

## Análisis bioinformático de proteínas involucradas en la transducción de señales en arqueas

Jaime Ramírez-Acosta<sup>1</sup>, Sofia Fonseca-Yepetz<sup>2</sup>, Carolina Lisette Moreno-Galván<sup>3</sup>, Isaac Serna Gutiérrez<sup>4</sup>, José Manolo Montufar Rodríguez<sup>5</sup>, Naurú Idalia Vargas-Maya<sup>6</sup>, Marcos Vicente-Gómez<sup>7</sup>, Felipe Padilla-Vaca<sup>8</sup>, Bernardo Franco<sup>9</sup>

<sup>1</sup>Departamento de Biología, División de Ciencias Naturales y Exactas, Universidad de Guanajuato.

<sup>5</sup>Colegio el Álamo, San Miguel Petapa, Guatemala.  
bfranco@correo.mx<sup>9</sup>

### Resumen

Los sistemas de transducción de señales son esenciales para que las células puedan percibir su entorno y responder de la manera más específica posible. El paradigma de estos sistemas de señalización en bacterias son los sistemas de dos componentes. Como su nombre lo indica, se tratan de sistemas de una proteína, típicamente anclada a la membrana plasmática que percibe una señal, se autofosforila y en una reacción de transfosforilación, activa a un segundo componente, una proteína citoplásmica que típicamente es un factor transcripcional. En la búsqueda de sistemas de señalización en arqueas, este trabajo se enfoca en analizar algunas cinasas de sistemas de dos componentes en arqueas Asgard, las cuales son consideradas los ancestros más cercanos a los eucariontes y extenso debate hay también de su origen. Los resultados obtenidos sugieren que, a diferencia de otras proteínas conocidas, en las arqueas Asgard las cinasas de sistemas de dos componentes son mayoritariamente citoplásmicas, pero con dominios (específicamente tipo PAS) que generalmente están involucrados en percibir oxígeno, luz o estado redox, siendo una peculiaridad en los sistemas de dos componentes. También se discute las implicaciones de esto en la evolución de las arqueas Asgard.

**Palabras clave:** Sistemas de dos componentes; Asgard arqueas; origen de células eucariotas; dominios PAS.

### Introducción

Todos los seres vivos requieren de mecanismos moleculares que les permitan percibir el ambiente para posteriormente activar procesos internos que les permitan responder adecuadamente. Las bacterias no son una excepción de esta necesidad. En estos organismos se tienen mecanismos de señalización altamente sofisticados que se han encontrado en la mayoría de las bacterias, en varias arqueas (Stock *et al.*, 2000), pero también en organismos eucariontes como son levaduras, hongos y plantas (Galperin *et al.*, 2018), aunque, hasta ahora, no se han encontrado en mamíferos (Schaller *et al.*, 2011).

Los sistemas de dos componentes están compuestos típicamente por una proteína anclada a la membrana celular llamada, cinasa sensora, que contiene al menos dos dominios, uno que percibe la señal específica que activa al sistema y otro citoplásmico con actividad de cinasa, es decir, que consume ATP para transfosforilar al segundo componente de estos sistemas: el regulador de la respuesta. El regulador de la respuesta es típicamente una proteína citoplásmica que tienen actividad de factor transcripcional. Existen ejemplos en los que el regulador de la respuesta tiene alguna actividad enzimática o bien, interacciona con otras proteínas, como es el caso del sistema de quimiotaxis (Padilla-Vaca *et al.*, 2017).

Típicamente, en los sistemas de dos componentes su mecanismo bioquímico de activación es como se muestra en la Figura 1.

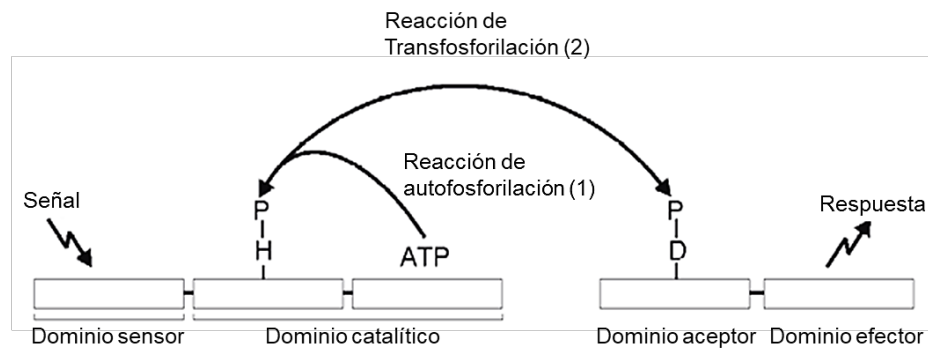


Figura 1. Mecanismo de activación de un sistema de dos componentes clásico (Tomado y modificado de Szurmant et al. (2007)). Se indican los dominios principales de la cinasa sensora (izquierda) y regulador de la respuesta (derecha).

Típicamente, la cinasa sensora, tiene al menos un cruce transmembranal que delimita el dominio sensor que queda en dirección del periplasma, y es en este dominio donde es percibida la señal. Una vez que percibe la señal, mediante un cambio en la conformación tridimensional de la proteína, permite que el dominio catalítico una e hidrolice ATP (1, Figura 1), lo que permite su autofosforilación en un residuo de histidina altamente conservado. En presencia del regulador de la respuesta, el fosfato es transferido por una reacción de transfosforilación (también llamado fosforelevo) al regulador de la respuesta en un residuo de aspártico también altamente conservado (2 en la Figura 1) (Szurmant *et al.*, 2007). Un cambio importante en la bioquímica de la transducción de señales entre bacterias y eucariontes es que en bacterias los residuos que son fosforilados en proteínas relacionadas con la transducción de señales son en los residuos de histidina y aspártico, mientras que en eucariontes esto ocurre en residuos de serina, treonina y tirosina. Claro que existen excepciones en ambos casos, pero no son tan comunes, pero no menos importantes e interesantes.

La arquitectura de los reguladores de respuesta no es tan diversa como lo es para las cinasas sensoras. Existen tres tipos de sistemas de dos componentes (Figura 2), que están las cinasas clásicas como la mostrada en la Figura 1 en la que se tiene un dominio sensor y un dominio catalítico. En la Figura 2 se muestran los dos tipos adicionales de sistemas de dos componentes, el primero, en el que la cinasa es híbrida, es decir, contiene un segundo dominio fosfotransmisor compuesto por un dominio aceptor con un residuo conservado de aspártico y un dominio fosfotransmisor secundario con una histidina conservada. Este segundo dominio es un mecanismo regulador ya que por un lado controla la velocidad de transferencia del grupo fosfato al regulador de la respuesta y también actúa como un regulador negativo ya que en ausencia de la señal que activa al sistema funciona como fosfatasa. Y el sistema multi pasos, en el que el dominio secundario fosfotransmisor es similar al sistema con cinasa híbrida, excepto que este segundo dominio son dos proteínas independientes.

