrC se duplicó con la posterior divergencia produciendo la existente vía NtrX-NtrY. Los análisis filogenéticos indican que la aparición de NtrX-NtrY probablemente produce “cross-talk” con otra vía de dos componentes no relacionada, el sistema PhoR-PhoB. En consecuencia, las proteínas PhoR y PhoB en α -proteobacterias acumularon sustituciones en sus residuos de especificidad para eliminar el “cross-talk”. Estas observaciones han demostrado que las proteínas independientes son más propensas a generar cross-talk y ramificaciones capaces de amplificar la señal debido a que varios RRs pueden ser modificados por un solo estímulo. Al mismo tiempo, el cross-talk actúa como una presión selectiva importante para los TCS que conduce a la diversificación de residuos de especificidad para producir vías aisladas, particularmente después de la duplicación de genes (Salvado et al. 2015).

Algunas vías de dos componentes no canónicas, particularmente las relacionadas con las denominadas cinasas híbridas no se enfrentan a la misma presión evolutiva para diversificar sus residuos de especificidad al igual que las vías canónicas. Las cinasas híbridas son HK que fosfotransfieren intramolecularmente de forma lateral a un dominio REC y su especificidad es impuesta por la proximidad espacial, y la proporción de dominios HK / RR es uno a uno. Esto significa que cada dominio HK sólo fosforila a un dominio RR. Por otra parte, las proteínas HK independientes también podrían ser más defectuosas y fosforilar RR no afines. De manera similar, los tipos de proteínas RR independientes pueden ser más propensos a la fosforilación por fuentes no afines. Tales eventos de fosforilación no afines son físicamente más difíciles de lograr en proteínas de tipo HK-RR. Por lo tanto, las proteínas híbridas representan un diseño que transduce señales con menor amplificación, pero mayor fidelidad que los dominios individuales (Salvado et al. 2015).

Aunque la evolución de las vías de dos componentes a menudo implica la producción de dos vías aisladas, hay casos en los que se añaden o integran los componentes de una vía a otra vía, dando como resultado una vía con una nueva arquitectura. Un ejemplo es la cascada de fosforrelevo en la esporulación de *B. subtilis* en el que cinco HK distintas pueden iniciar una vía de fosfotransferencia de cuatro pasos que culmina en la fosforilación de Spo0A. La arquitectura de este fosforrelevo está dictada por la preferencia de las cinasas para fosforilar Spo0F en relación con Spo0A (Figura 3A).

Otra alternativa para el origen de este tipo de fosforrelevo es que surgió de la unión de dos sistemas anteriormente independientes, de manera que el RR de una vía se convirtió en un fosfodonador para otra HK que perdió su capacidad para autofosforilarse. Cabe destacar que la histidina fosfotransferasa Spo0B, como muchas otras fosfotransferasas, se asemeja estructuralmente a una HK, pero ha perdido la capacidad de autofosforilarse. Existen otros escenarios posibles de cómo surgió el fosforrelevo de la esporulación pero sigue siendo incierto lo que ocurrió exactamente. Para esto, se necesitan estudios filogenéticos detallados, junto con mejores caracterizaciones de las vías de control de Spo0A en varios Firmicutes, el clado que contiene bacilos y clostridios. (Salazar and Laub 2015)

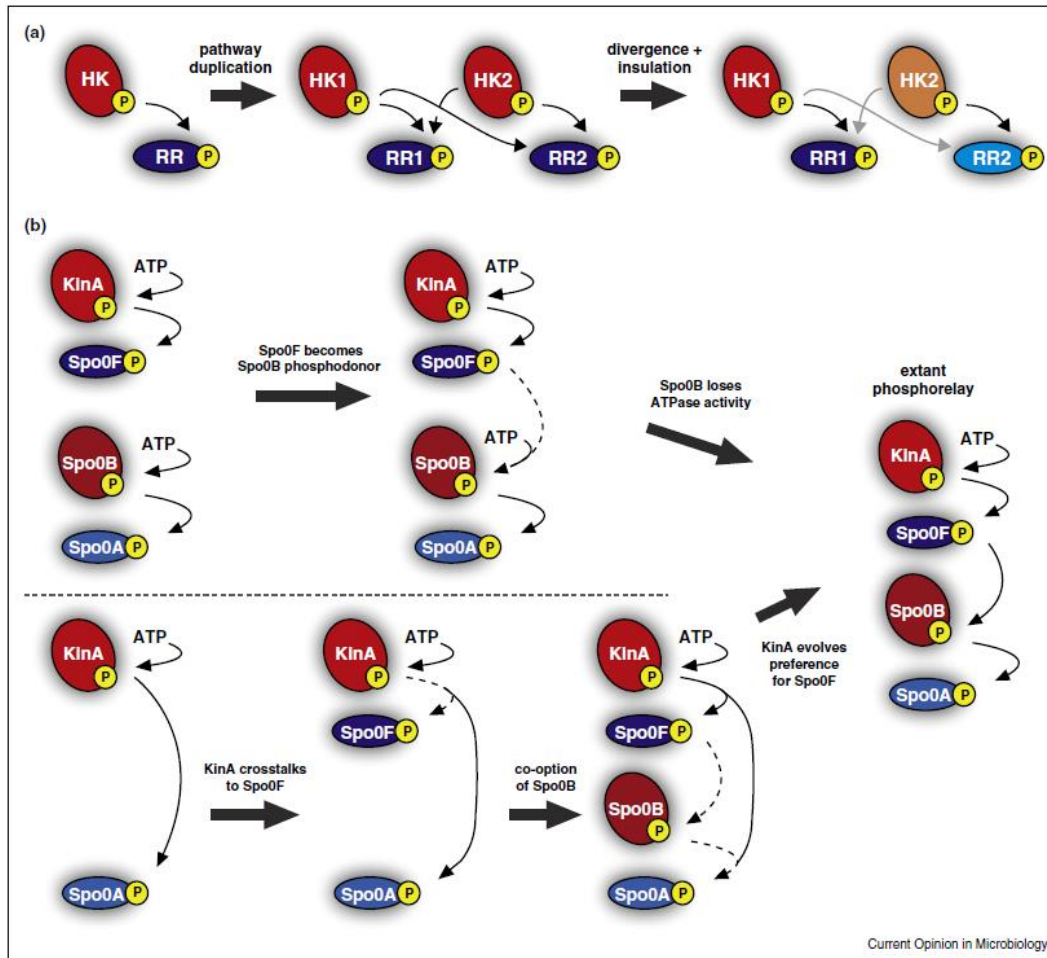


Figura 5. Evolución de los sistemas de dos componentes. a) sistema canónico de dos componentes: se expande y diversifica a través de la duplicación de genes y la subsecuente divergencia b) dos posibles modelos para el origen del fosforrelevo de la esporulación en *B subtilis*. (Salazar and Laub 2015)

Varios grupos de investigación han hecho esfuerzos por construir la historia evolutiva de esta vía de señalización. Stephenson y Hoch analizaron bioinformáticamente cómo cinco diferentes bacterias esporulantes de los géneros *Bacillus* y *Clostridia* podían cambiar sus vías de transducción de señales y adaptarse a ambientes extremadamente diferentes. Basándose en la comparación de secuencias de las proteínas del fosforrelevo, encontraron que los dominios de interacción entre proteínas están altamente conservados, mientras los dominios sensores son muy variables. Esto explica que la adaptación a nuevos ambientes resulta en una rápida evolución de los dominios sensores sin modificar las interacciones del resto de la vía. Encontraron que Spo0A está presente en todas las bacterias formadoras de esporas. Pero en el género *Clostridium*, se encontró que carecen del resto de los componentes de la cascada del fosforrelevo (Keith Stephenson and Hoch 2002b). Hoon y cols hacen un análisis más amplio en 24 especies de los mismos géneros pero incluyendo

a bacterias no esporulantes, y encontraron que la red de esporulación solo se encuentra en bacterias esporulantes, por lo que hacen una revisión completa de la red de esporulación tomando como modelo a *B. subtilis* y encontraron que la evolución de esta red no es al azar, sino que sigue la organización jerárquica de la propia vía; es decir, el conjunto de factores sigma, el regulador maestro Spo0A y la secuencia de activación de la vía están muy conservadas (80%) en todas las bacterias formadoras de esporas; los ciclos de retroalimentación se encuentran menos conservados (70%) y la conservación de los genes de esporulación están al final de la jerarquía (50%). Esto significa que para las bacterias esporulantes es más importante la conservación de la regulación que la presencia misma del gen estructural (de Hoon, Eichenberger, and Vitkup 2010).

Por otro lado, Galperin y cols (2012) descubrieron que los organismos formadores de esporas tenían genomas mayores que los no formadores de esporas, y resumieron que ~60 genes son aparentemente esenciales para la esporulación en *B. subtilis*. Este estudio introdujo la importante distinción entre los Firmicutes asporógenos (que no forman esporas) que codifican pocos genes de esporulación y las bacterias no esporógenas (o "condicionalmente no esporulantes") que codifican una gran número de genes de esporulación pero han perdido la capacidad de formar esporas debido a unas pocas mutaciones (relativamente recientes). No obstante, la definición de la lista de genes que son esenciales para la esporulación sigue siendo una tarea poco trivial y potencialmente útil (Galperin et al. 2012).

Finalmente, Jong y cols (2010) demostraron que la heterogeneidad de la esporulación entre células individuales se origina por la variabilidad en la expresión de genes del fosforrelevo, por la variación temporal en el desarrollo de las células y también proponen que la estocasticidad de la transcripción de estos genes podría ser un mecanismo regulador importante. Demostraron que el aumento de la expresión de fosfatasa previene la esporulación incluso si la transcripción de genes del fosforrelevo es alta, concluyendo así, que el flujo de fosfatos es el factor determinante para la decisión de esporular o no hacerlo (De Jong, Veening, and Kuipers 2010).

Por lo anterior, en este trabajo decidimos analizar los componentes que se sabe son responsables del flujo de fosfatos en la cascada de esporulación. Ya que la mayoría de los trabajos se centran únicamente en los miembros de los géneros *Bacillus* y *Clostridium*, se

planea extender la investigación a una amplia gama de bacterias para tratar de explicar cómo fue la evolución de cada componente, incluyendo algunas fosfatasa representativas que juegan un papel muy importante en la generación de ciclos de retroalimentación que regulan el inicio de la esporulación.

El modelo Clostridia

Los clostridios son bacterias anaeróbicas formadoras de endosporas que incluyen cepas de importancia para la salud humana y animal, degradación de la celulosa y producción de disolventes (Paredes, Alsaker, and Papoutsakis 2005). Si bien, se ha ganado mucha información del género *Bacillus*, todavía no se comprende completamente la formación de esporas en el género *Clostridium*. Mientras que algunos elementos durante la morfogénesis de esporas se conservan como Spo0A y sus vías corriente abajo, la orquestación genética, la regulación y la activación de la vía son considerablemente diferentes (Galperin et al. 2012). En los organismos de *Clostridium* no hay evidencia de un sistema de fosforrelevo para fosforilar Spo0A, esto deja preguntas pendientes acerca de cómo es iniciada la esporulación en este clado. Una hipótesis es que Spo0A se activa por fosforilación directa de las HKs huérfanas y otra posibilidad es la existencia de un sistema desconocido de fosforrelevo en *Clostridium* con características similares a Spo0F y Spo0B. Sin embargo, los esfuerzos computacionales para identificar tales componentes de fosforrelevo no han sido exitosos (Al-Hinai, Jones, and Papoutsakis 2015; Dürre 2011; Paredes, Alsaker, and Papoutsakis 2005).

En los últimos 5 años, varias HKs huérfanas que fosforilan directamente a Spo0A en *C. difficile*, *C. acetobutylicum*, *C. botulinum* y *C. thermocellum* han sido identificadas:

En *C. thermocellum* se identificaron tres HK (clo1313_0286, clo1313_2735 y clo1313_1942) que actúan como reguladores positivos de la esporulación, mientras que una cuarta HK (clo1313_1973) actúa como un regulador negativo. A diferencia de *Bacillus* o de otras especies de *Clostridium*, la eliminación de una sola quinasa que regula positivamente fue suficiente para abolir la esporulación en este organismo (Mearls and Lynd 2014). En *C. botulinum*, la HK CBO1120 demostró ser responsable de la fosforilación de Spo0A in vivo, y otras tres HKs CBO0336, CBO0340 y CBO2762 están relacionadas aunque no fue posible demostrarlo experimentalmente (Wörner et al. 2006). En *C. difficile* solo tres HKs CD2492, CD1492 y CD1579 han sido identificadas bioinformáticamente. De igual forma, en *C.*

acetobutylicum análisis bioinformáticos y experimentales indican que hay cinco HKs huérfanas codificadas por ca_c0323, ca_c0903, ca_c2730 y ca_c3319 que activan directamente a Spo0A, mientras que ca_c0437 causa un efecto negativo sobre la esporulación (Talukdar et al. 2015).

En resumen, del mismo modo que en *B. subtilis*, las HK identificadas también son de tipo huérfanas. En contraste las HK en Clostridium pueden jugar un papel en la regulación negativa, mientras que en Bacillus es controlada por la acción de fosfatasas. Al igual que en *B. subtilis*, Spo0A tiene un papel en la formación de biopelículas en *C. difficile* y *C. perfringens*. Adicionalmente, Spo0A de *C. difficile* es un regulador global de factores metabólicos y de virulencia fuera del proceso de esporulación, por lo que hay implicaciones potenciales en el tratamiento de cualquier fenotipo patógeno (Al-Hinai, Jones, and Papoutsakis 2015; Talukdar et al. 2015).

PLANTEAMIENTO DEL PROBLEMA

Los sistemas de dos componentes y la cascada del fosforrelevo funcionan como sensores ambientales que programan la expresión de genes de acuerdo a la composición del nicho ecológico en el que una bacteria reside. Se sabe que estos sistemas son muy versátiles y han seguido una estrategia evolutiva que permite a las bacterias adaptar las vías de señalización de forma rápida para sobrevivir en ambientes adversos, o bien tomar rutas de diferenciación celular altamente controladas.

Algunos estudios han encontrado que existen residuos de aminoácidos correspondientes a superficies de contacto que han resistido el cambio evolutivo; mientras que los dominios sensores de las H-cinasas muestran una gran variabilidad y rápida evolución manteniendo idénticas las interacciones proteína–proteína y proteína-DNA en la cascada del fosforrelevo (Galperin, Nikolskaya, and Koonin 2001). No obstante, poco se sabe acerca de los dominios sensores de las H-cinasas, así como la conservación a detalle de esta vía de señalización o sus posibles arquitecturas en otras bacterias más allá de los Firmicutes (Galperin et al. 2012).

A pesar de la tasa explosiva en que la secuenciación de genomas está creciendo (Land et al. 2015), el solo análisis de la secuencia no puede predecir con seguridad las funciones precisas de las regiones codificantes, por lo que el análisis experimental de la función génica sigue siendo indispensable. Por otro lado, la clasificación basada en la arquitectura de dominios puede ser engañosa cuando se usa sin un enfoque filogenómico exhaustivo, incluyendo un análisis del contexto genómico que haga posible visualizar la dinámica evolutiva que afecta la diversidad de TCSs como la duplicación de dominios, divergencia y eventos de fisión/fusión de genes de ésta compleja red de esporulación (Capra and Laub 2012; Stephens 1998).

Una de las características más importantes de esta cascada es la presencia de genes tipo "huérfanos" que por no encontrarse en operón, la identificación de candidatos como sus posibles proteínas asociadas es actualmente un reto (Ortet et al. 2015).

JUSTIFICACIÓN

La cascada del fosforrelevo en la esporulación de *Bacillus subtilis* ha sido un poderoso modelo de diferenciación celular ampliamente estudiado para evaluar los determinantes morfológicos, bioquímicos y genéticos de la esporulación, además de su evolución como un sistema de transducción de señales que ha permitido a las especies ocupar diferentes nichos ecológicos. A pesar de ello, todavía se desconoce cómo la mayoría de los linajes dentro de los Firmicutes regulan las señales de inicio de la esporulación y cómo ha evolucionado este sistema hacia nuevos esquemas de producción de esporas (Earl et al. 2010; de Hoon, Eichenberger, and Vitkup 2010; Hutchison, Miller, and Angert 2014).

Debido a que los componentes de esta cascada existen como ortólogos en diferentes especies, es posible relacionar su papel funcional en otros organismos y obtener algunos patrones de conservación. Algunos modelos han sido propuestos para tratar de explicar cómo ha sido la evolución de la red de esporulación, y se han centrado en bacterias Firmicutes esporulantes y algunas no esporulantes, también se sabe de la conservación de algunos ortólogos en otros géneros bacterianos como las Proteobacterias. Sin embargo, se hace necesario un panorama hacia otros clados filogenéticos que pueda trazar de forma integrativa la historia evolutiva de cada elemento de la vía, incluyendo las vías de regulación de la misma.

