Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Microbiology
ELSEVIER

Evolution and phyletic distribution of two-component signal transduction systems

Kristin Wuichet[1], Brian J Cantwell[1] and Igor B Zhulin[1,2]

Two-component signal transduction systems are abundant in prokaryotes. They enable cells to adjust multiple cellular functions in response to changing environmental conditions. These systems are also found, although in much smaller numbers, in lower eukaryotes and plants, where they appear to control a few very specific functions. Two-component systems have evolved in Bacteria from much simpler one-component systems bringing about the benefit of extracellular versus intracellular sensing. We review reports establishing the origins of two-component systems and documenting their occurrence in major lineages of Life.

Addresses
[1] Department of Microbiology, University of Tennessee, Knoxville, TN 37996, USA
[2] Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Corresponding author: Zhulin, Igor B (ijouline@utk.edu)

Current Opinion in Microbiology 2010, **13**:1–7

This review comes from a themed issue on
Cell regulation
Edited by Robert Bourret and Ruth Silversmith

1369-5274/$ – see front matter
Published by Elsevier Ltd.

DOI 10.1016/j.mib.2009.12.011

Introduction

Signal transduction systems regulate cellular processes in all living organisms. The evolutionary success of prokaryotes is dependent on the ability of these unicellular organisms to rapidly sense and respond to changes inside and outside their cell. Similarly, eukaryotes also must detect and transmit both internal and external (e.g. cell-to-cell) signals to adjust functions controlled by hormones, cytokines, mediators, and so on. Two-component regulatory systems (TCS) were the first major mode of signal transduction identified in bacteria [1]. Since the time of their discovery, many TCS have been identified and studied, primarily in prokaryotes [2••]. Advent of genomics led to massive identification of TCSs in microbial genomes [3,4•]. TCS have also been identified in eukaryotes [5]. It appears that their distribution in eukaryotes, specifically lower eukaryotes and plants, is significantly less abundant [2••,6]. Many bacterial TCSs have been shown to play a role in virulence [7,8], and the absence of the systems in mammals furthers interest in their study as potential drug targets [9]. Understanding the evolution of TCSs will expand knowledge about these thoroughly studied systems to derive both system-specific and universal principles of signal transduction. There have been several attempts to address the question of TCS origins and diversification [10••,11••,12,13]. Here, we summarize both advances and shortcomings of our current understanding in this important area.

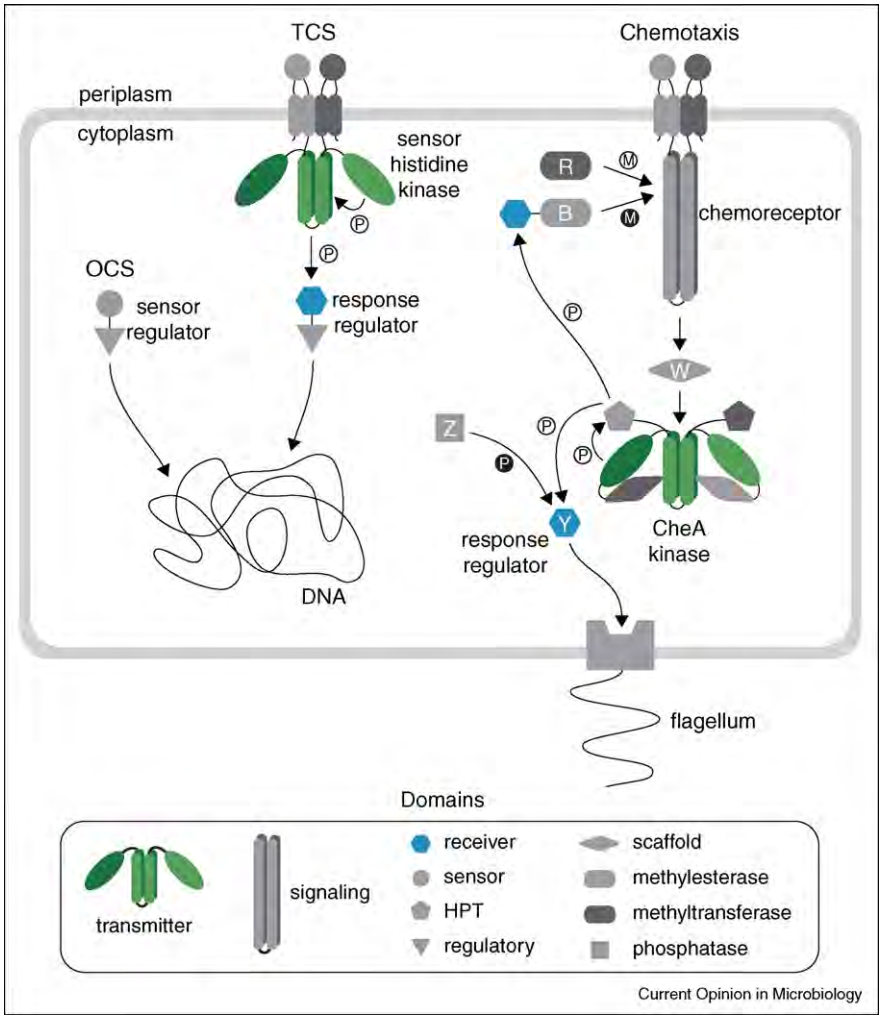
System design: one component, two components…