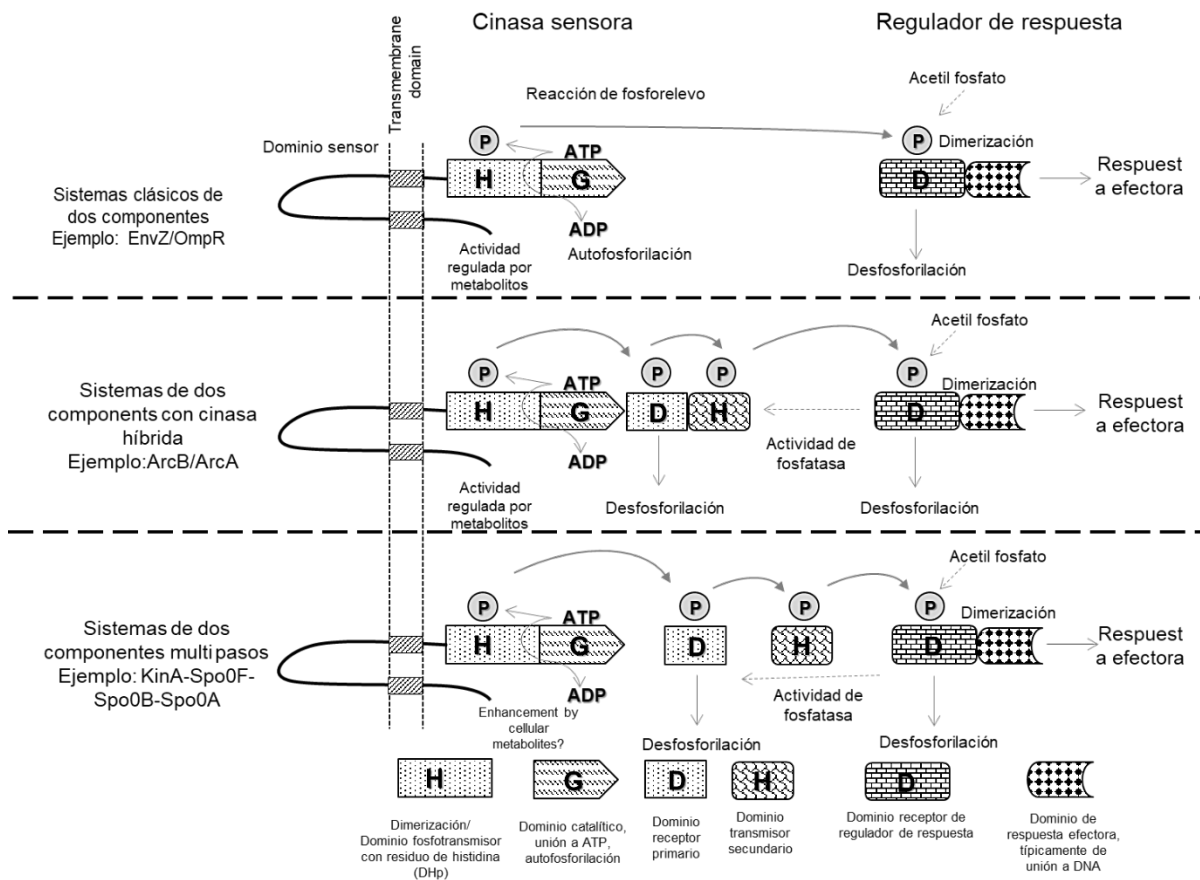


Figura 2. Estructura de los tipos de sistemas de dos componentes conocidos a la fecha. Se muestran dos tipos esenciales además del sistema clásico. En la parte superior se muestra un ejemplo de un sistema clásico compuesto por una cinasa de un solo dominio como el mostrado en la Figura 1. Al centro, se muestra un ejemplo en el que la cinasa es híbrida, la cual se caracteriza por tener dos módulos, uno que es de autofosforilación y un segundo de fosfotransferencia que tiene dos funciones, fosforilar al regulador de respuesta o defosforilarlo a través de actividad de fosfatasa en este dominio. El tercer caso es de los sistemas multicomponente, en el cual se caracteriza por tener el dominio fosfotransmisor como proteínas independientes, cuya función es también transferir el grupo fosfato como en un relevo, pero simplemente son sistemas que tienen péptidos independientes. Se indica la función de cada dominio en estas proteínas. Tomado y modificado de: Padilla-Vaca et al. (2017).

Las reacciones de fosfotransferencia pueden ocurrir entre dos subunidades de las mismas proteínas o bien, ocurrir sobre la misma proteína (Figura 3). Datos recientes sugieren que existen los dos tipos de mecanismos. En el primero, la reacción de autofosforilación es de tipo intermolecular, la cual es entre los dos monómeros de la cinasa sensora. Esto ocurre tanto en cinasas de tipo clásicas como híbridas y hay pocas evidencias de que ocurre de igual forma en las cinasas de multipasos aunque por la estructura que presentan lo más probable es que también así sea. En el caso de la pérdida de la señal, el dominio de fosfotransferencia de la cinasa funciona como una fosfatasa, en este caso, la reacción ocurre de manera intramolecular, es decir, sobre el mismo monómero (Teran-Melo et al., 2018).

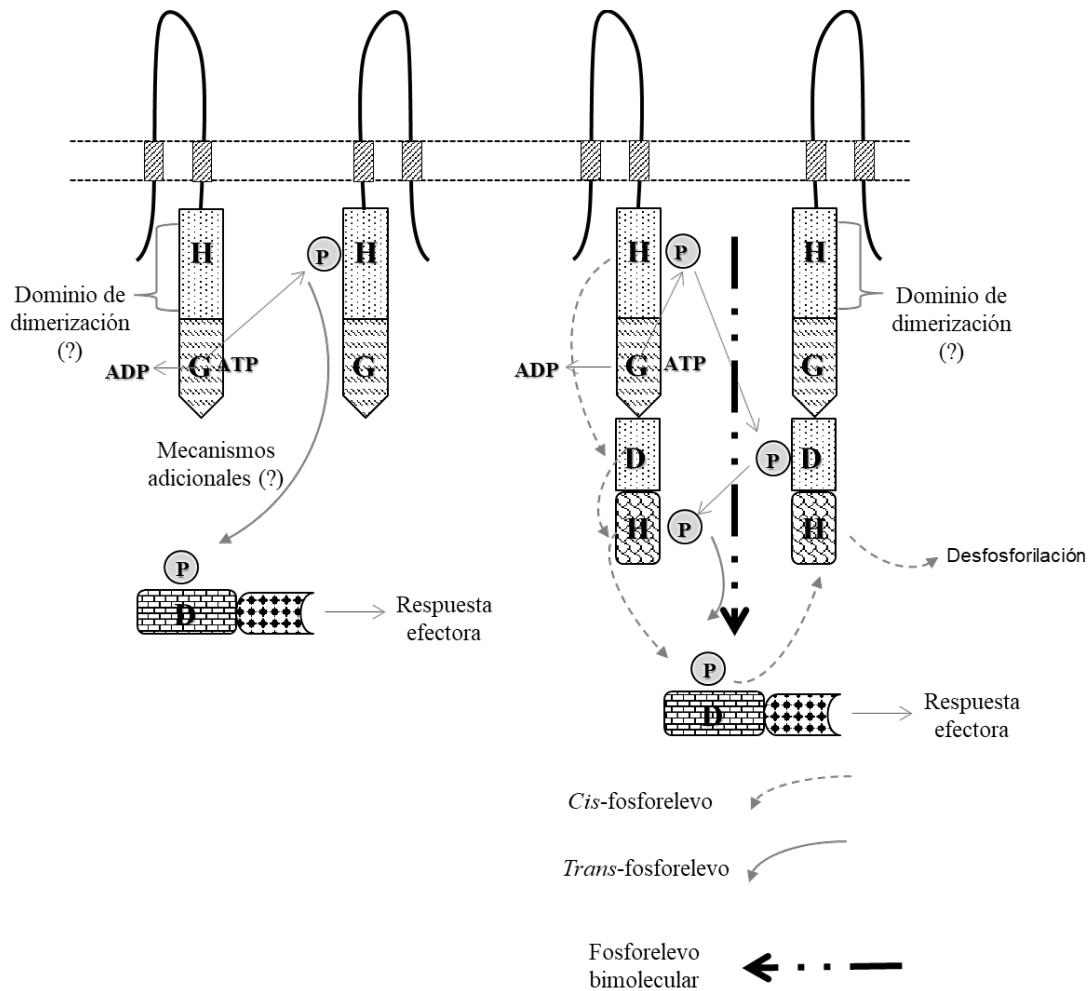


Figura 3. Las rutas de fosforilación en las cinasas clásicas son de manera intermolecular (trans), es decir, se transfiere el grupo fosfato de un monómero al otro. En el caso de las cinasas híbridas, también es una reacción intramolecular, con algunas excepciones reportadas, pero la reacción de defosforilación hasta nuestro conocimiento, es una reacción intramolecular (cis), es decir, sobre un mismo monómero. Tomado y modificado de: Padilla-Vaca et al. (2017).

