ANALYSIS OF AN ORPHAN TWO-COMPONENT SYSTEM RESPONSE REGULATOR (VARA) THAT CONTROLS THE EXPRESSION OF *RHODOCOCCUS EQUI* VIRULENCE GENE (*VAPA*)

By

IJEOMA ADAOHA AZUONYE

(Under the Direction of Mary Hondalus)

ABSTRACT

Virulence of the intracellular pathogen of foals and AIDS patients, *Rhodococcus equi*, requires the plasmid encoded virulence determinant VapA (virulence associated protein A). The role of a putative orphan two-component transcriptional response regulator, VarA (virulence associated regulator), in *vapA* expression was examined. A *varA* deletion mutant was shown to produce reduced amounts of VapA protein. In a murine *R. equi* infection model, the *varA* regulator mutant was attenuated to the same extent as a *vapA* mutant. The direct and specific binding of VarA to the *vapA* promoter region was established. Bioinformatic analysis of the *R. equi* genome yielded 20 putative histidine kinases. Of these, two were found to interact with VarA in yeast-two-hybrid (Y2H) direct mating experiments. A traditional Y2H library screen revealed a novel type of interaction between VarA and a serine-threonine kinase-like molecule. If confirmed outside of the yeast system, these interactions purport of a highly complex virulence regulatory network.

INDEX WORDS: *Rhodococcus equi*, virulence, two-component system, response regulator, histidine kinase, transcriptional activation, yeast two hybrid, gel shift assay.

ANALYSIS OF AN ORPHAN TWO-COMPONENT SYSTEM RESPONSE REGULATOR (VARA) THAT CONTROLS THE EXPRESSION OF *RHODOCOCCUS EQUI* VIRULENCE GENE (*VAPA*)

by

IJEOMA ADAOHA AZUONYE

B.S., University of Massachusetts-Amherst, 2004

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2008

ANALYSIS OF AN ORPHAN TWO-COMPONENT SYSTEM RESPONSE REGULATOR (VARA) THAT CONTROLS THE EXPRESSION OF *RHODOCOCCUS EQUI* VIRULENCE GENE (*VAPA*)

by

IJEOMA ADAOHA AZUONYE

Major Professor: Mary Hondalus

Committee: Russell Karls
David Peterson

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2008

DEDICATION

To the Lord, Jesus Christ, my savior, a constant source of inspiration, relief, and rebirth.

"I am the vine; you are the branches. If a man remains in me and I in him, he will bear much fruit; apart from me you can do nothing” (John 15:5).

To my father and mother who have always voiced that going to school either solves most of one’s problems in life or leads to the solution. It does!

For my husband, Michael Perry, I thank God for commanding your presence in my life, both before and after this program, and I thank you each day for a beautiful now.

ACKNOWLEDGEMENTS

I would like to extend a special thank you to Dr. Mary Hondalus for all of her support, guidance and patience through this project and for instilling in me a strong work ethic and sound advice for my future endeavors. My entire committee, including, Dr. Russell Karls, Dr. David Peterson and Dr. Mary Hondalus, has been an invaluable resource for learning, encouragement and critique. For this, I am continually grateful. I also want to thank my undergraduate mentors, Dr. Susan Leschine, Dr. Pablo Pomposiello, and Dr. Elizabeth Brainerd, all of whom were instrumental in my introduction to research in the basic sciences.

A sincere thanks to my fellow lab mates, Katherine Kroell, Yang Hong, Artem Rogovskyy, Garry Coulson, Vibhay Tripathi, Shelley Dickerson, David Quiceno and Melissa Morello, for all of their assistance and support throughout the years and for graciously enduring my less than perfect lab habits. Also, many thanks is due to Naomi Lucci, Adrienne Cottrell and Valerie Cadet-Wimberly, who are all wonderful friends and colleagues, for showing me the ropes and being of immense help both in and out of the lab. A huge thank you to all the friends I've made here in Athens, particularly Ignacio, Caree and Dionne. The power lunches and smiles shall continue baby!

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

BACKGROUND AND SIGNIFICANCE

Rhodococcus equi

Rhodococcus equi is a serious opportunistic intracellular respiratory pathogen primarily affecting young horses and immunocompromised people. First isolated in 1923 from pulmonary lung abscesses of diseased foals, (Magnusson, 1923), it is now well-documented that *Rhodococcus equi* infections typically result in an often-fatal pyogranulomatous pneumonia in susceptible hosts (11, 97, 145, 181). In addition, disseminated disease may occur wherein the bacterium can be cultured from various extrapulmonary tissues (45, 99, 131, 145). Approximately 50% of foals with *R. equi* pneumonia also develop an enteritis with diarrhea due to associated infection and inflammation of the intestines (39, 215). While foals and immunodeficient humans are the chief hosts, *R. equi* infection of swine is common in Asia, and cervical lymph node abcessation is the predominant clinical manifestation in this species (124, 165). Infection of a wide variety of other mammalian species, including cattle, goats, sheep, deer, cats, and dogs has been reported (97, 145, 174), but is more rare. In these less typical host species, *R. equi* disease generally presents as a wound infection that may or may not progress to abscessation, lymphadenopathy, fever, bacteremia, meningitis, and septic arthritis (145, 147, 174).

R. equi is a soil-derived organism that is red (salmon) pigmented hence the name "Rhodo" and is coccoid in morphology when cultured on solid medium (172). Since its original discovery by Magnusson et. al., *R. equi* has been classified under various taxa of the group,

Actinomycetales, previously as a *Corynebacterium* spp. or *Mycobacterium* spp. (25, 153, 174). In 1977 (72), the bacterium was named *Rhodococcus equi* and classified as a nocardiform actinomycete among the following genera, *Nocardia, Mycobacteria, Streptomyces*, and *Corynebacteria*, to which it is evolutionarily most closely related (97, 145, 174). The nocardiform actinomycetes are characterized by their G/C rich genomes and lipid/mycolic acid rich cell walls, and all are aerobic, gram positive and catalase positive (72-74).

R equi is widespread in the soil and is commonly present in the gastrointestinal tracts of grazing animals (1, 32, 221). While *R. equi*, an obligate aerobe, does not grow in the anaerobic intestinal environment of herbivores (17, 104, 173, 241), it is frequently detected in their feces (79, 174, 215). Previous studies indicate that *R. equi* is able to grow readily at temperatures that range from 15^{o}C to 37^{o}C (71, 104), characterizing the range of temperatures found in the general soil environment and the conditions found within fecal matter and in the intracellular environment. The optimal growth temperature for *R. equi* is 30^{o}C and it is unable to grow at temperatures below 10^{o}C (71, 104). Diagnoses of foal infections are more frequent during months of higher temperatures with wind and dry conditions acting as contributing factors (211). Inhalation of aerosolized dust particles coated with bacteria is the main route of *R. equi* infection (200).

Young horses are exceptionally susceptible to *R. equi* disease during their first six months of life with the majority of cases occurring in foals less than three months of age (10, 137, 251). Rare reports of *Rhodococcal* disease in adult horses are primarily co-incident with pre-existing immunodeficiencies or systemic illness (51, 64). With these exceptions, adult horses are immune to virulent *R. equi*. Unlike immunocompetent adult horses, foals are most severely immunologically naïve during the first three months of life and thus, are unable to mount an

effective immune response to virulent *R. equi* challenge (23, 100, 117, 140). Most foals are born during the spring and summer months, thus aligning the window of peak foal susceptibility with optimal environmental conditions for *R. equi* exposure (174). Effective clearance of the bacterium requires the induction of cell mediated immune responses including the release of interferon gamma (IFN-γ) which mediates activation of macrophages, triggering the generation of ROIs[1] and RNIs[2] (43, 93, 118) and direct cytotoxic effects. In addition to immunological naiveté, foals, like many other neonates, exhibit a TH_2-type cytokine bias which antagonizes the TH_1 type inflammatory responses that engender effective control and clearance of the organism (23, 33, 101, 140). Foals that are exposed early in life with a large and or continuous dose of *R. equi* are predisposed to development of active disease (174). While cell mediated responses are clearly required for protection against *Rhodococcal* disease, previous reports have suggested a role for antibody in the prevention of *R. equi* pneumonia (93, 100, 101, 132, 135, 166, 209). Even though the majority of immunocompetent adult horses have significant titers of anti-*R. equi* antibody, adoptive transfer of hyper-immune plasma to foals has proven to be only variably protective and an often unreliable and expensive method of controlling subsequent virulent *R. equi* challenge (33, 93, 101, 166, 174). Since the 1980s, with increases in HIV infection and vast improvements in diagnostic tools, *R. equi* infection has been more frequently diagnosed in humans and the organism has established itself in the recent decades as an emerging opportunistic pathogen (97, 153, 174, 175, 181, 190, 227). Human hosts are commonly post organ-transplant patients and cancer patients undergoing immunosuppressive therapies, and, more frequently, AIDS patients with severely low CD4 cytotoxic T- lymphocytes (11, 54, 77, 131, 138, 144, 153, 157, 175, 181, 190, 193, 195, 201, 206, 232). Susceptible humans, therefore,

[1] Reactive Oxygen Intermediates (Hydrogen peroxide, Superoxide).
[2] Reactive Nitrogen Intermediates (Peroxynitrite).

have reduced cellular immune function and are thus unable to mount effective clearance responses leaving them at risk of developing disease from a variety of opportunistic organisms. Despite appropriate antibiotic therapy, *R. equi* disease in people has a mortality of approximately 55% (86).

Pulmonary *R. equi* disease typically results in a pyogranulomatous bronchopneumonia with abscessation characterized by massive infiltration of neutrophils and macrophages. Shortly after inhalation, the bacterium is phagocytosed by the alveolar macrophages. At this stage, immunocompetent people and adult horses are able to suppress the replication of *R. equi* and clear the bacterium within a few days post infection (93, 94, 174, 209). Adequate immune responses include IFN-gamma-mediated activation of macrophages complete with the recruitment of neutrophils and cytotoxic killing mechanisms (23, 93, 94, 106, 107, 117, 118, 133). Intracellular organisms are then trafficking to lysosomal compartments for destruction (55, 174, 224). In contrast, in immunodeficient hosts with inadequate cellular immune functions, the bacterium is able to persist and replicate intracellularly within macrophages in unabated fashion (55, 224).

Virulence and Pathogenicity

R. equi pathogenesis remains poorly understood yet, since the early 90's, significant strides have been made. By analyzing sera of infected foals, Takai and others demonstrated that a 17-20kDa protein, later termed VapA (Virulence associated protein A), is expressed in large quantities by strains isolated directly from pneumonic lesions of afflicted foals (20, 68, 92, 97, 107, 169, 178, 210, 212, 214, 216) and is an immunodominant antigen. It was also discovered

that VapA is vulnerable to trypsin digestion and can be labeled with biotin, qualities that demonstrate localization of VapA to the cell surface (212, 214). Additionally, radiolabeling of VapA with [9,10-^{3}H] palmitate demonstrated that the protein is anchored to the cellular surface as is typical of modified proteins (145, 174). Surface expressed proteins are in a prime position to mediate host-pathogen interactions. Taken together, the characteristics that VapA exhibits are typical of bacterial virulence factors. The discovery of VapA was followed by another breakthrough in *R. equi* research. A large 80 - 85kDa virulence plasmid exclusive to virulent strains, was found to house the VapA gene (210). Furthermore, it was shown that disease causing strains cured of the virulence plasmid become avirulent in foals and mice and are quickly cleared (20, 28, 46). In addition, a 20 kDa protein, VapB, bearing high sequence homology to VapA was found in pig isolates of *R. equi*. VapA and VapB bear substantive amount of similarity to one another and are antigenically related, yet, they are not expressed by the same *R. equi* isolate. Isolates recovered from HIV-infected individuals express either VapA or VapB (35, 145, 151, 209). However, only VapA-expressing isolates have been found in foal isolates (209, 217, 231). These data suggest that VapA and VapB may be analogs of one another and are likely distinguished only by qualities that dictate host-specificity (145).

Sequencing of the virulence plasmid has contributed significantly to the current understanding of the pathogenicity of the organism. Based on sequence comparisons to known genes, the plasmid can be divided into three distinct regions organized by putative gene function: The conjugation region, replication and partitioning region and a 27.5 kDa pathogenicity island comprised of 27 open reading frames. As is the norm, the pathogenicity island is characterized by a significant change in G/C content (60.8%, compared to 66.6% for the rest of the plasmid), the presence of flanking tRNA-like genes, resolvase-like genes, and inverted repeats which

signify acquisition of the island by horizontal gene transfer (145, 222). Most importantly, VapA was found to reside within this island along with 6 homologues, VapC, D, E, F, G, H, and VapI (a pseudogene) a homologue of VapG that harbors a frame shift mutation (29, 170). The members of this novel 8-member gene family bear high similarity to one another at their C-terminal ends[3] but have no known homologues in other organisms. With the exception of VapA (which is surface anchored), six Vap proteins and four other ORFs (including ORF8) in the pathogenicity island are expressed along with signal sequences that allow for secretion of soluble proteins (28, 97, 145, 174). A study conducted by Jain and colleagues demonstrated that *vapA* is a virulence gene and that *vapC–vapF*, unlike *vapA*, were dispensable for growth of *R. equi* in mice (97, 109). While VapA has been implicated as a key virulence determinant, it is not likely to be the only gene encoded by the virulence plasmid that is essential to virulence (97, 98, 174). For instance, a plasmid cured *R. equi* strain that expressed a wildtype copy of VapA alone was insufficient in restoring replication of the strain in macrophages and virulence of the organism in mice and foals (109).

Some of the genes within the pathogenicity island (PI) are organized in operons and two regulators are identifiable within this region. Originally considered a mono-cistronic gene, recent studies have delineated that *vapA* is the first gene in a four member operon along with *vapI*, C, and D (29) (as shown in figure 3). Meijer and others discovered that *vapA* transcript is stable for a longer period of time than those of *vapI*, C, D, leading to differential expression of their gene products. Upstream of the *vapA* operon, a second operon houses five genes inclusive of *vapH* and the two regulatory genes (30, 174, 183). The first gene in the operon encodes VirR, a *LysR*-Type transcriptional regulator that exerts some control over *vapA* expression by binding to a

[3] VapF without its two frame-shift mutations would have C-terminal homology with other Vap genes.

promoter that is located directly upstream of *vapA*. Meijer and others demonstrated that a strain of *R. equi* lacking the *virR* operon was unable to produce VapA. Reconstitution with *vir*R alone was insufficient in restoring wildtype levels of VapA protein, whereas, reconstitution of the entire operon, was sufficient. Hence, VirR is not considered to be the sole regulator of *vapA* expression. These findings prompt significant research interest in the role of the second regulatory gene located at ORF8 and its relationship to VapA expression.

VarA (Virulence Associated Regulator A)

For the remainder of this work, the gene encoded at ORF8 on the virulence plasmid of *R. equi* will be referred to as VarA (Virulence Associated Regulator – A) for its role in regulating VapA expression. Prior to this study, Meijer and others have presented data that indirectly suggests a role for VarA in VapA expression and directly illustrates that both VirR and VarA are critical to *R. equi* virulence. Reconstitution experiments by Meijer and colleagues established that while VirR, the LysR-type transcriptional regulator exerts control over the production of VapA, reconstitution of the plasmid cured strain with the entire *virR-varA* operon was required for a return to wildtype levels of VapA (183). Ren and others demonstrated that both VirR and VarA deletion mutants were attenuated *in vivo* despite a significant upregulation of three to four other *vap* genes (including *vapA*) (179). Our studies now aim to define VarA as a two-component response regulator that directly regulates *vapA* expression.

VarA bears 30 - 65% sequence similarity to two-component system response regulators in other organisms. Two-component systems (TCS) are comprised of a membrane bound

histidine kinase typically encoded adjacent to a cytoplasmic response regulator. This class of regulatory system is the most abundant one in prokaryotic organisms (41, 75, 120, 205, 238). Upon detection of an environmental stimulus, the histidine kinase autophosphorylates at a conserved histidine residue and mediates phosphate transfer to a conserved aspartate residue on the response regulator. Phosphorylation activates the response regulator and enables transcriptional control by direct binding of the regulator to the promoter region of genes that encode appropriate effectors (22, 75, 82). Interestingly, there are no histidine kinase homologues encoded on the virulence plasmid and thus, the kinase partner of VarA must be encoded on the chromosome. This study seeks to adress two questions:

1.) Is VarA a direct regulator of VapA expression?

2.) Is VarA a two-component system response regulator?

Bacterial gene regulation systems

Prokaryotic gene regulation occurs by a variety of complex mechanisms (16, 21, 41, 134, 136). At the RNA level, transcription factors (TFs) interact with a specific DNA element in order to drive RNA-polymerase-directed transcription of gene message which may subsequently be translated to effector proteins (129). TFs that bind to cis-acting DNA elements are integral in the regulatory process and may function either as activators or a repressors of gene expression (96, 111, 129, 177, 185, 204, 243). Activators bind to promoter DNA sequence to stimulate activity of RNA polymerase. Whereas, repressors directly bind to RNA polymerase or to operator sequences to inhibit the initiation of transcription. During infection, an organism encounters a

variety of environmental stress stimuli that challenge its viability. Thus, in order to persist and cause disease within the host cells and tissues, bacteria require efficient adaptive and protective responses. TFs that are associated with virulence typically mediate the expression of virulence determinants that enable host colonization, evasion, persistence and toxin production among several others (36, 126, 146, 219). In fact, the disruption of TF genes that encode direct regulators has been an effective means of attenuating virulence of several pathogens (159, 189). These transcription events are tightly regulated and specifically coordinated by a series of signal transduction events that result in expression of a subset of genes that mediate the necessary adaptive responses (36, 47, 146, 237, 242). Host environmental cues such as changes in temperature, salinity, acidity, nutrient availability and antimicrobial species such as reactive oxygen and nitrogen moieties are examples of typical input signals. The activity of these TFs may occur either through direct or indirect contacts with the cis-acting DNA element at the promoter region: some bind to the DNA sequence directly, others interact with each other or to other TFs (96, 151).