As their name suggests, two-component signal transduction systems are composed of two dedicated proteins, a sensor and a regulator [14••], although some systems contain additional auxiliary components. Interestingly, there are other signal transduction systems that appear to be either simpler or more complex in their overall design than classical TCSs. The majority of sensing and signal transduction in prokaryotes appears to be carried out by so-called one-component systems (OCSs), single proteins that combine properties of both a sensor and a regulator [10••]. Usually, these properties reside in two distinct domains—sensory and regulatory (Figure 1), although they can be combined in a single domain or distributed between multiple domains within a single protein [4•,10••]. The relatedness between OCSs and TCSs lies in the fact that they share a repertoire of sensory and regulatory domains [10••].

The prototypical TCS employs four principal domains: sensory (also called input) and transmitter that reside in a sensor histidine kinase (HK) and receiver and regulatory (also called output) that reside in a response regulator (RR) (Figure 1). The transmitter and receiver communicate via phosphorylation [2••]. Many TCSs stray from this paradigm by employing multiple sensory domains (both extracellular and cytoplasmic) [15,16], and additional transmitter and regulatory domains, as well as intermediate components between the kinase and response regulator that usually comprise extended phosphorelays [2••,17]. Sensor kinases with C-terminal receiver domains are referred to as hybrid histidine kinases (HHK) [2••]. Similarly there are also hybrid response regulators (HRR) with C-terminal HK modules rather than typical output domains. HHKs and HRRs are ubiquitous, but still far outnumbered by HKs and RRs with standard architectures (Table 1). In addition to sensor, transmitter (kinase), receiver, and regulatory modules,

Please cite this article in press as: Wuichet K, et al. Evolution and phyletic distribution of two-component signal transduction systems, Curr Opin Microbiol (2010), doi:10.1016/j.mib.2009.12.011

Figure 1



Prokaryotic signal transduction paradigms. Extracellular sensing emerged with the advent of TCSs given that 73% of HKs are membrane associated compared to only 3% of OCSs [10••]. However, both OCSs and TCSs are predicted to predominantly regulate transcription [10••]. The chemotaxis system (represented by the canonical *E. coli* system) employs a number of additional components not reported in other TCSs that regulate methylation (M) (CheB and CheR), dephosphorylation (CheZ), and scaffolding (CheW).

histidine phosphotransfer (HPT) domains are commonly associated with TCSs. They often can be found fused to HHKs or as independent proteins to facilitate the phosphotransfer process [2••,17]. HPT domains can both receive phosphoryl groups from and transmit them to receiver domains [18]. As seen with HPT domains, receiver domains can also exist as independent proteins that lack additional output domains, which are members of extended phosphorelays, allosteric regulators, or other roles [19•]. The modular nature of these domains leads to a variety of architectures (Figure 2), resulting in diverse phosphorelay pathways.

At the apex of TCS complexity is the bacterial chemotaxis system, a specialized TCS that utilizes a dedicated chemoreceptor (sensor), a sensor-less HK (CheA protein), an RR protein that lacks a usual output domain (CheY protein), and a number of specialized auxiliary components (Figure 1). Genomic modularity of this system is considered elsewhere [20] and is beyond the scope of this review.

Phyletic distribution: abundance in prokaryotes and instances in eukaryotes

A genomic survey shows that TCSs are found in all three superkingdoms: Bacteria, Archaea, and Eukarya (Table 1). TCS components are present in 864 of 899 available completely sequenced bacterial genomes including all 21 phyla represented (Table 1). The only bacterial species lacking TCSs are pathogens (e.g. *Mycoplasma* species) and endosymbionts (e.g. *Amoebophilus* species) with severely reduced genomes (mistdb.com).