La organización interna de dominios en las proteínas, la topología de las cascadas de TCS y fosforrelevos, la manera en que los dominios interactúan entre sí y los mecanismos reguladores, juegan un papel importante en la variedad de respuestas y diseños de los circuitos. La reconstrucción de los orígenes de la cascada del fosforrelevo promete revelar la dinámica evolutiva que da forma e influye en todas las vías de dos componentes, incluidas sus conectividades y funciones (Fonslow et al. 2013; Salvado et al. 2015).

Con base en lo descrito anteriormente, en este trabajo se analizará la conservación y evolución de la cascada de fosforrelevo de la esporulación en múltiples bacterias, tanto esporulantes como no esporulantes; así como las proteínas de reciente conocimiento, que son de gran importancia regulando varios puntos de la cascada para retrasar la entrada de las células a esporulación, a menos que las condiciones adversas persistan.

PREGUNTA DE INVESTIGACIÓN

¿Cuál es la conservación y evolución en procariotas de cada componente de la cascada de fosforrelevo de la esporulación tomando como modelo a *B. subtilis*?

OBJETIVO GENERAL

Tomando las proteínas modelo de transducción de señales descritos en Bacilli y Clostridia para la esporulación, analizaremos su conservación y evolución en los diferentes phylum bacterianos mediante la búsqueda bioinformática de dominios homólogos.

OBJETIVOS PARTICULARES

- 1) Análisis de dominios de las proteínas modelo componentes del fosforrelevo en Bacilli y Clostridia:
 - Cinasas sensoras KinA, KinB, KinC, KinD y KinE.
 - Fosfotransferasas Spo0F y Spo0B.
 - Regulador de respuesta Spo0A.
 - Fosfatasas Spo0E, Sda, KipI y proteínas Rap.
- 2) Obtener secuencias aminoácidas de dominios homólogos a las proteínas modelo en procariotas utilizando la base de datos GeneContext Tool III.
- 3) Realizar comparaciones de los contextos genómicos en diferentes especies.
- 4) Evaluar la evolución de cada componente en base a su distribución filogenética.
- 5) Identificar la divergencia de la vía de señalización.
- 6) Integración de la información en base a los patrones encontrados.

HIPÓTESIS

La variación en la conservación filogenética de los componentes de la cascada del fosforrelevo en procariotas estará relacionada con la evolución de los sistemas de dos componentes hacia redes cada vez más complejas y con mayores niveles de control.

MÉTODOS

1. Programas bioinformáticos

- HAMMER, HammerScan matriz Pfam-A (Eddy 2010) alineador de secuencias a la matriz o base de datos de dominios Pfam
- Prank alineador múltiple de secuencias <http://www.ebi.ac.uk/goldman-srv/prank/>
- PhyML constructor de árboles filogenéticos por máxima verosimilitud
- iTOL visualizador de árboles filogenéticos <http://itol.embl.de/>

2. Base de datos

- Gene Context Tool III, computational Genomic Group, IBT-UNAM <http://operons.ibt.unam.mx/gctNG/> (versión de la base de datos no publicada), herramienta para visualizar el contexto genómico de un gen y sus ortólogos en todos los organismos actualmente secuenciados.

3. Desarrollo Experimental

1.- Obtención de arquitecturas de proteínas modelo

- I. Utilizando el nombre de la proteína obtener secuencias aminoácidas de cada proteína modelo en la base de datos Gene Context Tool III.
- II. Realizar HmmerScan con la matriz hammer Pfam-A y obtener las arquitecturas de las proteínas modelo de acuerdo a los “best hits”.

2.- Búsqueda de dominios homólogos en Bacteria

- I. Realizar búsqueda de dominios contenidos en las proteínas Spo0F (Response_reg), Spo0B (SPOB_a y SPOB_ab), Spo0A (Spo0A_C), Spo0E (Spo0E-like), Sda (sda), KipI (ASHI) y Rap (TPR_12) en la base de datos Gene Context Tool III y obtener las secuencias aminoácidas.
- II. Realizar HmmerScan con la matriz hammer Pfam-A a las secuencias obtenidas y obtener arquitecturas de dominios bajo los siguientes criterios: longitud no mayor al 100% de la proteína modelo, e-value < 0.001, misma cantidad de dominios.

3.- Análisis del contexto genómico

- I. De las proteínas obtenidas, seleccionar las que tengan el mismo COG que las descritas en *B. subtilis* y/o que se encuentren en el mismo contexto genómico de los modelos en al menos un gen.

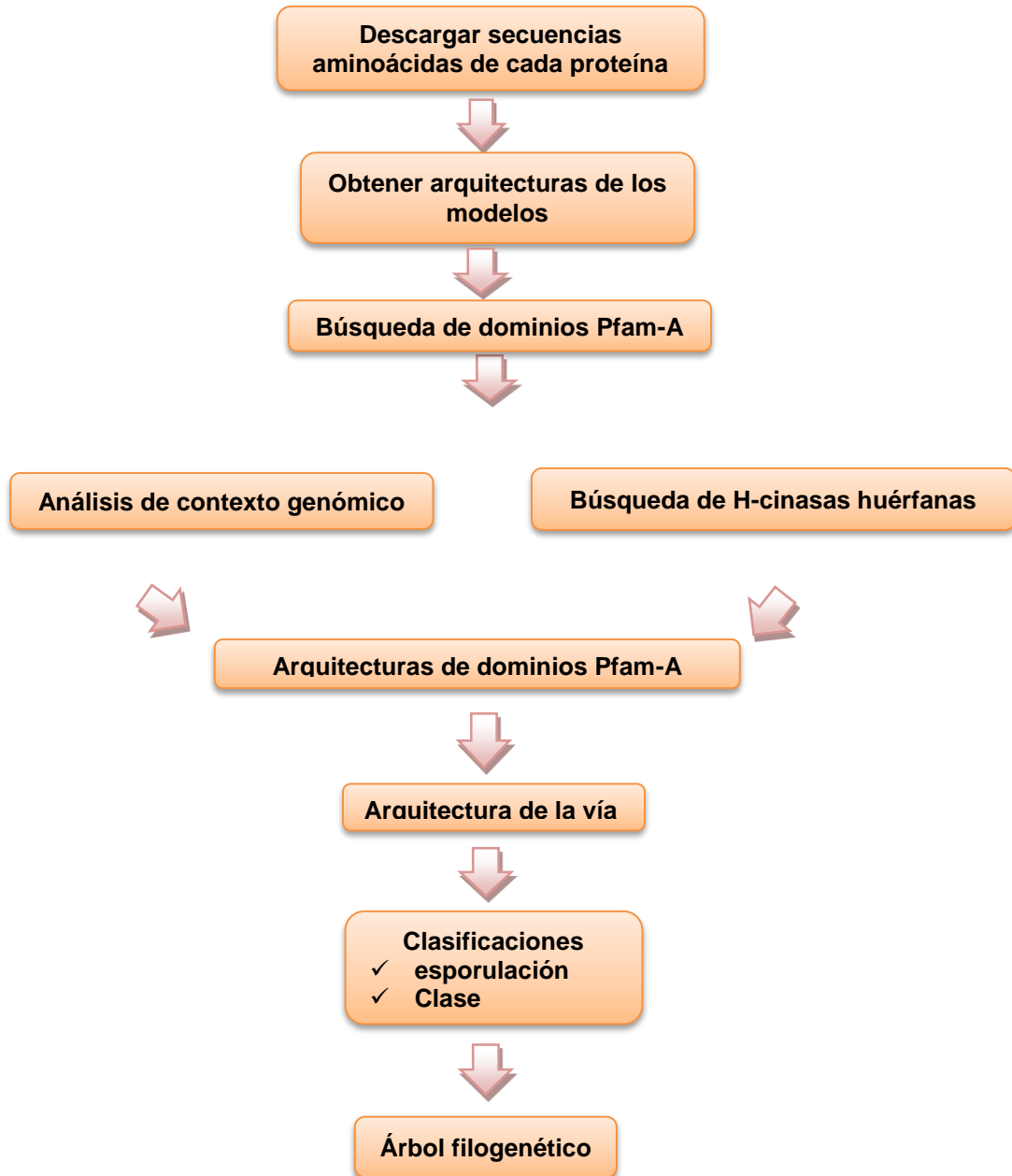
4.- Obtención de H-cinasas huérfanas

- I. Realizar búsqueda en procariotas del dominio HisKA en la base de datos Gene Context Tool III y obtener dos fastas: el primero que incluya el gen blanco y un gen del vecindario río abajo y río arriba, el segundo solo los genes blanco.
- II. Realizar HmmerScan con la matriz hammer Pfam-A a las secuencias de los genes vecinos y descartar los genes que contengan el dominio Response_reg. Las H-cinasas que no contengan en su vecindario al dominio Response_reg serán consideradas huérfanas.
- III. De las H-cinasas huérfanas obtenidas, realizar HmmerScan con la matriz hammer Pfam-A y descartar todas las que no se ajusten a los modelos bajo los siguientes criterios: longitud no mayor al 100% de la proteína modelo, e-value < 0.001, misma cantidad de dominios.

5.- Árbol filogenético

- I. Obtener secuencias nucleotídicas de la subunidad 16S ribosomal de los organismos que cuentan con al menos una de las proteínas analizadas en la base de datos Gene Context Tool III.
- II. Realizar alineamiento múltiple con Prank y obtener Árbol filogenético con PhyML.
- III. Clasificar los Firmicutes que contengan al menos una de las proteínas estudiadas en esporulantes y no esporulantes de acuerdo a lo reportado en la bibliografía como comprobado experimentalmente.
- IV. Realizar matriz para añadir todos los elementos analizados en la filogenia 16S.

ESQUEMA DE LA ESTRATEGIA EXPERIMENTAL



RESULTADOS

1.- Modelos de sistemas de transducción de señal en Bacilli y Clostridia

A pesar de la cercanía filogenética, fue necesario tomar un modelo de ambas clases bacterianas Bacilli y Clostridia debido a que la similitud a nivel de secuencia de sus proteínas difiere al grado de que el análisis por alineamientos de secuencia arroja resultados limitados que impiden construir de forma completa la historia evolutiva.

Mediante el análisis de dominios Pfam obtuvimos los dominios que conforman las arquitecturas de cada proteína, representados en la siguiente figura como cuadros en orden de N- a C- terminal.

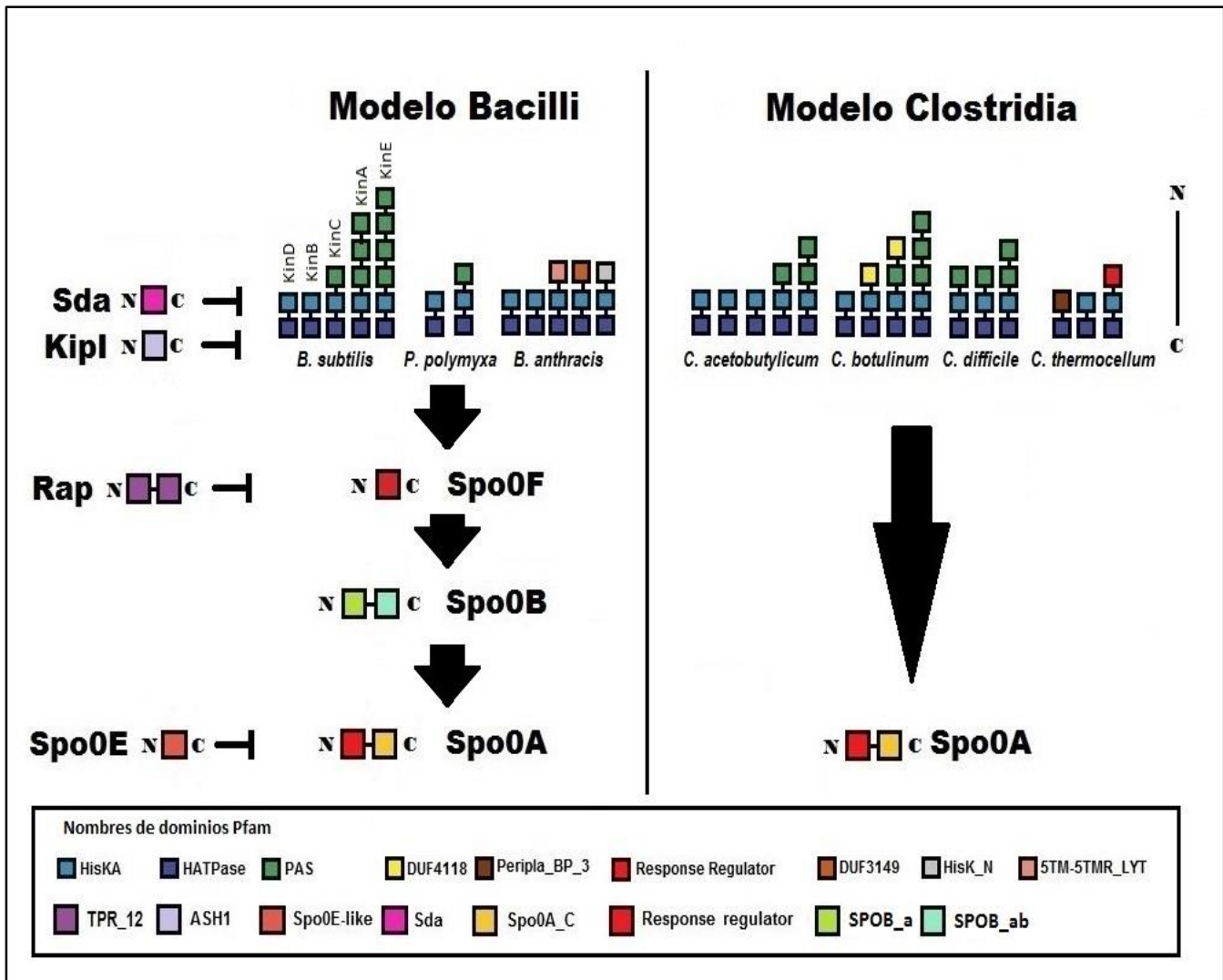


Figura 6. Arquitecturas de los modelos de sistemas de transducción de señales probados experimentalmente en Bacilli y Clostridia. Lado izq. Fosforrelevo en *B.subtilis*, *P.polimyxa* y *B.anthraxis*. Lado der. Sistema de dos componentes en *C.acetobutylicum*, *C.botulinum*, *C.difficile* y *C.thermocellum*. En color se representa el nombre de cada dominio según la clasificación Pfam.

Se observa que los únicos elementos que difieren en su arquitectura son las H-cinasas debido a la amplia diversidad de dominios sensores que existen en bacteria. Para contar con un mayor espectro de dominios sensores relacionados con la esporulación se incluyeron H-cinasas de las bacterias esporulantes más estudiadas experimentalmente: *B. subtilis*, *P. polymyxa* (Park, Park, and Choi 2012) y *B. anthracis* (Brunsing et al. 2005) de la clase Bacilli y en *C. acetobutylicum*, *C. botulinum*, *C. difficile* y *C. thermocellum* de la clase Clostridia. Se tomaron los nombres Pfam de los dominios de estos modelos para realizar la búsqueda de proteínas en Bacteria que cumplan con la misma arquitectura de dominios, e-value 0.001, con una longitud no mayor al modelo en un 100%, con el mismo COG de los modelos y con un contexto genómico similar al de los modelos en un rango de tres genes corriente abajo y tres genes corriente arriba.

2.- Contexto genómico de los genes de la esporulación en Bacilli y Clostridia

Utilizando la herramienta Gene Context Tool III se buscaron ortólogos de cada una de las proteínas modelo, preferentemente proteínas que contaran con la misma anotación o nombre, y se obtuvieron las gráficas resumiendo los contextos más representativos incluyendo tres genes corriente abajo y tres genes corriente arriba.