The complexity of some bacterial gene regulation systems may allow for multiple auxiliary TFs to substitute for one another in order to facilitate the eventual binding of a key regulator to cis-acting elements. For example, a *Pseudomonas syringae* strain encodes a sigma regulator factor, PsrA, which controls the regulation of genes that mediate epiphytic fitness, quorum-sensing signals, and interactions with its plant hosts (36). This type of regulation is indirect and is referred to as a regulator chain. In this context, the direct regulator, a TF that directly interacts with the DNA sequence at a promoter region, is a key component of bacterial gene regulation. In evaluating whether a TF is a direct regulator, sequence specific contacts

between the TF and DNA sequence must be demonstrated. A common method of evaluating this type of activity is the electrophoretic mobility shift assay (EMSA) (35, 192).

Electrophoretic Mobility Shift Assay

EMSA analysis is a well-established means of evaluating DNA sequence-specific protein interactions (35, 192). Purified protein of interest is incubated with linear DNA fragments (probe) that contain the putative binding site and the results of the binding assay is then examined on a non-denaturing polyacrylamide gel. The assay is based upon the observation that complexes of protein and DNA migrate through a gel matrix more slowly than free probe DNA fragments. Because migration of a protein-DNA complex through the gel is dependent on the sum of the molecular weights of both the DNA and the bound protein, a DNA-binding protein with a larger molecular weight retards probe migration to a greater extent (59). It is expected that some DNA-binding proteins exhibit cooperative binding and therefore bind to DNA in multimeric forms (31, 194). In such cases, binding results in two or more bands (within the same lane) that are indicative of these multimeric protein complexes. The gel shift assay can be performed by two methods. The protein can be radio-labeled and incubated with an unlabeled DNA probe. The reverse is the more conventional approach in which the DNA fragment is the radiolabeled component. Alternatively, DNA can be labeled by safer non-radioactive methods which can be as effective as radioactive detection methods (31, 163, 182, 192). Some of the limitations of this assay lie in the suitability of the probe sequence selected and in the optimization of experimental conditions. These critical parameters are listed as follows:

Probe Selection

The target DNA sequence is normally a linear fragment of less than 300 base pairs (bp) which contains a sequence(s) of interest. Selection of probe DNA sequence hinges upon the approximation of postulated binding sites based on the type of TF and the identification of DNA motifs to which it is known to bind. In the event that this type of information is unavailable, one can utilize a series of short sequences that surround cis-acting DNA elements upstream of the gene that is regulated to locate the sequences that show reduced migration through the gel matrix. However, to allow for the multiprotein complex assembly characteristic of many transcription factors, probe fragments longer than 100bp are better able to accommodate this type of protein:DNA interaction (31, 60, 113).

Nonspecific competitor

Unlabeled repetitive polymers of DNA such as poly(dI•dC) or poly(dA•dT) are typically included in the binding reactions to minimize the non specific binding of protein to the labeled target DNA. When added to the reaction along with the protein of interest and prior to the addition of labeled target DNA, they provide an excess of nonspecific sites that absorb proteins that will bind to any general DNA sequence. The result is a reduction in the observation of nonspecific DNA:protein interactions that generate false positive band shifts. The amount of carrier DNA can and should be titrated as an additional measure to create optimal binding conditions. A suitable amount of carrier DNA is that which is sufficient in preventing nonspecific proteins from binding to the probe, yet not too high as to prevent the specific binding of proteins that recognize

target sequences within the probe DNA. To ensure optimal results, the chosen type of carrier DNA sequence should bear little resemblance to the specific binding site.

Controls

To perform a thorough analysis of binding specificity, one must perform a series of control experiments to rule out false effects that can be attributed to nonspecific interactions that occasionally occur in EMSAs (35, 113, 180). For example, in the presence of DNA fragments containing mutant binding sites or unrelated DNA sequences, low intensity band shifts may be observed. The reverse is also possible, wherein an unrelated protein forms weak complexes with the target DNA sequence and generates a low intensity band. These low-intensity bands represent nonspecific, low-affinity interactions that are characteristic of DNA-binding proteins. The distinction between specific and nonspecific interactions can often be established by performing a series of control experiments including those described below.

Competition Mobility Shift Assay.

Most protein preparations contain both specific and nonspecific DNA-binding proteins despite stringent purification methods. Therefore, it is necessary to assess the sequence specificity of protein:DNA interactions. One method is by performing a competition binding assay. For a specific competition experiment, an excess of the same probe DNA fragment (unlabeled) can be used in binding reactions performed in the presence of the labeled probe DNA and the protein of interest. The observance of an abrogated gel shift is indicative of a specific interaction because unlabeled protein that binds to unlabeled DNA fragments is not amenable to detection.

Irrelevant DNA control

A second method of evaluating sequence specificity of DNA-protein complexes is by examining the behavior of the protein in the presence of a control probe sequence of similar size to the target probe DNA but dissimilar in sequence. This control is used in assessing the level of non-specific binding of the protein of interest to random DNA sequences.

Irrelevant Protein control

A third method of establishing specificity of binding is by evaluating the ability of a non-specific protein to shift the target probe DNA. An irrelevant protein control is typically of similar size and/or of similar molecular properties as the protein of interest.

Transcription factors

Predictions of protein class and function, made prior to experimentation and using computational methods, have become a widely accepted and invaluable resource for researchers. The identification of key sequence motifs and architectural markers are fundamental in making hypotheses about mechanisms of gene activity. One of the more prominent motifs found in TFs are HTH (helix-Turn-Helix) motifs which coordinate TF-DNA contacts and characterize the more highly-conserved DNA-binding domain of the protein (8). To a lesser degree, other DNA-binding motifs such as the zinc finger motifs (123), RNA-binding motifs (128, 139), and antiparallel β–sheets (103) may be represented at the DNA-binding domain. These regulatory

TFs almost invariably consist of two domains: an activating domain and a DNA-binding domain (103). Occasionally, a third, stimulus sensing domain is present (75, 120, 141, 150, 240). The location of each of these domains distinguishes the type and class of transcription factor and may also be predictive of activator or repressor activity (103). For example, LysR-type transcription factors typically consist of all three domains on a single molecule with the DNA-binding region at the N-terminus, and the activating domain at the C-terminus (183, 187, 245). This class of TF is the only known class of positive regulators that house an N-terminal DNA-binding domain (245). On the other hand, two-component system TFs house these domains on separate molecules that work in tandem (82, 134). The response regulator component houses the DNA-binding region at the canonical C-terminal region and the activating domain at the N-terminus. Meanwhile a third stimulus sensing domain is carried by a histidine kinase molecule which completes the two-component regulatory system (205). The VarA protein sequence is most similar to two-component system response regulators of other bacteria (145, 174). A DNA-binding domain at the C-terminus of the protein is detected. Based on previous studies that detail response regulator activity, this domain arrangement, more specifically the presence of a receiver domain, is predictive of the potential for VarA interaction with a membrane-bound histidine kinase partner. In two-component signal transduction systems, a sensor histidine kinase autophosphorylates upon detection of an environmental signal and transmits that signal through phosphorylation activation of the response regulator. Histidine kinases and their cognate response regulators typically exist as neighboring genes in many species (40, 48, 87), however, there are circumstances in which they do not. For example, there are response regulators (and histidine kinases) that are encoded as orphan genes in isolation of a cognate partner (90, 152) and others that in isolation can interact with paired response regulators or histidine kinases encoded

elsewhere (15, 61, 184). This is an example of crosstalk which often represents an intersection of different signaling pathways (15, 61). Along this line, searches of the *R. equi* virulence plasmid do not disclose the presence of any histidine kinase homologues, therefore, VarA which is encoded on the plasmid sans a cognate kinase is referred to as an orphan regulator. To efficiently address the mechanism by which VarA regulates the expression of *vapA*, it is necessary to decipher what character of regulator that VarA represents. This involves examining its predicted capacity as a two-component response regulator. In this context, we hypothesize that if *varA* encodes an orphan two-component response regulator, then it is likely to have a sensor histidine kinase partner that is chromosomally-encoded. This hypothesis necessitates an active search for putative chromosomally-encoded histidine kinases which can be accomplished by exploiting yeast two hybrid methodology.

Yeast Two hybrid Methodology

The yeast two hybrid (Y2H) methodology has been proven a powerful method of studying the interaction of proteins. With Y2H, it is possible to evaluate interaction pockets between specific amino acid regions (as in a random genomic library screen) or interactions between entire protein molecules (as in a direct screen of proteins suspected to interact) (37, 38, 66, 69, 164, 244). The technique is performed *in vivo* by taking advantage of the yeast galactose utilization (Gal4) transactivator protein, an endogenously expressed 881 amino acid protein (57, 130, 148, 225). Gal4 is composed of a DNA binding (aa 1 - 147) (119) and activation (aa 771 - 881) domain (24). In a hallmark experiment, Fields and Song demonstrated that the two separate Gal4 domains, which lack function alone, could each be fused to one of two proteins (X and Y)

and synthesized in yeast as chimeric proteins. An interaction between proteins X and Y brings the separate activation domain (AD) into close proximity to the DNA binding domain (BD), thereby reconstituting the function of the Ga14 transcription activator protein and driving the expression of downstream Gal4-responsive genes in yeast reporter strains (108, 203, 207, 218, 225). Because real interactions between proteins as small as 8 - 10 aa (143) and as large as 750 aa have been successfully investigated in a two-hybrid format (50, 69), the potential for identifying interacting proteins on a genomic scale is great (56, 218). In this approach, called a library screen, a known protein of interest (called bait) expressed as a Gal4 binding domain (BD)-fusion can be used to screen for potential interactions with random DNA genomic library fragments of various sizes and affinities for the protein of interest. These library fragments are expressed as random Gal4 activating domain (AD)-fusions (called prey) (110). The system can be used in the detection of protein-protein interactions with dissociation constant *(K_d)* of as low as 10^{-6}, and this has been substantiated in previous studies (49). While hydrophobic domains such as transmembrane regions may affect expression or translocation through the nuclear pores of the yeast cell where the interaction occurs, these domains may be less problematic when expressed as a small percentage of the overall protein and when contained within the protein itself (244). In cases where the interactions between two known proteins of interest are in question, it is possible to closely investigate this by taking a direct approach to the yeast two hybrid format (143). The latter direct mating approach hinges on the fact that yeast strains exist as haploid cells and can form diploids with cells of the opposite mating type when cultured in the vicinity of one another (37, 53). MatA, one mating type produces chemoattractants (called pheromones) that entice the cells of the opposite mating type (Matα) towards itself and vice versa. Each mating type can be independently transformed with a bait or prey plasmid expressing

the fusion genes of interest. Once the cells are mated, they form diploid cells that harbor both plasmids from which the fusion proteins are expressed and translocated to the nucleus where they may interact (37, 148, 225).

Yeast strain description and reporter gene descriptions

The tight regulation of the transcription from GAL-responsive promoters by GAL4 makes it a valuable tool for manipulating the expression of reporter genes in two-hybrid systems (110, 203). However, one of the drawbacks of the Y2H methodology lies is that a significant portion of interactions obtained from an initial genomic library screen result from nonspecific self activation of the prey vector. In this scenario, a randomly-inserted library fragment expresses a Gal4-activation-domain fusion protein on its own having DNA-binding capabilities and the ability to activate reporter gene activity in a nonspecific fashion (56). Several measures are taken to minimize such events. For instance, two-hybrid libraries are usually constructed in the AD vector rather than the DNA-BD vector (56, 66, 69, 244). Avoiding a fusion of random RNA polymerase activating proteins to the Gal4-DNA-BD, one can prevent the generations of library protein fusions that function as autonomous transcriptional activators (66, 110). However, this strategy alone is not sufficient to eliminate all nonspecific interactions. For this reason, a *Saccaromyces cerevisiae* strain of yeast, called pJ69, that encodes four different reporter genes (*ADE2*, *HIS3*, *MEL1, LACZ*) which are under control of three independent Gal4-responsive promoters (57, 110) is used to adjust the screening process to various levels of stringency. This yeast host has deletions of key nutritional genes (*ADE2, HIS3*) that are dependent on the specific binding of the Gal4 reponsive promoters (110). Furthermore, the addition of the colorigenic

substrate, 5-bromo-4-chloro-3-indolyl alpha-D-galactopyranoside (X-α-gal), of the *MEL1* gene product, alpha-galactosidase (207) , can be used as an additional measure of stringency testing as with the addition of the colorigenic substrate, 5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside (X-gal), of the Lacz gene product, beta galactosidase (57, 110). Additionally, since the prey plasmid carries encodes a leucine nutritional marker (Leu), and the bait a tryptophan nutritional marker (Trp), episomally produced tryptophan and leucine complements the Trp and Leu auxotropy of the yeast strains and facilitates plasmid selection in yeast (110). A high-stringency test would evaluate the growth of bait and prey on medium lacking adenine, histidine, leucine, and tryptophan and supplemented with either X-gal or X-α-gal. This level of stringency should eliminate the majority of nonspecific interactions along with some weaker legitimate interactions. The stringency can be reduced to evaluate weaker interactions at the expense of encountering more false positive interactions. Given this, it is necessary to re-evaluate the results of an initial screen by re-testing the observed phenotypes under the same conditions in a direct mating format (37). A large majority of initially observed interactions fail to re-interact and may be attributed to transcient nonspecific interactions, or true weak interactions (56). A re-testing of the initially observed phenotypes typically results in approximately 10% of initial fusion proteins that interact reproducibly. The inserts from these, can be analyzed by isolation from yeast, sequencing and sequence analysis by bioinformatics methods.

In applying the yeast two hybrid methodology to VarA of *R. equi*, we planned two approaches which could lead to overlapping results. Firstly, a random library of *R. equi* genes was screened for interactions with VarA. Secondly, we directly assessed VarA interaction with predicted *R. equi* histidine kinases. The availability of a completely sequenced *R. equi* genome

sequence facilitated the search for chromosomally-encoded histidine kinases (HKs). However, since the genome was not annotated, bioinformatic analyses were utilized to identify putative HK partners. With the use of bioinformatics, at the sequence level, these predictions are made by searching for conserved domains and motifs that are typically found in two-component system histidine kinases.

Domain organization and recognizable features of two-component system histidine kinases

In searching for the putative histidine kinase partner of VarA, the alternative approach requires an understanding of HK domain architecture. The archetypal HK consists of three recognizable domains in the following order: A less conserved sensory region including transmembrane domains, a highly-conserved histidine kinase phosphoacceptor domain (HisKA) and a highly-conserved ATPase domain (HATPase) that is specialized for auto-phosphorylation of HKs (78, 120, 162, 196, 238). Although not always present, several accessory domains are also characteristic of histidine kinases. An intracellular sensory linker domain involved in signal transduction is often present in HKs. Secondary structure prediction of this linker domain show the presence of a highly-conserved helix-turn-helix (HTH) fold even though the amino acid sequences are not well conserved. Thus, these domains are called HAMP domains because this common HTH fold was shared with histidine kinases, adenylyl cyclases, methyl-accepting chemotaxis proteins, and phosphatases (105). Mutational studies of the HAMP domain of BarA (223) Aer (27, 236), EnvZ (121, 249) and NarX (6) demonstrated that the HAMP domain is essential to signal transduction and is required for effective signal transduction by histidine kinases (6, 105, 250). Also found in the intracellular region are two other domains involved in

sensory functions called GAF domains (9, 95). GAF domains can bind small molecule co-factors such as 3',5' cyclic guanosine monophosphate and formate (95, 116) and play an integral role in signal transduction (9). At the extracellular region, on occasion, a cyclase/histidine kinase-associated sensing extracellular (CHASE) domain is found. CHASE domains are highly-conserved and are present in a wide variety of sensory transmembrane receptors including, histidine kinases, adenylate cyclases, predicted diguanylate cyclases, and either serine/threonine protein kinases or methyl-accepting chemotaxis proteins (5, 154, 250). It is postulated that these domains mediate the recognition of such stimuli as cytokines and short peptides that are important for the development an organism (250). There are several forms of CHASE domains (named CHASE2 – CHASE5) (154). CHASE3 domains are present in histidine kinases and are often found in combination with the HAMP domain and the intracellular sensing domain, GAF (154). On occasion, more specialized domains such as the universal stress protein (USP) domain superfamily is present within various conserved proteins, such as histidine kinases, that are found in bacteria (127). These USP domains are typically involved in adaptation to a wide variety of extracellular insults faced by stationary phase cells and have been linked to protection against DNA-damaging agents (127, 161). Additionally, because histidine kinase identity depends on the type of signal that it detects, domains that represent recognizable sequences of other well characterized histidine kinases are typically found at the N-terminal region of the sensor proteins. For example, a KdpD-like domain is annotated in protein domain search programs, SMART[4] and Interpro[5], as sequences that are highly similar to the N-terminal region of the well-studied osmosensitive histidine kinase sensor protein KdpD (12, 52, 88, 114, 233). With an understanding of histidine kinase domain organization and associated domains, it becomes easier

[4] Simple Modular Architectural Research Tool.
[5] A comprehensive protein domain search tool that integrates information from other related repositories.

to locate putative histidine kinases by comparing highly-conserved sequences of well-studied histidine kinases of bacteria that are related to *R. equi*. SMART and Interpro serve as repositories for information about protein families. With the use of these programs, a recognizable input sequence results in the calculated prediction of motif-based structural domains and functional sites along with a taxonomic breakdown of proteins that typically contain the domains, links to their sequences and primary literature (155, 171, 191). When the results of searches made with these bioinformatics tools are taken together with BLAST[6] protein identity predictions made by aligning the query sequence with close matches in other organisms (7), it becomes easier to undertake the manual annotation of putative *R. equi* histidine kinases to analyze the interaction of those putative kinases with the VarA protein. If VarA is indeed a two-component system response regulator, it is possible that it interacts with a histidine kinase protein that is chromosomally-encoded. In contrast with a large-scale screen of VarA interactions with the *R. equi* proteome, the direct Y2H mating approach is a more efficient way to investigate the histidine kinase partner of VarA.

[6] Basic Local Alignment Search Tool.