Please cite this article in press as: Wuichet K, et al. Evolution and phyletic distribution of two-component signal transduction systems, Curr Opin Microbiol (2010), doi:10.1016/j.mib.2009.12.011

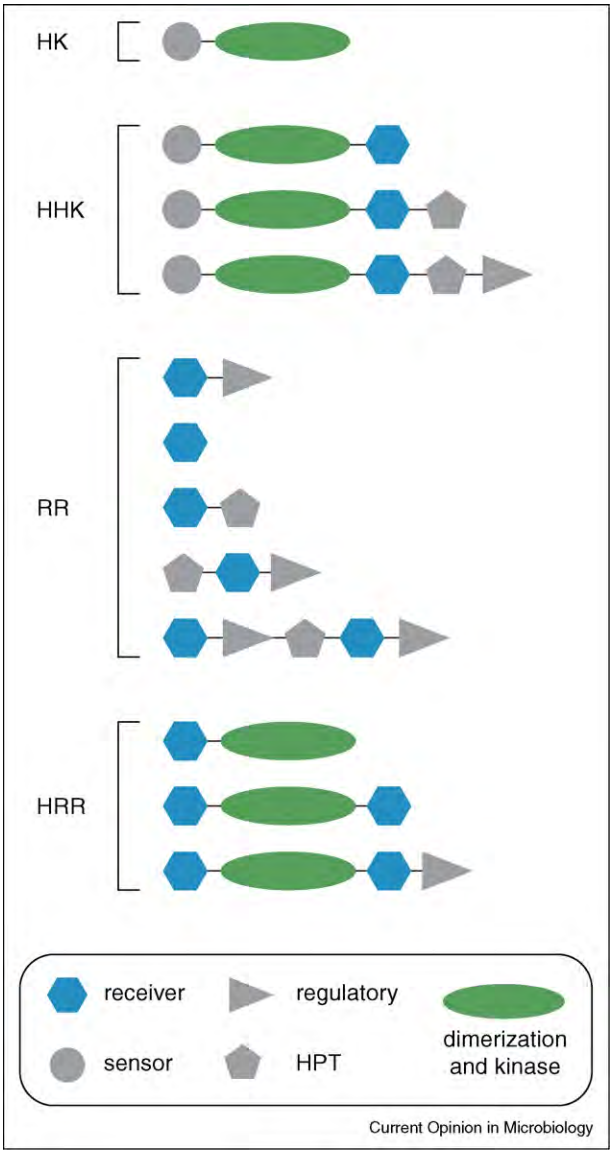
Table 1

Genomic and phyletic distribution of TCS components in three domains of Life

	Bacteria	Archaea	Eukarya
Genome occurrence			
Genomes	899	67	55
HK	857	30	1
RR	862	34	15
HHK	561	5	15
HRR	258	17	0
Other/HPT	220	2	15
Taxonomic distribution			
Taxonomic groups	21	5	5
HK	21	3	1
RR	21	2	3
HHK	16	1	3
HRR	15	1	0
Other/HPT	11	1	3
Total TCS counts			
HK	20,862	546	7
RR	26,962	304	80
HHK	4717	8	86
HRR	923	142	0
Other/HPT	329	4	18

Genome surveys have been carried out using the MiST2 database [4•] for Bacteria and Archaea and the genomic mode of the SMART 6 database [56] for Eukarya. CheA are not included in this survey. Domain architecture queries for HATPase_c, REC, and HPT domains were used to identify putative TCS components in SMART, which were then manually classified. BLAST queries using TCS sequences identified in Metazoan genomes in SMART support that they are the result of bacterial contamination and were subsequently excluded. TCS components are automatically classified in MiST2. In MiST2 the 'other' category of TCS components predominantly consists of stand-alone HPT proteins, but rarely includes proteins with architectures that do not fit the standard rules for TCS classification. HPT proteins are absent from Archaea. Bacteria and Archaea taxonomic groups are comprised of phyla as defined in the MiST database [4•]. Given the fewer representatives of eukaryotic phyla and their comparatively recent emergence in comparison to prokaryotic phyla, Eukarya taxonomic groups correspond to the more inclusive fungal, metazoan, plant, protist, and slime mold lineages.

Figure 2



Modularity of TCS domain architectures. A survey of TCS component domain architectures (mistdb.com) reveals extensive diversity, some of which is represented here. Receiver domains can be found tandemly repeated in many of these architectures. Similarly, there can be multiple sensor domains in RRs and HRRs in addition to HKs and HHKs. These components can be combined in different ways to form a variety of phosphorelay pathways [2••,6], which supports the importance of distinguishing between HHKs and HRRs when considering such possibilities [4•].

Previously TCSs were considered absent from the Tenericutes phylum [10••,11••], but the newly available genome of *Acholeplasma laidlawii* revealed multiple HKs and RRs (mistdb.com). In contrast to Bacteria, TCSs are present only in ~50% of genomes from Archaea. They appear to be entirely absent in Crenarchaeota, Korarchaeota, and Nanoarchaeota phyla. TCSs are found in 33 of the 42 genomes representing Euryarchaeota, and the one representative of Thaumarchaeota, *Nitrosopumilus maritimus* SCM1. Less than 30% of eukaryotic genomes encode TCSs (Table 1). Queries of the SMART database identify TCSs in fungal, plant, and slime mold lineages, but they are absent from metazoan and protist lineages, as suggested by an earlier genome survey [3].