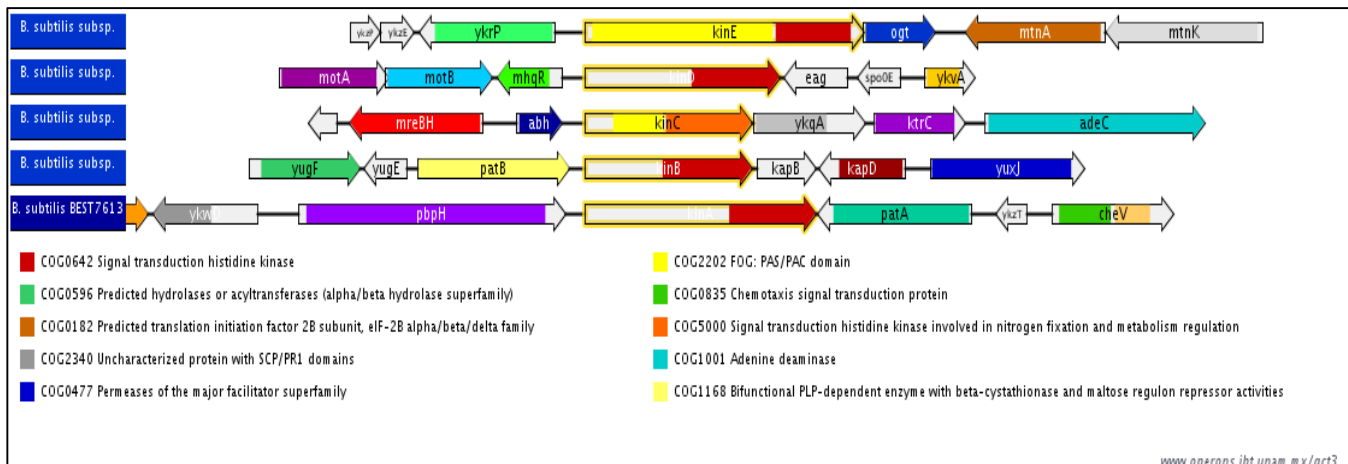


Figura 7.1 Contexto genómico de las cinco H-cinasas huérfanas de *Bacillus subtilis*, tres genes corriente abajo y corriente arriba, se representa con color el COG al que pertenece cada gen.

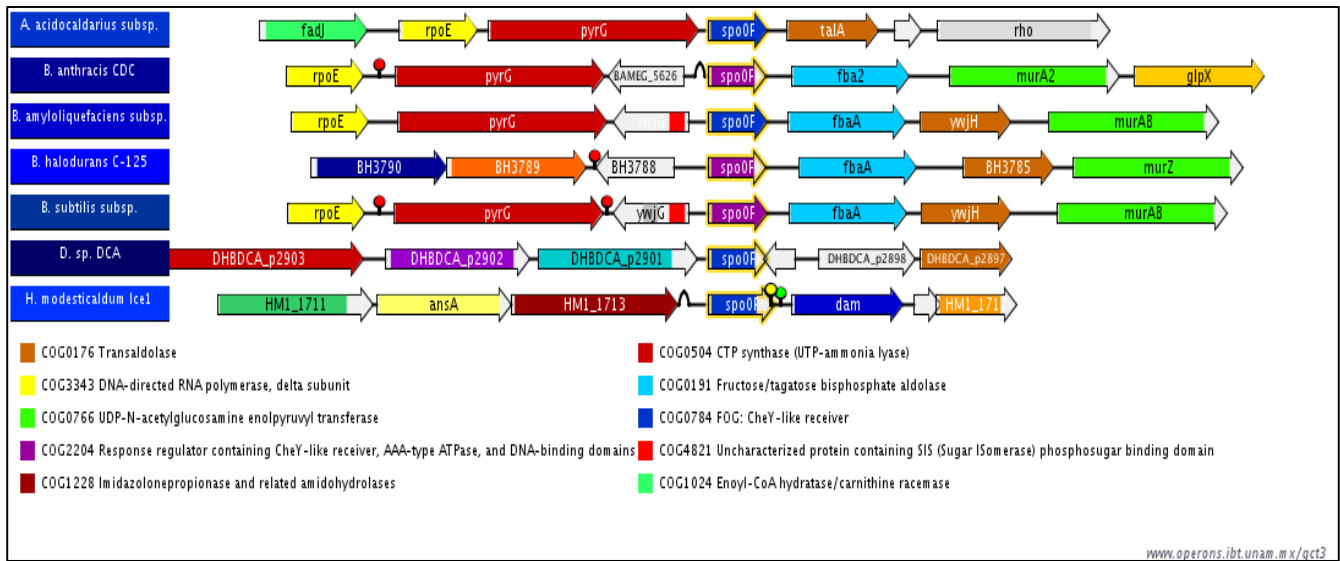


Figura 7.2 Contexto genómico del regulador SpoOF en Bacilli y Clostridia, tres genes corriente abajo y corriente arriba, se representa con color el COG al que pertenece cada gen.

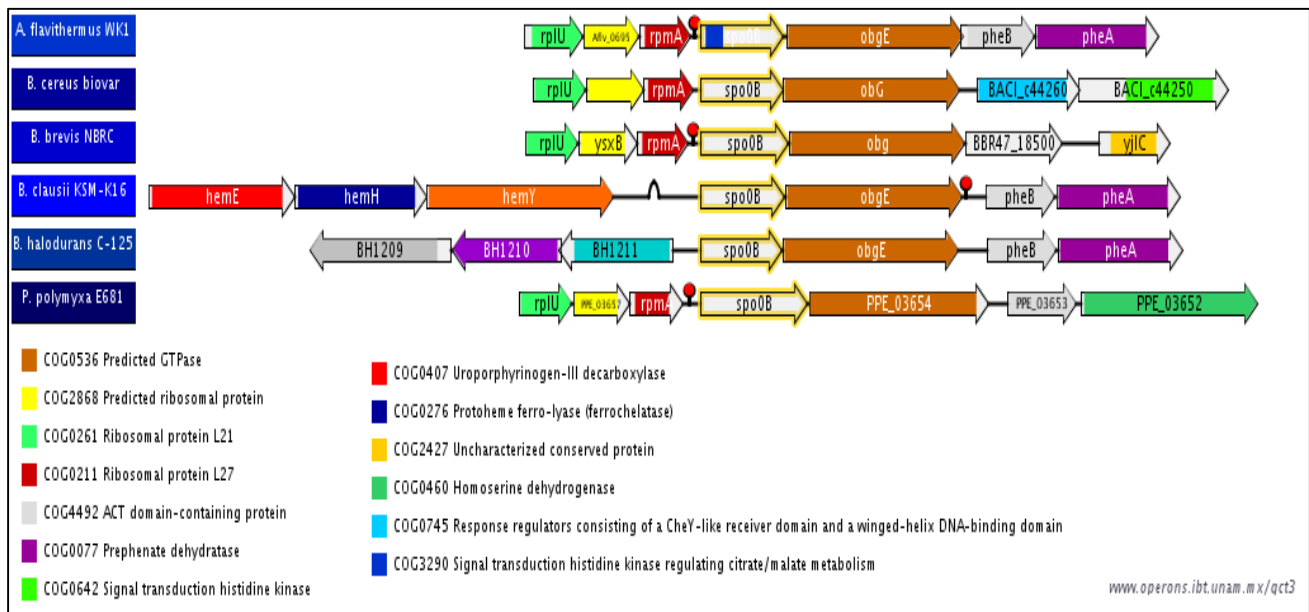


Figura 7.3 Contexto genómico de la fosfotransferasa SpoOB en Bacilli, tres genes corriente abajo y corriente arriba, se representa con color el COG al que pertenece cada gen.

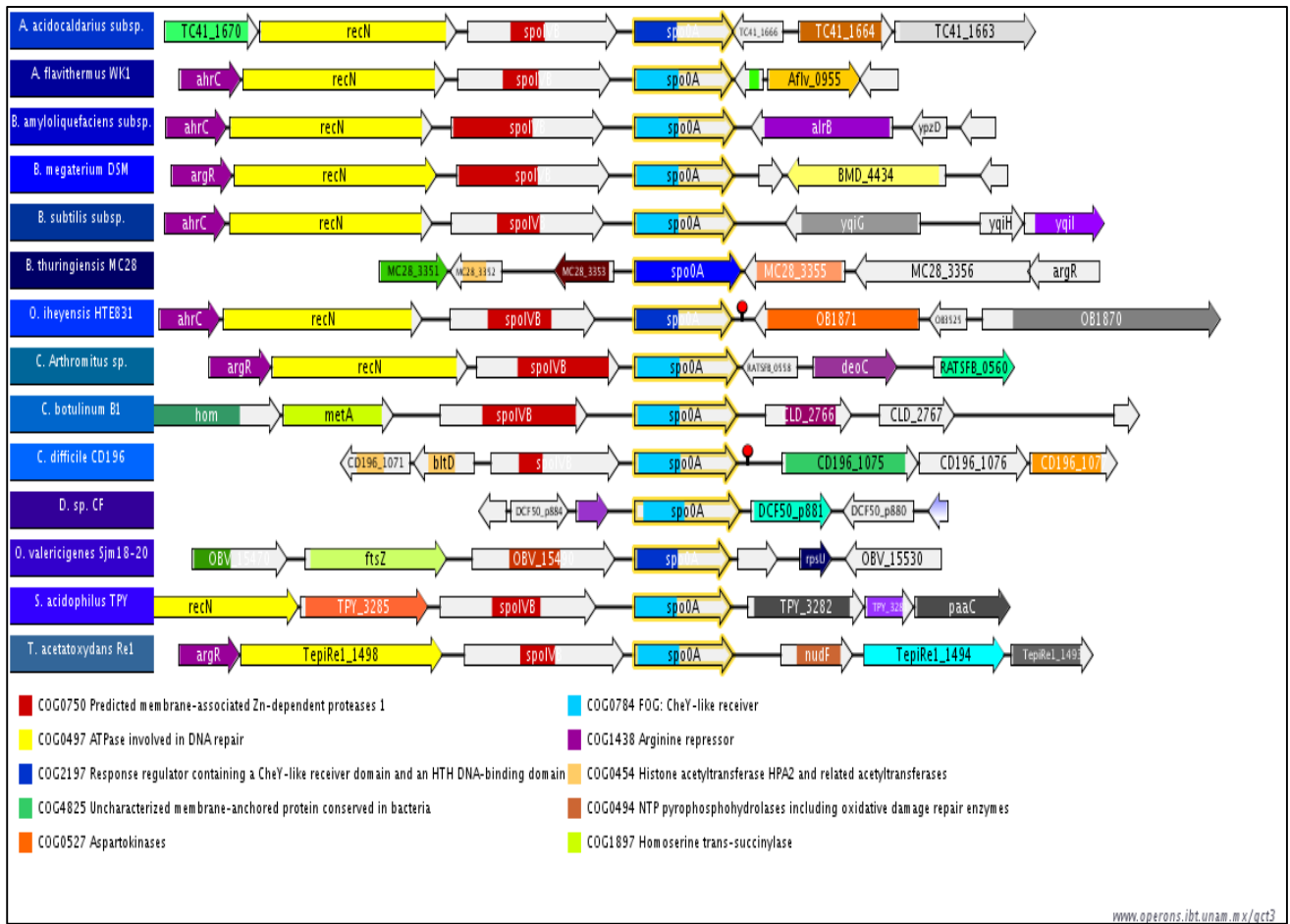


Figura 7.4 Contexto genómico del regulador de respuesta Spo0A en Bacilli y Clostridia, tres genes corriente abajo y corriente arriba, se representa con color el COG al que pertenece cada gen.

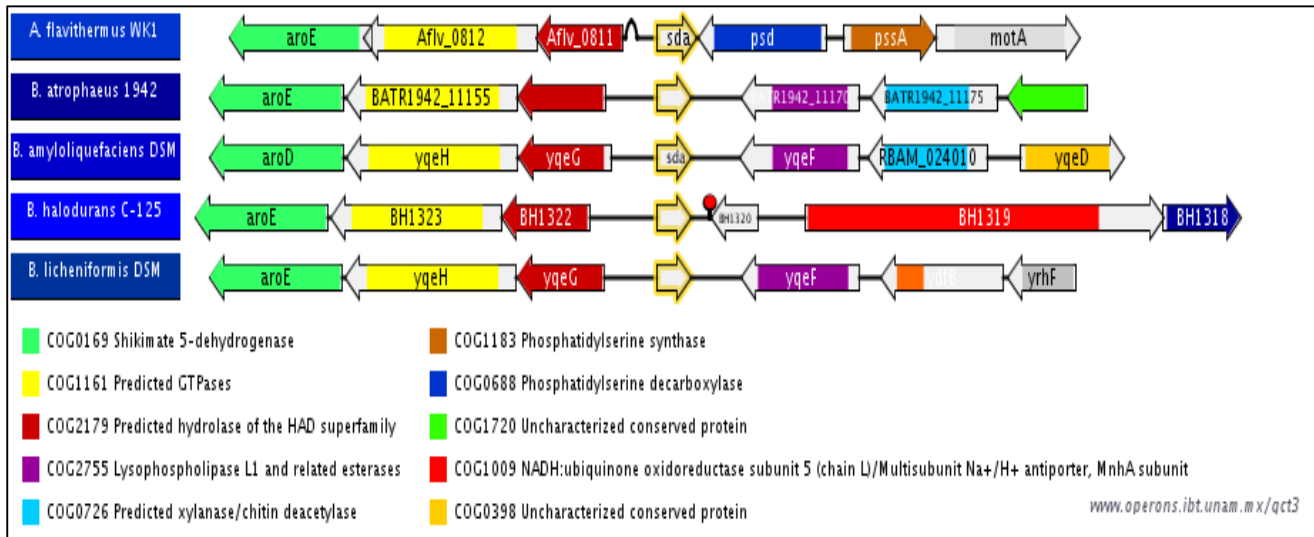


Figura 7.5 Contexto genómico de la fosfatasa Sda en Bacilli, tres genes corriente abajo y corriente arriba, se representa con color el COG al que pertenece cada gen.

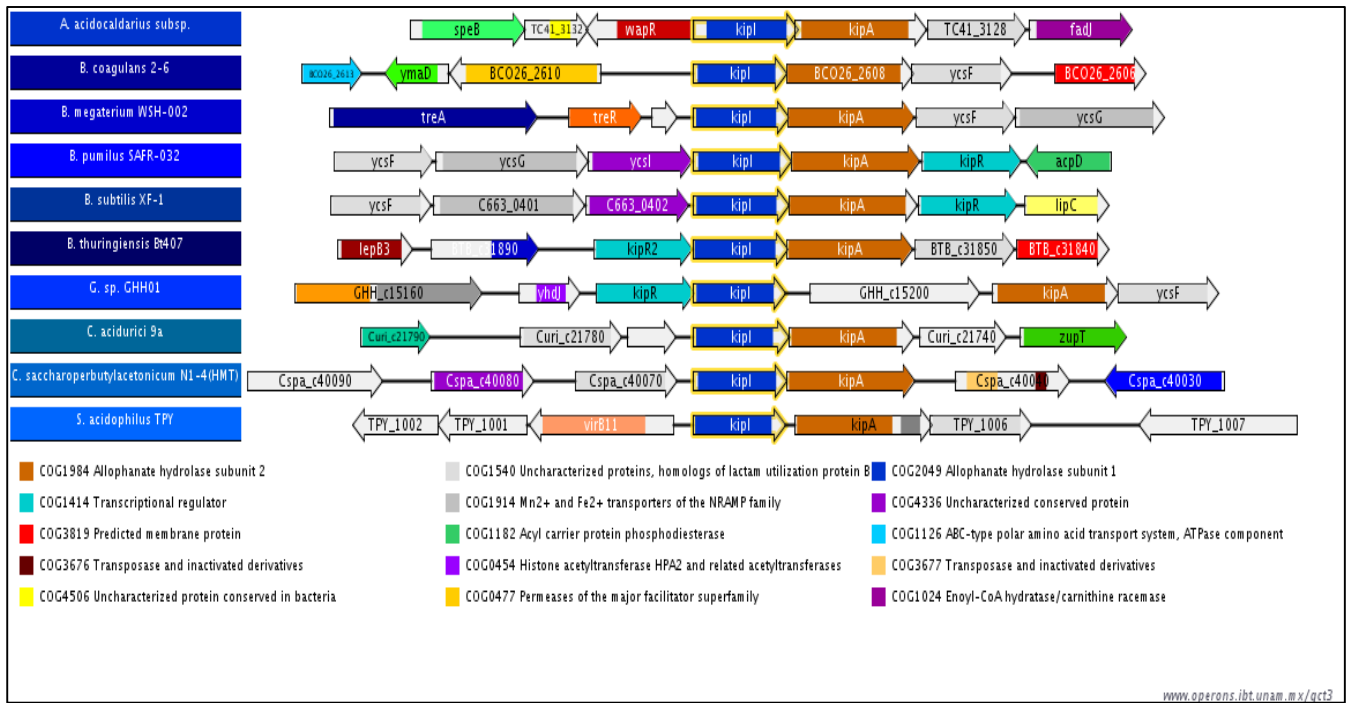


Figura 7.6 Contexto genómico de la fosfatasa Kipl en Bacilli y Clostridia, tres genes corriente abajo y corriente arriba, se representa con color el COG al que pertenece cada gen.