CHAPTER 2

RESEARCH DESIGN AND MATERIALS

2.1 Strains and culture conditions

Plasmids, primers and bacterial strains, used in this work are detailed in (Appendix A-C). Chemically competent *E. coli* Tam1 cells were obtained from Active Motif (Carlsbad, CA) and were used in all *E. coli* based transformations. Unless otherwise noted, *E. coli* strains were grown at 37^{o}C in Luria-Burtani (LB) medium with shaking at 200-250 rpm, or on a roller apparatus. Where appropriate, antibiotics were included in the growth media at the following concentrations (μg/ml): Carbenicillin, 100; kanamycin, 50. Yeast strains were cultured at 30^{o}C in either YPD (Clontech, Palo Alto, CA) or SD (Synthetic Dropout selection media) from Difco laboratories (Detroit, MI) supplemented with appropriate CSM (Complete Supplement Mix) (Bio101, Vista, Calif.) amino acids (see media formulations – Appendix D).

2.2 Murine infection studies

All experiments were conducted in compliance with approved Institutional Animal Care and Usage Committee (IACUC) protocols. Female SCID mice were obtained from The Jackson Laboratory (Bar Harbor, ME) and infected at approximately 8 weeks of age. Frozen aliquots of the bacterial strains with known titers were thawed, cultured for 1 h at 37°C and after centrifugation, suspended in phosphate-buffered saline (PBS). Even though the titers of the frozen bacterial aliquots had been determined previously, the titer of the inoculum was

reconfirmed at the time of injection by dilution plating of the injection stock. Mice were intravenously (i.v.) injected in the tail veins with approximately 3 x 10^5 colony forming units (CFU) of either wild type *R. equi* 103+, the *vapA* deletion mutant, the *varA* deletion mutant or the *varA* deletion mutant complemented with an extrachromosomal copy of *varA*. To determine organ burden and monitor bacterial clearance, 4 mice per group per time point were humanely sacrificed at various times post-infection. Organs were aseptically removed postmortem and homogenized. Serial dilutions of the homogenates were plated onto brain heart infusion (BHI) agar (Difco Laboratories, Detroit, MI, USA) and plates incubated for 36-48 h. The number of bacterial colonies recovered was recorded.

2.3 DNA manipulations

For the purposes of this study and with the exception of those plasmids synthesized by Celtek Genes (Nashville, TN), purified genomic DNA from *Rhodococcus equi* strain 103+, commonly referred to as the wild type, was utilized as a template for polymerase chain reaction (PCR) amplification of *R. equi* genes. All PCR reactions were performed using a BioRad thermal cycler (Hercules, CA) and *Pfu* turbo polymerase (Stratagene, La Jolla, CA) in a total volume of 25 µl. Unless otherwise stated, 6% dimethyl sulfoxide (DMSO) was added to increase primer specificity on the G/C rich *R. equi* DNA. This typically resulted in significantly greater yields of amplified product. *Pfu* turbo buffer, supplied by the manufacture, contained magnesium and was used in all reactions at the recommended volume per 25 µl reaction. Dideoxynucleotide triphosphate (dNTPs) (Invitrogen Inc., Carlsbad, Calif.) were used at a final concentration of 200 µM. Primer working stocks were all 10 mM and used at a final concentration of 10 µM. Agarose

gel purification of PCR fragments were done by standard methods (115). All restriction enzymes and T4 DNA ligase were purchased from New England Biolabs (Beverly, MA) and used according to the manufacturer's instructions. In cloning procedures, the products of restriction digestion were separated by gel electrophoresis on 0.8% agarose gel prestained with Syber safe DNA intercalating dye (Invitrogen Inc., Carlsbad, Calif.), and visualized under ultraviolet light. Sequencing confirmation of constructs was performed by the UGA Sequencing Facility. Sequence analyses was aided by manual inspection of associated Chromas files for correct reading frame and integrity of base pairs relative to the expected gene product as defined by the 5,043,170 bp *R. equi* genomic sequence available at the Wellcome Trust Sanger Institute website (http://www.sanger.ac.uk/Projects/R_equi/). Sequence analyses were further facilitated by the use of Clone Manager software (Science and Educational Software, Cary, NC). Those plasmids that proved correct (in terms of restriction banding pattern and DNA sequence) were frozen at -80^{o}C as reference stocks. Approximately 1-5 micrograms of plasmid DNA or DNA ligation product was incubated with 50 microliters of Tam1 competent cells on ice for 30 minutes. Cells were heat-shocked in a 42^{o}C water bath for exactly 30 seconds and allowed to cool on ice for two minutes. Subsequently, 250 microliters of SOC medium, equilibrated to room temperature, was added to the heat shocked cells as recovery medium. Transformation reactions were then incubated at 37^{o}C with shaking at ~250 rpm for at least 1 hour. The cells were then plated on LB selection medium supplemented with the appropriate antibiotic.

2.4 Electrophoretic mobility shift assay

The *var*A gene was cloned as a MBP (maltose binding protein)–tag fusion in T7-RNA polymerase promoter-driven pIVEX vector (Roche, Alameda, CA), to allow for *in vitro* expression using the cell free Rapid Translation System (RTS) (Roche). The protein was generated, purified by nickel affinity chromatography and kindly provided by Dr. Adrie Steyn, University of Alabama (UAB). The total protein concentration was determined according to the Bradford assay using the Coomassie Plus Assay Kit (Pierce Chemical Co., Rockford, IL) with bovine serum albumin (BSA) as a standard. To detect sequence-specific DNA–protein interactions, a digoxygenin-endlabeled synthetic oligonucleotide (248 bp) comprising the previously determined and published *vapA* promoter region (P*vapA*) was amplified by polymerase chain reaction using the following primer set: forward primer: 5'-CACGGCAGTTGTTCGCAGGTT-3', reverse primer: 5'-TTAGCTCACCGCGGCACTCTC-3'. As an irrelevant DNA control probe, a 249 bp sequence was PCR-amplified from the pUC19 vector (Active Motif, Carlsbad, CA) using the following primer set; forward primer, 5'-ATACCTGTCCGCCTTTCTCCC-3', and reverse primer, 5'-TAGCACCGCCTACATACCTC-3'. Maltose binding protein was obtained from Biomeda (Foster City, CA) at a concentration of 0.2 mg/ml and was used in 640 ng and 940 ng quantities in a total volume of 10 µl. Electrophoretic mobility shift (EMSA) assays were performed as specified in the DIG Gel Shift Kit, 2nd generation (Roche Applied Science) using the reagents provided with the kit, at the recommended final concentrations. Probe DNA was labeled with Digoxygenin (DIG-ddUTP) at 37^{o}C for 15 minutes. In 4 µl of the provided gel shift binding buffer, various quantities of protein were then incubated with 1 µl of poly-dIdC nonspecific competitor DNA (0.1 µg/µl), and 2ul of

labeled P*vapA* (0.4 ng/µl) or labeled irrelevant probe DNA (0.4 ng/µl) for a 15 min incubation at room temperature (15 - 25^{o}C). The binding mixture was then loaded on a 6% NOVEX retardation gel (Invitrogen Inc., Carlsbad, CA) and electrophoresis was conducted at 80 V for 1.25 h. EMSA gels were electro-blotted and crosslinked onto positively-charged nylon membranes for 60 minutes at 30 V and 300 mA (Hybond-N^{+}; Amersham, Braunschweig, Germany) using the Novex X-cell II Module system (Invitrogen Inc., Carlsbad, CA). DNA was detected by chemiluminescence with an alkaline-phosphatase-conjugate anti-Digoxygenin antibody.

2.5 Yeast two hybrid methods

All yeast two hybrid experiments were performed in *Saccromyces cerevisiae* strain pJ69 described elsewhere in detail (110). The pJ69 matα and pJ69 matA strains were a kind gift from Dr. Adrie Steyn (UAB).

Backbone vectors and Control Plasmids

The following vectors which were previously described (Clontech) were a generous gift from Dr. Adrie Steyn (UAB): pVA3-1, which encodes a fusion of the murine p53 protein (amino acids 72 – 390) with the Gal4-BD, and pTD1, which encodes a fusion of the SV40 large T antigen (amino acids 87 – 708) with the Gal4-AD, served as positive controls throughout the yeast two hybrid library screening and direct mating approach as well as in the co-immunoprecipitation studies. pLAM5'-1, which encodes the human Lamin C protein (amino

acids 66 – 230) fused with the Gal4- BD, served as a negative control throughout these studies. lamin C neither forms complexes nor interacts with most other proteins.

Construction of the Bait Plasmids

Two separate in-frame fusions of *varA* were produced by directly cloning PCR amplified whole *varA* and partial *varA* into pGBKT7 (Clontech) as 5'*NcoI* 3' *SalI* fragments. To maintain the integrity of the reading frames, with respect to the Gal4 Binding Domain, nucleotides "gt" were added after the engineered *NcoI* restriction site on the forward primer. This creates a valine with the lone "T" nucleotide that would have otherwise shifted the frame. All plasmid constructs were confirmed via restriction digests and were sequenced (UGA Sequencing Facility). The sequences are free of mutations and frameshifts.

pIA13

Based on previous annotations of the *R. equi* virulence plasmid (28), the full length of the *varA* gene encoded at ORF8, is 756 bp long. The full length *varA* gene minus its stop codon was PCR amplified using the forward primer 5- ATAGTCGACATGCCGTGGCCCTGCCAA-3' and the reverse primer 5'-ATGCCATGGTTATGGAGTCCATTCGCGGTAT-3'. The region amplified represents virulence plasmid gene coordinates 8534 – 9290. The fragment was then gel purified and digested with restriction enzymes *Nco*I and *Sal*I. The resultant fragment was cloned in-frame with amino acids 1 – 147 of the Gal4 DNA binding domain in pGBKT7 (Clontech) to create pIA13.

pIA14

To avoid the possible autoactivity of the bait plasmids expressing the whole VarA fusion protein, in addition to the *VarA-whole* gene construct (pIA13), a plasmid was constructed in which only the minimal region projected to interact with histidine kinases, the receiver domain of the VarA protein, was cloned in-frame with the Gal4 DNA binding domain. Protein domain searches of the full VarA sequence using the Interpro program (http://www.ebi.ac.uk/Interpro/) delineated that the first 357 bases (virulence plasmid gene coordinates 8534 – 8891) represent the receiver domain of the regulator based on sequence alignments against a repository of other well-studied two-component response regulator sequences. The partial length *varA* sequence minus its stop codon was PCR-amplified using the forward primer 5'-ATAGTCGACGATACGCAATGCCAGCTCAC-3' and the reverse primer 5'-ATGCCATGGTTATGGAGTCCATTCGCGGTAT-3'. The fragment was then gel purified and digested with *NcoI* and *SalI* . The resultant fragment was cloned in-frame with amino acids 1 – 147 of the Gal4 DNA binding domain in pGBKT7 (Clontech) in order to create pIA14. Because the conserved aspartate residue is located in the receiver domain the omission of the last 399 bases of *varA* was not projected to have any negative effect on possible interactions of the resultant partial VarA protein with putative histidine kinases in a *R. equi* library screen.

Construction of the R. equi Yeast Two Hybrid Library

The library was made and provided by Shruti Jain (formerly of Harvard School of Public Health). Briefly, total *R. equi* genomic DNA was isolated and tested for the presence of the virulence plasmid by PCR using primers to genes located on the virulence plasmid (109). The

genomic DNA was partially and separately digested with *Aci*I and *Hpa*II. Upon agarose gel electrophoresis of the partial digests, fragments ranging in size from 400 to 2000 bp were cut out and eluted. The prey vector, pGADT7, was digested with *Cla*I and dephosphorylated using calf intestinal phosphatase. Complete dephosphorylation (and therefore minimal background) was confirmed by setting up self ligation reactions. Then the cut and dephosphoylated vector was separately ligated to the *Aci*I and *Hpa*II genomic DNA fragments and the reactions individually electroporated into *E. coli* Electromax DH10B cells (Invitrogen, Carlsbad, CA). A total of approximately 3 x 10^6 transformants were obtained from the *Aci*I fragment ligation reaction, while approximately 1.5 x 10^6 clones resulted when using the *Hpa*II fragment ligation reaction for transformation. *Aci*I and *Hpa*II fragment transformants were individually pooled and total plasmid DNA, representing the individual *Aci*I and *Hpa*II prey libraries, were isolated from each pool and stored separately.

Tests for Auto Activation

Self-activation tests of bait vectors were performed on SD –Ade –His –Trp interaction medium. Competent pJ69 matα were transformed separately with pIA13, pIA14, and pLam5'-1 (harboring a fusion of the human lamin C protein to the Gal4 DNA binding domain) served as a negative control for the autoactivation screen. Transformants were selected on SD –Trp plates. Three well-isolated colonies were selected at random from each of the three plates and streaked onto thirds of an SD-Trp plate. After four days of growth, large loopfuls of each clone were streaked onto thirds of SD –Ade –His –Trp plates. A lack of observed growth on interaction

medium even after two weeks of culturing at 30°C confirmed the absence of auto activation of the *ADE2* and *HIS3* reporter genes.

Co-transformation of bait and prey plasmids

In a single reaction, a liter of log phase (OD_{600} 0.65) pJ69A cells cultured in YPD broth were made competent and co-transformed via the lithium acetate method (67, 188) with 100 μg pIA13, 100 μg pIA14, and 200 μg of prey constructs harboring random fusions of a ~0.4 kB – 2 kB *R. equi* genomic library. Transformants were selected directly on high-stringency solid medium (SD –Ade –His –Leu –Trp). After 10 days of incubation at 30°C, 54 colonies were visible on interaction medium, SD –Ade –His –Leu –Trp.

Analysis and Verification of Putative positive Clones with Retesting of Phenotypes.

A well-isolated single colony of each putative interacting clone was repeatedly cultured in 5 ml of SD–Leu broth to relieve selective pressure on the bait vector. After four serial passages of each colony in liquid medium, a 100 μl sample was streaked for isolation on SD – Leu plates. Ten to twentyfour individual colonies were replica plated onto SD – Trp in order to test for the absence of growth on medium devoid of tryptophan supplementation. Those colonies that grew on SD – Leu plates but not on SD – Trp were designated cured of the bait plasmid. One of these was chosen at random and frozen with 10% glycerol at -80°C, and stored as the reference stock. To perform confirmatory matings, yeast strains of the opposite mating type, matα (pJ69α), were transformed individually with VarA Gal4 DNA binding domain fusion

constructs. 5 μg of pIA13, 5 μg of pIA14 and 5 μg of pLAM5'-1 the negative control plasmid (Clontech, Palo Alto, CA) were introduced into matα cells along with 0.1 μg of salmon sperm DNA via the lithium acetate method (67, 188). pJ69 matα transformants harboring these bait transformants were plated on SD – Trp medium. Loopfuls of pJ69 matA strains harboring only the 54 individual putative positive prey vectors were mated pairwise with pJ69 matα pIA13 transformants, with pJ69 matα pIA14 and with pJ69 matα pLam5'-1 transformants on YPD selective medium. After 6-12 hours of incubation at 30^{o}C, individual mating pairs were scraped off the YPD plates and expanded on SD –Leu –Trp to select for resultant diploid cells. A lawn of cells typically resulted. A loopful of each mating pair was taken from the SD – Leu –Trp plates and re-tested on high-stringency interaction medium identical to that used in the original screen (SD – Ade – His – Leu – Trp). Mating pairs were also plated on higher stringency medium (SD – Ade – His – Leu – Trp supplemented with 5-bromo-4-chloro-3-indolyl-α-D-galactopyranoside (X-α-gal)) allowing for easy detection of reporter activity from the *MEL1* encoded alpha galactosidase gene which metabolizes X-α-gal to yield blue colonies as upon growth. Putative positive prey vectors that yielded colonies on high-stringency interaction medium when mated with VarA-whole (pIA13) or VarA-partial (pIA14) but did not grow when mated with the pLAM5'-1 negative control were classified as re-confirmed putative interactions and set aside for additional levels of scrutiny in yeast. Of 54 original putative positive results, the 9 clones that re-tested positive on high-stringency medium were subjected to growth in the presence of the compound 3' amino-1'2'4-triazole. While the bait plasmids had been analyzed for auto activity, there are several reports of leaky expression from the *HIS3* reporter gene (26, 42), which increases the likelihood of encountering false positive results. The addition of 3-AT to high-stringency selection medium serves as a means for decreasing leaky expression of the *HIS3* gene

by inhibition of the product of the His biosynthetic gene (26, 42) thereby causing histidine starvation. Cells expressing the *HIS3* gene as a result of a putative interaction and Gal4 activation are more likely to overcome this starvation effect by way of *HIS3* reporter activity. However, we noted that 3-AT also negatively impacted the growth of all cells, including known strong interactors, pTD1 and pVA3.

Liquid alpha-galactosidase assay

As described in the Yeast Protocols Handbook (Clontech Laboratories, Inc), colonies expressing a pair of proteins being analyzed were plated on SD-Ade-His-Leu-Trp + X-α-gal interaction medium. After 4-5 days of growth, three 2 - 3 mm colonies from each plate, were selected for use in the liquid alphagalactosidase assay. A volume of 5ml of liquid synthetic dropout (SD) medium containing the appropriate dropout supplements was inoculated with the colonies and incubated overnight for approximately 16 - 18 hrs with shaking (250 rpm). The following day, each culture was vortexed vigorously for 0.5 - 1 min to disperse clumps and the optical density was determined by spectrophotometric readings at $O.D._{600}$ using a SmartSpec (BioRad Laboratories Inc., Hercules, CA). In Assay # 1, the $O.D._{600}$ of each cell suspension as 0.5 - 1. Assay # 2 consisted of mating pairs that were grown in the presence of 3-AT on SD-Ade-His-Leu-Trp + X-α-gal interaction medium and inoculated into appropriate SD medium. All colonies inoculated generated log-phase cultures of $O.D._{600}$ = 0.5-1. 1ml of each culture was centrifuged at 14,000 rpm for 2 min to pellet the cells. The cell-free supernatant was then carefully transferred to a clean microcentrifuge tube and stored at room temperature for use in the colorimetric assay. The colorimetric assay was then performed in a 96-well UV-transparent,

flat-bottom microplate (Corning Costar Corp., Cambridge, MA) using the 200 µl format exactly as described in the Yeast Protocols Handbook. The assay buffer consisted of 2X sodium acetate (EMD Chemicals, Inc., Gibbstown, NJ) and 1X *p*-nitrophenyl-α-d-galactopyranoside (PNP-α-Gal) solution of 100 mM (Sigma), and 10X Stop Solution of 1 M Na_2CO_3 (Sigma-Aldrich, Inc). Absorbance units were measured at 405 nM using a model 680 microplate reader (BioRad Laboratories Inc.). The α–galactosidase units were calculated as milliunits/ (ml x cell) using the following formula:

α-galactosidase [milliunits/(ml x cell)] = OD_{405} x V_f x 1,000/[(ε x b) x t x Vi x OD_{600}]

t = elapsed time (in min) of incubation

(Assay # 1 = 1150 min) (Assay # 2 = 1140min).