The distribution of HHKs and HRRs in Bacteria, Archaea, and Eukarya reveal differences in their predominant TCS signal transduction mechanisms. HHKs make up less than 20% of TCS kinases in bacteria, and they are found in only 16 of the 21 phyla (Table 1). Surprisingly, more than 90% of eukaryotic TCS kinases are HHKs, and they are present in all three lineages that have TCSs. Only 1% of the TCS kinases in Archaea are HHKs, and they are found exclusively in Euryarchaeota. HRRs are entirely absent from Eukarya and comprise

Please cite this article in press as: Wuichet K, et al. Evolution and phyletic distribution of two-component signal transduction systems, Curr Opin Microbiol (2010), doi:10.1016/j.mib.2009.12.011

only 3% of the RRs in Bacteria; however, they make up more than 30% of the RRs in Archaea. While multiple TCSs employing HHKs have been studied in both Bacteria and Eukarya [6,17], there is an astonishing lack of experimental information on HRRs. This is particularly relevant to Archaea (see Table 1).

Given the vast numbers of TCSs in Bacteria, which have been shown to control a variety of cellular processes [21], it is nearly impossible to review them in a traditional way. Spectacularly, there is little to no information on TCSs in Archaea, but experimental studies have revealed interesting aspects of TCSs in eukaryotes. Unicellular fungi typically encode a small number of histidine kinases [22,23•], while the genome sequences of various filamentous fungi have revealed between 11 and 21 histidine kinases [23•,24,25,26•]. All of the kinases identified in fungi are hybrid and in all cases, these organisms encode a single HPT protein [23•,26•]. The fungi typically encode two RRs, but some of the filamentous fungi possess a third response regulator as well [26•]. The slime mold *Dictyostelium discoideum* encodes 21 two-component proteins. These include 14 HHKs and 6 RRs [27]. Similarly to all fungi [23•,26•], *D. discoideum* encodes only a single HPT protein, indicating that the sensor kinases may integrate multiple signals through a single phosphorelay [27]. Compared to the lower eukaryotes, plants encode a greater number and diversity of phosphorelay proteins. Analysis of the *Arabidopsis thaliana* and *Oryza sativa* genomes identified 14–16 HKs, 5 HPT proteins, and 32 RRs in each [28,29]. In contrast to fungi and *D. discoideum*, HHKs represent approximately half of the plant TCSs rather than all, but the importance of HPT proteins in these systems is evident by their expansion. This is in stark contrast with prokaryotes that possess HHKs, fewer than 40% of which encode an HPT protein in their genomes (Table 1), which is consistent with the results of a previous survey [6].

Evidence for Bacterial origins and horizontal gene transfer

There are multiple evolutionary forces driving the emergence and expansion of TCS diversity that can be visualized using computational approaches. TCS components are often classified on the basis of the sequence similarity of their conserved domains using either simple BLAST hits [30], or more robust tree-based [11••] and Hidden Markov Model [4•] methods. Domain architecture analysis is yet another way to classify TCS components [15]. Phylogenetic trees alone can reveal potential relationships between the classes [6,11••]. Similarly, varying degrees of domain architecture complexity can hint at an evolutionary history [10••]. The abundance and distribution of TCS systems and their classes can be analyzed further to determine evolutionary relationships [6,10••]. These approaches have revealed a general consensus on the origins of TCSs. OCSs are predicted to be the progenitors of TCSs due to their low complexity, higher abundance, and wider variety of sensory (input) and regulatory (output) domains in comparison to TCSs [10••]. Distribution analysis supports that at least some OCSs were present in the Last Universal Common Ancestor [10••]. Phylogenetic analysis indicated that TCSs originated in Bacteria and were radiated to Archaea and Eukarya via multiple lateral transfer events [11••]. This observation is also supported by the greater abundance and wider distribution of TCS in Bacteria in comparison to Archaea [10••] and Eukarya (Table 1).

The co-evolutionary relationship between cognate HKs and RRs is evident in the congruent phylogenetic trees built from multiple sequence alignments of these proteins [11••]. Cognate HK and RR proteins are often encoded adjacent to each other in genomes, which further supports their co-evolutionary relationship [4•,10••]. Large-scale efforts to deduce interacting HKs and RRs in multiple genomes have relied on genome context [31,32•]. However, many systems utilize orphan HKs and RRs that are not encoded adjacent to their cognate partners [4•], and recent experimental studies support their importance in multiple organisms [33–36]. Furthermore, genome context and phylogenetic methods do not always agree when predicting cognate HKs and RRs [11••]. Domain architecture-based classification can be totally misleading when used without phylogenetic support. This is illustrated by stunning differences in the domain architectures of orthologous HKs that control nitrogen fixation (FixL) in Rhizobia [37–39].