Las rutas de fosforilación han sido ampliamente estudiadas usando proteínas producidas de manera recombinante y con ellas hacer experimentos de fosforilación *in vitro*. El estudio clásico bioquímico de estos sistemas involucra el uso de ATP radiactivo (en la posición  $\gamma$  con  $^{32}\text{P}$ ), lo que lo hace sumamente peligroso, ahora hay avances con métodos químicos que permiten el retardo de la migración en gel mediante compuestos conocidos como Phos-tag (Kinoshita-Kikuta *et al.*, 2022), una molécula que une grupos fosfato y que permite retardar la migración de fosfo-proteínas en geles desnaturizantes comunes (SDS-PAGE). Sea cual sea la técnica empleada, estos ensayos han permitido elucidar los mecanismos de fosfotransferencia así como algunos ejemplos de moléculas que inhiben su actividad, siendo estos últimos de interés para controlar infecciones de organismos bacterianos patógenos (Ma y Phillips-Jones, 2021). El gran reto para esto es lograr reconstituir en sistemas *in vitro* cinasas completas y ancladas a membranas artificiales para poder evaluar su actividad ante la presencia de agentes inhibidores, por lo que resulta un reto poder producirlas de manera funcional. Un ejemplo es el hallazgo de que la siamycina I funciona como un inhibidor del sistema de dos componentes de percepción del quorum del patógeno oportunista *Enterococcus faecalis*, que controla la formación de biopelículas y la producción de proteasas que forman parte de su arsenal de factores de virulencia (Ma y Phillips-Jones, 2021).

Los sistemas de señalización son diversos, incluso dentro de la misma especie. En el caso del sistema de quimiotaxis en el género de *Geobacter*, podemos encontrar múltiples cinasas (como es el caso de *Geobacter bremensis*) o ninguna como es el caso de *Geobacter soli* (Silva y Salgueiro, 2021), lo que sugiere que los sistemas de dos componentes son muy diversos y complejos y que estudiar en todos los organismos posibles abre la oportunidad de encontrar nuevas formas de percibir el ambiente.

El sistema de señalización que fue el primer paradigma de los sistemas de transducción de señales en bacterias está involucrado en un proceso de reconocimiento del medio ambiente vital para la supervivencia que es la quimiotaxis (Falke *et al.*, 1997). Sin embargo, se han identificado muchos sistemas de señalización en bacterias, de los cuales en la mayoría de los casos no han sido caracterizados en detalle ni se conoce su función o el estímulo para el cual responden.

Una de las principales razones por las que el estudio estructural y funcional de sistemas de dos componentes es importante obedece a la estrecha relación que tiene con la evolución de la vida en la tierra, ya que son sistemas de señalización que surgieron primero y luego fueron transferidos y evolucionaron en otros organismos, incluso en eucariotes. Adicionalmente, son rutas de señalización altamente específicas, es decir, una cinasa en general tiene un solo regulador de respuesta (Yamamoto *et al.*, 2005). Esta especificidad ha sido un tema muy importante en la evolución de los sistemas de dos componentes, ya que se ha demostrado que su origen fue por duplicación génica y luego, mediante la acumulación de pocas mutaciones (aspecto sumamente interesante, ya que se esperaría fueran muchas mutaciones), adquirieron especificidad por un regulador de respuesta y diversidad (Nocedal y Laub, 2022), estableciendo rutas de señalización altamente especializadas. El principal reto es lograr conocer la enorme diversidad de estos sistemas en el universo de organismos procariontes.

Una característica evolutiva también importante es la relación que hay entre el tamaño del genoma del organismo y el número de sistemas de dos componentes. En la Figura 4 se muestra una gráfica en la que se puede mostrar una relación curiosa al respecto, entre más pequeño es el genoma, menor es el número de sistemas de dos componentes y viceversa. Sin embargo, en la naturaleza existen excepciones muy interesantes e importantes de estudiar ya que pueden dar información adicional de cómo estos sistemas han ido evolucionando desde el origen de la vida en nuestro planeta.

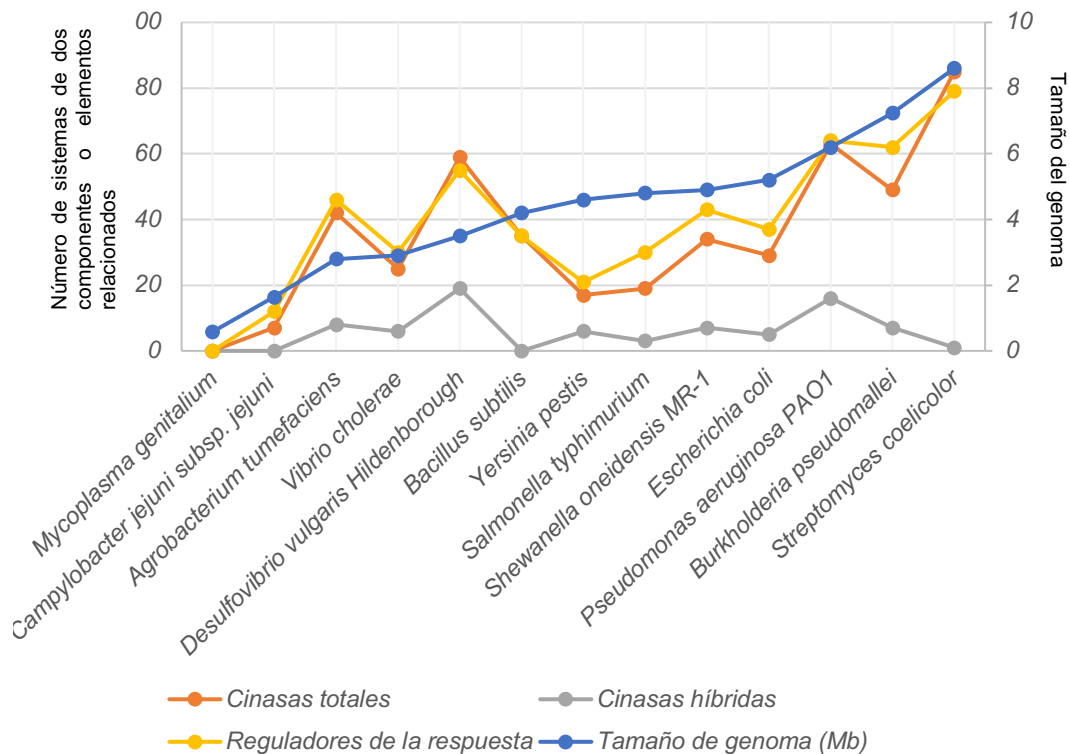


Figura 4. Existe una correlación entre el tamaño del genoma bacteriano y el número de sistemas de señalización. En la gráfica se muestra en el eje izquierdo el número de cinzas totales, híbridas y reguladores de la respuesta contra el tamaño del genoma (eje de la derecha). Aunque la muestra es pequeña, se muestra esta tendencia a tener un mayor número de sistemas de señalización entre más grande es el genoma. Datos tomados de Zhang *et al.* (2006) y Ma y Phillips-Jones (2021).

En el presente proyecto, se pretende abordar mediante herramientas bioinformáticas, aspectos estructurales y filogenéticos de algunas de las proteínas relacionadas con la percepción de señales ambientales en Lokiarchaeotas, un grupo de bacterias aisladas de las profundidades del océano ártico y que se piensa son el ancestro más directo a los eucariontes. Estas bacterias colectivamente se han asignado como el superphylum Asgardchaeta (Spang *et al.*, 2015). Recientemente se ha logrado tener secuencias genómicas de algunas especies de Lokiarchaeotas, lo que permite estudiar algunos aspectos importantes de estas.

Pero ¿por qué estudiar estos organismos? Sin duda, desde su descubrimiento han sacudido el árbol de la vida, ya que tienen características sumamente peculiares, la más importante es la posibilidad de tener (o haberla perdido) la capacidad de fagocitar o remodelar sus membranas, siendo esta cualidad muy relevante para el origen de los eucariontes ya que está íntimamente relacionado con la formación de los organelos celulares (Watson, 2019; Spang *et al.*, 2015) y esto ha sido confirmado mediante la compleja tarea de aislar y caracterizar estos organismos que son de crecimiento extremadamente lento en el laboratorio (Imachi *et al.*, 2020). Por tal motivo, la caracterización detallada de estos organismos no es tarea fácil en el laboratorio, por lo que las herramientas computacionales son una oportunidad para tratar de inferir algunos aspectos de la fisiología de estas bacterias.