V_f = final volume of assay (200 µl).

V_i = volume of culture medium supernatant added (16 µl).

OD_{600} = optical density of overnight culture.

ε x b = *p*-nitrophenol molar absorbtivity at 410 nm x the light path (cm).

= 10.5^7 (ml/µmol) for 200-µl format.

Bioinformatic searches for histidine kinases

Using histidine kinase (HK) sequences of known histidine kinases (DevS, KdpD, MprB, MtrB, PhoR, PrrB) of *Mycobacterium tuberculosis*, protein domain searches were performed using the SMART (Simple Modular Architecture Research Tool) database. Using SMART,

permutations of the following three domains that are typically found within these histidine kinases (HKs) were located: transmembrane domain, the dimerization/ HK phosphoacceptor domain (HisKA) which contains a conserved histidine residue surrounded by an **XHZLRTPL** motif called the H-box **(**where X is usually an S but can vary, and Z is usually either a D or an E residue), and an ATP-binding region that is unique to HKs, referred to as the HATPase domain. The HisKA domain of approximately 8 to 20 amino acids was found in all of the aforementioned *M. tuberculosis* HK sequences used in searching against the *R. equi* genome sequence (http://www.sanger.ac.uk/Projects/R_equi/). From the *M. tuberculosis* HK sequences (DevS, KdpD, MprB, MtrB, PhoR, PrrB) only residues that were shown in SMART to compose of H-Box were selected and used in BLASTp searches against the *R. equi* genome. Sequences earning high similarity and identity scores were extended 1000 bp in frame in each direction and put through the SMART query to reveal putative HKs possessing the three domains that are characteristic of HKs. Furthermore, a stretch of sequence with high scoring matches that resemble the receiver domains of typical TCS response regulators was always identified upstream or downstream of the putative HKs. In this manner, 20 independent putative HK sequences and their putative 20 response regulator partners were hit upon a multitude of times (at least 4 times) with each *M. tuberculosis* sequence and at several different conserved regions of the putative kinase. The gene coordinates of these putative proteins were marked according to the numbering provided by the Wellcome Trust Sanger Institute. The fact that the putative response regulators (RRs) were found were in the same orientation and within approximately 2500 bp from the HK was key in solidifying a claim that these 20 putative HK and response regulator homologs are likely 20 paired putative Two-component Systems. Another method was

[7] According to Beer Lambert's law, Clontech has calculated this value as the proportionality constant (ε X b) using a Corning Costar UV-transparent flat-bottom plate.

used in an attempt to locate any missed HKs. The amino acid residues comprising the receiver domains of known *M. tuberculosis* response regulator partners of (DevS, KdpD, MprB, MtrB, PhoR, PrrB) were used in BLASTp searches against the *R. equi* genome. Strong hits were located on FASTA and 1 kb of upstream and downstream sequence was analyzed by SMART. No additional HKs were identified by this method. In sum, 20 putative paired HKs were found in sequences starting from bp 110854 – 4815737 on both the positive and complementary strands of the 5.04 Mb sequence, covering almost the entire length of the *R. equi* genome. With the exception of To examine the possibility of crosstalk between the 20 putative HKs and the VarA regulator, a plan was devised to individually clone all of the putative histidine kinases into the bait vector, transform yeast and perform direct mating experiments with yeast carrying prey plasmids into which VarA-whole or the receiver domain only had been cloned.

Cloning of HKs into the Bait Vector

The start and stop codons of each entire histidine kinase (HK) were annotated on the genome sequence obtained from the Sanger *R. equi* genomic file. HK ORF calling was done with the help of SMART, Interpro protein domain searches, and NCBI BLASTp searches as an additional measure of surety. In addition, manual searches for start codons were performed based on the frame in which the HK domains were found. It was also helpful to ensure that there were no overlapping ORFS, keeping in mind that on occasion, genes that are co-transcribed may overlap at the stop codon of one ORF and the start of a new ORF. The primers for amplification of the putative HKs were designed in such a way as to omit the N-terminal regions of the HK that participates in the sensing domain (namely the transmembrane regions) in order to reduce to

possibility of obtaining false negatives in a yeast two hybrid screen. As interactions occur within the nucleus of the yeast cell, it is necessary for the translated fusion genes to be translocated into the nuclear compartment via nuclear localization sequences that are produced at the N-terminal ends of the fusion proteins. Removing the hydrophobic rich, sensing domain and associated transmembrane regions would obviate the risk of a missed interaction due to HK fusions that become trapped in the nuclear pores of the host yeast cell. Structural analyses of each HK was performed via the TMpred software (http: //www.ch.embnet.org/software/TMPRED_form.html) which predicts possible transmembrane regions. Immediately after the last predicted transmembrane region and in-frame with the predicted protein, a forward primer was designed having an *Nco*I restriction site and nucleotides "GT" were added prior to the beginning of the *R. equi* sequence in order to maintain the proper frame with the Gal4 Binding domain while creating an artificial glycine site as a result. Glycine is a simple nonpolar/neutral amino acid, so it was thought that such a change would not likely alter the final properties of the fused VarA protein.

Cloning of VarA into prey vector

Whole and partial-length VarA was synthesized with restriction sites 5' *Xho*I 3' *Xma*I and cloned into the prey vector, pGADT7 by Celtek Genes and sequenced. Both constructs were verified by restriction digestion. Both cloned regions were confirmed for correct reading frame and integrity of bases by manual analysis of Chromas files and with the help of Clone Manager compare 2 sequences tool.

CHAPTER 3

RESULTS

3.1 VarA is virulence determinant

R. equi is a soil borne organism, yet it is capable of intracellular survival in mammalian hosts. To make the transition to a parasitic lifestyle, the organism must sense its surroundings and respond appropriately via expression of genes necessary for continued existence in the host. One of the genes required for host survival is *vapA*, the precise function of which is presently unknown. Proper regulation of *vapA* expression is critical to *R. equi* pathogenesis. The expression of *vapA* is induced by oxidative stress, by acidic pH, and by low concentrations of iron and magnesium which are all environments likely to be encountered in macrophages and *in vivo* (19, 20, 178, 208, 214). At the outset of my project, it was unclear how these various environmental cues were sensed and relayed to the cytosolic transcriptional machinery of the bacterium. The gene is one of the two regulatory genes located within the pathogenicity island of the *R. equi* virulence plasmid, both of which have been shown to be necessary for optimal *vap*A expression (183). A *varA* deletion mutant constructed by Jain and Hondalus (unpublished) was found to be significantly impaired in its ability to produce VapA protein (Figure 1). The latter *in vitro* observation led us to predict that the *var*A mutant would behave similarly to a *vap*A mutant strain *in vivo*. To address the effect of *varA* deletion on *R. equi* virulence, we challenged severe combined immunodeficient (SCID) mice, lacking B and T-cells with the *varA* mutant and compared its clearance to that of wild type strain 103+. By sacrificing mice at several times post challenge and plating dilutions of organ homogenates, bacterial burden was

determined. This murine model is unocompromised, and therefore, wild type *R. equi* replicates well and establishes a chronic infection in these mice, which would be lethal if allowed to run its course (133). In contrast to wild type *R. equi*, the *varA* deletion mutant failed to replicate in SCID mice and was cleared with similar kinetics as that of the *vapA* mutant (Figure 2A, 2B). Complementation of the *varA* mutant with an extra-chromosomal copy of the *var*A gene restored bacterial burden to that of wild type, confirming that the *in vivo* replication defect of the mutant was due to the deletion of *varA* and was not a polar effect on downstream genes. Taken together, the *in vivo* study shows the necessity of *varA* for *R. equi* virulence and the western analysis implies that *vapA* expression is greatly reduced in the absence of *varA* because *varA* regulates its expression. Based on its similarity to known two-component response regulators, VarA is predicted to contain a C-terminal helix-turn-helix motif, the foremost motif found in regulatory proteins known to bind DNA. Thus, we hypothesize that VarA controls *vap*A expression by binding directly to the *vapA* promoter rather than by regulating *vap*A expression indirectly via contact with an intermediate transcription factor.

3.2 VarA is a direct regulator of VapA production

To establish whether or not VarA is a direct regulator of *vap*A expression, electrophoretic mobility shift assays (EMSA) were performed. These experiments aimed to verify the *varA* gene predictions of a C-terminal DNA-binding domain that enables protein:DNA contacts at the promoter region of *vap*A. In these assays, the protein of interest (rVarA) was analyzed, along with the necessary controls for making assessments of binding specificity. Earlier studies of the *R. equi* virulence plasmid located the position of the *vapA* promoter and

determined the start sites of transcription and translation (183). That information allowed for the PCR-amplification of the 249 bp *vapA* promoter probe (Figure 3) utilized in the DNA binding assays. Recombinant VarA (rVarA) produced as an N-terminal maltose binding protein (MBP) fusion was available for these studies. Maltose binding fusion tags are used to increase the stability of the RNA and improve solubility of the protein during the over-expression process (220). These tags can be cleaved after purification of the desired protein (112). In this case, however, attempts to cleave the MBP tag were unsuccessful, necessitating the use of the MBP-tagged protein in the DNA binding experiments. Since protein domain search programs as used to determine that the VarA DNA-binding domain was C-terminally located, we reasoned that the MBP tag might not hinder DNA-binding. In addition, as a means of assessing possible binding artifacts resulting from the presence of the MBP tag, conditions were included wherein MBP alone was used as an irrelevant protein control. Incubation of the 249 bp *vapA* promoter probe (P*vapA*) with rVarA at molar ratios of approximately 1 : 4100 and 1 : 6000 produced band shifts (Figure 4, lanes 1 and 2) while incubation of P*vapA* with purified MBP at the same molar ratios did not (Figure 4, lanes 3 and 4). A probe only reaction served as a negative control (Figure 4, lane 5). This result establishes that rVarA binds directly to P*vapA*, and that the observed binding is not the result of false positive artifacts mediated by the presence of the uncleavable 5' MBP tag. To further examine the specificity of the binding reaction, additional experiments also examined the binding of rVarA to an irrelevant DNA probe, a 248 bp fragment of pUC19 DNA (New England Biolabs). As previously, band shifts are observed in reactions of rVarA incubated with P*vapA* at a DNA: protein molar ratio of 1 : 1400 and 1 : 2100 (Figure 5, lane 1, 2, 3), while rVarA failed to bind irrelevant probe DNA at the same DNA : protein molar ratios (Figure 5, lane 5, 6, 7). Incubation of increasing levels of rVarA with P*vap*A yielded two slowly migrating

complexes (Figure 5, lane 3). At lower amounts of rVarA, the more rapidly migrating complex was predominant (Figure 5, lane 1, 2). In sum, these data show that VarA binds directly and specifically to the *vapA* promoter.

3.3 The *R. equi* chromosome encodes at least 20 independent putative TCS

Two-component system (TCS) signal transduction systems typically consist of a histidine kinase (HK) and a response regulator (RR) partner that work in tandem to enable bacteria to, respectively, sense and response to a variety of environmental stimuli (205). Based on this traditional model of signal transduction, coupled with high sequence similarity of VarA to well-studied RR of other bacterial systems, speculation arose that VarA activity is mediated by phospho-transfer from a HK partner (149, 150, 240). The absence of a gene encoding a HK on the virulence plasmid led us to postulate that if VarA is indeed a two-component system RR, it likely has a partner and that partner is likely chromosomally-encoded. Therefore, two genomic approaches were undertaken to locate the putative HK partner(s) of VarA. The availability of a sequenced *R. equi* genome facilitated this search. However, to date, this genome sequence remains unannotated. In the first of these approaches, extensive bioinformatic analysis was performed to identify and clone all of the putative HKs of *R. equi*. Predictions of putative *R. equi* HK genes based on their similarities to known HKs in sequence, domain organization and neighboring gene analyses was performed as detailed in the Introduction. A thorough search (as detailed in Research Design and Materials) revealed 20 putative two-component system HK encoded on the *R. equi* chromosome. Membrane-spanning helices were detected in all 20 putative HKs within the N-terminal extracellular sensory region. Also detected in various HKs

were recognizable signal peptides. Each putative HK was neighbored by a RR homologue found within approximately 2500bp upstream or downstream of each HK gene and in the same orientation. These 20 putative two-component systems of *R. equi* are detailed in Figures 6A-T.

3.4 VarA interacts with two chromosomally-encoded putative histidine kinases in yeast

Having discovered 20 putative chromosomally-encoded kinases that each appeared to be paired with a cognate putative RR, we contemplated the possibility of regulator crosstalk between VarA and one or more of these chromosomally-encoded kinases. To determine which of the 20 putative HKs is responsible for activating VarA by phosphorylation, the yeast two hybrid (Y2H) methodology was used. Of the 20 putative *R. equi* HKs located, 17 were cloned as in-frame Gal4 DNA-binding domain (Gal4-BD) fusions for expression in a mating type α pJ69 yeast strain. Although attempted, the remaining three could not be cloned (Table 1). Once constructed and prior to use in experimentation, each of the plasmids was analyzed for auto-activation of reporter genes. One, HK18, was found to autoactivate, therefore, further analysis of this fusion was suspended. All others were confirmed to be free of nonspecific reporter activity. These HK bait plasmids were then individually analyzed in pair-wise direct mating experiments against a MatA pJ69 yeast strain carrying prey plasmids expressing whole length VarA (VarA-whole) or just the receiver domain of VarA (VarA-partial) as a fusion with the Gal4 activation domain (AD/VarA-W or AD/VarA-P respectively). Minimal non-specific interaction of both VarA fusions occurred using the well established Y2H control pLam5-1 which expresses human lamin C protein as a Gal4 binding domain fusion (as described in Research Design and Materials). Positive results were recorded in cases where mating of the bait with prey plasmids

resulted in growth on interaction medium (SD-Ade-His-Leu-Trp) while failing to grow when mated with empty prey plasmid (a negative control which is used in examining auto-activity of the HK plasmids). As detailed (Table 1), of the sixteen HKs that were examined in direct matings, two (HK1 and HK16) yielded positive results in yeast even under conditions of increased stringency in selection medium. An illustration of the positive interaction of VarA with HKI is provided in Figure 7. In these mating experiments, a Gal4-BD[8]/HK1 fusion protein encodes amino acids 702 - 1392 of a previously described *R. equi* PhoR HK homologue (Figure 12) and lacks the N-terminal putative trans-membrane regions, amino acids 1 to 701. A second fusion expresses Gal4-BD[9]/HK16 containing amino acids 958 - 1959 of an *R. equi* HK-like protein (Figure 13) (lacking the N-terminal putative trans-membrane regions, amino acids 1 to 957. When each of these gene fusions independently mated with the VarA-W and Var-P activation domain fusion proteins, the result was transcriptional activation of the reporter genes (*HIS3, ADE2, MEL1*) expressed from three independent Gal4-responsive promoters. The AD/VarA-W interaction in yeast is shown (Figure 7) while AD/VarA-P is not. Because the level of stringency used initially was insufficient to completely eliminate all background nonspecific transactivation when AD/VarA-W or AD/VarA-P were co-expressed with a negative control plasmid, pLAM5'-1, an additional level of stringency was added and the interactions of HK1 and HK16 to the two VarA fusions were reassessed. A common culprit of false positives derives from the *HIS3* reporter gene which is known to be 'leaky' in that is there is a low level of constitutive expression that can skew results (42, 56). By plating bait and prey co-transformants on high-stringency interaction medium supplemented with 3-amino-1,3,4- triazole (3-AT), a

[8] DNA-Binding domain

[9] DNA-Binding domain

competitive inhibitor of imidazoleglycerol-phosphate dehydratase, the *HIS3* gene product and an enzyme involved in the biosynthesis of histidine (91, 122, 239), the yeast cells become starved for histidine. This starvation can be overcome only through high level expression of *HIS3*, as would result from a true interaction and strong trans-activation events (37, 110). As depicted in Figures 8 and 9, supplementation of high-stringency medium with 40 μg/ml of X-α-gal (the colorigenic substrate for the *MEL1* gene product alpha-D-galactosidase) and 3 mM 3-AT completely eliminated the background growth observed in co-expression of the negative control vector pLAM5'-1 with AD/VarA-W or AD/VarA-P. Even under the enhanced stringency, co-expression of BD/HK1 and BD/HK16 fusion constructs with the VarA-Gal4-AD constructs resulted in significant growth which is consistent with true interactions. Since 3-AT essentially starves the cells for histidine, there was a notable amount of reduced viability in all mating pairs examined including the established positive control interacting pair pTD1 and pVa3 (Figure 8A, 9A) The remaining 14 HK Gal4-BD fusions were also individually mated with AD/VarA-W or AD/VarA-P but failed to grow on interaction medium even under lower stringency conditions, indicating a lack of interaction (data not shown). Table 1 displays the qualitative results obtained from these direct mating experiments.

Quantitative Analysis of HK1 interaction with VarA in yeast.