TCS evolution in eukaryotes

Substantial phylogenetic evidence in light of the known pattern of distribution of eukaryotic TCSs led to a conclusion that the TCSs in eukaryotes represent multiple independent horizontal gene transfer events from bacteria to eukaryotes that occurred at a time well after the split into the three superkingdoms [11••]. Many genes in plant nuclear genomes are thought to have been transferred from the chromosomes of cyanobacterial symbionts during the evolution of the chloroplast, with as much as 18% of the protein-coding genes of *A. thaliana* inherited from cyanobacteria [40]. The CSK protein of *A. thaliana* is a HK, which is encoded by a nuclear gene and targeted to the chloroplast. The cognate RR of the CSK is plastid encoded, and phylogenetic analysis of this family of proteins identified the cyanobacterial HK, Hik2, as a common ancestor [41••].

TCSs in eukaryotes have undergone a number of evolutionary innovations distinct from those seen commonly in bacterial systems. One such feature is the integration of TCS signaling pathways with other signaling systems typical of eukaryotes. The output of the fungal TCSs feeds into a mitogen-activated protein kinase (MAPK) pathway. This is best studied in *Saccharomyces cerevisiae*,

Please cite this article in press as: Wuichet K, et al. Evolution and phyletic distribution of two-component signal transduction systems, Curr Opin Microbiol (2010), doi:10.1016/j.mib.2009.12.011

where the Sln1-Ypd1-Ssk1 phosphorelay activates the Ssk2 MAPK leading to control of genes for osmotic stress [42]. Similar pathways are found in other fungal species studied [26]. Although RRs in yeast can be regulated by phosphorylation, one has been shown to affect gene expression in a manner independent of aspartate phosphorylation that is typical of prokaryotes [43]. The RdeA-RegA phosphorelay in *D. discoideum* have been shown to regulate cAMP levels [44] and the DokA HK can be regulated by serine phosphorylation [45]. The *A. thaliana* ethylene receptors interact with the serine-threonine kinase CTR1 [46] and mediates responses to ethylene independent of histidine-aspartate phosphotransfer [47]. The plant cytokinin TCSs appear to function similarly to the common bacterial systems and directly regulate gene expression through phosphorylation of a DNA-binding RR [48]. However, even in this case there are significant differences, as cytokinin signaling seems to involve considerable crosstalk with multiple HKs and RRs communicating through a common pool of HPT proteins [49]. The phytochrome receptor proteins of plants contain a divergent HK domain that possesses serine-threonine kinase activity [50]. A family of RRs in *A. thaliana* is not phosphorylated [51], but some play important roles in the regulation of the circadian clock [52].

Conclusions

Comparative genomics methods prove to be effective for describing the diversity of TCS components and tracing their evolution. Such information is currently inaccessible by experimental methods alone. While domain architecture and genome context analyses can visualize the evolutionary forces effecting TCS diversity, such as domain duplication, domain/gene birth and death, gene duplication/divergence, and gene fission/fusion events [6,12,15,53,54•], there are still many unanswered questions about the evolution of these systems that cannot be addressed without an exhaustive phylogenomic approach (Box 1). Special effort must be made to be comprehensive, particularly to reduce erroneous conclusions that can arise from analyses using only a single method. Similarly, analyses utilizing only a single genome [55] can only provide limited information that may not be applicable to the larger landscape of TCS evolution and function. Current studies are primarily descriptive and rely on a single methodology. Future efforts utilizing systems-level phylogenomics are required in order to properly classify and understand complete signal transduction pathways.

Box 1 Future directions and unanswered questions

- What levels of complexity can be achieved in extended phosphorelays, particularly those involving HRRs? What is the exact function of HRR?
- How can we efficiently assign orphan HKs/HHKs and RRs/HRRs to their cognate partners, especially in the light of the exponential growth of genomic data?
- How often orthologous HKs and RRs perform different functions in different organisms? Defining orthology is the main approach to protein function prediction. How careful should we be in using orthology for function prediction in signal transduction proteins?
- Sensory specificity of HKs remains a single less-studied topic in two-component signal transduction. We should develop high-throughput computational and experimental methodologies to address this problem.
- Various high-throughput computational and experimental methods are now being used to predict targets for transcriptional regulators, including RRs. A comprehensive comparison of these methods and a unified approach must be developed in the nearest future.

Acknowledgements

This work was supported by NIH grant GM072285 and funds from the BioEnergy Science Center which is supported by the Office of Biological and Environmental Research in the DOE Office of Science.