En el presente trabajo, se muestran los resultados de un análisis exhaustivo de las proteínas anotadas como cinasas sensoras provenientes del metagenoma de fondo marino (Levasseur *et al.*, 2017) en el que se han identificado proteínas relacionadas a transducción de señales en *Candidatus Lokiarchaeum* sp GC14\_75, de los cuales hay 14 cinasas, una proteína hipotética que contiene los dominios necesarios para ser una cinasa sensora y 4 reguladores de la respuesta y contrastarlas con las cinasas sensoras de *Candidatus Prometheoarchaeum syntrophicum* MK-D1 cuyo genoma está completamente secuenciado (Imachi *et al.*, 2020) y hay 12 secuencias identificadas. De este organismo no hay reguladores de la respuesta anotados como tales.

En la Figura 5 se muestra un alineamiento de las cinasas de estos organismos (de cada uno entre sí), dejando ver que es claro que el dominio carboxilo terminal está altamente conservado en *C. Lokiarchaeum* sp, pero en el caso de *C. Prometheoarchaeum* son mucho más conservadas entre sí, incluso su secuencia permitió en la anotación del genoma tener su identificación con homólogos conocidos, lo que sugiere que son proteínas que tienen la capacidad de hacer la reacción de fosfotransferencia en ambos casos pero quizá la señal sea diferente. Como se puede ver en la Figura 5, no están muy conservadas entre sí en el caso de *C. Lokiarchaeum* sp, lo que sugiere que cada una puede responder a una señal diferente. Dada la naturaleza del ambiente en el que viven estos microorganismos y su importancia con respecto al origen de los eucariontes, es importante conocer algunos aspectos básicos de sus sistemas de señalización, si están conservados con otros organismos o si bien, están conservados exclusivamente en estas arqueas.

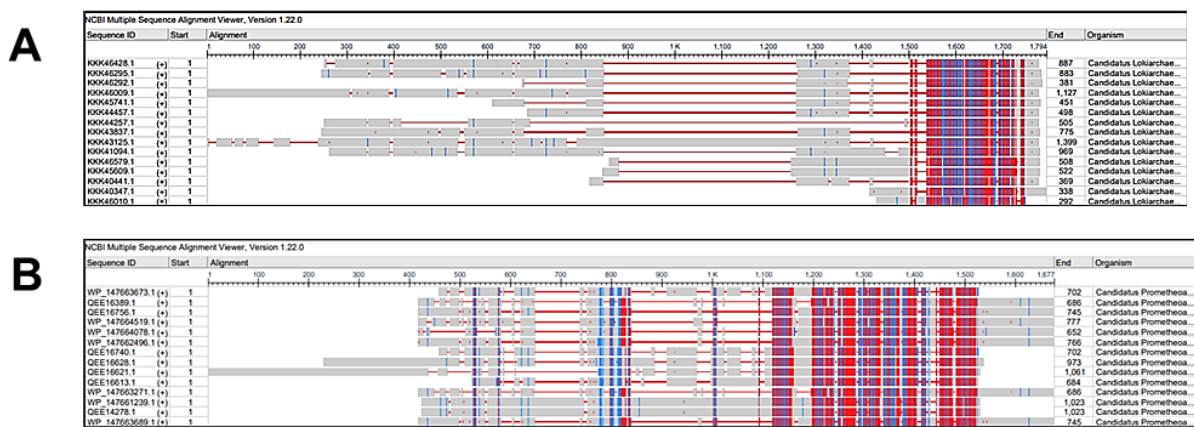


Figura 5. Alineamiento COBALT de las proteínas cinasas sensoras de *C. Lokiarchaeum* sp (panel A) y *C. Prometheoarchaeum* (panel B). Los dominios amino terminal en ambos casos son lo menos conservado en los dos casos, lo que sugiere pueden estar involucradas en la percepción de señales diferentes.

Por todo lo anterior, se busca identificar dominios estructurales que sugieran ser los responsables de la percepción de la señal en estos organismos y resultado de esto, identificar posibles cinasas homólogas que sugieran la posible señal que estas proteínas sean capaces de percibir.

## Resultados

Para analizar todas las cinasas de estos organismos se tomaron las siguientes estrategias. Primero, comparar todas las secuencias entre sí, tomando en cuenta los dos ejemplos de organismos y tipos de secuenciación. El hecho que una secuencia sea un metagenoma, implica riesgos de generar secuencias quiméricas que pueden constituir un artefacto en la interpretación de la homología de estos genes. El segundo paso fue generar modelos de alta calidad mediante el algoritmo AlphaFold2 (Jumper *et al.*, 2021) y comparar las estructuras de ellas. Finalmente, utilizando también modelos de AlphaFold2 depositados en la base de datos AlphaFold Protein Structure Database (<https://alphafold.ebi.ac.uk/>) del microorganismo modelo *Escherichia coli* como referentes específicos para las cinasas sensoras. Desde que se obtuvieron las secuencias, en el caso de *C. Prometheoarchaeum* la anotación automática de su genoma las identificó como homólogos algunas cinasas ya conocidas, como son el caso de ArcB y AtoS.

El análisis general de las secuencias reveló en todos los casos que el amino terminal es muy variable y que precisamente corresponde al modulo de detección de la señal. En muchos casos, se encontraron uno o varios dominios PAS. Este dominio ha sido reconocido como un importante módulo en proteínas que está relacionado con la detección de luz, oxígeno, estado redox (Vreede *et al.*, 2003). Estos dominios son el claro ejemplo de por qué el modelo estructural es tan importante, son dominios con poca conservación de secuencia, pero con una alta conservación estructural, donde el plegamiento  $\alpha/\beta$  es típica en estos dominios y de hecho fácilmente identificables. Los resultados de todos los dominios identificados en estas proteínas se resumen en la Tabla 1.