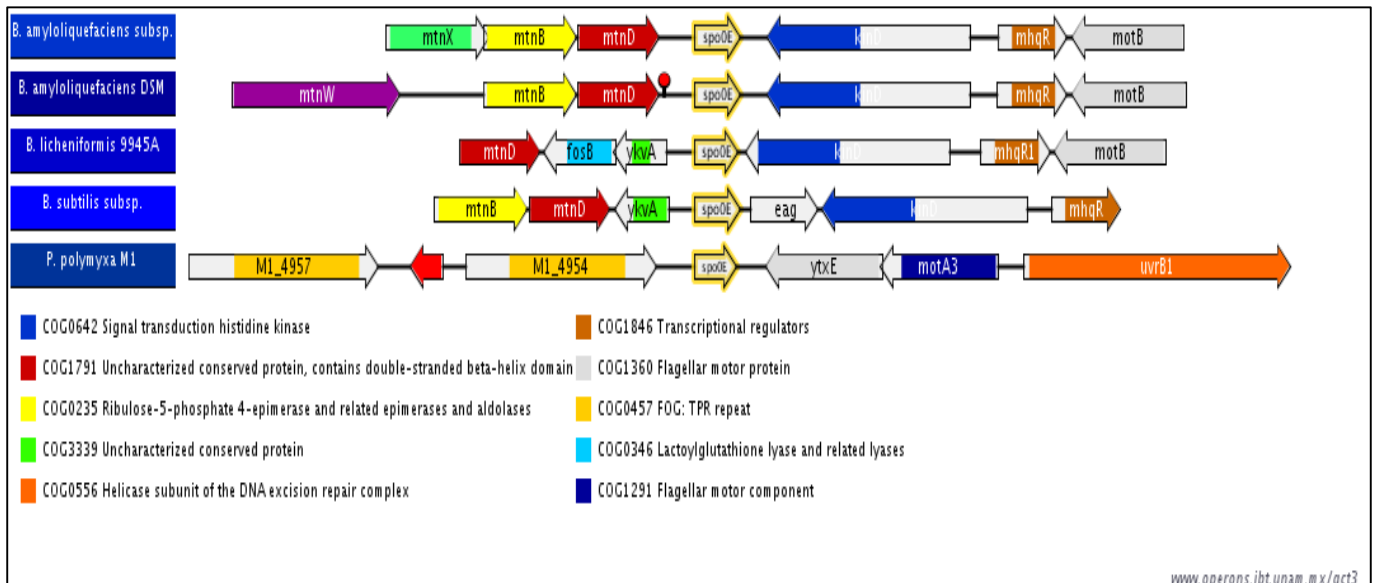


Figura 7.7 Contexto genómico de la fosfatasa Spo0E en Bacilli, tres genes corriente abajo y corriente arriba, se representa con color el COG al que pertenece cada gen.

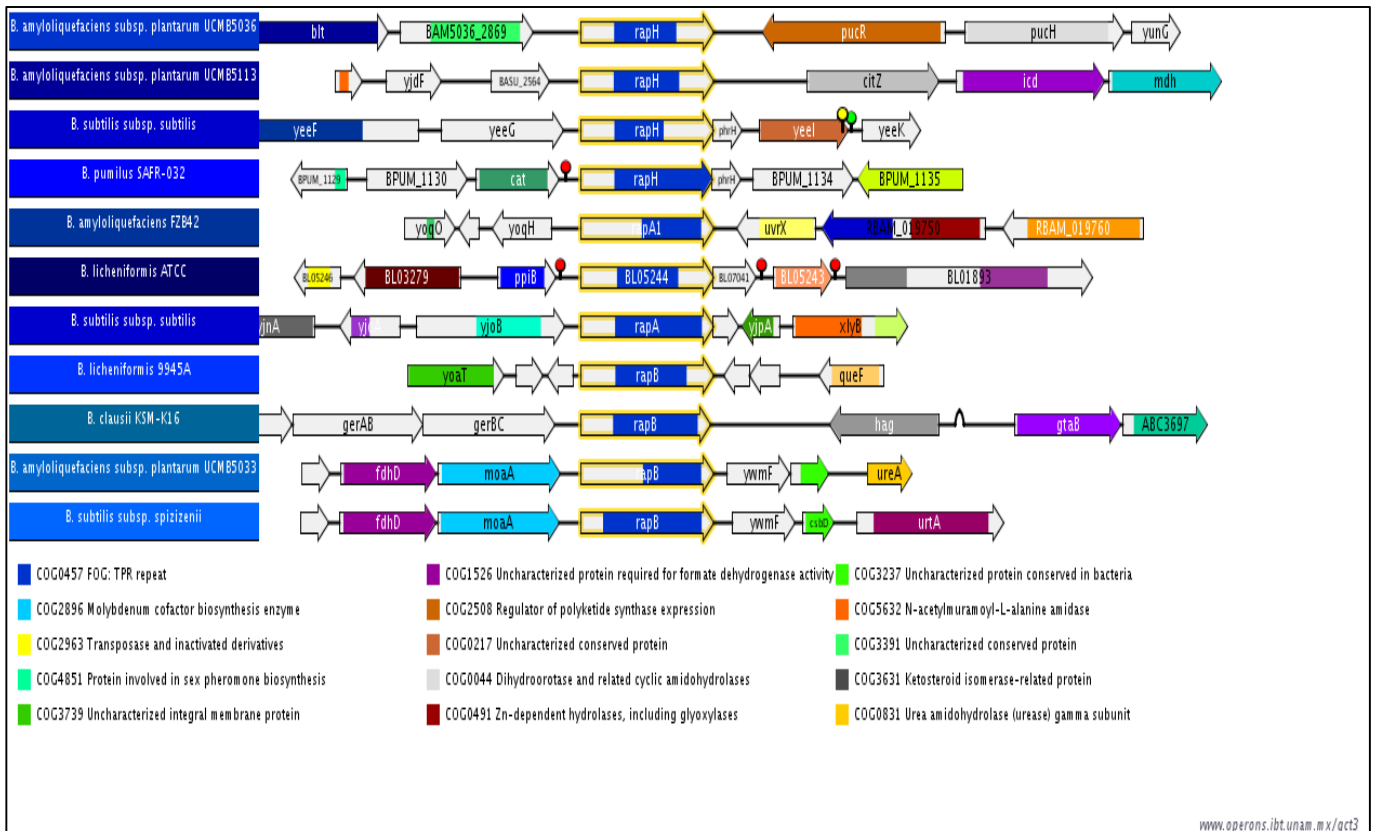


Figura 7.8 Contexto genómico de las fosfatasa Rap en Bacilli y Clostridia, tres genes corriente abajo y corriente arriba, se representa con color el COG al que pertenece cada gen.

La cantidad de genes variables en cada resumen son diferentes para cada proteína, por ejemplo, la proteína Spo0E (Figura 7.7) tiene un contexto genómico menos variable que el vecindario de Spo0A (figura 7.4).

En el contexto de las H-cinasas de *B. subtilis* se observa que ninguna de las cinco tiene en su vecindario un COG de regulador de respuesta, por lo tanto, todas son de tipo huérfanas (Figura 7.1). El contexto de Spo0F en Bacilli sí es conservado mientras que en Clostridia la variabilidad del contexto aumenta. Esta proteína está relacionada con el COG0784 y el COG2204 que tienen la misma función de dominio receptor del regulador de respuesta (Figura 7.2). Spo0B no tiene un COG que lo represente, con excepción de un caso que cuentan con el COG3290 de HK reguladora del metabolismo del citrato/ malato, esto es indicativo de una posible relación evolutiva. Spo0B cuenta con un contexto más conservado, es decir, los genes aledaños son poco variables (Figura 7.3). Spo0A tiene un contexto genómico muy diverso y definido corriente arriba, contrario al de corriente abajo, y también aumenta su variabilidad hacia las Clostridias. Curiosamente esta proteína cuenta con los COG0784 receptor del regulador de respuesta y COG2197 dominio HTH de unión

a DNA (Figura 7.4), esto nos dice que las proteínas Spo0A con el COG0784 podrían estar relacionadas evolutivamente con Spo0F.

El contexto corriente arriba de Sda está totalmente conservado mientras que corriente abajo es parcialmente conservado y no cuenta con un COG funcional (Figura 7.5). Kipl cuenta el COG2049 de alofanato hidrolasa subunidad 1, tiene un contexto muy variable, y los genes aledaños pueden estar localizados tanto corriente arriba como corriente abajo. El único gen que se conserva en todos los casos y en el mismo lugar es su proteína reguladora KipA (Figura 7.6). Spo0E tampoco tiene un COG que lo represente, suele estar junto a la H-cinasa KinD pese a una excepción que no cumple con el contexto (Figura 7.7). Las proteínas Rap así como las H-cinasas huérfanas no cuentan con un contexto genómico definido, los genes aledaños son tan variables que solo se puede considerar que las proteínas cuenten con el mismo COG0457 de dominio TPR repetido, solo algunas de ellas tienen contiguo un pequeño gen correspondiente a sus respectivos pentapéptidos reguladores (Figura 7.8).

Estos modelos se utilizaron como filtros para conservar sólo los hits que cumplieran con pertenecer al mismo COG como un indicativo de que los genes estén relacionados con la misma función, además de contar con al menos un gen aledaño similar al de los modelos anteriores.

3.- Distribución en Bacteria de dominios homólogos a los modelos de sistemas de transducción de señales en Firmicutes

En la siguiente figura se muestra la presencia o ausencia a través de los diferentes phylum bacterianos de dominios homólogos que cumplen con las características de los modelos propuestos. En el tamaño circular está representado el número de bacterias que sí cuentan con la proteína homóloga en relación al total de bacterias secuenciadas en el correspondiente phylum, en color está representada la moda estadística de las copias paralogas de dicha proteína por bacteria y las barras representan, en proporción, el total de bacterias secuenciadas en cada phylum.

Distribución de dominios homólogos a la cascada de fosforrelievo de la esporulación

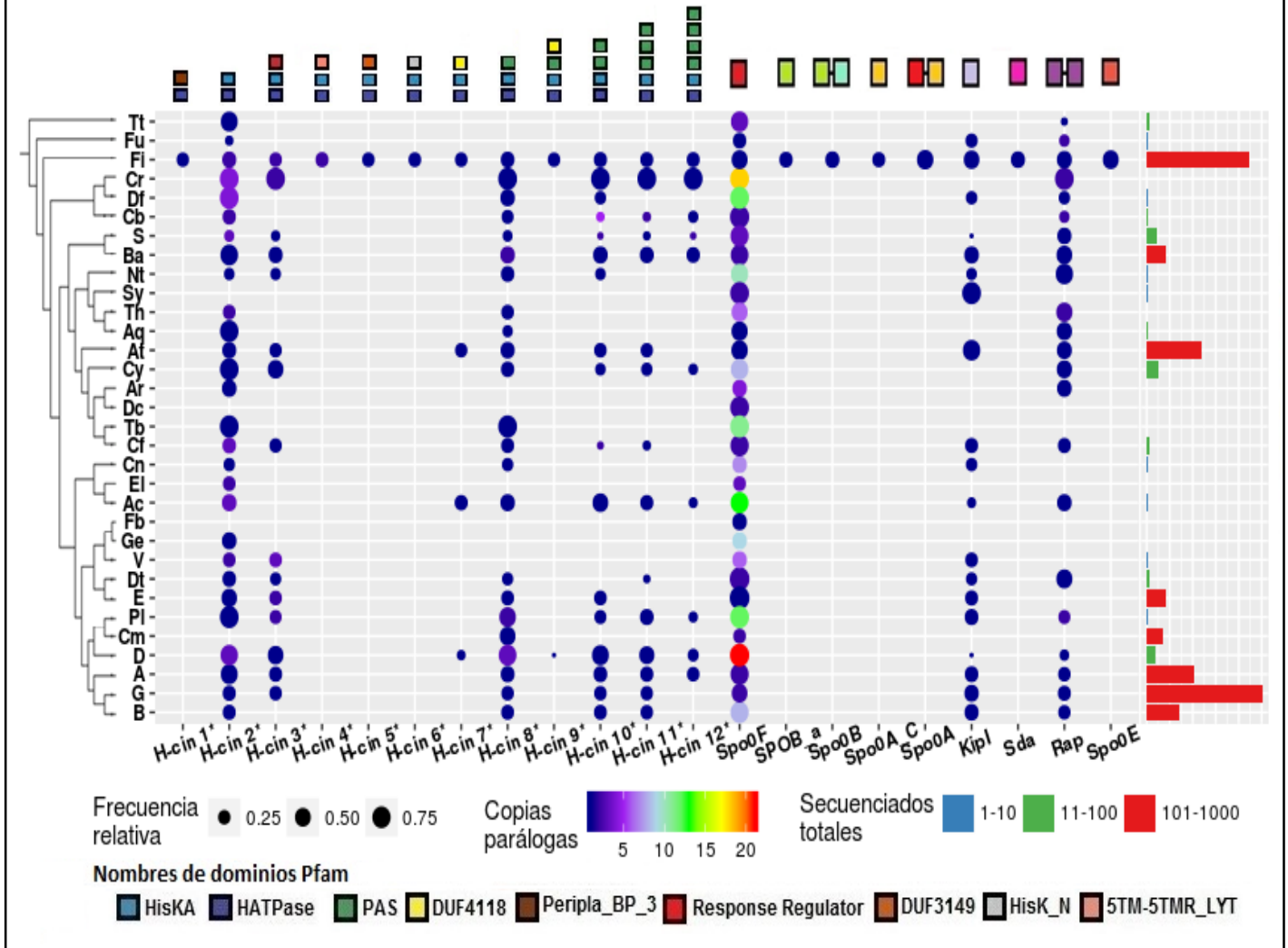


Figura 8. Distribución en procariontes de dominios homólogos a los sistemas de transducción de señales descritos en Bacilli y Clostridia. Las barras representan la cantidad de organismos secuenciados por Phylum, el tamaño representa la cantidad de bacterias que cuentan con la proteína en relación al total de secuenciados en dicho phylum. Lado izq. Adjunto árbol 16s de los phyla bacterianos. En color se representa la moda estadística de las copias parálogos por bacteria en cada phylum.

Se consideró un total de 3220 organismos secuenciados y a pesar de que los phylum son heterogéneos en su total de organismos secuenciados, en general, la abundancia de cada proteína prevalece por debajo del 75% en la mayoría de los grupos. El hecho de que no hay conservación absoluta dentro de un solo phylum bacteriano demuestra la gran diversidad de fenotipos que existe dentro de ellos mismos.

Se encontraron 12 arquitecturas de HKs de Bacillus y Clostridia distintas que han sido caracterizadas experimentalmente como parte de una cascada de esporulación. En la figura

anterior se observa que existen homólogos de las HKs distribuidos a lo largo de casi todos los phylum bacterianos con ~5 copias parálogas, pocas son las familias que no conservan una HK homóloga y es posible que la transferencia de varias HK empezara a partir de Firmicutes, con excepción de la arquitectura 2 que resulta ser mucho más ancestral. En el análisis de dominios encontramos muy conservados los dominios HisKA, HTPase_c y PAS en Bacteria, y otros dominios sensores no distribuidos DUF4118, DUF3149, HisK_N, 5TM-5TMR_LYT y Peripla_BP_3 presentes únicamente en Firmicutes. Las funciones de algunos de los dominios variables están relacionadas con el sensado de diferentes sustratos y con estructuras de anclaje a la membrana. En general, las arquitecturas son sencillas y la que existe en mayor abundancia es la número 8 que contiene un dominio PAS. La arquitectura 2 parece ser la más ancestral quizá porque no cuenta con dominios sensores, o puede que los tenga entre los dominios TM para captar luz.

Las H-cinasas 8, 10, 11 y 12 se encuentran ampliamente distribuidas, y en su arquitectura conservan dominios PAS en tándem, los cuales son muy ancestrales y ampliamente distribuidos en el reino bacteriano. Las arquitecturas 1, 4, 5, 6, 7 y 9 se encuentran exclusivamente en Firmicutes.

Las fosfotransferasas Spo0B y Spo0A están distribuidas únicamente dentro de Firmicutes; curiosamente, encontramos que ambas proteínas pueden estar presentes en una forma reducida que se ignora si es funcional o no, pero son de gran importancia puesto que nos dan información valiosa de su evolución. El dominio Response_reg o REC tiene una amplia distribución en numerosos phyla, además de existir en hasta 20 copias parálogas, lo que nos habla de lo recurrido y versátil que puede ser este dominio.

Las fosfatasas Kipl y Rap están ampliamente distribuidas en casi todos los phylum, inclusive en donde no hay H-cinasas homólogas, lo que nos dice que pueden tener otra función aparte de ser fosfatasas, o bien, Kipl tiene la plasticidad de ser fosfatasa de varios tipos de H-cinasas. En contraste, Sda que inhibe a la misma H-cinasa que Kipl sí es exclusiva de Firmicutes, al igual que Spo0E que regula a Spo0A.