References and recommended reading

Paper of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Nixon BT, Ronson CW, Ausubel FM: **Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes ntrB and ntrC**. *Proc Natl Acad Sci USA* 1986, **83**:7850-7854.

2. •• Stock AM, Robinson VL, Goudreau PN: **Two-component signal transduction**. *Ann Rev Biochem* 2000, **69**:183-215.
This is by far the most comprehensive, authoritative review on TCS.

3. Wolanin PM, Thomason PA, Stock JB: **Histidine protein kinases: key signal transducers outside the animal kingdom**. *Genome Biol* 2002, **3**: REVIEWS3013.

4. • Ulrich LE, Zhulin IB: **The MiST2 database: a comprehensive genomics resource on microbial signal transduction**. *Nucleic Acids Res* 2009 doi: 10.1093/nar/gkp940.
This paper describes a continuously updated database that classifies TCS components in all prokaryotic genomes. It includes the first attempts to delineate evolutionary relationships between TCS proteins on the basis of domain orthology and also catalogs HHK and HRR proteins as distinct types of TCS components.

5. Alex LA, Simon MI: **Protein histidine kinases and signal transduction in prokaryotes and eukaryotes**. *Trends Genet* 1994, **10**:133-138.

6. Zhang W, Shi L: **Distribution and evolution of multiple-step phosphorelay in prokaryotes: lateral domain recruitment involved in the formation of hybrid-type histidine kinases**. *Microbiology* 2005, **151**:2159-2173.

7. Moon K, Gottesman S: **A PhoQ/P-regulated small RNA regulates sensitivity of Escherichia coli to antimicrobial peptides**. *Mol Microbiol* 2009.

8. Gooderham WJ, Hancock RE: **Regulation of virulence and antibiotic resistance by two-component regulatory systems in *Pseudomonas aeruginosa***. *FEMS Microbiol Rev* 2009, **33**:279-294.

9. Schreiber M, Res I, Matter A: **Protein kinases as antibacterial targets**. *Curr Opin Cell Biol* 2009, **21**:325-330.

10. •• Ulrich LE, Koonin EV, Zhulin IB: **One-component systems dominate signal transduction in prokaryotes**. *Trends Microbiol* 2005, **13**:52-56.

Please cite this article in press as: Wuichet K, et al. Evolution and phyletic distribution of two-component signal transduction systems, Curr Opin Microbiol (2010), doi:10.1016/j.mib.2009.12.011

While individual proteins that could both sense an input and effect an output had been identified previously in experimental studies, this paper was the first to establish the dominance of OCSs as the major mode of signal transduction in prokaryotes and their placement as the progenitors of TCSs.

11. •• Koretke KK, Lupas AN, Warren PV, Rosenberg M, Brown JR: **Evolution of two-component signal transduction**. *Mol Biol Evol* 2000, **17**:1956-1970.
This paper marks the first major attempt to classify TCS components in multiple genomes using a phylogenetic approach. The results demonstrate the co-evolution of HKs and RRs as well as the importance of using multiple methods to establish cognate pairs rather than genome context alone.

12. Whitworth DE, Cock PJ: **Evolution of prokaryotic two-component systems: insights from comparative genomics**. *Amino Acids* 2009, **37**:459-466.

13. Pertseva MN, Shpakov AO: **The prokaryotic origin and evolution of eukaryotic chemosignaling systems**. *Neurosci Behav Physiol* 2009, **39**:793-804.

14. •• Kofoid EC, Parkinson JS: **Transmitter and receiver modules in bacterial signaling proteins**. *Proc Natl Acad Sci USA* 1988, **85**:4981-4985.
This paper uncovered the universal modularity of TCS.

15. Ashby MK: **Distribution, structure and diversity of ''bacterial'' genes encoding two-component proteins in the Euryarchaeota**. *Archaea* 2006, **2**:11-30.

16. Zhulin IB, Nikolskaya AN, Galperin MY: **Common extracellular sensory domains in transmembrane receptors for diverse signal transduction pathways in bacteria and archaea**. *J Bacteriol* 2003, **185**:285-294.

17. Appleby JL, Parkinson JS, Bourret RB: **Signal transduction via the multi-step phosphorelay: not necessarily a road less traveled**. *Cell* 1996, **86**:845-848.

18. Kato M, Mizuno T, Shimizu T, Hakoshima T: **Insights into multistep phosphorelay from the crystal structure of the C-terminal HPt domain of ArcB**. *Cell* 1997, **88**:717-723.

19. • Jenal U, Galperin MY: **Single domain response regulators: molecular switches with emerging roles in cell organization and dynamics**. *Curr Opin Microbiol* 2009, **12**:152-160.
This review provides an insightful view from both experimental and computational angle on a particular group of RR.