**Tabla 1.** Principales dominios identificados en arqueas Asgard

C. Lokiarcheaum (Metagenoma)	
Número de acceso	Dominios*
KKK45741.1	PAS (2); BaeS; AdeS_HK; HK_VicK; PRK15347; HATPase_c (2)
KKK40441.1	PAS; BaeS; AdeS_HK; HK_VicK; HATPase_c (2)
KKK44257.1	VicK; HK_Walk; PRK11360; HATPase_c (3)
KKK46537.1	BaeS; AdeS_HK; HK_Walk
KKK46579.1	Heme pocket; PAS (4); BaeS; HATPase_c (3); AdeS_HK; HK_Walk
KKK43125.1	Heme pocket (4); REC_NtrC1-like; ActR_PrrA_rreg; AtoC; PAS (13); GAF_2; BaeS; AdeS_HK; HK_VicK
KKK45609.1	PAS_3; BaeS; HK_walk; AdeS_HK;
KKK46428.1	BaeS; HK_Walk; Sensor Box 1 (3); PAS (10); AdeS_HK ; Heme pocket (3)
KKK46295.1	Heme pocket (2); HATPase_EvgS-ArcB-TorS; HK_Walk; PAS (9); PAS_3 (2); AdeS_HK
KKK44457.1	Heme pocket; HK_Walk; HK_VicK; HATPase_EvgS-ArcB-TorS; PAS (4) Sensory Box 1
KKK40347.1	BaeS ; BaeS_Smes; PRK10364; HATPase; HK_Walk; HK_VicK; AdeS_HK
KKK46010.1	HisKA ; REC; CheY; PRK11100; PRK10365; KdpD
KKK46009.1	Heme Pocket (3); PAS (12) (contiene 12 PAS en 5 superfamilias); BaeS; AdeS_HK; HATPase
KKK41094.1	Putative Active Site (3); Sensory Box (3); Heme Pocket (3); ATP Binding Site (1); ATP Lid (1); Mg-Binding Site (1) GxG Motif (1); Contiene 7 PAS en 3 superfamilias; BaeS (1); BaeS_Smes (1); HK-Walk (1); HK-VicK (1); AdeS_HK (1); HATPase (1); HATPase_c (2); PRK09303 (1)
KKK43837.1	Putative Active Site (1); Sensory Box (1); Heme Pocket (1); ATP Binding Site (1); ATP Lid (1); Mg-Binding Site (1); GxG Motif (1); Contiene 3 PAS en 1 superfamilia; BaeS (1); BaeS_Smes (1); HK-Walk (1); HK-VicK (1); AdeS_HK (1); HATPase (1); HATPase_c (2); PRK11360 (1)
KKK46292.1	Sensory Box (1); ATP Binding Site (1); ATP Lid (1); Mg-Binding Site (1); GxG Motif (1); PAS (1); BaeS (1); BaeS_Smes (1); HK-Walk (1); HK-VicK (1); AdeS_HK (1); HATPase (1); HATPase_c (2); PRK09303 (1)
C. Prometheoarchaeum (Genoma Secuenciado)	
QEE16613.1	Putative active site (1); Sensory box (1); Heme Pocket (1); ATP Binding Site (1); ATP Lid (1); Mg-Binding Site (1); GxG Motif (1); PAS (4); Contiene 4 PAS en 2 superfamilias; Periplasmic Binding Protein Type 2 (1); BaeS (1); BaeS_Smes (1); HK-Walk (1); HK-VicK (1); TMAO_TorS (1); AdeS_HK (1); HATPase_Walk_EvgS-ArcB-TorS like (1); HATPase_c (2)
QEE16628.1	Putative active site (2); Sensory box (2); Heme Pocket (2); ATP Binding Site (1); ATP Lid (1); Mg-Binding Site (1); GxG Motif (1); Contiene 7 PAS en 3 superfamilias; BaeS (1); BaeS_Smes (1); HK-Walk (1); HK-VicK (1); TMAO_TorS (1); AdeS_HK (1); HATPase_Walk_EvgS-ArcB-T (1); HATPase_c (2); PRK11107 (1)
WP_147663689.1	Putative Dimer Interface (1); Sensory Box (2); ATP Binding Site (1); ATP Lid (1); Mg-Binding Site (1); Metal Binding Site (1); GxG Motif (1); Active site (1); Contiene 2 PAS en 2 superfamilias; BaeS (1); BaeS_Smes (1); HK-Walk (1); HK-VicK (1); TMAO_TorS (1); AdeS_HK (1); HATPase_AtoS like (1); HATPase_c (2); REC (4 que comprende 1 CheY, 2 REC y 1 Response_reg; PRK11360 (AtoS) (1)
WP_147663673.1	Heme Pocket (2); PAS (7); contiene 7 PAS en 3 superfamilias; TMAO_TorS ; BaeS; PRK11091; AdeS_HK; HATPase
QEE16740.1	Heme Pocket (2); contiene 7 PAS en 3 superfamilias; BaeS ; TMAO_TorS ; PRK11091; AdeS_HK; HATPase
WP_147664519.1	PAS (1); BaeS; AdeS_HK; HATPase_c (2)
WP_147662496.1	PAS (8); BaeS; PRK13557; HATPase_c (2); Heme pocket (2)
WP_147664078.1	Heme pocket (2); PAS (3); BaeS; HATPase_c (2); AdeS_HK; HK_Walk
QEE16621.1	Heme pocket (2); PAS (9); GAF_2; TMAO_torS; HK_VicK; HK_Walk; AdeS_HK; HATPase_c (2)
QEE16756.1	PAS (3); BaeS; PRK11360; HATPase_c (2); CheY; REC (2); ActR_PrrA_rreg; Metal binding site
QEE16389.1	PBP2_HisK; HK_Walk; HK_VicK
WP_147661239.1	PAS (11); BaeS

\* Dominios identificados por Conserved Domain Search, NCBI. Entre paréntesis se indica el número de dominios identificados de un tipo específico.

Respecto a la comparación de las secuencias de las cinasas de cada organismo muestra dos aspectos interesantes. En la Figura 1 panel A se muestra un análisis filogenético de las cinasas de los dos tipos de secuencias de arqueas Asgard comparando con todas las cinasas de *E. coli*, como prototipo de sistemas de dos componentes y que la mayoría se han caracterizado funcionalmente.

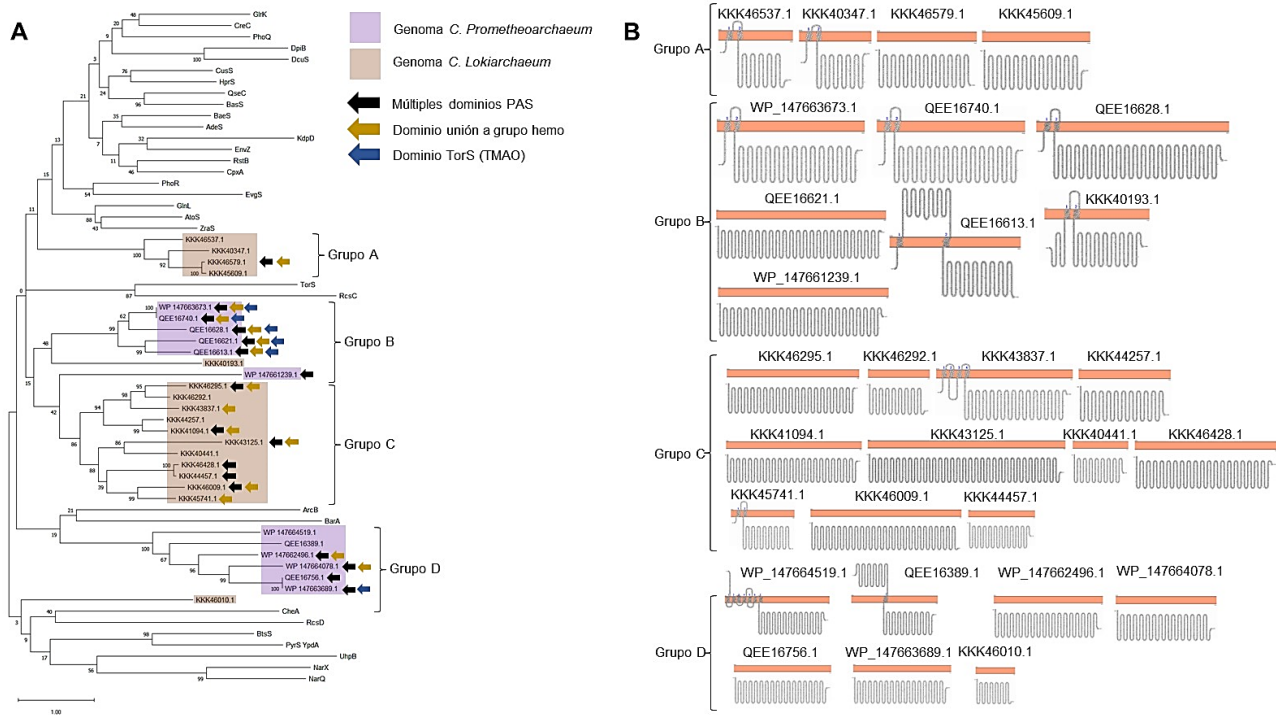


Figura 1. Análisis filogenético y topológico de las cinasas sensoras de dos organismos Asgard. En el panel A se muestra un análisis filogenético entre las cinasas de organismos Asgard y las 30 cinasas sensoras de *E. coli*. Con los datos mostrados en la Tabla 1 se indican las cinasas con múltiples dominios PAS (flecha negra), dominios de unión a un grupo hemo (flechas amarillas) y dominios similares a TorS de unión a TMAO (flechas azules). Se indican los 4 clústeres identificados por homología. En el panel B se muestra la topología de las cinasas sensoras por grupos. Panel A fue generado con MEGA versión 11.0.13 (Tamura *et al.*, 2021) usando 500 iteraciones de bootstrap. Panel B fue generado con Protter (Omasits *et al.*, 2013).