Having established that HK1 interacts with VarA in yeast, on indicator plates, we next attempted to quantify the level of reporter gene activity with standard liquid alpha galactosidase assays. Because alpha-D-galactosidase is secreted into the extracellular space, alpha-galactosidase activity is easily assayed by adding a colorigenic substrate into the culture

supernatant (see Research Design and Materials for details). We first evaluated reporter activity of cells grown in the absence of 3-AT to ensure efficient growth of cells to exponential phase while expecting some background levels of reporter activity due to the leaky *HIS3* gene. As shown in Figure 10, BD/HK1 mated with AD/VarA-W resulted in a significant level of alpha-D-galactosidase activity comparable to that observed with the positive control mating of pTD1 and pVA3 which, respectively, produce known strong interactors, p53 and SV40 T antigen. Consistent with qualitative results of mating pLAM5'-1 with AD/VarA-W or AD/VarA-P some background levels of reporter activity resulted but remained insignificant in comparison to BD/HK1 mated with AD/VarA-W and the positive control mating. The results obtained with the BD/HK1 mated with AD/VarA-P followed the same trend as with the AD/VarA-W. However, although activity was above background levels, the AD/VarA-P was significantly reduced to one-eight that of AD/VarA-W reporter gene activity (Figure 10). A repeat analysis of alpha-galactosidase reporter activity was conducted with the culturing of yeast cells in the presence of 3-AT with the expectation that 3-AT would eliminate nonspecific reporter gene. As in the initial assay, BD/HK1 mated with AD/VarA-W, BD/HK1 mated with AD/VarA-P and the positive control mating retained notable galactosidase activity (Figure 11). Although, quantitatively, the level of activity of all mating pairs including the positive control, PTD1/pVA3, was reduced by 50% or greater in the presence of 3-AT, which was previously observed to decrease cell viability. Notably, the BD/HK1 and AD/VarA-W mating resulted in reporter activity that was approximately 75% reduced in the presence of 3-AT. Similarly, reporter activity of the positive control interaction (pTD1 and pVA3) was reduced by 50% in the presence of 3-AT. In sum, these data suggest that, in yeast, HK1 (Figure 12) specifically binds VarA. Furthermore, this interaction becomes significantly weaker when limited to the predicted receiver domain of the

molecule. Because of the difficulty observed in obtaining synchronous log phase cells for the alpha galactosidase assay, the interaction between VarA and HK16 (Figure 13) was not analyzed using this method.

3.5 Yeast two hybrid (Y2H) *R. equi* library screen identifies putative interacting partners of VarA

Concurrent with the direct mating approach, a traditional screen of the *R. equi* genome was conducted in Y2H format. To avoid the generation of auto active clones, we cloned *varA* and not the library gene fragments into the Gal4-BD. These bait plasmids produce whole length VarA (VarA-whole) or just the receiver domain of VarA (VarA-partial) as a fusion with the Gal4 DNA-binding domain (BD/VarA-W or BD/VarA-P). Neither bait plasmid autoactivated when mated with empty prey vector (pGADT7) (Data not shown). A random *R. equi* genomic library of gene fragments ranging from approximately 400 - 2000 bp was previously created by S. Jain (formerly of Harvard School of Public Health) and was available for use. A large scale co-tranformation of BD/VarA-W, BD/VarA-P and AD/library into a pJ69 matA strain of *Saccaromyces cerevisiae* was performed and then the cells were directly plated on high-stringency interaction medium (SD-Ade-His-Leu-Trp) to select for both transformants and for strong putative interactions at the expense of eliminating weaker legitimate. A total of 54 putative positive colonies grew on high-stringency interaction medium. The high probability of obtaining false-positive results in library screens necessitated a second more direct screening of the 54 putative positives to assess the integrity of those interactions. All 54 putative colonies were cured of the bait plasmids expressing VarA-W and VarA-P (see Research Design and

Materials for detail). The interaction phenotypes were then re-tested by independent pairwise direct matings of a BD/VarA-W, BD/VarA-P expressing matA yeast strain with each of the original 54 putative positive clones. The negative control plasmid (BD/pLam5'-1) was also mated pairwise with each of the 54 putative clones in order to assess nonspecific interactions. Putative positive library inserts that re-interacted with the VarA fusions and not pLam5-1 under the same conditions as the initial screen (SD-Ade-His-Leu-Trp) and or under a higher level of stringency (SD-Ade-His-Leu-Trp +X-α-gal +3-AT) were judged to be true interactors in yeast. As detailed (Table 2), of the 54 re-tested putative clones, 9 re-interacted under the same high-stringency conditions as in the original screen, demonstrating reproducible results and the integrity of the screen. Of those 9 true interactors, all interacted with bait expressing the BD/VarA-W. Only 3 (C21, C28 and F21) interacted with both BD/VarA-W and BD/VarA-P (table 2). When assessed for interaction with VarA in the presence of 3-AT, 5 of the 9, (A27, B24, C21, C28, F21) yielded growth and blue color lending credence that these are strong interactions (Table 2). Of all 9 mating pairs that were re-assessed, it is notable that clone C28 is the only prey plasmid with an insert that yielded positive results when assessed with both BD/VarA-W and BD/VarA-P and also with and without the addition of 3-AT (Table 2).

VarA interacts with a serine-threonine kinase in yeast

To identify the genes whose products interacted with VarA in yeast, a previously described lyticase-based cell disruption method (13, 81) was used in isolating the plasmids from yeast that expressed interacting inserts. The tough cell wall of yeast has been cited as a deterrent in purification of plasmids. Of the 9 plasmids that reproducibly interacted in yeast, 5 were

isolated and their inserts analyzed by bidirectional sequencing of inserts. Despite attempted recovery, the remaining 4, plasmids are yet to be isolated from yeast. Table 3 details the results of sequencing and BLAST analyses of recovered inserts and the open reading frames of *R. equi* in which they are found. One of these, A24, is a 251 bp hit insert sequence that is encoded within the 1536 bp ORF00265 (gene coordinates 194291 – 195952) and is predicted to have a 'GTG' start codon downstream of a strong predicted ribosome binding site. A BLASTp search revealed that it bears high sequence similarity to an NHL repeat protein in the bacterium, *Solibacter usitatus* (Table 3, Figure 14). The NHL repeats, named after NCL-1, HT2A and Lin-41, are transmembrane sensory receptors that are found in serine/threonine kinase (STK) proteins in a diverse range of pathogenic bacteria. With an intracellular N-terminal kinase domain and extracellular C-terminal sensor domain these STKs play a role in signal transduction by autophosphorylating in response to external stimuli and phosphotransfer to other proteins. An alignment against various evolutionary distant serine threonine kinase proteins of other organisms including, *Rattus norvegicus*, *Drosophila*, *Solibacter usitatus* and a plant derived STK was performed. As anticipated, the sequences exhibit significant sequence variability leading to lower similarity scores (not shown). The analysis was extended to the pairwise comparison of a well-established STK1 of *Streptococcus agalactiaie* and Pkn8 of *Myxococcus xanthus* which have been shown to phosphorylate response regulator proteins (85, 176, 226). It is notable that the most similar of all STKs analyzed pairwise, was that of A24 with the solibacter NHL repeat containing protein with a score of 24 while that of two well-studied bacterial STKs generated a pairwise alignment score of only 14 (Data not shown).

VarA interacts with three other R. equi proteins

Two other library fragments that interacted with VarA in yeast were found within clones B24 and C25. Firstly, clone B24 encodes an *R. equi* library insert of 353 bp that exists within the 2199 bp ORF02813 (gene coordinates 2054140 – 2056338) and bears high sequence identity (87% (1310 bp/1496 bp)) to FtsK DNA translocase proteins of other bacteria. The highest scoring hits were found in the related, *Rhodococcus RHA1*, *Mycobacteria, Streptomyces* and *Nocardi*a species. Secondly, clone C25 encodes an *R. equi* library insert of 551 bp that exists within the 1356 bp ORF00411 (gene coordinates 314816 – 316171) and bears high sequence identity (80% (1063 bp/1326 bp)) to cytosine-5'-methyltransferase proteins of other bacteria. The highest scoring hits were also found in the related, *Rhodococcus RHA1*, *Mycobacteria, Streptomyces* and *Nocardi*a species. With strong prediction scores, these insert sequences are likely to encode an FtsK DNA translocase and a cytosine-5'-methyltransferase, respectively. The *R. equi* insert sequences found within clones 21 and 27 were identical or sibling (SIB) sequences of 146 bp within the 777 bp ORF05293 (gene coordinates 3888719 – 3889495) (Table 3). In BLASTp analyses, this SIB had strong similarity scores to hypothetical proteins found in various bacteria, including related actinomycetes species, which revealed the strongest similarity scores (data not shown). That the detection of SIBs was reproducible in a genome wide screen by Y2H format attests to the integrity of the screening process while alluding to the integrity of the interaction observed between VarA and the putative 777 bp hypothetical protein in yeast. Nevertheless, these interactions remain to be confirmed outside of yeast in a biochemical pull-down assay.

Table 1: Yeast two hybrid direct mating plasmids and results of direct mating experiments

Putative positive HK plasmids	Non-interacting plasmids	Plasmids not constructed	Autoactive plasmids
pIA15 (HK1)	pIA16 (HK2)	pIA17 (HK3)	pIA32 (HK18)
pIA30 (HK16)	pIA18 (HK4)	pIA21 (HK7)	
	pIA19 (HK5)	pIA27 (HK13)	
	pIA20 (HK6)		
	pIA22 (HK8)		
	pIA23 (HK9)		
	pIA24 (HK10)		
	pIA25 (HK11)		
	pIA26 (HK12)		
	pIA28 (HK14)		
	pIA29 (HK15)		
	pIA31 (HK17)		
	pIA33 (HK19)		
	pIA34 (HK20)		

Table 2: Qualitative results of yeast two hybrid library screen VarA-positive interactions.

key
+ Some growth observed
++ Significant growth observed
- No growth observed

	VarA-W		VarA-P		pLam5'-1	
	- 3-AT	**+ 3 mM 3-AT**	**- 3-AT**	**+ 3 mM 3-AT**	**- 3-AT**	**+ 3 mM 3-AT**
A24	+	-	-	-	-	-
B24	+	+	-	-	-	-
A27	++	+	-	-	-	-
B25	+	-	-	-	-	-
C21	+	+	+	-	-	-
C25	+	-	-	-	-	-
C27	+	-	-	-	-	-
C28	+	+	+	+	-	-
F21	+	-	+	+	-	-

Table 3: BLASTp analyses of five VarA-interacting *R. equi* library screen inserts.

* Sibling clones (Sibs) C21 and C27 are identical sequences that reproducibly interacted independently of one another.

Clone name	Length of clone insert	Length of entire ORF	Highest scoring matches and predicted protein (BLASTp)	% Identity	E-value
A24	251bp	1536bp	*Solibacter usitatus* NHL-repeat containing protein	(37%)	$2e^{-14}$
B24	353bp	2199bp	*Rhodococcus sp. RHA1* DNA translocase (FtsK)	(87%)	0.0
C21* C27*	146bp	777bp	*Rhodococcus sp. RHA1* Hypothetical protein	(86%)	$2e^{-22}$
C25	551bp	1356bp	*Rhodococcus sp. RHA1* Putative Cytosine –5'– methyltransferase	(80%)	9e-33

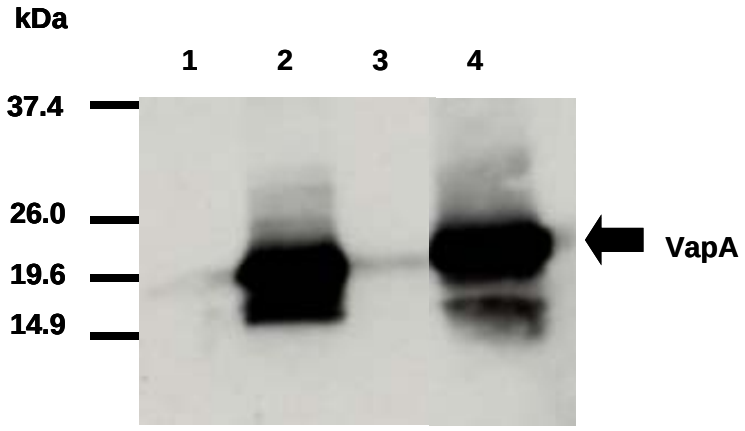


Figure 1

A *varA* mutant produces reduced amounts of VapA protein. Western blot showing significantly reduced amounts of VapA protein in the Δ*varA* mutant (lane 3). VapA is restored to wild type 103+ levels (lane 2) when a wild type copy of the mutated gene is provided (lane 4). The virulence plasmid-cured strain 103- served as a negative control (lane 1).

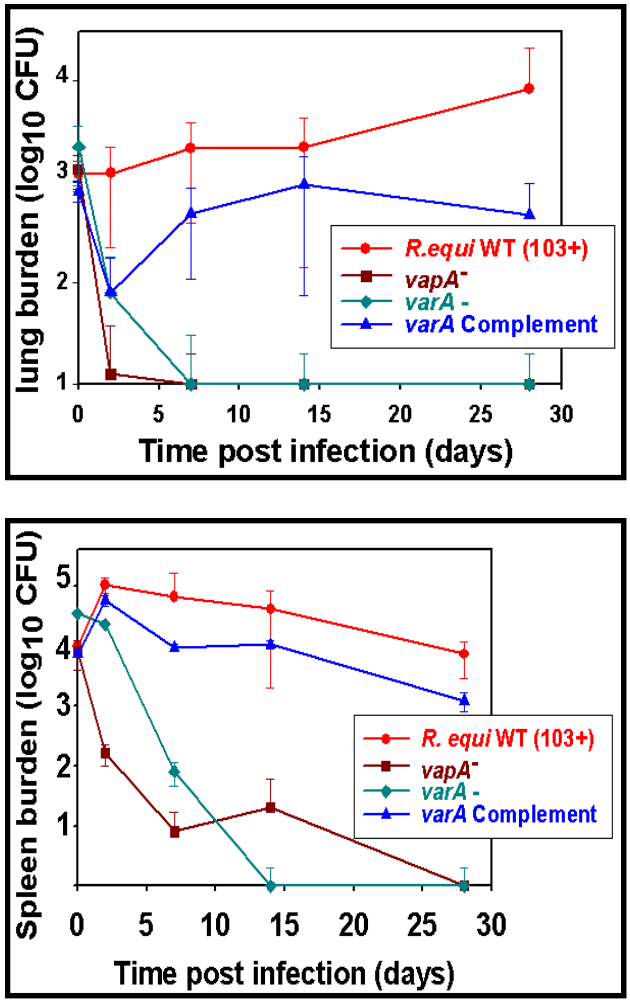


Figure 2

A *varA* deletion mutant is attenuated in a SCID mouse challenge model of *R. equi* infection. Severe Combined Immunodeficient (SCID) mice were challenged (i.v.) with 3 x 10^5 CFU of the indicated strain of *R. equi*. At various times post challenge, 4 mice per group were sacrificed and bacterial burdens were determined by serial dilution plating

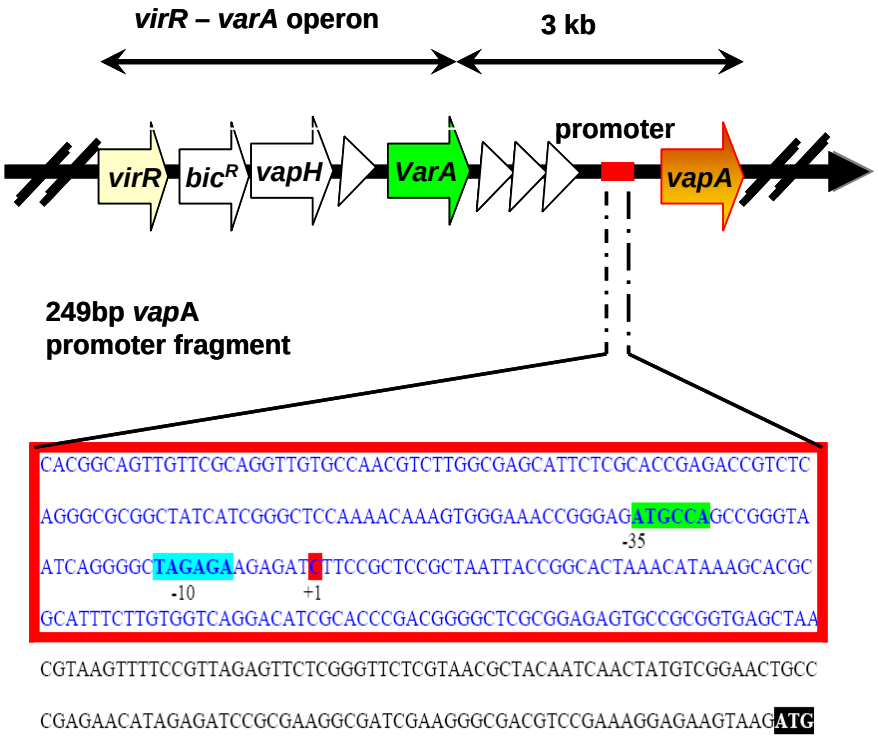


Figure 3

A Schematic diagram depicting the location of *varA* and *vapA* on the *R. equi* virulence plasmid. The -10 and -35 regions of the *vapA* promoter are indicated as is the transcriptional (+1) and translational start sites (highlighted ATG). The 249 bp promoter probe fragment used in EMSA experiments is boxed in red.

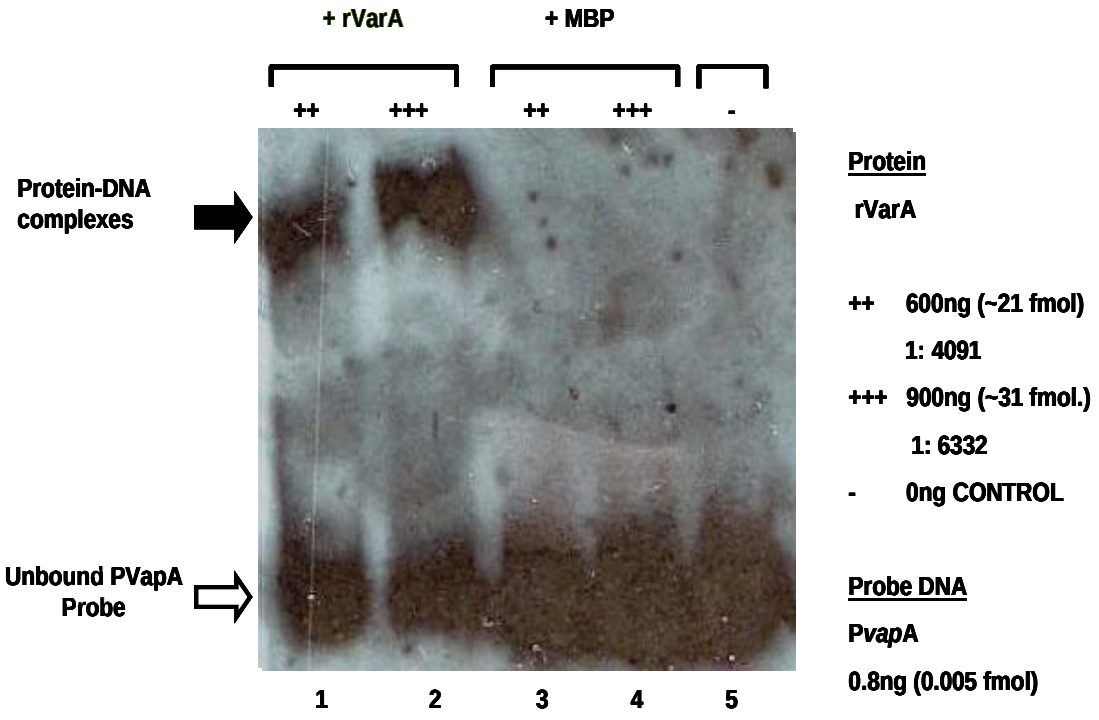


Figure 4

rVarA but not an irrelevant protein binds to the *vapA* promoter. A 249 bp digoxigenin-labeled DNA fragment (P*vapA*) encompassing the *vapA* promoter was incubated in the presence of varying amounts of rVarA (lanes 1 and 2) or purified MBP (lanes 3 and 4). Mobility in the complete absence of protein (lane 5) served as the negative (probe only) control. DNA : protein molar ratios are indicated to the right of the figure.