En general, no existe otro phylum fuera de Firmicutes en donde se encuentren las cuatro fosfatasas, así como el resto de los componentes.

4.- H-cinasas huérfanas en procariotas

Las HK huérfanas es una característica que no todos los phyla comparten, inclusive no todas las bacterias dentro un solo phylum las conservan. El total de HK tipo huérfanas por bacteria en cada phylum se representa en la siguiente figura, cada punto indica la cantidad de HK que tiene una bacteria. Encontramos que algunos phyla no cuentan con HK huérfanas como Fibrobacteres y Dictioglomi, y no todos los organismos de un phylum tienen HK huérfanas, la mayoría de estos cuentan con alrededor de 20 HK huérfanas pero hay phyla que pueden contener hasta 80 HK huérfanas como las Cyanobacterias. La cantidad de HK huérfanas está relacionada con la presencia de fosforrelevos o TCS más sofisticados en donde varias HKs pueden fosforilar a un solo RR o bien una HK puede fosforilar a varios RRs.

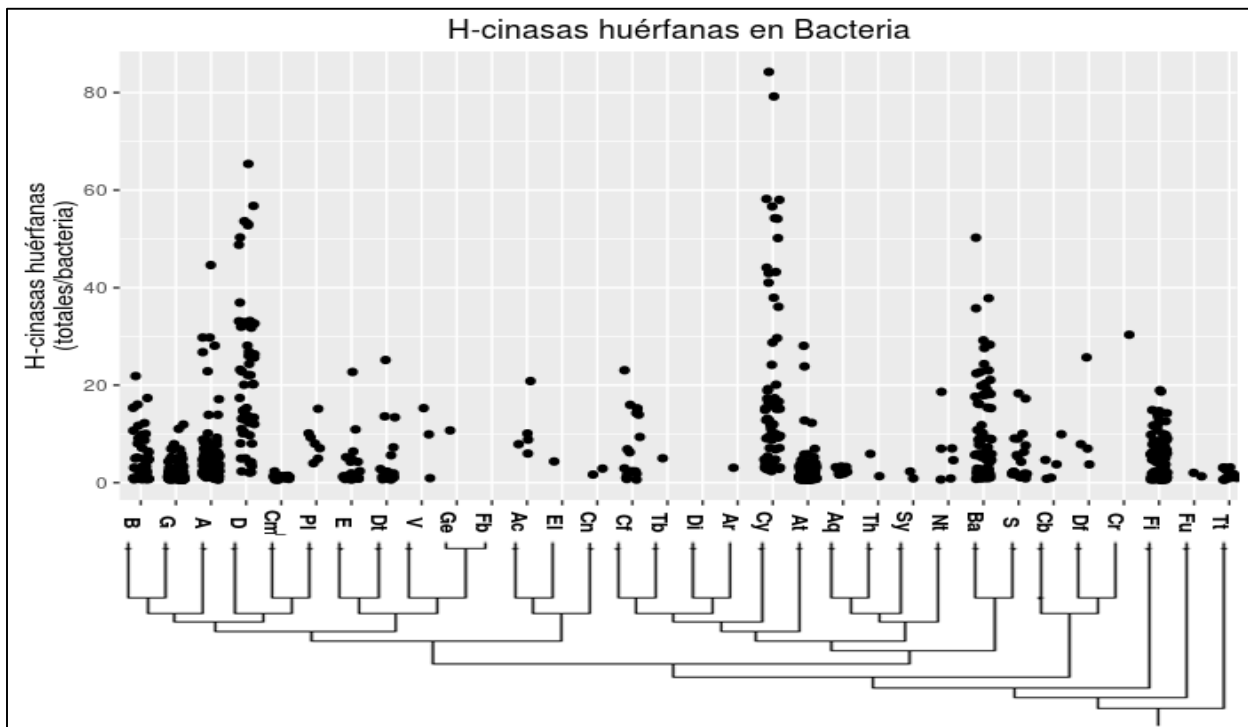


Figura 9. Total de H-cinasas huérfanas por bacteria en cada phylum ordenados por distancia evolutiva 16S, cada punto representa una bacteria, la mayoría de ellas conservan de 1-20 HK huérfanas, la minoría conservan de 20-80 HK huérfanas, hay pocos phylum con bacterias que no cuentan con HK huérfanas.

Las clases del phylum Firmicutes como Bacilli, Clostridia si conservan hasta 20 HK huérfanas, en menor grado Negativicutes y Erysipelotrichia, y en las clases Tenericutes y Thermolithobacteria no fue posible detectar HK huérfanas, reflejando con ello la gran variabilidad de fenotipos dentro del mismo phylum. En la siguiente gráfica se representa el total de HK huérfanas en cada bacteria. Se observa que las bacterias no esporulantes (azul) tienden a conservar pocas H-cinasas huérfanas, así como las bacterias que no tienen genes

de fosforrelevo y, que por lo tanto no esporulan, tienden a conservar menos de cinco HK huérfanas, mientras las bacterias esporulantes tienden a adquirir mayor número de HK huérfanas. La clasificación de “esporulantes” se refiere a organismos en los que experimentalmente se documentó la formación de esporas, los “no esporulantes” se refiere a los organismos que han sido caracterizados experimentalmente como no formadores de esporas, y el caso de los organismos que no cuentan con documentación experimental se clasificaron como “desconocido”.

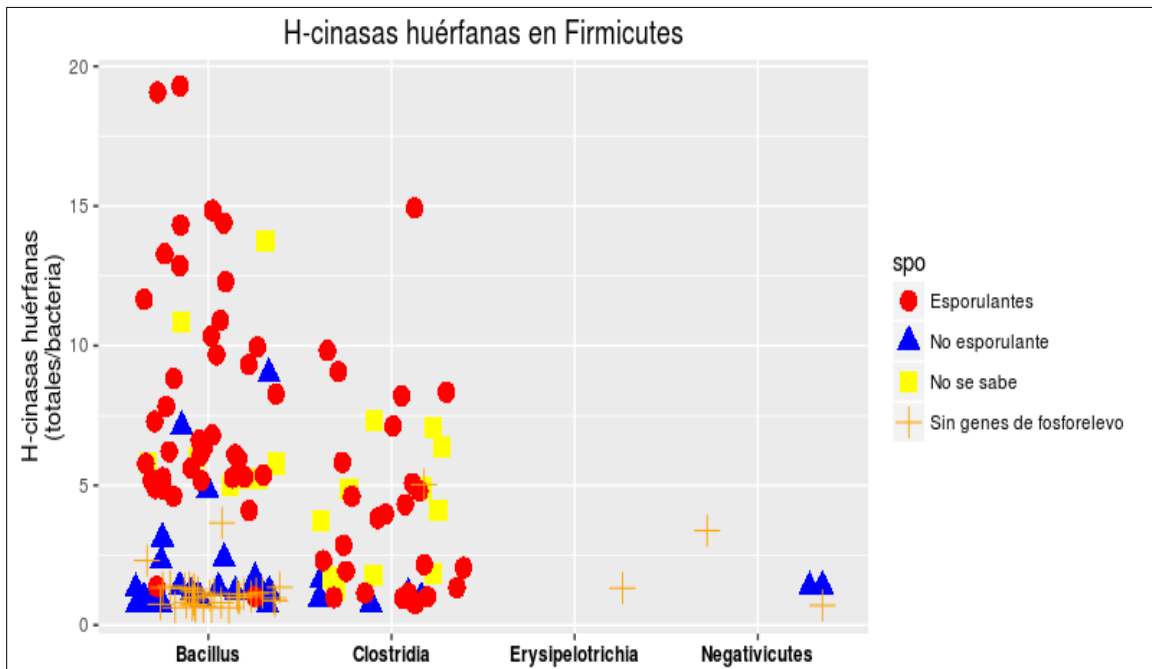


Figura 10. H-cinasas huérfanas por Clase en Firmicutes. Total de H-cinasas huérfanas por bacteria en cada Clase, cada punto representa una bacteria. Se clasifican como esporulantes o no esporulantes, no se sabe y bacterias que tienen H-cinasas huérfanas pero no tienen ningún otro elemento de la cascada de fosforrelevo.

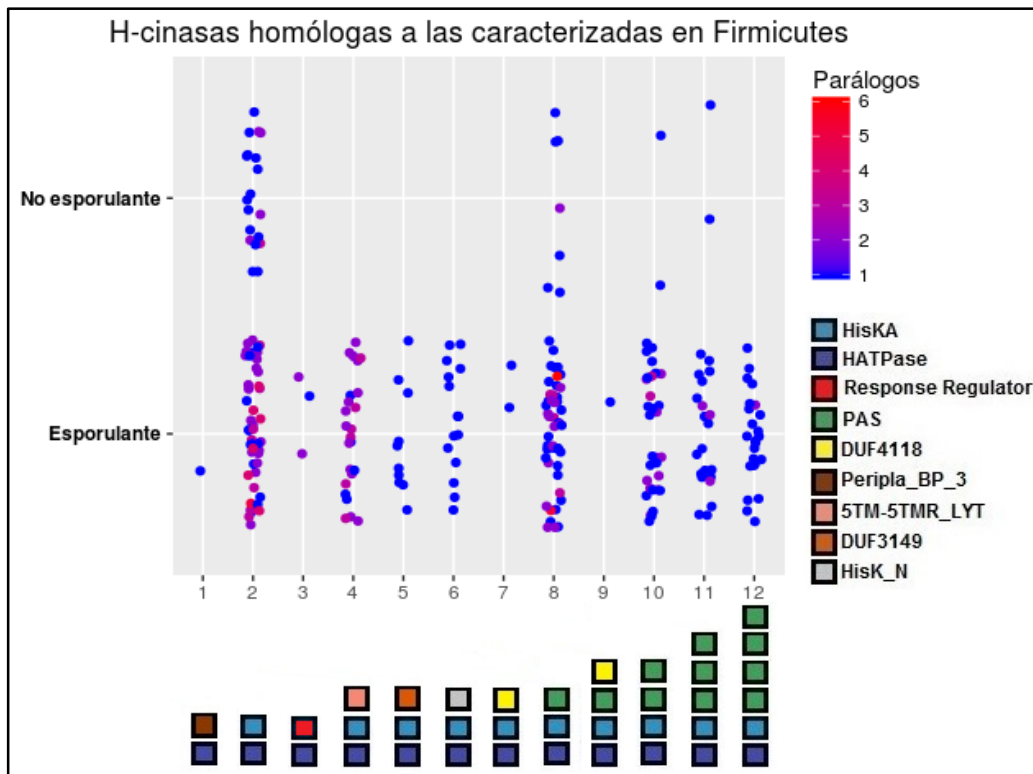


Figura 11. Distribución en Bacilli y Clostridia de las H-cinasas homólogas a las modelos, clasificadas de acuerdo a su capacidad de esporular. En color se representan las copias parálogas de la H-cinasa por bacteria.

Encontramos también que el grupo de bacterias esporulantes se encuentra enriquecido con las arquitecturas de HK caracterizadas esporulantes, algunas de ellas en varias copias parálogas por lo que varias de ellas se pueden considerar como candidatas necesarias para iniciar la esporulación. Aunque el grupo de bacterias no esporulantes también cuenta con las arquitecturas 2, 8, 10 y 11 en menor cantidad, estas arquitecturas son muy distribuidas y por lo que es más probable que estén relacionadas con otro tipo de vías distintas a la esporulación.

5.- Conservación de los elementos del fosforrelevo en Firmicutes

Los resultados de la búsqueda por dominios en diferentes Clases de Firmicutes se encuentran clasificados en la siguiente gráfica como grupos de organismos que conservan desde el subdominio Spo0A_C hasta los que conservan los tres componentes de la cascada completa, así como sus combinaciones.

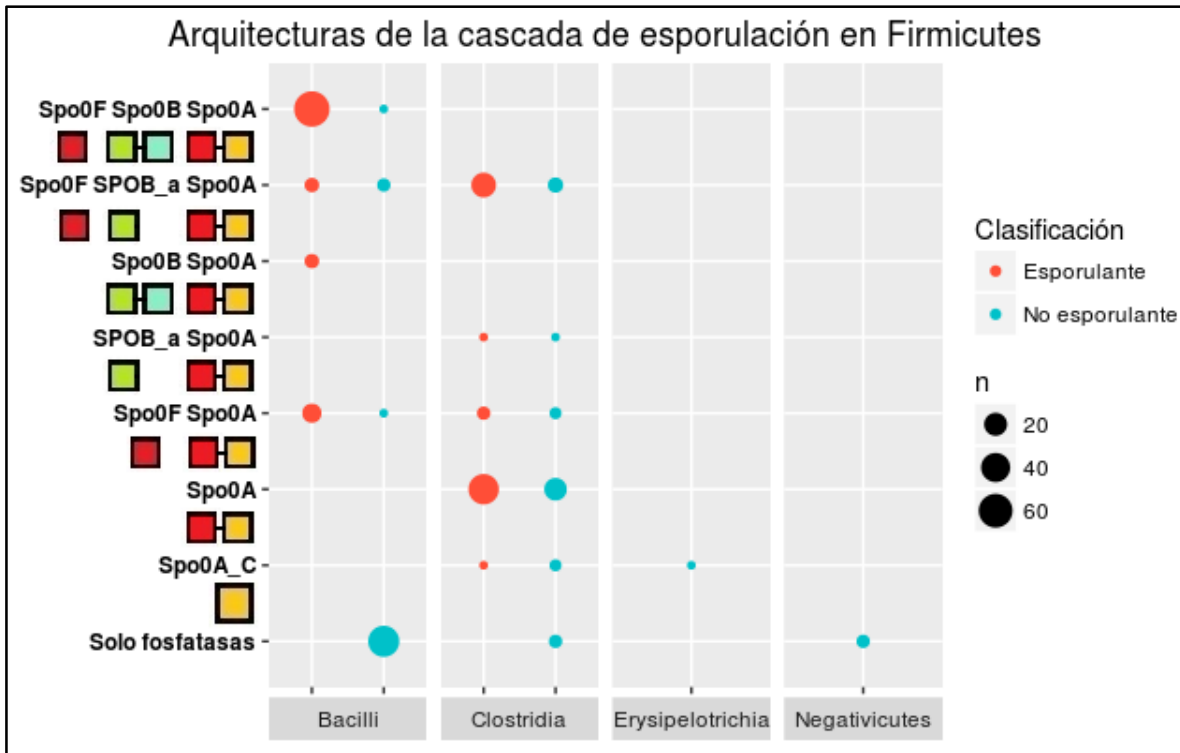


Figura 12. Conservación de los elementos del fosforrelevo en Firmicutes. Diversidad de arquitecturas de la vía completa en las clases Bacilli, Clostridia, Erysipelotrichia y Negativicutes. Clasificación de esporulantes en color. La densidad del punto representa la proporción de organismos por clase que cuentan con elementos de la vía.

Esto nos permite observar los grupos bacterianos que van adquiriendo el resto de los componentes hasta llegar a formar la vía completa de esporulación presente en *B. subtilis*. Adicionalmente podemos observar en color cuáles de ellos adquieren la capacidad de esporular, así como el mínimo de componentes necesarios para ello. Es interesante como una Clostridia puede esporular con la proteína reducida Spo0A_C pues el resto no puede hacerlo, la mayoría de las Clostridias cuentan solo con Spo0A. No hay Bacillus que solo tengan Spo0A_C o Spo0A, varios de ellos tienen al menos a Spo0F, Spo0B o SPOB_a. No hay Clostridias que cuenten con la cascada completa pero si hay Clostridias y Bacillus con la cascada semicompleta (con SPOB_a) que sí son capaces de esporular, se desconoce si en fosforrelevo o en TCS.