El genoma secuenciado de *C. Prometheoarchaeum* muestra un agrupamiento cercano a dos tipos de cinasas sensoras con función conocida, ya sea TorS o ArcB/BarA. En el primer caso, es una cinasa necesaria en *E. coli* para la percepción de ambientes ácidos (Li *et al.*, 2022) y la detección del osmoregulador y aceptor final de electrones trimetilamina N-óxido o TMAO por sus siglas en inglés (J. Moore y Hendrickson, 2012; Moore y Hendrickson, 2009). Para las cinasas ArcB y BarA, ArcB está relacionada con la detección del estado redox de las quinonas (Georgellis *et al.*, 2001) y BarA con la detección de acetato (Alvarez *et al.*, 2021). Algo semejante se observa con el genoma de *C. Lokiarchaeum*, aunque la mayoría de las cinasas derivadas de su metagenoma parecen formar dos clústeres, uno cercano a la cinasa AtoS, la cual responde a ácidos grasos de cadena corta, siendo el acetoacetato el principal activador; regulando genes requeridos para la síntesis y metabolismo de poliaminas, histamina, regulación de la producción de poli-(R)-hidroxibutirato, y la regulación por calcio (Kyriakidis y Tiligada, 2009). Así mismo, tiene también un homólogo cercano, ZraS, cuya señal son los niveles altos de zinc en el medio y tienen como respuesta elevada también a la limitación de fosfato y estrés por etanol (Choudhary *et al.*, 2020). Un segundo grupo forma un clúster más independiente que por las características de sus secuencias que la mayoría tienen un dominio PAS, podrían ser cinasas cuyo papel sea específico en el ambiente que estos organismos viven.

Un punto importante de analizar en estas cinasas es su topología, ya que una cinasa de *C. Lokiarchaeum* (KKK46010.1) mostró una homología cercana a CheA, una cinasa relacionada con la quimiotaxis que funcionalmente es citoplásmica, aunque en este caso es mucho más corta. En la Figura 1 panel B se muestra la topología por grupos filogenéticos (Figura 1 panel A). Un aspecto que llama la atención de la topología es que la mayoría de las cinasas encontradas parecen no tener dominios que las anclen a la membrana celular, lo que sugiere que su papel puede estar relacionado con la detección intracelular de metabolitos. Adicionalmente, las cinasas en el grupo B, algunas están anotadas como homólogos a ArcB (QEE16740.1, QEE16628.1, QEE16621.1 y QEE16613.1), los casos de QEE16621.1 y QEE16613.1 tienen una topología atípica con respecto a ArcB, la primera no logramos encontrar cruces transmembranales y la segunda tiene



un dominio periplásmico demasiado grande para este tipo de cinasas, cuyo módulo de recepción de la señal está en el citoplasma (Malpica *et al.*, 2004). La presencia en algunas de ellas de varios dominios PAS sugieren una estrecha relación con el estado redox celular, ya que en la profundidad oceánica en la que habitan la única fuente de luz sería derivada de actividad volcánica.

Finalmente, el análisis estructural permitió confirmar la presencia de múltiples dominios PAS en estas cinasas. En la Figura 2 se muestra el alineamiento estructural entre las cinasas analizadas por grupos (Figura 1) y algunos dominios PAS determinados experimentalmente y el modelo del dominio PAS de ArcB.

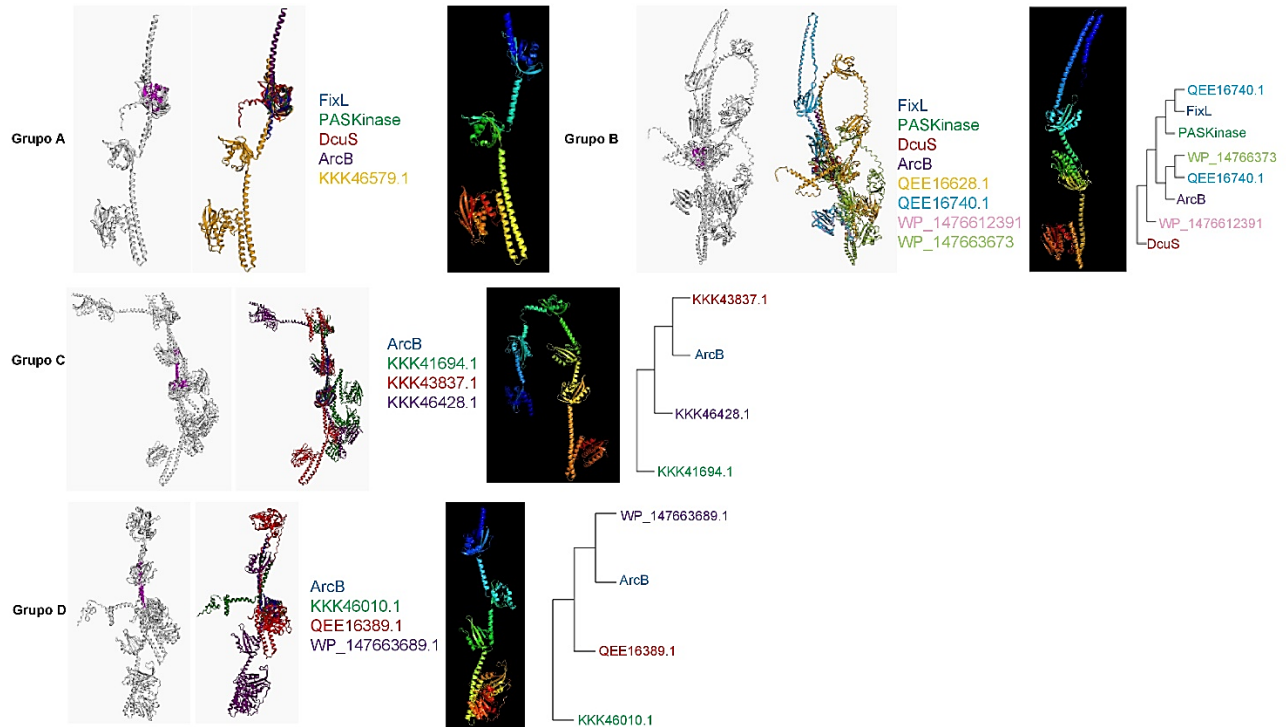


Figura 2. Análisis estructural de los dominios PAS de algunas cinasas de los grupos filogenéticos identificados. Se indica el grupo al que pertenecen. En magenta se muestran las regiones idénticas y en color la zona que cada PAS usado como referencia está ubicado. Grupo A, mayor similitud con FixL, se muestra la estructura de la cinasa con esquema de color arcoíris (azul N-terminal, rojo C-terminal). Grupo B, se muestra QEE16740.1 como referencia en esquema de color arcoíris. Grupo C se muestra KKK46428.1 como referencia en esquema de color arcoíris, en el alineamiento se distorsiona para posicionar con el PAS alineado. Solamente se analizaron contra ArcB dado que no hay homología con el resto de los dominios PAS. Grupo D, se muestra como referencia WP\_147663689, solamente se hizo alineamiento con el PAS de ArcB ya que no hay homología con las otras referencias. Se muestra la homología del dominio PAS en cada caso. Referencias: FixL (PDB 1DRM), ArcB (generado por nuestro grupo), PAS Kinase (PDB 1LL8), DcuS (PDB 2WON). Valores RMSD: Grupo A, 2.01; Grupo B, 1.73, Grupo C, 1.11 y Grupo D, 0.9.

Los resultados del análisis estructural sugieren que las cinasas del Grupo A y B son más conservados con respecto a sus contrapartes bacterianas, mientras que para el grupo C y D son más divergentes, lo que sugiere que pueden participar en la percepción de alguna señal específica del ambiente que habitan estas bacterias. Se requiere un análisis más detallado para encontrar las diferencias entre estos dominios PAS para poder predecir a qué señal responden. La presencia en algunas de estas proteínas de sitios de unión a grupos hemo, sugiere que pueden participar en la percepción de estado redox celular.