20. Wuichet K, Alexander RP, Zhulin IB: **Comparative genomic and protein sequence analyses of a complex system controlling bacterial chemotaxis**. *Methods Enzymol* 2007, **422**:1-31.

21. Parkinson JS, Kofoid EC: **Communication modules in bacterial signaling proteins**. *Annu Rev Genet* 1992, **26**:71-112.

22. Bahn YS, Kojima K, Cox GM, Heitman J: **A unique fungal two-component system regulates stress responses, drug sensitivity, sexual development, and virulence of *Cryptococcus neoformans***. *Mol Biol Cell* 2006, **17**:3122-3135.

23. • Catlett NL, Yoder OC, Turgeon BG: **Whole-genome analysis of two-component signal transduction genes in fungal pathogens**. *Eukaryot Cell* 2003, **2**:1151-1161.
This study provides a phylogenomic analysis of two-component systems in a number of filamentous fungal pathogens in comparison to the model organism *Neurospora*. This study established the conservation of the downstream phosphorelay proteins among fungi despite variable numbers of HK proteins.

24. Pott GB, Miller TK, Bartlett JA, Palas JS, Selitrennikoff CP: **The isolation of FOS-1, a gene encoding a putative two-component histidine kinase from *Aspergillus fumigatus***. *Fungal Genet Biol* 2000, **31**:55-67.

25. Virginia M, Appleyard CL, McPheat WL, Stark MJ: **A novel 'two-component' protein containing histidine kinase and response regulator domains required for sporulation in *Aspergillus nidulans***. *Curr Genet* 2000, **37**:364-372.

26. • Bahn YS: **Master and commander in fungal pathogens: the two-component system and the HOG signaling pathway**. *Eukaryot Cell* 2008, **7**:2017-2036.
This comprehensive review describes in detail the well-studied fungal two-component system and its interaction with the MAPK-mediated HOG stress response signaling pathway.

27. Anjard C, Loomis W: **Histidine kinases of dictyostelium**. In *Histidine Kinases in Signal Transduction.* Edited by Inouye M, Dutta R. Academic Press; 2003:421-438.

28. Hwang I, Chen HC, Sheen J: **Two-component signal transduction pathways in *Arabidopsis***. *Plant Physiol* 2002, **129**:500-515.

29. Pareek A, Singh A, Kumar M, Kushwaha HR, Lynn AM, Singla-Pareek SL: **Whole-genome analysis of *Oryza sativa* reveals similar architecture of two-component signaling machinery with *Arabidopsis***. *Plant Physiol* 2006, **142**:380-397.

30. Ashby MK: **Survey of the number of two-component response regulator genes in the complete and annotated genome sequences of prokaryotes**. *FEMS Microbiol Lett* 2004, **231**:277-281.

31. Burger L, van Nimwegen E: **Accurate prediction of protein–protein interactions from sequence alignments using a Bayesian method**. *Mol Syst Biol* 2008, **4**:165.

32. • Weigt M, White RA, Szurmant H, Hoch JA, Hwa T: **Identification of direct residue contacts in protein–protein interaction by message passing**. *Proc Natl Acad Sci USA* 2009, **106**:67-72.
A superior methodology to predict contact sites between HK and RR, which is applicable to other systems employing protein–protein interactions.

33. Durham-Colleran MW, Verhoeven AB, van Hoek ML: ***Francisella novicida* forms in vitro biofilms mediated by an orphan response regulator**. *Microb Ecol* 2009.

34. Wang W, Shu D, Chen L, Jiang W, Lu Y: **Cross-talk between an orphan response regulator and a noncognate histidine kinase in *Streptomyces coelicolor***. *FEMS Microbiol Lett* 2009, **294**:150-156.

35. Pan X, Ge J, Li M, Wu B, Wang C, Wang J, Feng Y, Yin Z, Zheng F, Cheng G *et al.*: **The orphan response regulator CovR: a globally negative modulator of virulence in *Streptococcus suis* serotype 2**. *J Bacteriol* 2009, **191**:2601-2612.

36. Higgs PI, Jagadeesan S, Mann P, Zusman DR: **EspA, an orphan hybrid histidine protein kinase, regulates the timing of expression of key developmental proteins of *Myxococcus xanthus***. *J Bacteriol* 2008, **190**:4416-4426.

37. D'Hooghe I, Michiels J, Vanderleyden J: **The *Rhizobium etli* FixL protein differs in structure from other known FixL proteins**. *Mol Gen Genet* 1998, **257**:576-580.

38. Boesten B, Priefer UB: **The C-terminal receiver domain of the *Rhizobium leguminosarum* bv. viciae FixL protein is required for free-living microaerobic induction of the fnrN promoter**. *Microbiology* 2004, **150**:3703-3713.