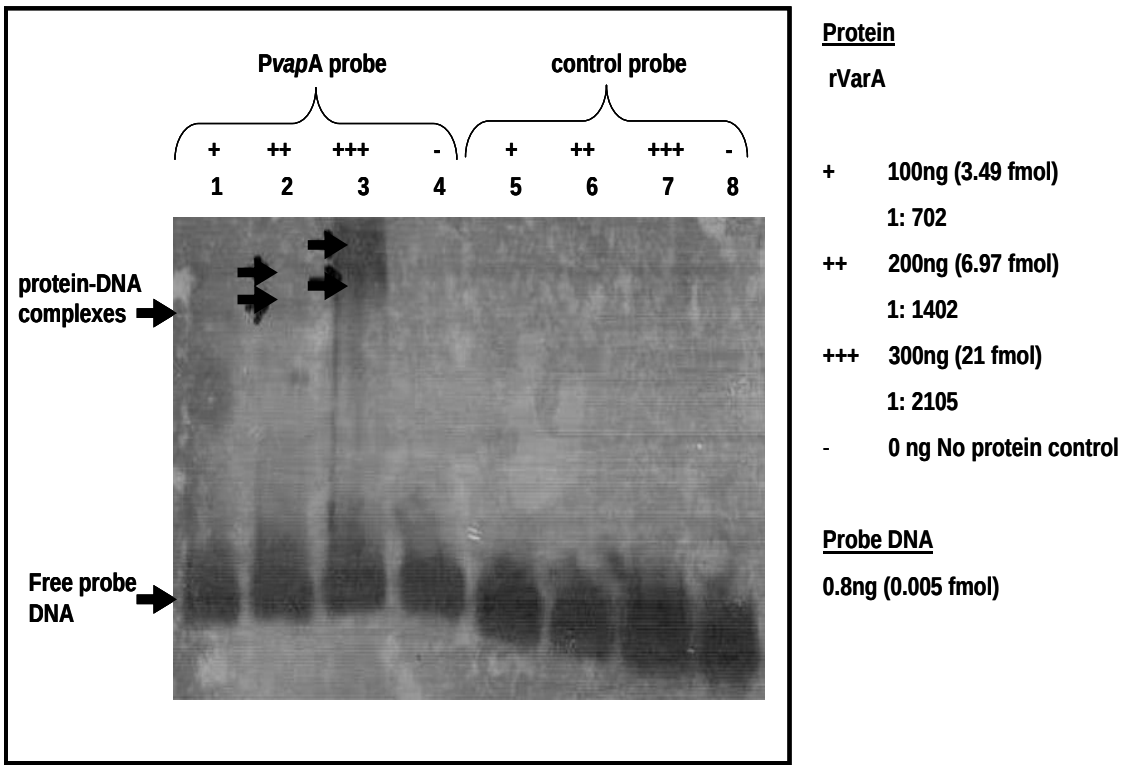


Figure 5

rVarA does not bind irrelevant DNA. Varying amounts of rVarA was incubated with a 249 bp digoxigenin labeled DNA fragment (P*vapA*) (lanes 1, 2, and 3) encompassing the *vapA* promoter or a similarly labeled 248 bp fragment of pUC19 DNA (lanes 5, 6 and 7). Migration the complete absence protein served as negative controls (lanes 4 and 8). DNA : protein molar ratios are indicated to the right of the figure.

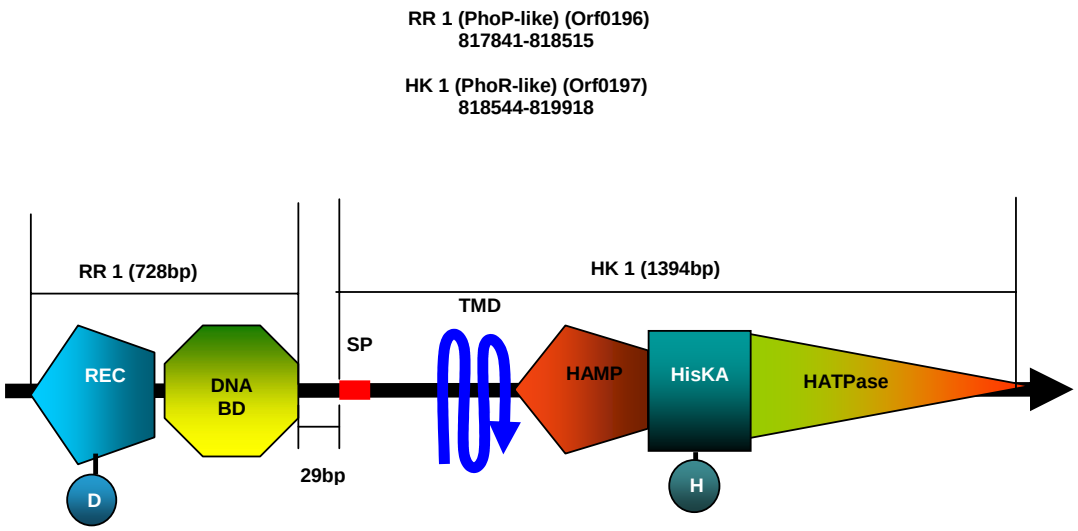


Figure 6A.

Schematic of putative two-component system 1 (TCS1): The following gene predictions have been selected using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Response regulator 1 (RR1) is represented as Orf01096 on the sense strand of the Sanger database *Rhodococcus equi* genomic sequence with the gene coordinates 817841 - 818515. The 242-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR1 bears significant sequence similarity to PhoP response regulators of other bacteria. Just 29 bp downstream of RR1 and in the same orientation is histidine kinase 1 (HK1) which is represented as Orf01097 on the Sanger database *R. equi* genomic sequence with the gene coordinates 818544 - 819918. The putative 464-aa protein is predicted to consist of a 27 bp signal peptide and a maximum of 4 transmembrane domain (TMD) helices at the N-terminus. This is followed by an intracellular sensing domain known as the HAMP for HKs, adenylyl cyclases, methyl binding proteins and phosphatases. A HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain are also predicted. The BLAST calculates that HK1 bears significant sequence similarity to PhoR HKs of other bacteria.

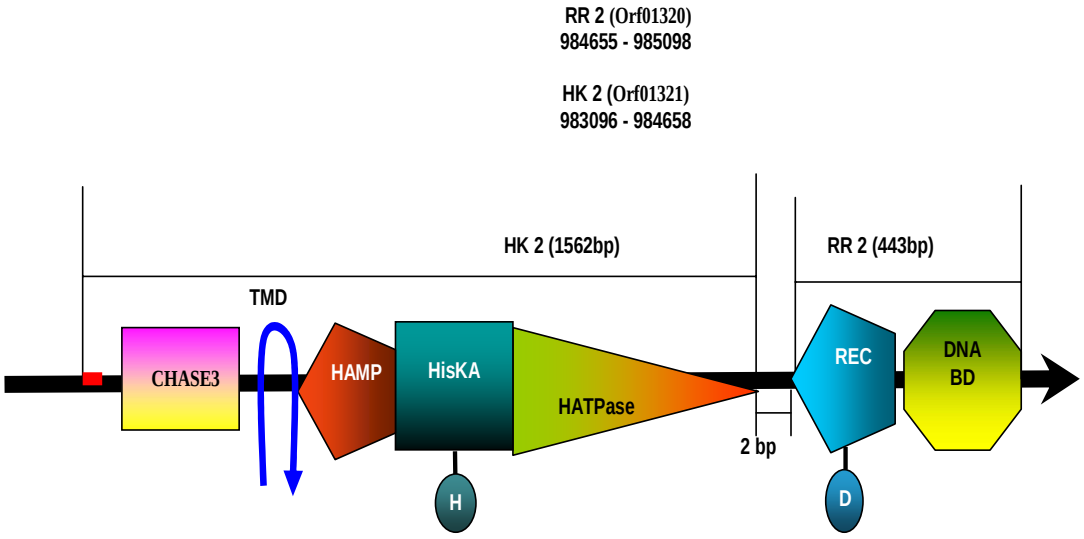


Figure 6B.

Schematic of putative two-component system 2 (TCS2): The following gene predictions have been selected using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Histidine kinase 2 (HK2) is represented as Orf01321 on the sense strand of the Sanger database *R. equi* genomic sequence with the gene coordinates 983096 - 984658. At the N- terminus, the putative 522-aa protein is predicted to consist of a 37 bp signal peptide (SP) and cyclase/histidine kinase-associated sensing extracellular (CHASE3) domain. Completing the extracellular portion of the molecule is a maximum of 2 transmembrane domain (TMD) helices. This is followed by an intracellular sensing domain known as the HAMP for HKs, adenylyl cyclases, methyl binding proteins and phosphatases. A HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain is also predicted. The BLAST calculates that HK2 bears significant sequence similarity to un-identified multisensor HKs of other bacteria. Just 2 bp downstream of HK2 and in the same orientation is response regulator 2 (RR2) which is represented as Orf01320 on the Sanger database *R. equi* genomic sequence 984655 - 985098. The 147-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC). The BLAST calculates that RR2 bears significant sequence similarity to response regulator genes of other bacteria.

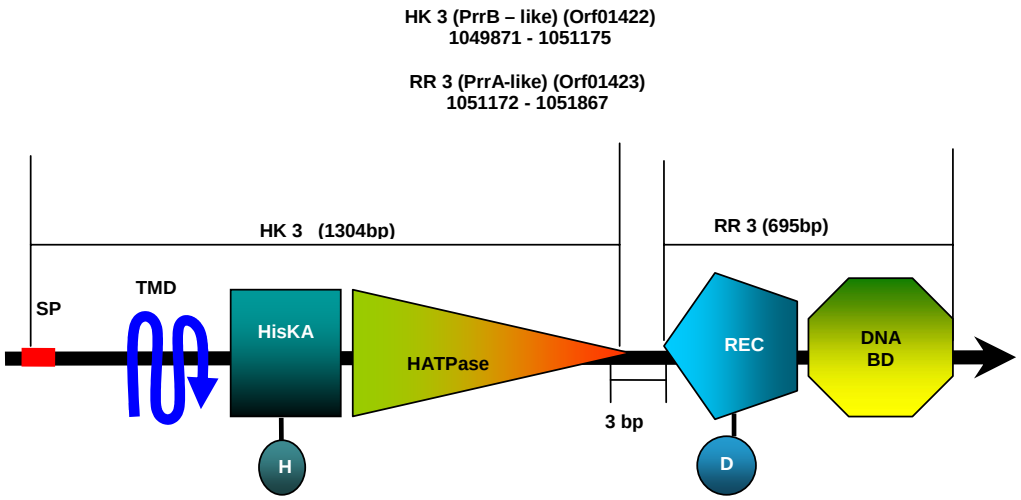


Figure 6C.

Schematic of putative two-component system 3 (TCS3): The following gene predictions have been selected using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Histidine kinase 3 (HK3) is represented as Orf01422 on the antisense strand of the Sanger database *R. equi* genomic sequence with the gene coordinates 1049871 - 1051175. The putative 464-aa protein is predicted to consist of a maximum of 4 transmembrane domain (TMD) helices at the N-terminus. A HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain is also predicted. The BLAST calculates that HK3 bears significant sequence similarity to PrrB histidine kinases of other bacteria. Just 3 bp downstream of HK3 and in the same orientation is response regulator 3 (RR3) which is represented as Orf01423 on the Sanger database *R. equi* genomic sequence with the gene coordinates 1051172 – 1051867. The 231-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR3 bears significant sequence similarity to response PrrA regulator genes of other bacteria.

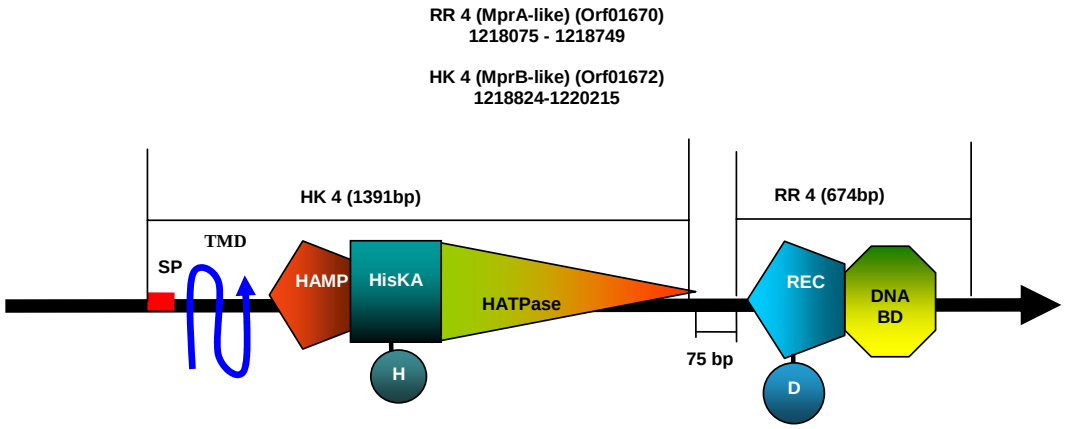


Figure 6D.

Schematic of putative two-component system 4 (TCS4): The following gene predictions have been selected using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Histidine kinase 4 (HK4) which is represented as Orf01672 on the sense strand of the Sanger database *Rhodococcus equi* genomic sequence with the gene coordinates 1218824 - 1220215. The putative 450-aa protein is predicted to consist of a 31 bp signal peptide and a maximum of 3 transmembrane domain (TMD) helices at the N-terminus. This is followed by an intracellular sensing domain known as the HAMP for HKs, adenylyl cyclases, methyl binding proteins and phosphatases. A HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain are also predicted. The BLAST calculates that HK4 bears significant sequence similarity to MprB HKs of other bacteria. Just 75 bp downstream of HK4 and in the same orientation is response regulator 4 (RR4) is represented as Orf01670 on the Sanger database *R. equi* genomic sequence with the gene coordinates 1218075 – 1218749. The 224-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR4 bears significant sequence similarity to mprA response regulators of other bacteria.

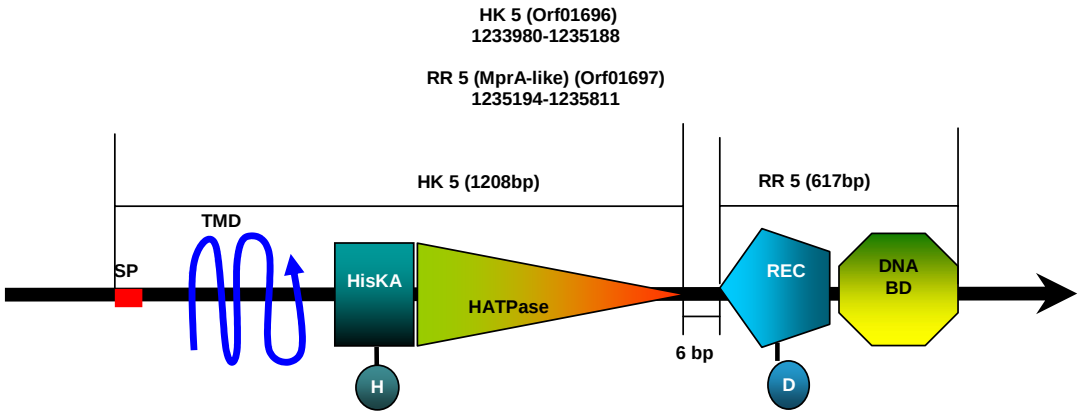


Figure 6E.

Schematic of putative two-component system 5 (TCS5): The following gene predictions have been selected using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Histidine kinase 5 (HK5) is represented as Orf01696 on the sense strand of the Sanger database *R. equi* genomic sequence with the gene coordinates 1233980 - 1235188 on. The putative 402-aa protein is predicted to consist of a 36 bp signal peptide and a maximum of 5 transmembrane domain (TMD) helices at the N-terminus. A HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain is also predicted. The BLAST calculates that HK5 bears significant sequence similarity to HKs of other bacteria. Just 6 bp downstream of HK5 and in the same orientation is response regulator 5 (RR5) which is represented as Orf01697 on the Sanger database *R. equi* genomic sequence with the gene coordinates 1235194 - 1235811. The 224-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR5 bears significant sequence similarity to response MprA regulator genes of other bacteria.