## Discusión

Encontrar las condiciones adecuadas para poder estudiar proteínas de organismos tan complejos de crecer como las arqueas Asgard en el laboratorio resulta muy complicado. Gracias a técnicas computacionales, al menos se pueden inferir algunas características y abordar preguntas evolutivas de este tipo de proteínas, teniendo una pequeña aproximación a su posible función y evolución.

Las cinasas tienen características particulares como es que la mayoría en el genoma secuenciado de *C. Prometheoachaeum* presentan arreglo híbrido (Singh *et al.*, 2021), lo que sugiere que estas proteínas en organismos tan complejos como las arqueas Asgard siguen teniendo mecanismos evolutivos similares al resto de las bacterias.

En el presente trabajo se encontraron características relevantes en sistemas de dos componentes de dos datos genómicos diferentes, uno ambiental (meta genoma) y otro directamente aislado y crecido para secuenciar su genoma y caracterizar otros aspectos fisiológicos. Destaca de manera importante que las proteínas del genoma totalmente secuenciado pertenecen a dos tipos de sistemas de dos componentes ampliamente caracterizados en bacterias, el sistema AtoS y ArcB. El sistema AtoS está involucrado en el metabolismo de ácidos grasos de cadena corta en la presencia de acetoacetato y estimula la síntesis de la molécula de reserva poli-R-3-hidroxitirato (Theodorou *et al.*, 2006). Este resultado está en concordancia con la evolución metabólica de las arqueas con el origen de los eucariontes, específicamente en la formación de organelos, ya que la dependencia metabólica entre estas bacterias ha originado la especulación sobre la dependencia de una arquea productora de H<sub>2</sub> con una deltaproteobacteria capaz de producir metabolitos necesarios para el hospedero (la Asgard arquea), sobre todo por la producción de metabolitos como propionato, acetato y otros compuestos de cadena corta, y una capacidad de oxidar sulfatos (López-García y Moreira, 2020), dando como resultado una relación estrecha que eventualmente culminó con la fagocitosis o endocitosis de la deltaproteobacteria acompañante. El que unas cinasas de un organismo totalmente secuenciado y crecido en el laboratorio sean de la familia de AtoS, sugiere fuertemente que es parte del sistema de señalización para establecer las relaciones metabólicas que han sostenido la teoría del origen por simbiosis (compartir nutrientes entre dos o más organismos) de los eucariontes.

Por otra parte, los resultados obtenidos con los datos metagenómicos hay que tomarlos con cautela, ya que los métodos computacionales para su análisis pueden crear artefactos (específicamente quimeras) con las secuencias, como se ha demostrado con las proteínas ribosomales (Garg *et al.*, 2020).

Finalmente, la importancia del estudio en todos sus aspectos de las arqueas Asgard resulta de mucha importancia dada su proximidad evolutiva con los eucariontes. En el mes de junio, se publicó la ruta completa para el remodelaje de la membrana mediada por ubiquitina y por primera vez se encontró esta molécula en bacterias (Hatano *et al.*, 2022), siendo un paso muy importante para establecer la evolución de una ruta muy importante en eucariontes. El trabajo de Hatano y colaboradores (2022) adicionalmente sugiere que las rutas en eucariontes para la reparación de membrana mediada por ubiquitina involucraron la duplicación de genes en estas arqueas, resultando en la diversificación de estos genes y adquiriendo nuevas funciones que en eucariontes se establecieron como los mecanismos principales.

Finalmente, la propuesta más sólida para el origen de los eucariontes es directamente bacteriano. La propuesta considera que de las bacterias, un grupo de bacterias específico PVC (Planctomycetes-Verrucomicrobia-Chlamydiae) ha puesto en consideración un modelo de un origen gradual de las arqueas y posteriormente de los eucariontes (Devos, 2021). En la Figura 3 se muestra la propuesta de Devos (2021). Este trabajo aporta un contexto evolutivo para los sistemas de dos componentes, específicamente para las cinasas sensoras, con un origen cercano a las bacterias, lo que está en concordancia con esta propuesta y también con un grupo de cinasas que han evolucionado específicamente en estos organismos.

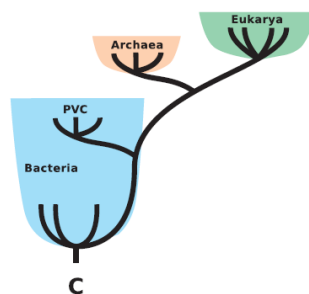


Figura 3. Propuesta de Devos (2021) para el origen de las arqueas y eucariontes. En azul se muestra el grupo de bacterias dentro de las cuales se encuentran ancestros posibles de las arqueas, los organismos PVC se proponen como intermediarios en este proceso. En rojo claro el grupo de arqueas y en verde los eucariontes. Tomado de Devos (2021).

## Conclusiones

Con base en los resultados obtenidos, los sistemas de señalización en arqueas Asgard son estructuralmente similares a otros bien caracterizados de bacterias e incluso su secuencia primaria permite agrupar a algunos de ellos con sistemas de dos componentes con funciones ya descritas, así también como un grupo totalmente nuevo con características específicas de estos organismos. Esto abre las posibilidades de estudiar este tipo de cinasas en un contexto ecológico particular en el que estas arqueas viven, especialmente con relación al origen de los eucariontes y la dependencia metabólica de los dos tipos de organismos.

## Referencias

- Alvarez, A. F., Rodríguez, C., González-Chávez, R., & Georgellis, D. (2021). The *Escherichia coli* two-component signal sensor BarA binds protonated acetate via a conserved hydrophobic-binding pocket. *Journal of Biological Chemistry*, 297(6), 101383. <https://doi.org/10.1016/j.jbc.2021.101383>
- Choudhary, K. S., Kleinmanns, J. A., Decker, K., Sastry, A. V., Gao, Y., Szubin, R., Seif, Y., & Palsson, B. O. (2020). Elucidation of Regulatory Modes for Five Two-Component Systems in *Escherichia coli* Reveals Novel Relationships. *mSystems*, 5(6). <https://doi.org/10.1128/msystems.00980-20>
- Devos, D. P. (2021). Reconciling Asgardarchaeota Phylogenetic Proximity to Eukaryotes and Planctomycetes Cellular Features in the Evolution of Life. *Molecular Biology and Evolution*, 38(9), 3531–3542. <https://doi.org/10.1093/molbev/msab186>
- Falke, J. J., Bass, R. B., Butler, S. L., Chervitz, S. A., & Danielson, M. A. (1997). THE TWO-COMPONENT SIGNALING PATHWAY OF BACTERIAL CHEMOTAXIS: A Molecular View of Signal Transduction by Receptors, Kinases, and Adaptation Enzymes. *Annual Review of Cell and Developmental Biology*, 13(1), 457–512. <https://doi.org/10.1146/annurev.cellbio.13.1.457>
- Galperin, M. Y., Makarova, K. S., Wolf, Y. I., & Koonin, E. V. (2018). Phyletic Distribution and Lineage-Specific Domain Architectures of Archaeal Two-Component Signal Transduction Systems. *Journal of Bacteriology*, 200(7). <https://doi.org/10.1128/jb.00681-17>
- Garg, S. G., Kapust, N., Lin, W., Knopp, M., Tria, F. D. K., Nelson-Sathi, S., Gould, S. B., Fan, L., Zhu, R., Zhang, C., & Martin, W. F. (2020). Anomalous Phylogenetic Behavior of Ribosomal Proteins in Metagenome-Assembled Asgard Archaea. *Genome Biology and Evolution*, 13(1). <https://doi.org/10.1093/gbe/evaa238>
- Georgellis, D., Kwon, O., & Lin, E. C. C. (2001). Quinones as the Redox Signal for the Arc Two-Component System of Bacteria. *Science*, 292(5525), 2314–2316. <https://doi.org/10.1126/science.1059361>
- Hatano, T., Palani, S., Papatziomou, D., Salzer, R., Souza, D. P., Tamarit, D., Makwana, M., Potter, A., Haig, A., Xu, W., Townsend, D., Rochester, D., Bellini, D., Hussain, H. M. A., Ettema, T. J. G., Löwe, J., Baum, B., Robinson, N. P., & Balasubramanian, M. (2022). Asgard archaea shed light on the evolutionary origins of the eukaryotic ubiquitin-ESCRT machinery. *Nature Communications*, 13(1). <https://doi.org/10.1038/s41467-022-30656-2>
- Imachi, H., Nobu, M. K., Nakahara, N., Morono, Y., Ogawara, M., Takaki, Y., Takano, Y., Uematsu, K., Ikuta, T., Ito, M., Matsui, Y., Miyazaki, M., Murata, K., Saito, Y., Sakai, S., Song, C., Tasumi, E., Yamanaka, Y., Yamaguchi, T., . . . Takai, K. (2020). Isolation of an archaeon at the prokaryote–eukaryote interface. *Nature*, 577(7791), 519–525. <https://doi.org/10.1038/s41586-019-1916-6>
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., . . . Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2>
- Kinoshita-Kikuta, E., Koike, T., & Kinoshita, E. (2022). Recent advances in the Phos-tag technique focused on the analysis of phosphoproteins in a bacterial two-component system. *Journal of Proteomics*, 252, 104429. <https://doi.org/10.1016/j.jprot.2021.104429>
- Kyriakidis, D. A., & Tiligada, E. (2009). Signal transduction and adaptive regulation through bacterial two-component systems: the *Escherichia coli* AtoSC paradigm. *Amino Acids*, 37(3), 443–458. <https://doi.org/10.1007/s00726-009-0241-z>