39. David M, Daveran ML, Batut J, Dedieu A, Domergue O, Ghai J, Hertig C, Boistard P, Kahn D: **Cascade regulation of nif gene expression in *Rhizobium meliloti***. *Cell* 1988, **54**:671-683.

40. Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D: **Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus**. *Proc Natl Acad Sci USA* 2002, **99**:12246-12251.

41. •• Puthiyaveetil S, Kavanagh TA, Cain P, Sullivan JA, Newell CA, Gray JC, Robinson C, van der Giezen M, Rogers MB, Allen JF: **The ancestral symbiont sensor kinase CSK links photosynthesis with gene expression in chloroplasts**. *Proc Natl Acad Sci USA* 2008, **105**:10061-10066.
This work describes identification of a nuclear encoded HK gene in *Arabidopsis* whose protein product is exported to the chloroplast. The protein is shown to be evolutionarily related to plastid encoded HKs in red algae and to cyanbacterial HKs.

42. Maeda T, Wurgler-Murphy SM, Saito H: **A two-component system that regulates an osmosensing MAP kinase cascade in yeast**. *Nature* 1994, **369**:242-245.

Please cite this article in press as: Wuichet K, et al. Evolution and phyletic distribution of two-component signal transduction systems, Curr Opin Microbiol (2010), doi:10.1016/j.mib.2009.12.011

43. Li S, Ault A, Malone CL, Raitt D, Dean S, Johnston LH, Deschenes RJ, Fassler JS: **The yeast histidine protein kinase, Sln1p, mediates phosphotransfer to two response regulators, Ssk1p and Skn7p**. *EMBO J* 1998, **17**:6952-6962.

44. Thomason PA, Traynor D, Cavet G, Chang WT, Harwood AJ, Kay RR: **An intersection of the cAMP/PKA and two-component signal transduction systems in *Dictyostelium***. *EMBO J* 1998, **17**:2838-2845.

45. Oehme F, Schuster SC: **Osmotic stress-dependent serine phosphorylation of the histidine kinase homologue DokA**. *BMC Biochem* 2001, **2**:2.

46. Gao Z, Chen YF, Randlett MD, Zhao XC, Findell JL, Kieber JJ, Schaller GE: **Localization of the Raf-like kinase CTR1 to the endoplasmic reticulum of *Arabidopsis* through participation in ethylene receptor signaling complexes**. *J Biol Chem* 2003, **278**:34725-34732.

47. Stepanova AN, Alonso JM: **Arabidopsis ethylene signaling pathway**. *Sci STKE* 2005, **2005**:cm4.

48. Hwang I, Sheen J: **Two-component circuitry in *Arabidopsis* cytokinin signal transduction**. *Nature* 2001, **413**:383-389.

49. Oka A, Sakai H, Iwakoshi S: **His-Asp phosphorelay signal transduction in higher plants: receptors and response regulators for cytokinin signaling in *Arabidopsis thaliana***. *Genes Genet Syst* 2002, **77**:383-391.

50. Yeh KC, Lagarias JC: **Eukaryotic phytochromes: light-regulated serine/threonine protein kinases with histidine kinase ancestry**. *Proc Natl Acad Sci USA* 1998, **95**:13976-13981.

51. Makino S, Kiba T, Imamura A, Hanaki N, Nakamura A, Suzuki T, Taniguchi M, Ueguchi C, Sugiyama T, Mizuno T: **Genes encoding pseudo-response regulators: insight into His-to-Asp phosphorelay and circadian rhythm in *Arabidopsis thaliana***. *Plant Cell Physiol* 2000, **41**:791-803.

52. Matsushika A, Makino S, Kojima M, Mizuno T: **Circadian waves of expression of the APRR1/TOC1 family of pseudo-response regulators in *Arabidopsis thaliana*: insight into the plant circadian clock**. *Plant Cell Physiol* 2000, **41**:1002-1012.

53. Cock PJ, Whitworth DE: **Evolution of prokaryotic two-component system signaling pathways: gene fusions and fissions**. *Mol Biol Evol* 2007, **24**:2355-2357.

54. • Galperin MY: **Structural classification of bacterial response regulators: diversity of output domains and domain combinations**. *J Bacteriol* 2006, **188**:4169-4182.
A thorough classification of RR output domains showing their incredible diversity.

55. Lavin JL, Kiil K, Resano O, Ussery DW, Oguiza JA: **Comparative genomic analysis of two-component regulatory proteins in *Pseudomonas syringae***. *BMC Genomics* 2007, **8**:397.

56. Letunic I, Doerks T, Bork P: **SMART 6: recent updates and new developments**. *Nucleic Acids Res* 2009, **37**:D229-D232.

Please cite this article in press as: Wuichet K, et al. Evolution and phyletic distribution of two-component signal transduction systems, Curr Opin Microbiol (2010), doi:10.1016/j.mib.2009.12.011