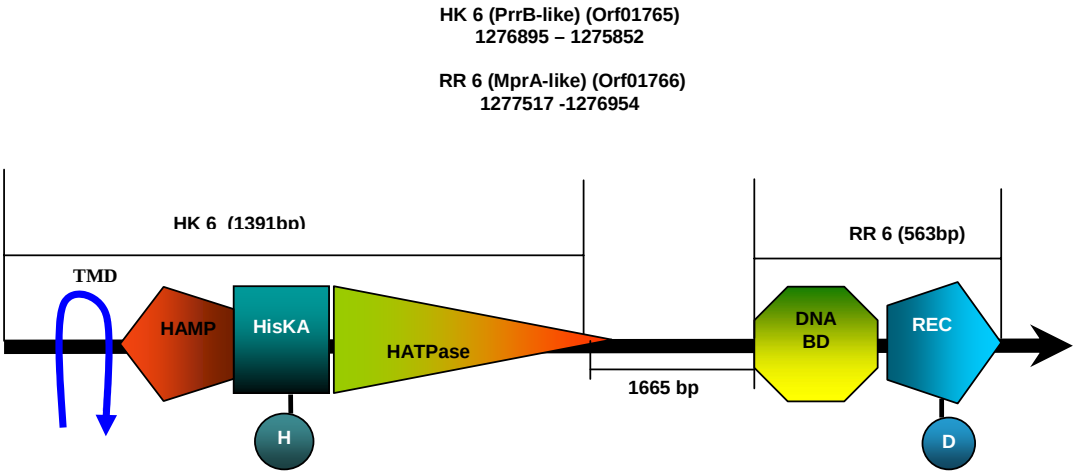


Figure 6F.

Schematic of putative two-component system 6 (TCS6): The following gene predictions have been selected using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Histidine kinase 6 (HK6) is represented as Orf01765 on the antisense strand of the Sanger database *R. equi* genomic sequence with the gene coordinates 1276895 – 1275852. The putative 347-aa protein is predicted to consist of a maximum of 2 transmembrane domain (TMD) helices at the N-terminus. This is followed by an intracellular sensing domain known as the HAMP for HKs, adenylyl cyclases, methyl binding proteins and phosphatases. A HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain is also predicted. The BLAST calculates that HK6 bears significant sequence similarity to PrrB HKs of other bacteria. At 1665 bp downstream of HK6 and in the same orientation is response regulator 6 (RR6) which is represented as Orf01766 on the Sanger database *R. equi* genomic sequence with the gene coordinates 1277517 - 276954. The 187-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR6 bears significant sequence similarity to MprA response regulator genes of other bacteria.

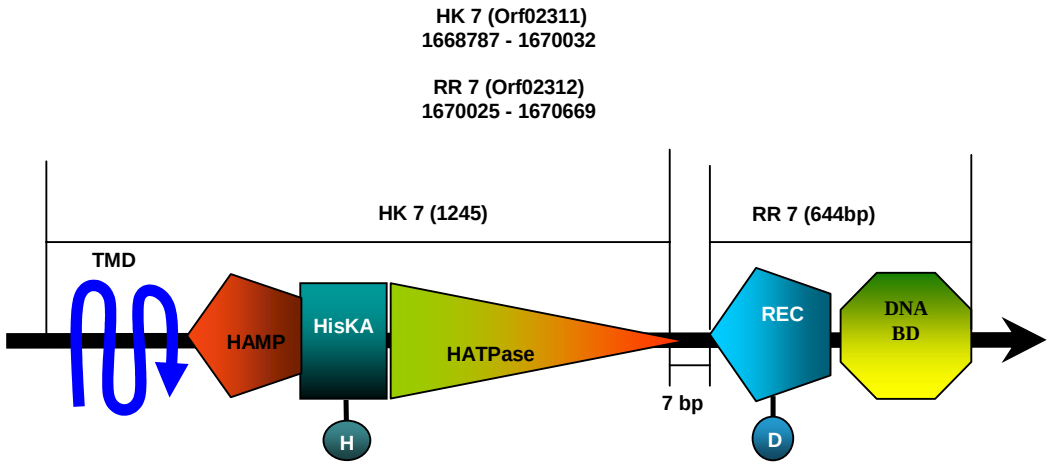


Figure 6G.

Schematic of putative two-component system 7 (TCS7): The following gene predictions have been selected using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Histidine kinase 7 (HK7) is represented as Orf02311 on the sense strand of the Sanger database *R. equi* genomic sequence with the gene coordinates 1668787 – 1670032. The putative 347-aa protein is predicted to consist of a maximum of 4 transmembrane domain (TMD) helices at the N-terminus. This is followed by an intracellular sensing domain known as the HAMP for HKs, adenylyl cyclases, methyl binding proteins and phosphatases. A HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain is also predicted. The BLAST calculates that HK7 bears significant sequence similarity to HKs of other bacteria. Just 7-bp downstream of HK7 and in the same orientation is response regulator 7 (RR7) which is represented as Orf02312 on the Sanger database *R. equi* genomic sequence with the gene coordinates 1670025 - 1670669. The 224-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR7 bears significant sequence similarity to response regulator genes of other bacteria.

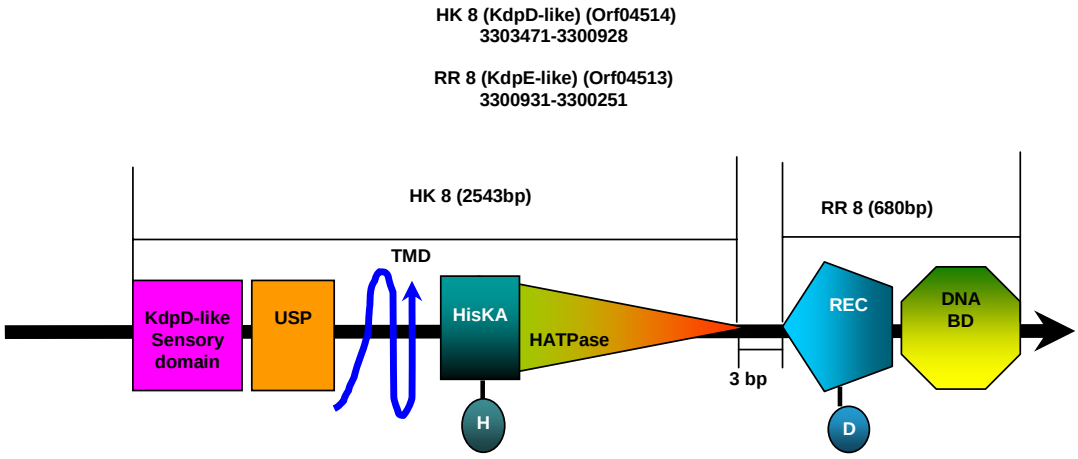


Figure 6H.

Schematic of putative two-component system 8 (TCS8): The following gene predictions have been settled upon using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Histidine kinase 8 (HK8) is represented as Orf04514 on the sense strand of the Sanger database *R. equi* genomic sequence with the gene coordinates 3303471 - 3300928 on the antisense strand of the *R. equi* genome. The putative 847-aa protein is predicted to consist of a KdpD-like sensory domain which is an osmosensitive K^+ channel that is present in the KdpD HK. A universal stress protein (USP) domain which is thought to respond to a wide variety of stressors that damage DNA is also detected. Also at the N-terminus is a maximum of 3 predicted transmembrane domain (TMD) helices. This is followed by a HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain. KdpD HKs always contain a USP domain in addition to the conserved KdpD sensor domain, HisKA and HATPase domains. The BLAST calculates that HK8 bears significant sequence similarity to KdpD HKs of other bacteria. Just 3 bp downstream of HK8 and in the same orientation is response regulator 8 (RR8) which is represented as Orf04513 on the sense strand of the Sanger database *R. equi* genomic sequence with the gene coordinates 3300931 - 3300251. The 226-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR8 bears significant sequence similarity to KdpE response regulator genes of other bacteria.

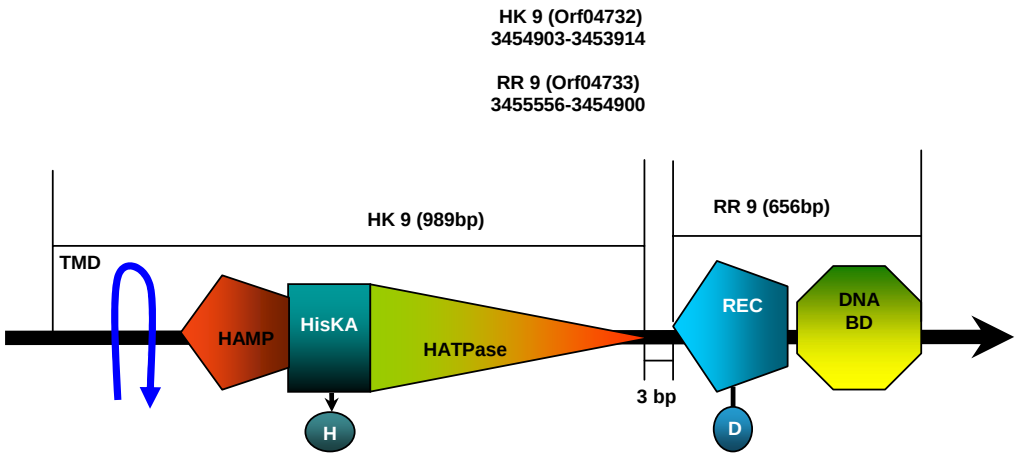


Figure 6I.

Schematic of putative two-component system 9 (TCS9): The following gene predictions have been selected using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Histidine kinase 9 (HK9) is represented as Orf04732 on the antisense strand of the Sanger database *R. equi* genomic sequence with the gene coordinates 3454903-3453914. The putative 349-aa protein is predicted to consist of a maximum of 2 transmembrane domain (TMD) helices at the N-terminus. This is followed by an intracellular sensing domain known as the HAMP for HKs, adenylyl cyclases, methyl binding proteins and phosphatases. A HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain is also predicted. The BLAST calculates that HK9 bears significant sequence similarity to HKs of other bacteria. Just 3 bp downstream of HK9 and in the same orientation is response regulator 9 (RR9) which is represented as Orf04733 on the Sanger database *R. equi* genomic sequence with the gene coordinates 3455556-3454900. The 218-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR9 bears significant sequence similarity to response regulator genes of other bacteria.

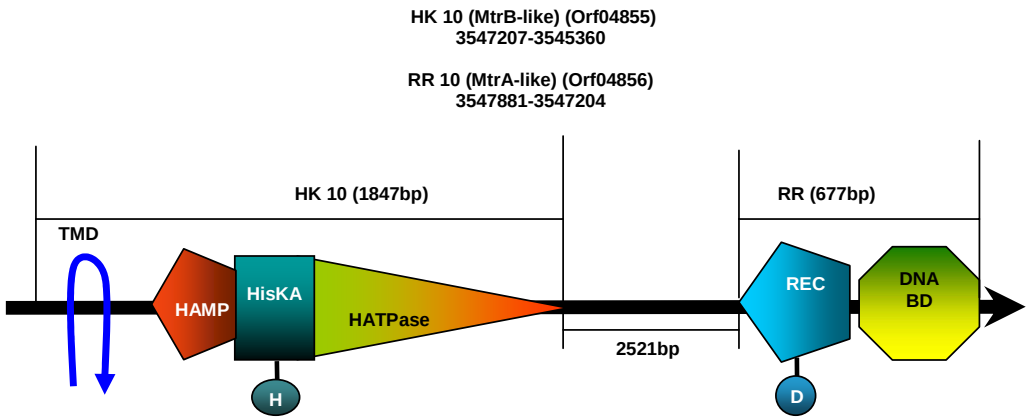


Figure 6J

Schematic of putative two-component system 10 (TCS10): The following gene predictions have been selected using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Histidine kinase 10 (HK10) is represented as Orf04855 on the antisense strand of the Sanger database *R. equi* genomic sequence with the gene coordinates 3547207 - 3545360. The putative 615-aa protein is predicted to consist of a maximum of 2 transmembrane domain (TMD) helices at the N-terminus. This is followed by an intracellular sensing domain known as the HAMP for HKs, adenylyl cyclases, methyl binding proteins and phosphatases. A HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain is also predicted. The BLAST calculates that HK10 bears significant sequence similarity to MtrB HKs of other bacteria. At 2521 bp downstream of HK10 and in the same orientation is response regulator 10 (RR10) which is represented as Orf04856 on the Sanger database *R. equi* genomic sequence with the gene coordinates 3547881-3547204. The 225-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR10 bears significant sequence similarity to MtrA response regulator genes of other bacteria.

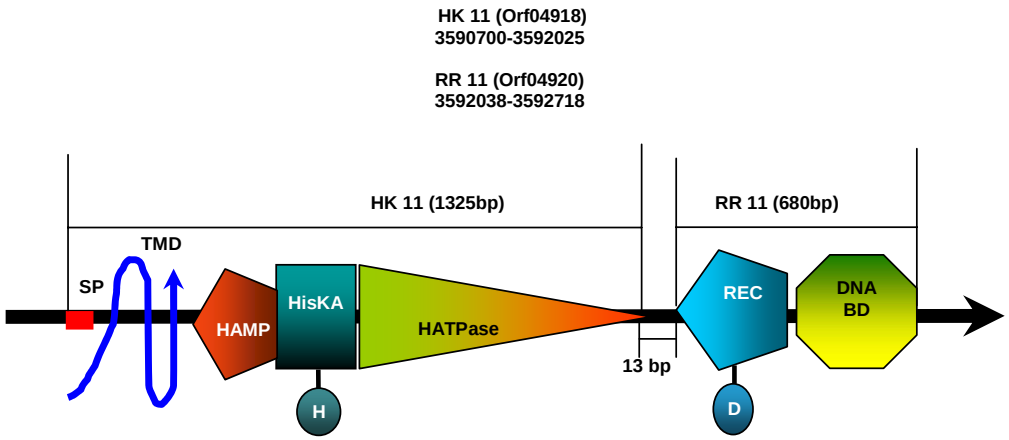


Figure 6K.

Schematic of putative two-component system 11 (TCS11): The following gene predictions have been selected using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Histidine kinase 11 (HK11) is represented as Orf04918 on the antisense strand of the Sanger database *R. equi* genomic sequence with the gene coordinates 3590700 - 3592025. The putative 427-aa protein is predicted to consist of a 21 bp signal peptide (SP) and a maximum of 3 transmembrane domain (TMD) helices at the N-terminus. This is followed by an intracellular sensing domain known as the HAMP for HKs, adenylyl cyclases, methyl binding proteins and phosphatases. A HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain is also predicted. The BLAST calculates that HK11 bears significant sequence similarity to HKs of other bacteria. Just 13 bp downstream of HK11 and in the same orientation is response regulator 11 (RR11) which is represented as Orf04920 on the Sanger database *R. equi* genomic sequence with the gene coordinates 3592038 - 3592718. The 226-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR11 bears significant sequence similarity to response regulator genes of other bacteria.

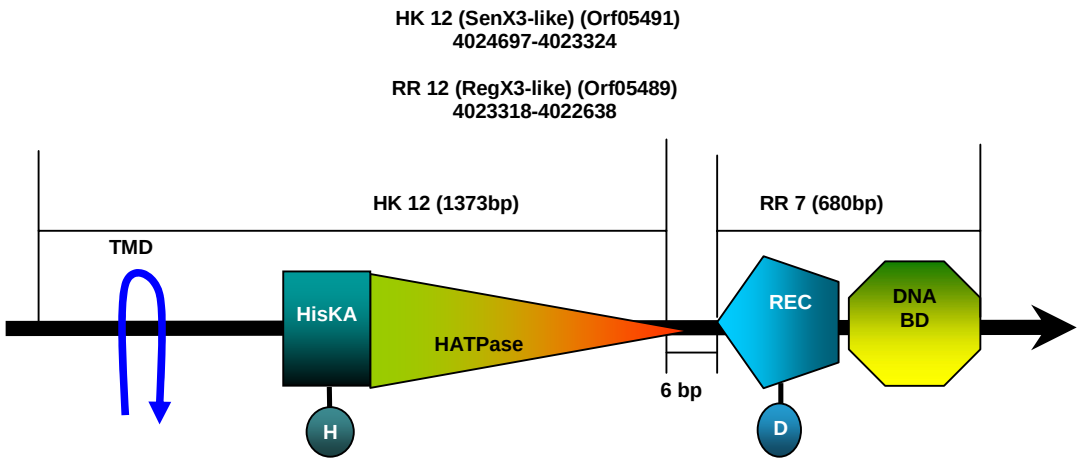


Figure 6L.

Schematic of putative two-component system 12 (TCS12): The following gene predictions have been selected using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Histidine kinase 12 (HK12) is represented as Orf05491 on the antisense strand of the Sanger database *R. equi* genomic sequence with the gene coordinates 4024697 - 4023324. The putative 457-aa protein is predicted to consist of a maximum of 2 transmembrane domain (TMD) helices at the N-terminus. This is followed by a HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain is also predicted. The BLAST calculates that HK12 bears significant sequence similarity to SenX3 HKs of other bacteria. Just 6 bp downstream of HK12 and in the same orientation is response regulator 12 (RR12) which is represented as Orf05489 on the Sanger database *R. equi* genomic sequence with the gene coordinates 4023318 - 4022638. The 226-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR12 bears significant sequence similarity to RegX3 response regulator genes of other bacteria.

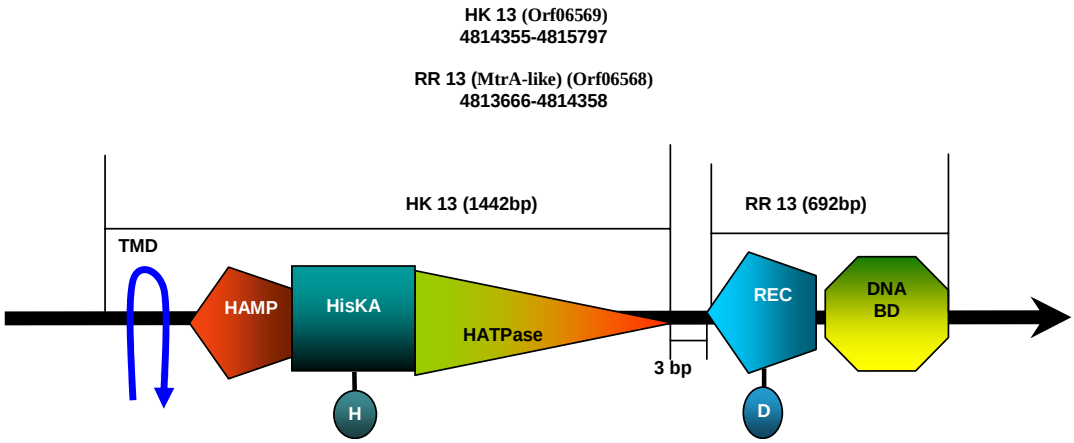


Figure 6M.