- Levasseur, A., Merhej, V., Baptiste, E., Sharma, V., Pontarotti, P., & Raoult, D. (2017). The Rhizome of Lokiarchaeota Illustrates the Mosaicity of Archaeal Genomes. *Genome Biology and Evolution*, 9(10), 2635–2639. <https://doi.org/10.1093/gbe/evx208>
- Li, G., Morigen, y Yao, Y. (2022). TorR/TorS Two-Component system resists extreme acid environment by regulating the key response factor RpoS in *Escherichia coli*. *Gene*, 821, 146295. <https://doi.org/10.1016/j.gene.2022.146295>
- López-García, P., & Moreira, D. (2020). Cultured Asgard Archaea Shed Light on Eukaryogenesis. *Cell*, 181(2), 232–235. <https://doi.org/10.1016/j.cell.2020.03.058>
- Ma, P., y Phillips-Jones, M. K. (2021). Membrane Sensor Histidine Kinases: Insights from Structural, Ligand and Inhibitor Studies of Full-Length Proteins and Signalling Domains for Antibiotic Discovery. *Molecules*, 26(16), 5110. <https://doi.org/10.3390/molecules26165110>
- Malpica, R., Franco, B., Rodriguez, C., Kwon, O., & Georgellis, D. (2004). Identification of a quinone-sensitive redox switch in the ArcB sensor kinase. *Proceedings of the National Academy of Sciences*, 101(36), 13318–13323. <https://doi.org/10.1073/pnas.0403064101>
- Moore, J., y Hendrickson, W. (2012). An Asymmetry-to-Symmetry Switch in Signal Transmission by the Histidine Kinase Receptor for TMAO. *Structure*, 20(4), 729–741. <https://doi.org/10.1016/j.str.2012.02.021>
- Moore, J. O., y Hendrickson, W. A. (2009). Structural Analysis of Sensor Domains from the TMAO-Responsive Histidine Kinase Receptor TorS. *Structure*, 17(9), 1195–1204. <https://doi.org/10.1016/j.str.2009.07.015>
- Nocedal, I., y Laub, M. T. (2022). Ancestral reconstruction of duplicated signaling proteins reveals the evolution of signaling specificity. *eLife*, 11. <https://doi.org/10.7554/elife.77346>
- Omasits, U., Ahrens, C. H., Müller, S., & Wollscheid, B. (2013). Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics*, 30(6), 884–886. <https://doi.org/10.1093/bioinformatics/btt607>
- Padilla-Vaca, F., Mondragon-Jaimes, V., & Franco, B. (2017). General Aspects of Two-Component Regulatory Circuits in Bacteria: Domains, Signals and Roles. *Current Protein & Peptide Science*, 18(10). <https://doi.org/10.2174/1389203717666160809154809>
- Schaller, G., Shiu, S. H., & Armitage, J. (2011). Two-Component Systems and Their Co-Option for Eukaryotic Signal Transduction. *Current Biology*, 21(9), R320–R330. <https://doi.org/10.1016/j.cub.2011.02.045>
- Silva, M. A., y Salgueiro, C. A. (2021). Multistep Signaling in Nature: A Close-Up of *Geobacter* Chemotaxis Sensing. *International Journal of Molecular Sciences*, 22(16), 9034. <https://doi.org/10.3390/ijms22169034>
- Singh, D., Gupta, P., Singla-Pareek, S. L., Siddique, K. H., & Pareek, A. (2021). The Journey from Two-Step to Multi-Step Phosphorelay Signaling Systems. *Current Genomics*, 22(1), 59–74. <https://doi.org/10.2174/1389202921666210105154808>
- Spang, A., Saw, J. H., Jørgensen, S. L., Zaremba-Niedzwiedzka, K., Martijn, J., Lind, A. E., van Eijk, R., Schleper, C., Guy, L., & Ettema, T. J. G. (2015). Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature*, 521(7551), 173–179. <https://doi.org/10.1038/nature14447>
- Stock, A. M., Robinson, V. L., & Goudreau, P. N. (2000). Two-Component Signal Transduction. *Annual Review of Biochemistry*, 69(1), 183–215. <https://doi.org/10.1146/annurev.biochem.69.1.183>
- Szurmant, H., White, R. A., & Hoch, J. A. (2007). Sensor complexes regulating two-component signal transduction. *Current Opinion in Structural Biology*, 17(6), 706–715. <https://doi.org/10.1016/j.sbi.2007.08.019>
- Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, 38(7), 3022–3027. <https://doi.org/10.1093/molbev/msab120>
- Teran-Melo, J. L., Peña-Sandoval, G. R., Silva-Jimenez, H., Rodriguez, C., Alvarez, A. F., & Georgellis, D. (2018). Routes of phosphoryl group transfer during signal transmission and signal decay in the dimeric sensor histidine kinase ArcB. *Journal of Biological Chemistry*, 293(34), 13214–13223. <https://doi.org/10.1074/jbc.ra118.003910>

- Theodorou, M. C., Panagiotidis, C. A., Panagiotidis, C. H., Pantazaki, A. A., & Kyriakidis, D. A. (2006). Involvement of the AtoS-AtoC signal transduction system in poly-(R)-3-hydroxybutyrate biosynthesis in *Escherichia coli*. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1760(6), 896–906. <https://doi.org/10.1016/j.bbagen.2006.01.020>
- Vreede, J., van der Horst, M. A., Hellingwerf, K. J., Crielaard, W., & van Aalten, D. M. (2003). PAS Domains. *Journal of Biological Chemistry*, 278(20), 18434–18439. <https://doi.org/10.1074/jbc.m301701200>
- Watson, T. (2019). The trickster microbes that are shaking up the tree of life. *Nature*, 569(7756), 322–324. <https://doi.org/10.1038/d41586-019-01496-w>
- Yamamoto, K., Hirao, K., Oshima, T., Aiba, H., Utsumi, R., & Ishihama, A. (2005). Functional Characterization in Vitro of All Two-component Signal Transduction Systems from *Escherichia coli*. *Journal of Biological Chemistry*, 280(2), 1448–1456. <https://doi.org/10.1074/jbc.m410104200>
- Zhang, W., Culley, D. E., Wu, G., & Brockman, F. J. (2006). Two-Component Signal Transduction Systems of *Desulfovibrio vulgaris*: Structural and Phylogenetic Analysis and Deduction of Putative Cognate Pairs. *Journal of Molecular Evolution*, 62(4), 473–487. <https://doi.org/10.1007/s00239-005-0116-1>