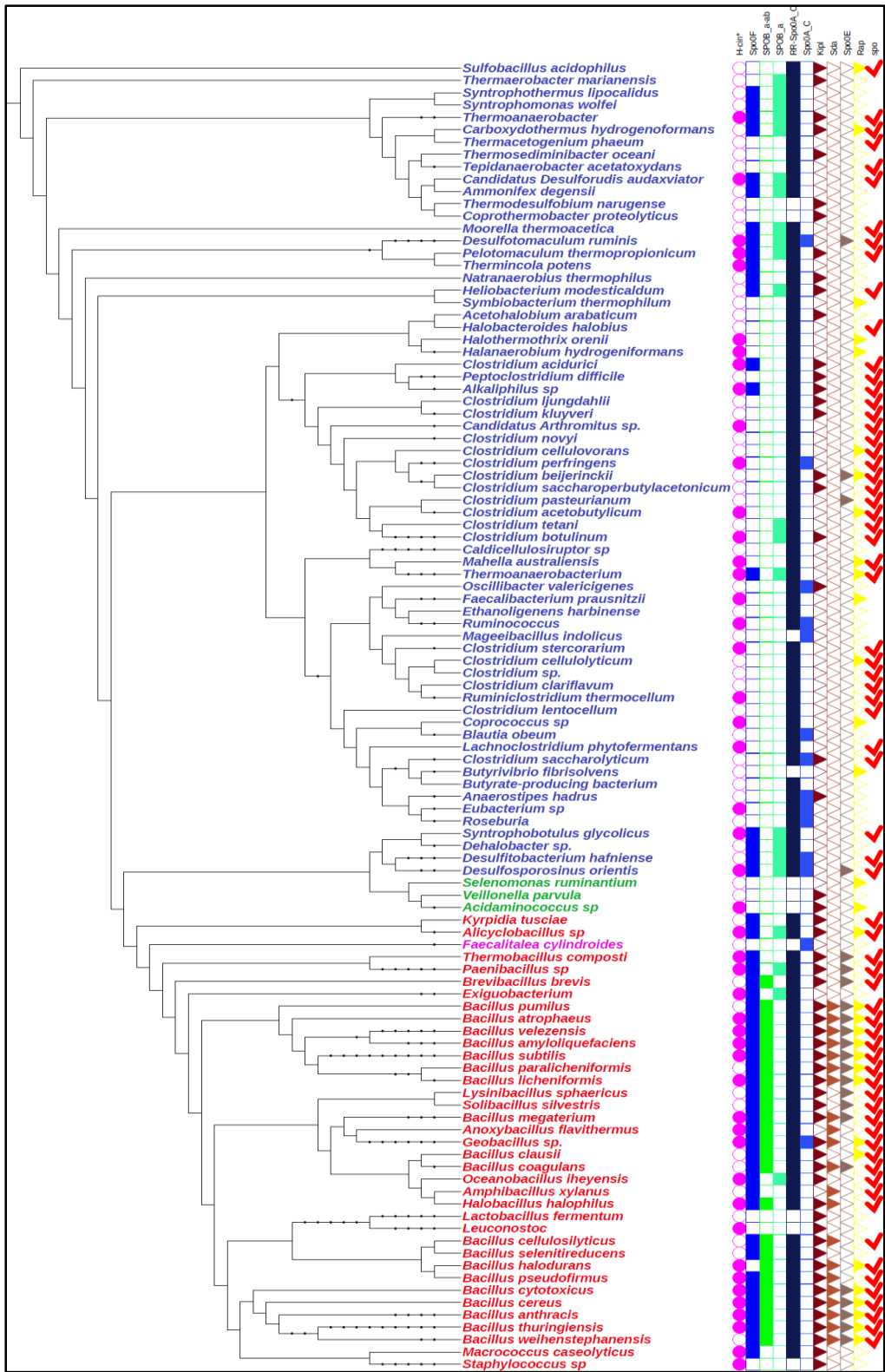


Figura 13. Filogenia 16S ribosomal de la conservación de los elementos de la cascada de la esporulación en las Clases Bacilli (rojo), Clostridia (azul), Erisipelotrichia (rosa) y Negativicutes (verde), el ○ rosa representa H-cinasas huérfanas, en □ spo0F (azul rey), □ Spo0B (verde) y □ Spo0A (azul marino), en △ las diferentes fosfatasa. Clasificación de esporulantes en ✓ roja solo organismos esporulantes.

Se grafica árbol filogenético 16S ribosomal por máxima verosimilitud donde podemos observar a Clostridia en azul, Negativicutes en verde, Erisipelotrichia en Rosa y Bacillus en Rojo. Del lado derecho se grafica la presencia o ausencia de cada componente de la cascada de esporulación y con paloma se representan todos los organismos que sí esporulan. Estos resultados coinciden con los árboles 21S ribosomales de Kunisawa y cols. Las bacterias termohalófilas anaeróbicas o Halanaerobiales se ramifican del resto de las especies de Clostridia. Un clado formado por las clases Negativicutes, Bacilli y Erysipelotrichia se coloca como una posición hermana a las especies de Clostridia (Kunisawa 2015).

Las cinasas se conservan en menor cantidad en Clostridia que en Bacillus. No todos los organismos esporulantes conservan H-cinasas de acuerdo a los modelos, esto nos dice que deben usar otras arquitecturas aún desconocidas.

En la gráfica se observa que las Clostridias son las más ancestrales y en su mayoría conservan un sistema de dos componentes. Sin embargo, eso no se cumple en todos los casos, ya que algunas de ellas conservan a Spo0F y aunque no cuentan con Spo0B en su forma completa, si cuentan con la proteína en su forma reducida Spo0B_a. Las Clostridias también cuentan con la proteína reducida Spo0A_C como una copia paráloga a la proteína completa. En contraste, los Bacillus en su mayoría cuentan con Spo0B y Spo0A completas y pocas excepciones cuentan con estas proteínas reducidas. Spo0F se conserva poco en Clostridia pero también hay Bacillus que no conservan Spo0F. Es importante resaltar que la presencia de Spo0A en su forma más completa es condición necesaria pero no suficiente para esporular.

Kipl y las proteínas Rap son muy conservadas a lo largo de todo el árbol aunque disminuyen hacia las Clostridias. Sda solo está presente en algunos Bacillus y Spo0E está escasamente presente en Clostridia pero abundante en la mayoría de géneros Bacillus.

Los clados no esporulantes como Negativicutes, Erisipelotrichia, Staphylococcus, Lactobacillus y Leuconostoc conservan escasos elementos de la cascada, puntualmente no cuentan con Spo0A. Es claro que no todos los Bacillus conservan todos los elementos, especialmente los no esporulantes y los más cercanos a Clostridia, y los géneros que conservan todos los elementos únicamente son *B. pumilus*, *B. atrophaus*, *B. velezensis*, *B. amylolicuefaciens*, *B. subtilis*, *B. licheniformis*, *B. paralicheniformis*, *B. welhenstphanensis*, *B. thuringensis*, *B. anthraxis*, *B. cereus* y *B. citotoxicus*.

DISCUSIÓN

El sesgo de información hacia los organismos más estudiados como *B. subtilis* y *C. acetobutylicum* aunado al escaso trabajo experimental para caracterizar proteínas esporulantes ha resultado en pocos modelos de estudio; y a pesar de la tasa explosiva en que las bases de datos están creciendo (Land et al. 2015), predecir con seguridad las funciones precisas de las regiones codificantes en un genoma con solo analizar su secuencia no es tarea fácil (Stephens 1998). Algunas limitaciones del modelo de *B. subtilis*, como la poca conservación de secuencias (~20%) entre los organismos más cercanos, hace que la búsqueda de homólogos a las H-cinasas KinA-KinE no sea suficiente para cubrir todo el espectro de posibles H-cinasas involucradas en las cascadas de esporulación, lo cual hace necesaria la predicción de H-cinasas como candidatas a proteínas relacionadas con la esporulación. Como alternativa, Vogel y cols. (2004) proponen que si se observa la misma combinación de dominios en dos proteínas diferentes es probable que mantengan la misma función y que hayan evolucionado a partir de un mismo ancestro por duplicación y no por ensamblado independiente o diferentes rutas de recombinación (Vogel et al. 2004). Pero la clasificación basada en la arquitectura de dominio puede ser totalmente engañosa cuando se usa sin el apoyo filogenético, mientras que si se complementa con el análisis del contexto genómico se pueden visualizar algunos procesos evolutivos que afectan la diversidad de TCS, como la duplicación, nacimiento de dominios, divergencia y eventos de fisión o fusión de genes (Capra and Laub 2012; Stragier 2002). Narikawa y cols. (2004) realizaron un análisis de los dominios PAS no conservados entre Cianobacterias que les permitió identificar algunos rasgos evolutivos como la duplicación o el “shuffling”; aunque los dominios especie-específicos, los adquiridos por transferencia lateral de genes o los que se perdieron en otros linajes, no se pueden descartar mediante este análisis (Narikawa et al. 2004). Es importante mencionar que una de las características más importantes de esta cascada de esporulación es que sus componentes tanto en *Bacillus* como en *Clostridia* se encuentran como genes huérfanos, es decir, HK y RR no están contiguos en el genoma; esto abre la puerta a estudios de coevolución en sistemas de tipo huérfanos que determinar bioinformáticamente siguen siendo un reto (Cheng et al. 2014). Por todo lo anterior, la metodología empleada en este trabajo arroja resultados más amplios que permiten obtener un mejor panorama de la conservación en el resto de bacterias a partir de la búsqueda de dominios homólogos, en lugar de proteínas homólogas.

Se sabe que las HK están presentes en todo el reino bacteriano, pero la capacidad de formar endosporas data desde antes del aumento de O₂ en la atmósfera. Algunas teorías han considerado que la esporulación se estableció a principios de la generación de divergencia en la evolución, inclusive se propone que el último ancestro común era una bacteria esporulante didérmica que más tarde fue perdiendo dichas características en muchos linajes dentro de los Firmicutes. Debido a que el phylum Firmicutes está más cerca de la raíz del árbol de la vida, la mayoría de las teorías proponen que la diversidad observada en este phylum podría ser explicada por pérdidas y no por ganancias génicas (Tocheva, Ortega, and Jensen 2009). Actualmente este proceso de esporular parece estar restringido a los Firmicutes, Gram-positivos que pertenecen a las clases Bacillus y Clostridia (Stragier 2002). Estas clases se separaron hace aproximadamente 2,7 mil millones de años y 2,3 mil millones de años respectivamente, y aunque forman endosporas en un proceso muy similar de esporulación sus diferencias son de suma importancia biológica; por ejemplo, la limitación de nutrientes no es la principal señal para el inicio de la esporulación en Clostridias, sino las condiciones ácidas en el medio ambiente (Dürre 2011). Estudios comparativos han confirmado que la cascada de fosforrelevo arriba de Spo0A no se conserva, particularmente entre Bacilos y Clostridias. Sin embargo, ~60 genes se han determinado como indispensables para la esporulación, siendo Spo0A y los factores sigma (σ_H , σ_F , σ_E , σ_G , y σ_K) universalmente conservados en los formadores de esporas (Galperin et al. 2012). De igual forma se han distinguido bacterias que parecen recientemente haber perdido la capacidad de esporular, y otras especies que no se ha observado que esporulen parecen tener la capacidad de esporular en base a la presencia de las firmas de esporulación mencionadas (Hutchison, Miller, and Angert 2014).

Se sabe que del 1- 2% de un proteoma procariota se compone de proteínas TCS pero solo algunos phylum tienen un pequeño porcentaje de especies en donde todas sus HKs y RRs están separados en el genoma: 20 de 2066 especies de Firmicutes, 2 de 635 en Actinobacteria, 6 de 235 en Bacteroidetes, 11 de los 2.246 en Gammaproteobacteria, 48 de 451 en Alphaproteobacteria, 7 de cada 366 en Betaproteobacteria, 4 de cada 108 en Chlamydias y 3 de 118 en Cyanobacterias (Salvado et al. 2015). Como se esperaba, los phylum antes mencionados son los más representativos de nuestros datos (Figura 9) donde representamos las bacterias que tienen al menos una HK huérfana, siendo las Cyanobacterias las que contienen hasta 80 HKs huérfanas por bacteria. Esto es de suma importancia pues los grupos con HKs huérfanas tienen la capacidad de diseñar la mayor

cantidad de fosforrelevos. En *B. subtilis* la disposición de los genes huérfanos KinA y Spo0F en lados opuestos del cromosoma juega un papel funcional importante para generar pulsos de Spo0A-P necesarios para asegurar la regulación adecuada del programa de esporulación en coordinación con la replicación del ADN (Narula et al. 2015).

En este trabajo decidimos evaluar únicamente la presencia o ausencia en cada phylum de los homólogos a H-cinasas esporulantes que cumplan con ser huérfanas, para identificar aquellas H-cinasas candidatas que aún no están caracterizadas como esporulantes y observar cómo ha sido la transmisión de sus arquitecturas a través de los phylum. Los resultados (Figura 8) muestran que Firmicutes es un phylum muy ancestral a partir del cual se transmitieron estas H-cinasas hacia los demás phylum con excepción de la arquitectura 2 que está presente desde los ancestros de estos. La distribución de las arquitecturas indica que los dominios PAS (en verde) son los más ancestrales y distribuidos, esto quiere decir que varios de los estímulos para iniciar la esporulación están relacionados con la cantidad de luz y la concentración de O₂ que es de gran importancia para medir la densidad de población en el medio, una de las principales señales para iniciar la esporulación. Las arquitecturas estudiadas resultan ser poco complejas en comparación con las Cianobacterias que pueden llegar a contar con más de 13 dominios PAS en tándem (Figura 6). En el clado de las Clostridias las H-cinasas con dominios TM y PAS no suelen estar fuertemente conservadas (Dürre 2011), lo que concuerda con la menor distribución de las H-cinasas modelo en este clado (Figura 13).

Por otro lado, las arquitecturas 1, 4, 5, 6 y 7 contienen dominios específicos de Firmicutes, ya que no son conservados en ningún otro phylum, sería interesante observar que tan relacionadas están estas arquitecturas específicas con el medio ambiente. También observamos que el grupo de bacterias esporulantes está enriquecido con todas estas proteínas modelo (Figura 11). Aunque es razonable esperar que muchos genes implicados en el proceso de esporulación sean altamente específicos para formadores de endosporas, este no es siempre el caso; algunos de estos genes están presentes en muchos no formadores de esporas y parecen haber evolucionado a funciones especiales en bacterias esporulantes como algunas H-cinasas, Spo0F y las fosfatasa Kipl y Rap (Figura 13) (Stragier 2002). Análisis comparativos de los dominios no conservados con el resto de H-cinasas huérfanas hacen falta para trazar a detalle la historia evolutiva de estas proteínas.

Es de importancia destacar que dentro de cada phylum las especies muestran una amplia variedad de organizaciones y fenotipos; si bien la presencia y ausencia de los elementos del fosforrelevo es importante, también lo es la conectividad de los circuitos entre proteínas, así como la organización de dominios en las proteínas que generan una amplia variedad de arquitecturas en TCS y fosforrelevo (Capra and Laub 2012; Salvado et al. 2015). De acuerdo a un censo realizado por Salvado y cols. (2015), proteínas con dominios HK y RR están presentes en el proteoma del 100% de las especies en los siguientes filos: Aquificae, Chlorobi, Verrucomicrobia, Chloroflexi, Cianobacterias, Deferribacteres, Deinococcus-Thermus, Dictyoglomi, Acidobacteria, Nitrospirae, Planctomycetes, Epsilonproteobacteria, Espiroquetas, Thermodesulfobacteria, y Thermotogae. Y en el resto de los phylum solo un pequeño porcentaje de las especies no cuenta con proteínas HK y/o RR. Por otro lado, las Deltaproteobacterias son el grupo con el porcentaje promedio más alto de proteínas TCS, en contraste con Tenericutes y Chlamydias que cuentan con el porcentaje más bajo (Salvado et al. 2015). Estos datos coinciden con nuestros resultados ya que los phylum mencionados son los más representativos y en particular las Deltaproteobacterias cuentan con el mayor número de copias parálogas de Spo0F (~20), lo que implica un alto número de fosforrelevo en su genoma.

El número de copias parálogas se consideró importante, porque si las duplicaciones génicas son altas los organismos evolucionan rápidamente y logran adaptarse con mayor éxito a diversos ambientes o diferentes presiones evolutivas. La proteína con más copias parálogas es Spo0F y su amplia distribución implica la existencia de fosforrelevo del tipo HK-RR-HPt-RR. Este tipo de fosforrelevo donde los elementos son independientes se cree que se generan porque varias H-cinasas pueden fosforilar al mismo RR, o bien, una H-cinasa puede fosforilar a más de un RR (Salazar and Laub 2015).

Como ha sido descrito en trabajos anteriores, Spo0B resulta exclusivo de la clase Bacilli (Figura 13), lo que indica un evento importante de especiación porque en Clostridia el análogo en función que recibe y transfiere el grupo fosforilo corresponde al dominio HPt. En su mayoría este dominio se encuentra en combinación con otros dominios TCS y en menor cantidad como dominio independiente únicamente en Firmicutes, Tenericutes, Actinobacteria, Bacteroidetes y Espiroquetas (Kato et al. 1997; Salvado et al. 2015). Curiosamente, Spo0B no comparte una similitud significativa de secuencia con los dominios HPt aun cuando ambos surgen de un dominio catalítico de una HK (Capra and Laub 2012;

Stragier 2002); inclusive la similitud entre los propios homólogos de ambas proteínas es de 20-35%, lo que hace todavía más difícil su reconocimiento. Mattoo y cols. (2008) encontró que Spo0B de *Bacillus anthracis* evolucionó a partir de una citrato / malato HK. Observaron también que el locus que abarca Spo0B tiene genes implicados en biosíntesis ribosomal, síntesis de proteínas, desarrollo, esporulación y división celular (*obg*, *rpmA*, *prp* y *rplU*), estos genes concuerdan con nuestros resultados y son de gran importancia para reconocer ortólogos de esta proteína. Mattoo sugiere que Spo0B ha estado bajo mucha presión evolutiva, pues ha adquirido muchas características nuevas en *B. anthracis*, por lo cual la búsqueda de ortólogos no se adapta a la definición convencional de estos. Incluso propone que la patogenicidad puede ser una de las principales razones para estos cambios (Mattoo et al. 2008).

Poco se sabe acerca de la presencia de Spo0B y Spo0F en Clostridium y se ha vuelto poco trivial concluir si los elementos fueron ganados o perdidos. Nuestros resultados indican que a pesar de que en el resto de bacterias se fue perdiendo la capacidad de esporular, el fosforrelevo implicó una ganancia de elementos que solo sucedió en un clado dentro de Bacilli. Tres razones apoyan esta teoría i) se sabe que evolutivamente el fosforrelevo resulta ser más sofisticado que un TCS, ii) la regulación en las ancestrales Clostridias es por las mismas HKs mientras que en Bacilli no solo se ganaron transferasas, sino también fosfatasas y iii) se ha demostrado que los formadores de esporas tienden a tener genomas más grandes (mayor a 2300 Kb) que los no formadores de esporas como los lactobacilos, los estafilococos y los estreptococos (Galperin et al. 2012). Nosotros proponemos que Spo0B tuvo que ser creado en el brinco hacia un fosforrelevo, y que su formación empezó desde las Clostridias pues varias de ellas conservan el dominio SPOB_a (Figura 13), en contraste los bacilos casi no cuentan con la proteína reducida. Curiosamente la mayoría de los organismos que cuentan con SPOB_a también cuentan con Spo0F, lo que deja abierta la pregunta de si estas proteínas actúan como un fosforrelevo en Clostridia. Igualmente, hay Bacillus que no conservan Spo0F y podrían funcionar como un sistema de dos componentes. Sin embargo, hacen falta estudios experimentales para comprobar estas posibilidades, pero es de particular interés porque hasta la fecha no se ha descrito la presencia de Spo0B completa en Clostridia, con excepción de un gen similar encontrado en *C. tetani* (Dürre 2011), y su posible origen ha sido poco estudiado.

Tal y como lo describe Brown y cols (1994) nuestros resultados reiteran que el regulador Spo0A no se encuentra fuera de Firmicutes, ni siquiera en parientes cercanos a *Bacillus*

como *Listeria* y *Staphylococcus* (Figura 13), lo que sugiere que la regulación de la actividad genética durante la transición a la fase estacionaria mediada por Spo0A es exclusiva de bacterias Gram-positivas formadoras de endosporas (Brown et al. 1994). Se sabe que en las clases Bacillus y Clostridia esta proteína participa tanto en TCS (Clostridium) como en fosforrelevo (Bacillus) y ambas clases han desarrollado varios tipos de diferenciación celular regulados por Spo0A, por lo que es probable que su presencia en Firmicutes que no esporulan esté relacionada con otros procesos de diferenciación celular como la producción de biopelículas o producción de toxinas fratricidas (Lopez, Vlamakis, and Kolter 2009) y en Clostridia con procesos de virulencia o síntesis de disolventes (Paredes, Alsaker, and Papoutsakis 2005; Pettit et al. 2014). En el árbol filogenético (Figura 13) se muestra que los géneros que no conservan Spo0A coincidentemente no tienen la capacidad de esporular, no obstante, algunos de ellos sí cuentan con la misma proteína en su forma reducida Spo0A_C, pero su presencia no les restablece la capacidad de esporular, lo que sugiere que probablemente no sea funcional. Algunas otras bacterias esporulantes conservan esta forma reducida como un parálogo que pudo surgir por la necesidad de crear más diseños de las vías de diferenciación celular.

En los conjuntos de presencia o ausencia de proteínas (Figura 12) encontramos grupos que cuentan con Spo0A, pero no Spo0F, así como grupos que cuentan con Spo0A y Spo0F, pero no Spo0B. Esto nos lleva a la conclusión de que Spo0A y Spo0F pertenecieron a TCS separados que sólo en *Bacilli* fue necesario unir en fosforrelevo a través de Spo0B, tal vez, para lograr mayor regulación sobre Spo0A al decidir entre varios destinos celulares.

De acuerdo con Salazar y cols. (2015) el fosforrelevo de *B. subtilis* evolucionó a partir de una antigua vía canónica de dos componentes. Una de las posibilidades que proponen es que una vía de esporulación ancestral involucraba H-cinasas que activaban directamente a una ancestral Spo0A con la posterior integración de los dos componentes involucrados en el medio. Tal escenario habría requerido el cambio en la especificidad de la H-cinasa a preferir el nuevo regulador Spo0F (57,000 veces más eficiente cinéticamente) en vez de H-cinasa⇒Spo0A, Spo0F a preferir Spo0B y Spo0B a preferir Spo0A (1.1 millones de veces más eficiente cinéticamente) (Grimshaw et al. 1998). Varios experimentos se han realizado con las proteínas de *C. botulinum* y *C. acetobutylicum* donde varias HK pueden fosforilar directamente a Spo0A de *B. subtilis* pero no a Spo0F; mientras que KinA de *B. subtilis* no puede fosforilar a Spo0A de *Clostridium*. Estas observaciones pueden implicar que las especies Bacilli ganaron los componentes intermedios Spo0F y Spo0B. Sin embargo,

también podría ser que el fosforrelevo estuvo presente en un ancestro común de los bacilos y clostridios con la pérdida subsecuente de los elementos intermedios de la vía en clostridios. Además, las H-cinasas de Clostridia no son homólogos obvios de las H-cinasas KinA-E de Bacilos, lo que sugiere que las especies de *Clostridium* pudieron haber perdido todos los componentes excepto Spo0A y luego adaptar otras H-cinasas para fosforilar directamente a Spo0A debido a que las señales para la iniciación de la esporulación son diferentes en ambos organismos. Varios investigadores proponen que este tipo de variaciones en HKs están vinculados a los distintos hábitats y capacidades metabólicas específicas de cada formador de esporas (Al-Hinai, Jones, and Papoutsakis 2015; Salazar and Laub 2015; Stragier 2002).

Nuestros resultados nos permiten sugerir que ancestralmente Spo0F perteneció a otro fosforrelevo o TCS debido a su amplia presencia en bacterias no esporulantes. El hecho de que Spo0A y Spo0A_C estén presentes en Clostridias no esporulantes significa que esta proteína perteneció a un TCS HK –Spo0A que estuvo relacionado con otra ruta de diferenciación celular, o bien, se perdió la capacidad de esporular tras la pérdida de otras firmas esporulantes abajo en la cascada. La filogenia hace evidente que los Bacillus adquieren Spo0A y Spo0A_C de las Clostridias. Por otro lado, algunos Bacillus y Clostridias generan SPOB_a, aunque no se sabe si es funcional. Y solo un grupo de Bacillus añade SPOB_ab concluyendo la formación de Spo0B completa para construir la vía completa (Figura 14).

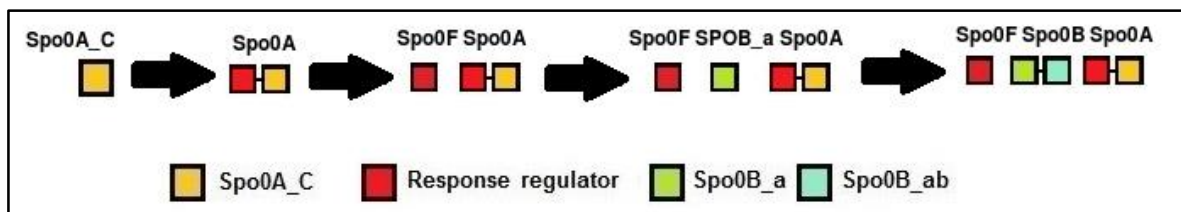


Figura 14.- Formación de la cascada del fosforrelevo en Firmicutes. Los pasos en la evolución de la cascada de fosforrelevo empieza a partir del dominio Spo0A_C, los dominios Response_reg se van añadiendo consecutivamente; de igual forma SPOB_a se añade a la cascada y culmina en la proteína completa Spo0B que únicamente se logra en Bacilli.

En el caso de los elementos de regulación, Mitrophanov y cols. (2008) realizaron una búsqueda de homólogos en algunas bases de datos y notaron que Sda y Kipl, ambas represoras de KinA, muestran diferentes distribuciones filogenéticas. Sda está bien conservado sólo entre especies de Bacillus (Burkholder, Kurtser, and Grossman 2016),

mientras que los homólogos de Kipl se encuentran en una amplia gama de bacterias Gram-positivas y Gram-negativas. Esto sugiere que algunas fosfatasa pudieron haber sido adquiridas lateralmente y haber conservado sus propiedades funcionales a pesar de la divergencia a nivel de secuencia (Mitrophanov and Groisman 2008). Stephens y cols. argumentan que los posibles inicios de Kipl y su regulador KipA están relacionados con una enzima de la vía del nitrógeno (Stephens 1998). En concordancia con nuestros resultados (Figura 9) Kipl está ampliamente distribuido en casi todos los phylum, curiosamente también en donde no hay homólogos de H-cinasas, esto nos hace pensar que en cada transferencia de esta proteína su función se adaptó a diferentes H-cinasas o bien, adoptó otra. Las proteínas Rap muestran resultados similares aunque en la literatura han sido encontradas solo en *Bacillus* (Stragier 2002). Estas proteínas están más relacionadas con el sentido de metabolitos propios del vecindario por lo que su transferencia ha sido necesariamente amplia. No así, Sda que inhibe a la misma H-cinasa que Kipl y se distribuye sólo en algunos *Bacillus*. También encontramos a Spo0E solo en Firmicutes, por lo que sí es específico de la esporulación, además es interesante que puede existir en varias copias parálogas lo que implica que Spo0A tenga bastante regulación y la regulación que aportan estas copias parálogas contribuyan a la heterogeneidad de las poblaciones en *Bacillus*.

En general, encontramos que no existe phylum en donde se encuentren las cuatro fosfatasa, con la excepción de Firmicutes, lo que implica que las fosfatasa de *B. subtilis* Sda y Spo0E no fueron adquiridas ni han sido transferidas al resto de los phylum, y que su presencia este probablemente relacionada con los numerosos destinos celulares, inclusive hay varios phylum que no cuentan con ningún tipo de fosfatasa. Esto resalta el grado de sofisticación que implica su presencia, pues cada fosfatasa que se añade a la vía actúa como un punto de control que funciona como la expresión booleana "AND", debido a que los RR son fosforilados sólo si las HK afines realizan su actividad cinasa "Y" si la acción de las fosfatasa se inhibe. Este hecho es de suma importancia biológica en la mayoría de los *Bacillus* pero no en *Clostridia* donde la regulación es proporcionada por las mismas HKs. Además, se ha comprobado que ni Spo0A ni Spo0F o σ^H son limitantes de la velocidad en la que la esporulación se lleva a cabo en *B. subtilis*, sino que estas proteínas generan fluctuaciones estocásticas sobre los niveles de Spo0A~P mediado por el flujo de fosfatos en el fosforrelevo y que la base para la heterogeneidad de poblaciones se genera por el ruido o drenado del flujo de fosfatos mediado por las fosfatasa en puntos clave de control

(Chastanet et al. 2010; De Jong, Veening, and Kuipers 2010; Narula et al. 2012; Stragier 2002; Vishnoi et al. 2013).

El hecho de que el total de proteínas exista sólo en Firmicutes significa que solo en este phylum existe una presión evolutiva tal que conduce a múltiples destinos celulares los cuales, además, requieren varios puntos de control. Dicha distribución de las proteínas probablemente refleja adaptaciones distintivas de estos organismos a sus nichos ecológicos específicos. (Galperin et al. 2012) Sin embargo, análisis experimental de la función génica sigue siendo una necesidad crítica que es indispensable se realice en conjunto con métodos genómicos comparativos para poder describir la diversidad de los componentes de TCS y trazar su evolución (Capra and Laub 2012; Stephens 1998).

CONCLUSIONES

Aunque es evidente que la capacidad de esporular se extingue por pérdidas génicas en Firmicutes y en otros phylum, la perfección del fosforrelevo indica ganancias solo en el clado de los organismos que conservaron la capacidad de esporular. La conservación de los dominios homólogos a cada componente de la cascada de fosforrelevo de la esporulación de *B. subtilis* y Clostridia reveló una variedad de arquitecturas distintas de esta vía de señalización. Varias de las H-cinasas han sido transmitidas al resto de los phyla, mientras otras son arquitecturas que todavía no han sido compartidas más allá de Firmicutes. La evolución del resto de la vía de señalización se observa a partir de las Clostridias como un sistema de dos componentes de la forma Kin-Spo0A, después algunas Clostridias y Bacillus conservan Spo0F y las proteínas reducidas SPOB_a y Spo0A_C, y culmina en redes como el fosforrelevo con complejos niveles de control en forma de fosfatasas especialmente en las especies del clado Bacillus.

PERSPECTIVAS

- Añadir la clasificación de cada bacteria de acuerdo a los destinos celulares que son capaces de tomar puede revelar la evolución de Spo0A en relación con las necesidades de la población.
- Ampliar este tipo de análisis para observar la conservación de los factores sigma, pues son firmas importantes para el compromiso de la esporulación (Narula et al. 2012) que se encuentran abajo en la cascada y se sabe son variables entre Clostridia y Bacilli.

BIBLIOGRAFÍA

- Al-Hinai, Mohab A, Shawn W Jones, and Eleftherios T Papoutsakis. 2015. "The Clostridium Sporulation Programs: Diversity and Preservation of Endospore Differentiation." *Microbiology and molecular biology reviews : MMBR* 79(1): 19–37. <http://mmbbr.asm.org/content/79/1/19.full>.
- Ashby, Mark K. 2006. "Distribution, Structure and Diversity Of 'bacterial' genes Encoding Two-Component Proteins in the Euryarchaeota." *Archaea (Vancouver, B.C.)* 2(1): 11–30. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2685588&tool=pmcentrez&rendertype=abstract>.
- Barba-Ostria, Carlos A. 2014. "Los Sistemas de Dos Componentes: Circuitos Moleculares Versátiles." *Tip* 17(1): 62–76. <http://www.sciencedirect.com/science/article/pii/S1405888X14703201>.
- Bick, Matthew J et al. 2009. "How to Switch Off a Histidine Kinase: Crystal Structure of Geobacillus Stearothermophilus KinB with the Inhibitor Sda." *Journal of Molecular Biology* 386(1): 163–77.
- Brown, David P. et al. 1994. "Characterization of spo0A Homologues in Diverse Bacillus and Clostridium Species Identifies a Probable DNA-Binding Domain." *Molecular Microbiology* 14(3): 411–26.
- Browne, Hilary P et al. 2016. "Culturing of 'unculturable' Human Microbiota Reveals Novel Taxa and Extensive Sporulation." *Nature* 533(7604): in press. <http://www.nature.com/doi/10.1038/nature17645>.
- Brunsing, Ryan L. et al. 2005. "Characterization of Sporulation Histidine Kinases of Bacillus Anthracis." *Journal of Bacteriology* 187(20): 6972–81.
- Burbulys, D., K. a. Trach, and J. a. Hoch. 1991. "Initiation of Sporulation in B. Subtilis Is Controlled by a Multicomponent Phosphorelay." *Cell* 64(3): 545–52.
- Burkholder, William F, Iren Kurtser, and Alan D Grossman. 2016. "Replication Initiation Proteins Regulate a Developmental Checkpoint in Bacillus Subtilis." *Cell* 104(2): 269–79. [http://dx.doi.org/10.1016/S0092-8674\(01\)00211-2](http://dx.doi.org/10.1016/S0092-8674(01)00211-2).
- Capra, Emily J, and Michael T Laub. 2012. "Evolution of Two-Component Signal Transduction Systems." *Annual Review of Microbiology* 66(1): 325–47. <http://www.ncbi.nlm.nih.gov/pubmed/22746333>.
- Chastanet, Arnaud et al. 2010. "Broadly Heterogeneous Activation of the Master Regulator for Sporulation in Bacillus Subtilis." *Proceedings of the National Academy of Sciences of the United States of America* 107(18): 8486–91.
- Cheng, R R, F Morcos, H Levine, and J N Onuchic. 2014. "Toward Rationally Redesigning Bacterial Two-Component Signaling Systems Using Coevolutionary Information." *Proc Natl Acad Sci U S A* 111(5): E563-71. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3918776/pdf/pnas.201323734.pdf>.
- Cunningham, Katherine A, and William F Burkholder. 2009. "The Histidine Kinase Inhibitor Sda Binds near the Site of Autophosphorylation and May Sterically Hinder Autophosphorylation and Phosphotransfer to Spo0F." *Molecular Microbiology* 71(3): 659–77.
- Devi, S. N. et al. 2015. "In Vivo Functional Characterization of the Transmembrane Histidine Kinase KinC in Bacillus Subtilis." *Microbiology* 161: 1092–1104. <http://mic.sgmjournals.org/cgi/doi/10.1099/mic.0.000054>.
- Driks, A. 2002. "Overview: Development in Bacteria: Spore Formation in Bacillus Subtilis." *Cellular and Molecular Life Sciences* 59(3): 389–91.
- Dürre, Peter. 2011. "Ancestral Sporulation Initiation." *Molecular Microbiology* 80(3): 584–87.
- Earl, Ashlee M et al. 2010. "NIH Public Access." 16(6): 1–11.
- Eddy, Sean R. 2010. "HMMER User's Guide." *Hmmer Manual* (March): 0–93.

- Eswaramoorthy, P., and M. Fujita. 2010. "Systematic Domain Deletion Analysis of the Major Sporulation Kinase in *Bacillus Subtilis*." *Journal of Bacteriology* 192(6): 1744–48. <http://jb.asm.org/cgi/doi/10.1128/JB.01481-09>.
- Eswaramoorthy, Prahathes et al. 2011. "Expression Level of a Chimeric Kinase Governs Entry into Sporulation in *Bacillus Subtilis*." *Journal of Bacteriology* 193(22): 6113–22.
- Fabret, Céline, Victoria a Feher, and James a Hoch. 1999. "Two-Component Signal Transduction in *Bacillus Subtilis* : How One Organism Sees Its World MINIREVIEW Two-Component Signal Transduction in *Bacillus Subtilis* : How One Organism Sees Its World †." *Journal of bacteriology* 181(7): 1975–83.
- Fonslow, Bryan R et al. 2013. "Biological Insights from Structures of Two-Component Proteins." 10(1): 54–56.
- Fujita, M, and R Losick. 2005. "Evidence That Entry into Sporulation in *Bacillus Subtilis* Is Mediated by Gradual Activation of a Master Regulator." *Genes Dev.* 19: 2236–2244.
- Gallego del Sol, Francisca, and Alberto Marina. 2013. "Structural Basis of Rap Phosphatase Inhibition by Phr Peptides." *PLoS Biology* 11(3): 14–16.
- Galperin, Michael Y et al. 2012. "Genomic Determinants of Sporulation in Bacilli and Clostridia: Towards the Minimal Set of Sporulation-Specific Genes." *Environmental Microbiology* 14(11): 2870–90.
- Galperin, Michael Y, Anastasia N Nikolskaya, and Eugene V Koonin. 2001. "Novel Domains of the Prokaryotic Two-Component Signal Transduction Systems." *FEMS microbiology letters* 203(1): 11–21. <http://www.ncbi.nlm.nih.gov/pubmed/11557134>.
- Garrity, George M, Julia a Bell, Timothy G Lilburn, and East Lansing. 2004. "Taxonomic Outline of the Prokaryotes." *Bergey's Manual of Systematic Bacteriology* 2(May): 1–399.
- González-Pastor, José Eduardo. 2011. "Cannibalism: A Social Behavior in Sporulating *Bacillus Subtilis*." *FEMS Microbiology Reviews* 35(3): 415–24.
- Grimshaw, Charles E et al. 1998. "Synergistic Kinetic Interactions between Components of the Phosphorelay Controlling Sporulation in *Bacillus Subtilis*." *Biochemistry* 37(5): 1365–75.
- Henry, Jonathan T, and Sean Crosson. 2011. "Ligand-Binding PAS Domains in a Genomic, Cellular, and Structural Context." *Annual review of microbiology* 65: 261–86. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3298442&tool=pmcentrez&rendertype=abstract>.
- Hoch, J. A., and K. I. Varughese. 2001. "Keeping Signals Straight in Phosphorelay Signal Transduction." *Journal of Bacteriology* 183(17): 4941–49.
- de Hoon, Michiel J.L., Patrick Eichenberger, and Dennis Vitkup. 2010. "Hierarchical Evolution of the Bacterial Sporulation Network." *Current Biology* 20(17): R735–45. <http://linkinghub.elsevier.com/retrieve/pii/S0960982210007773>.
- Hutchison, Elizabeth A, David A Miller, and Esther R Angert. 2014. "Sporulation in Bacteria : Beyond the Standard Model." *Microbiology Spectrum* (May 2016): 1–15.
- Ireton, Keith, and Alan D Grossman. 1994. "A Developmental Checkpoint Couples the Initiation of Sporulation to DNA Replication in *Bacillus Subtilis*." *The EMBO journal* 13(7): 1566–73. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=394986&tool=pmcentrez&rendertype=abstract>.
- Jiang, Min, Weilan Shao, Marta Perego, and James A Hoch. 2000. "Multiple Histidine Kinases Regulate Entry into Stationary Phase and Sporulation in *Bacillus Subtilis*." *Molecular Microbiology* 38(3): 535–42.
- De Jong, Imke G., Jan Willem Veening, and Oscar P Kuipers. 2010. "Heterochronic Phosphorelay Gene

- Expression as a Source of Heterogeneity in *Bacillus Subtilis* Spore Formation." *Journal of Bacteriology* 192(8): 2053–67.
- Jung, Kirsten, Luitpold Fried, Stefan Behr, and Ralf Heermann. 2012. "Histidine Kinases and Response Regulators in Networks." *Current Opinion in Microbiology* 15(2): 118–24. <http://dx.doi.org/10.1016/j.mib.2011.11.009>.
- Kato, Masato, Takeshi Mizuno, Toshiyuki Shimizu, and Toshio Hakoshima. 1997. "Insights into Multistep Phosphorelay from the Crystal Structure of the C-Terminal HPT Domain of ArcB." *Cell* 88(5): 717–23.
- Kunisawa, Takashi. 2015. "Evolutionary Relationships of Completely Sequenced Clostridia Species and Close Relatives." *International Journal of Systematic and Evolutionary Microbiology* 65(11): 4276–83.
- Land, Miriam et al. 2015. "Insights from 20 Years of Bacterial Genome Sequencing." *Functional & integrative genomics* 15(2): 141–61. <http://www.ncbi.nlm.nih.gov/pubmed/25722247> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4361730>.
- Lewis, R J et al. 2000. "Domain Swapping in the Sporulation Response Regulator Spo0A." *Journal of molecular biology* 297(3): 757–70.
- López, Daniel, and Roberto Kolter. 2010. "Extracellular Signals That Define Distinct and Coexisting Cell Fates in *Bacillus Subtilis*." *FEMS Microbiology Reviews* 34(2): 134–49.
- Lopez, Daniel, Hera Vlamakis, and Roberto Kolter. 2009. "Generation of Multiple Cell Types in *Bacillus Subtilis*." *FEMS Microbiology Reviews* 33(1): 152–63.
- Mattoo, Abid R. et al. 2008. "Spo0B of *Bacillus Anthracis* - A Protein with Pleiotropic Functions." *FEBS Journal* 275(4): 739–52.
- McLoon, Anna L. et al. 2011. "Spatial Regulation of Histidine Kinases Governing Biofilm Formation in *Bacillus Subtilis*." *Journal of Bacteriology* 193(3): 679–85.
- Mearls, Elizabeth B., and Lee R. Lynd. 2014. "The Identification of Four Histidine Kinases That Influence Sporulation in *Clostridium Thermocellum*." *Anaerobe* 28: 109–19.
- Mitrophanov, Alexander Y AY, and EA Eduardo a Groisman. 2008. "Signal Integration in Bacterial Two-Component Regulatory Systems." *Genes & development* 22(19): 2601–11. <http://genesdev.cshlp.org/content/22/19/2601>. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2751022&tool=pmcentrez&rendertype=abstract>.
- Mongkolthanaruk, Wiyada. 2012. "Classification of *Bacillus* Beneficial Substances Related to Plants, Humans and Animals." *Journal of Microbiology and Biotechnology* 22(12): 1597–1604.
- Narikawa, Rei, Shinobu Okamoto, Masahiko Ikeuchi, and Masayuki Ohmori. 2004. "Molecular Evolution of PAS Domain-Containing Proteins of Filamentous Cyanobacteria through Domain Shuffling and Domain Duplication." *DNA Research* 11(2): 69–81.
- Narula, Jatin et al. 2015. "Chromosomal Arrangement of Phosphorelay Genes Couples Sporulation and DNA Replication." *Cell* 162(2): 328–37.
- Narula, Jatin, Seram N Devi, Masaya Fujita, and Oleg A Igoshin. 2012. "Ultrasensitivity of the *Bacillus Subtilis* Sporulation Decision." *Proceedings of the National Academy of Sciences of the United States of America* 109(50): E3513-22. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3528541&tool=pmcentrez&rendertype=abstract>.
- Nasir, Arshan, and Gustavo Caetano-anollés. 2015. "A Phylogenomic Data-Driven Exploration of Viral Origins

- and Evolution.” (September).
- Nicholson, W. L. 2002. “Roles of Bacillus Endospores in the Environment.” *Cellular and Molecular Life Sciences* 59(3): 410–16.
- Ortet, Philippe et al. 2015. “P2CS: Updates of the Prokaryotic Two-Component Systems Database.” *Nucleic Acids Research* 43(D1): D536–41.
- Paredes, C J, K V Alsaker, and E T Papoutsakis. 2005. “A Comparative Genomic View of Clostridial Sporulation and Physiology.” *Nature Reviews Microbiology* 3(12): 969–78. <http://www.ncbi.nlm.nih.gov/pubmed/16261177>.
- Park, Soo Young, Seung Hwan Park, and Soo Keun Choi. 2012. “Characterization of Sporulation Histidine Kinases of Paenibacillus Polymyxa.” *Research in Microbiology* 163(4): 272–78. <http://dx.doi.org/10.1016/j.resmic.2012.02.003>.
- Pettit, Laura J et al. 2014. “Functional Genomics Reveals That Clostridium Difficile Spo0A Coordinates Sporulation, Virulence and Metabolism.” *BMC genomics* 15(1): 160. <http://bmcbgenomics.biomedcentral.com/articles/10.1186/1471-2164-15-160>.
- Phillips, Z E V, and M A Strauch. 2002. “Cellular and Molecular Life Sciences Bacillus Subtilis Sporulation and Stationary Phase Gene Expression.” 59: 392–402.
- Piggot, Patrick J. 1996. “Spore Development in Bacillus Subtilis.” *Current Opinion in Genetics and Development* 6: 531–37.
- Rowland, Susan L. et al. 2004. “Structure and Mechanism of Action of Sda, an Inhibitor of the Histidine Kinases That Regulate Initiation of Sporulation in Bacillus Subtilis.” *Molecular Cell* 13(5): 689–701.
- Salazar, Michael E, and Michael T Laub. 2015. “Temporal and Evolutionary Dynamics of Two-Component Signaling Pathways.” *Current opinion in microbiology* 24: 7–14. <http://www.ncbi.nlm.nih.gov/pubmed/25589045>.
- Salvado, Baldiri, Ester Vilaprinyo, Albert Sorribas, and Rui Alves. 2015. “A Survey of HK, HPT, and RR Domains and Their Organization in Two-Component Systems and Phosphorelay Proteins of Organisms with Fully Sequenced Genomes.” *PeerJ* 3: e1183. <https://peerj.com/articles/1183>.
- Schultz, Joachim E, and Janani Natarajan. 2013. “Regulated Unfolding: A Basic Principle of Intraprotein Signaling in Modular Proteins.” *Trends in Biochemical Sciences* 38(11): 538–45. <http://dx.doi.org/10.1016/j.tibs.2013.08.005>.
- Sonenshein, Abraham L. 2000. “Control of Sporulation Initiation in Bacillus Subtilis.” *Current Opinion in Microbiology* 3(Figure 1): 561–66.
- Stephen Spiro and Ray Dixon. 2010. *Sensory Mechanisms in Bacteria: Molecular Aspects of Signal Recognition*.
- Stephens, C. 1998. “Bacterial Sporulation: A Question of Commitment?” *Current biology : CB* 8: R45–48.
- Stephenson, K, and J a Hoch. 2001. “PAS-A Domain of Phosphorelay Sensor Kinase A: A Catalytic ATP-Binding Domain Involved in the Initiation of Development in Bacillus Subtilis.” *Proceedings of the National Academy of Sciences of the United States of America* 98(26): 15251–56.
- Stephenson, Keith, and James a Hoch. 2002a. “Two-Component and Phosphorelay Signal-Transduction Systems as Therapeutic Targets.” *Current opinion in pharmacology* 2(5): 507–12. <http://www.ncbi.nlm.nih.gov/pubmed/15032730>.
- Stephenson, Keith, and James a. Hoch. 2002b. “Evolution of Signalling in the Sporulation Phosphorelay.” *Molecular Microbiology* 46(2): 297–304.
- Stewart, Richard C. 2010. “Protein Histidine Kinases: Assembly of Active Sites and Their Regulation in Signaling

- Pathways." *Current Opinion in Microbiology* 13(2): 133–41. <http://dx.doi.org/10.1016/j.mib.2009.12.013>.
- Stragier, Patrick. 2002. "A Gene Odyssey: Exploring the Genomes of Endospore-Forming Bacteria." In *Bacillus Subtilis and Its Closest Relatives*, Washington, DC: ASM Press, 519–25.
- Stragier, Patrick, and Richard Losick. 1996. "MOLECULAR GENETICS OF SPORULATION IN BACILLUS SUBTILIS." *Annu. Rev. Genet* 30: 297–341.
- Talukdar, Prabhat K. et al. 2015. "Updates on the Sporulation Process in Clostridium Species." *Research in Microbiology* 166(4): 225–35.
- Taylor, Barry L, and Igor B. Zhulin. 1999. "PAS Domains: Internal Sensors of Oxygen, Redox Potential, and Light." *Microbiology and molecular biology reviews: MMBR* 63(2): 479–506. <http://mmbbr.asm.org/content/63/2/479.full>.
- Tocheva, Elitza I, Davi R Ortega, and Grant J Jensen. 2009. "Sporulation, Bacterial Cell Envelopes and the Origin of Life." *bioRxiv* 266(August): 131–40. <http://biorxiv.org/content/early/2016/02/29/041624.abstract%5Cnhttp://link.springer.com/10.1007/978-1-4020-9636-5>.
- Tojo, Shigeo, Kazutake Hirooka, and Yasutaro Fujita. 2013. "Expression of kinA and kinB of Bacillus Subtilis, Necessary for Sporulation Initiation, Is under Positive Stringent Transcription Control." *Journal of Bacteriology* 195(8): 1656–65.
- Varughese, Kottayil Iype. 2002. "Molecular Recognition of Bacterial Phosphorelay Proteins." *Current Opinion in Microbiology* 5(2): 142–48.
- Vishnoi, Monika et al. 2013. "Triggering Sporulation in Bacillus Subtilis with Artificial Two-Component Systems Reveals the Importance of Proper Spo0A Activation Dynamics." *Molecular Microbiology* 90(1): 181–94.
- Vlamakis, Hera et al. 2013. "Sticking Together: Building a Biofilm the Bacillus Subtilis Way." *Nature Reviews Microbiology* 11(3): 157–68. <http://www.nature.com/doi/10.1038/nrmicro2960>.
- Vogel, Christine et al. 2004. "Structure, Function and Evolution of Multidomain Proteins." *Current Opinion in Structural Biology* 14(2): 208–16.
- Vreede, Jocelyne et al. 2003. "PAS Domains. Common Structure and Common Flexibility." *Journal of Biological Chemistry* 278(20): 18434–39.
- Wang, L. et al. 2001. "Dissection of the Functional and Structural Domains of Phosphorelay Histidine Kinase A of Bacillus Subtilis." *Journal of Bacteriology* 183(9): 2795–2802.
- Wang, Ling, Roberto Grau, Marta Perego, and James A. Hoch. 1997. "A Novel Histidine Kinase Inhibitor Regulating Development in Bacillus Subtilis." *Genes and Development* 11(19): 2569–79.
- Wörner, Kristina, Hendrik Szurmant, Christina Chiang, and James A. Hoch. 2006. "Phosphorylation and Functional Analysis of the Sporulation Initiation Factor Spo0A from Clostridium Botulinum." *Molecular Microbiology* 59(3): 1000–1012.
- Wu, R et al. 2013. "Insight into the Sporulation Phosphorelay: Crystal Structure of the Sensor Domain of Bacillus Subtilis Histidine Kinase, KinD." *Protein science: a publication of the Protein Society* 22(5): 564–76. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3649258&tool=pmcentrez&rendertype=abstract>.