Schematic of putative two-component system 13 (TCS13): The following gene predictions have been selected using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Histidine kinase 13 (HK13) is represented as Orf06569 on the sense strand of the Sanger database *R. equi* genomic sequence with the gene coordinates 4814355-4815797. The putative 525-aa protein is predicted to consist of a maximum of 2 transmembrane domain (TMD) helices at the N-terminus. This is followed by an intracellular sensing domain known as the HAMP for HKs, adenylyl cyclases, methyl binding proteins and phosphatases. A HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain is also predicted. The BLAST calculates that HK13 bears significant sequence similarity to HKs of other bacteria. Just 3 bp downstream of HK13 and in the same orientation is response regulator 13 (RR13) which is represented as Orf06568 on the Sanger database *R. equi* genomic sequence with the gene coordinates 4813666-4814358. The 226-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR13 bears significant sequence similarity to MtrA response regulator genes of other bacteria.

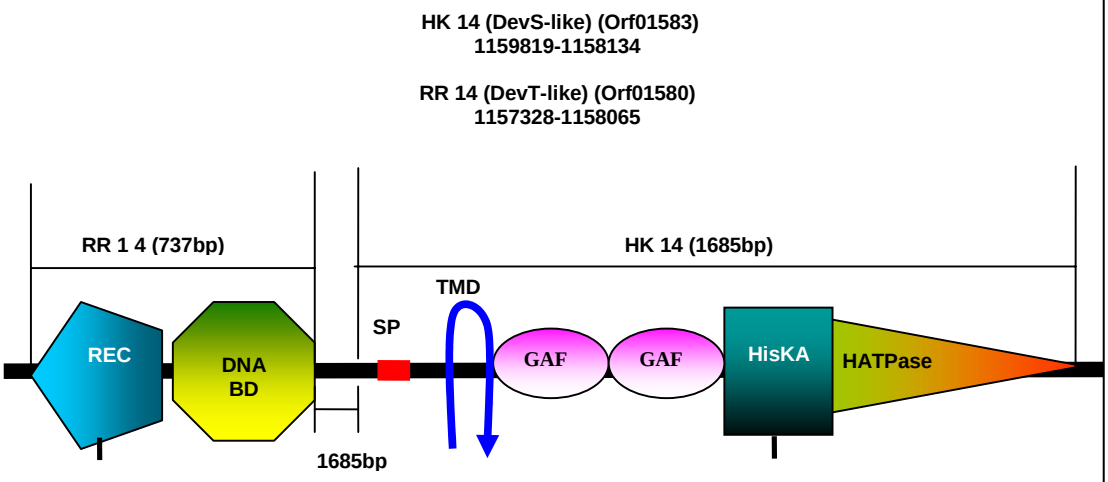


Figure 6N.

Schematic of putative two-component system 14 (TCS14): The following gene predictions have been selected using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Histidine kinase 14 (HK14) is represented as Orf01583 on the antisense strand of the Sanger database *R. equi* genomic sequence with the gene coordinates 1159819 - 1158134. The putative 619-aa protein is predicted to consist of a maximum of 2 transmembrane domain (TMD) helices at the N-terminus. This is followed by 2 conserved intracellular sensing domains known as the GAF which are involved in signaling and are found in hundreds of other signaling and sensory proteins from all three kingdoms of life. A HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain is also predicted. The BLAST calculates that HK14 bears significant sequence similarity to DevS HKs of other bacteria. At 1685 bp upstream of HK14 and in the same orientation is response regulator 14 (RR14) which is represented as Orf01580 on the Sanger database *R. equi* genomic sequence with the gene coordinates 1157328 - 1158065. The 245-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR14 bears significant sequence similarity to DevT response regulator genes of other bacteria.

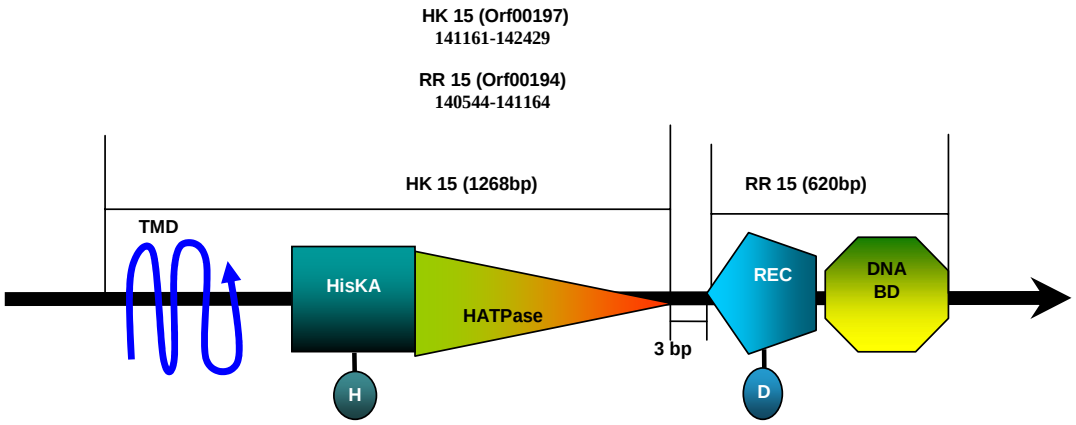


Figure 6O.

Schematic of putative two-component system 15 (TCS15): The following gene predictions have been selected using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Histidine kinase 15 (HK15) is represented as Orf00197 on the antisense strand of the Sanger database *R. equi* genomic sequence with the gene coordinates 141161 - 142429. The putative 608-aa protein is predicted to consist of a maximum of 5 transmembrane domain (TMD) helices at the N-terminus. This is followed by a HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain. The BLAST calculates that HK15 bears significant sequence similarity to HKs of other bacteria. Just 3 bp downstream of HK15 and in the same orientation is response regulator 15 (RR15) which is represented as Orf00194 on the Sanger database *R. equi* genomic sequence with the gene coordinates 140544 - 141164. The 206-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR15 bears significant sequence similarity to response regulator genes of other bacteria.

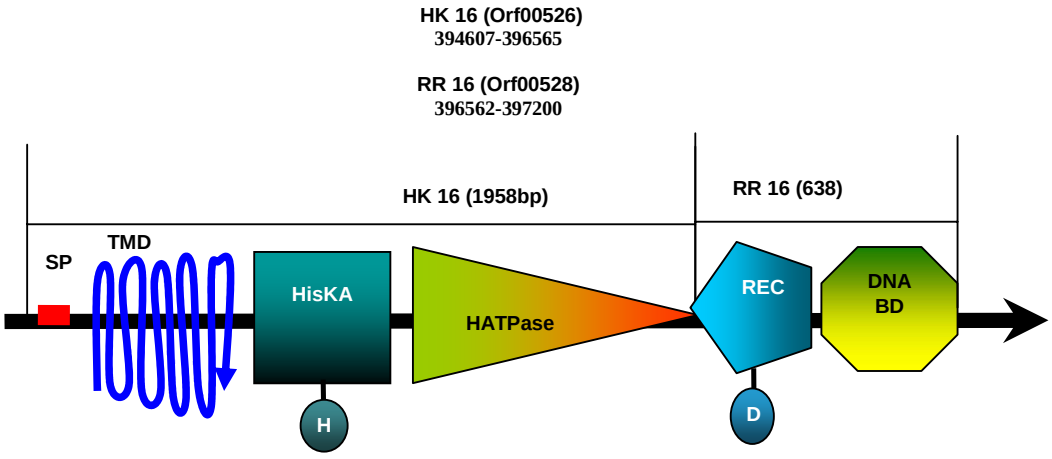


Figure 6P.

Schematic of putative two-component system 16 (TCS16): The following gene predictions have been selected using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Histidine kinase 16 (HK16) is represented as on the sense strand of the Sanger database *R. equi* genomic sequence with the gene coordinates 394607 - 396565. The putative 651-aa protein is predicted to consist of a 35 bp signal peptide (SP) and a maximum of 10 transmembrane domain (TMD) helices at the N-terminus. This is followed by a HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain. The BLAST calculates that HK16 bears significant sequence similarity to HKs of other bacteria. In the same orientation and with a start codon that is predicted to overlap the stop codon of HK16 is response regulator 16 (RR16) which is represented as Orf00528 on the Sanger database *R. equi* genomic sequence with the gene coordinates 396562 - 397200. The 212-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR16 bears significant sequence similarity to response regulator genes of other bacteria.

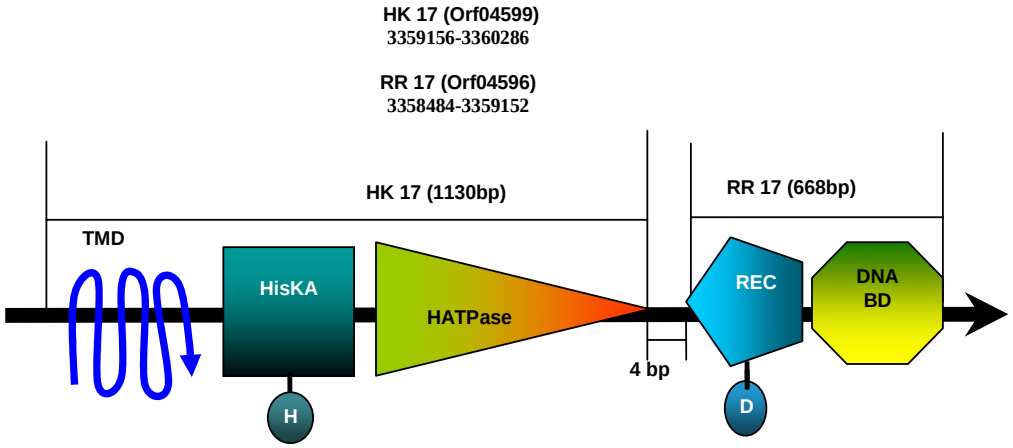


Figure 6Q.

Schematic of putative two-component system 17 (TCS17): The following gene predictions have been selected using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Histidine kinase 17 (HK17) is represented as Orf04599 on the Sanger *Rhodococcus equi* database with the gene coordinates 3359156 - 3360286 on the antisense strand of the *R. equi* genome. The putative 498-aa protein is predicted to consist of a maximum of 6 transmembrane domain (TMD) helices at the N-terminus. This is followed by a HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase). The BLAST calculates that HK17 bears significant sequence similarity to HKs of other bacteria. Just 4 bp downstream of HK17 and in the same orientation is response regulator 17 (RR17) which is represented as Orf04596 on the Sanger database *R. equi* genomic sequence with the gene coordinates 3358484 - 3359152. The 222-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR17 bears significant sequence similarity to response regulator genes of other bacteria.

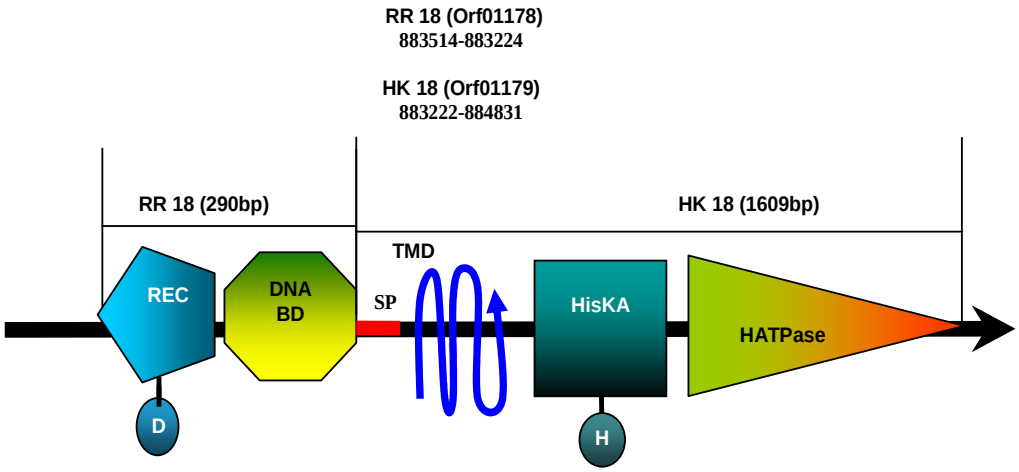


Figure 6R.

Schematic of putative two-component system 18 (TCS18): The following gene predictions have been selected using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Response regulator 18 (RR18) is represented as Orf01178 on the antisense strand of the Sanger database *R. equi* genomic sequence with the gene coordinates 883514 - 883224. The 222-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR18 bears significant sequence similarity to response regulator genes of other bacteria. In the same orientation as RR18 is histidine kinase 18 (HK18) which is represented as Orf01179 on the Sanger database *R. equi* genomic sequence with the gene coordinates 883222 - 884831. The putative 532-aa protein is predicted to consist of a maximum of 5 transmembrane domain (TMD) helices at the N-terminus. This is followed by a HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain. The BLAST calculates that HK18 bears significant sequence similarity to HKs of other bacteria.

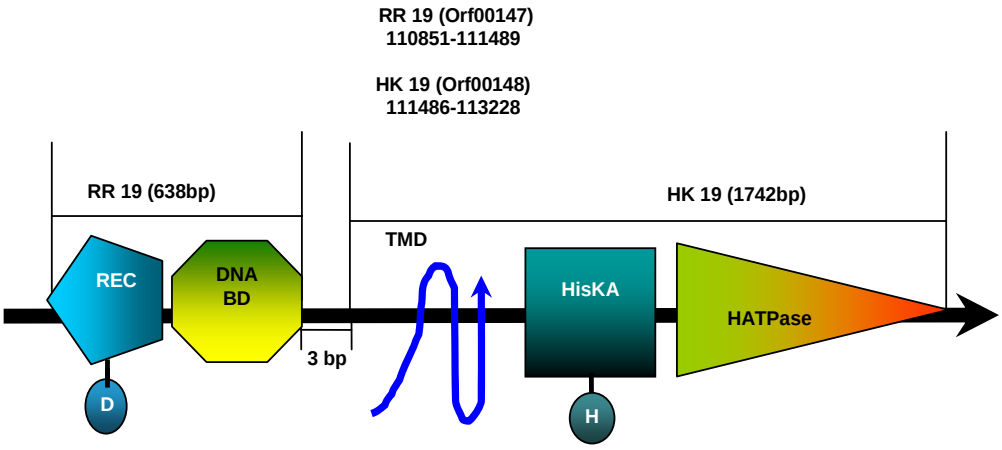


Figure 6S.

Schematic of putative two-component system 19 (TCS19): The following gene predictions have been selected using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Response regulator 19 (RR19) which is represented as Orf00147 on the sense strand of the Sanger database *R. equi* genomic sequence with the gene coordinates 110851 - 111489. The 212-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR19 bears significant sequence similarity to response regulator genes of other bacteria. Just 3 bp downstream and in the same orientation as RR19 is histidine kinase 19 (HK19) which is represented as Orf01148 on the Sanger database *R. equi* genomic sequence with the gene coordinates 111486 - 113228. The putative 370-aa protein is predicted to consist of a maximum of 3 transmembrane domain (TMD) helices at the N-terminus. This is followed by a HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain. The BLAST calculates that HK19 bears significant sequence similarity to HKs of other bacteria.

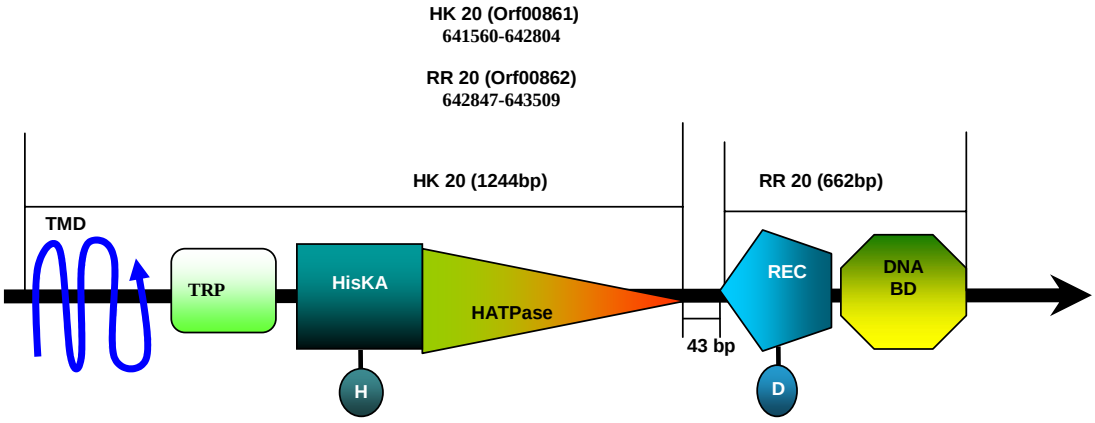


Figure 6T.

Schematic of putative two-component system 20 (TCS20): The following gene predictions have been settled upon using a combination of the Simple Modular Architectural Research Tool (SMART), the Basic Local Alignment Search Tool (BLAST), a TM pred program used in the prediction of transmembrane domain locations. Histidine kinase 20 (HK20) is represented as Orf00861 on the sense strand of the Sanger database *R. equi* genomic sequence with the gene coordinates 641560-642804. The putative 413-aa protein is predicted to consist of a maximum of 5 transmembrane domain (TMD) helices at the N-terminus. A domain that is found in tryptophan (TRP) associated membrane proteins of actinomycetes such as *Mycobacteria*, *Streptomyces* and *Corynebacteria* is also detected. This is followed by a HK phosphoacceptor (HisKA) domain and a HK ATPase (HATPase) domain. The BLAST calculates that HK20 bears significant sequence similarity to HKs of other bacteria. Just 43 bp downstream of HK20 and in the same orientation is response regulator 20 (RR20) which is represented as Orf00862 on the Sanger database *R. equi* genomic sequence with the gene coordinates 642847 - 643509. The 220-aa putative protein is predicted to be composed of an N-terminal receiver domain (REC) and a C-terminal DNA-binding domain (DNA BD) complete with helix-turn-helix motifs. The BLAST calculates that RR20 bears significant sequence similarity to response regulator genes of other bacteria.

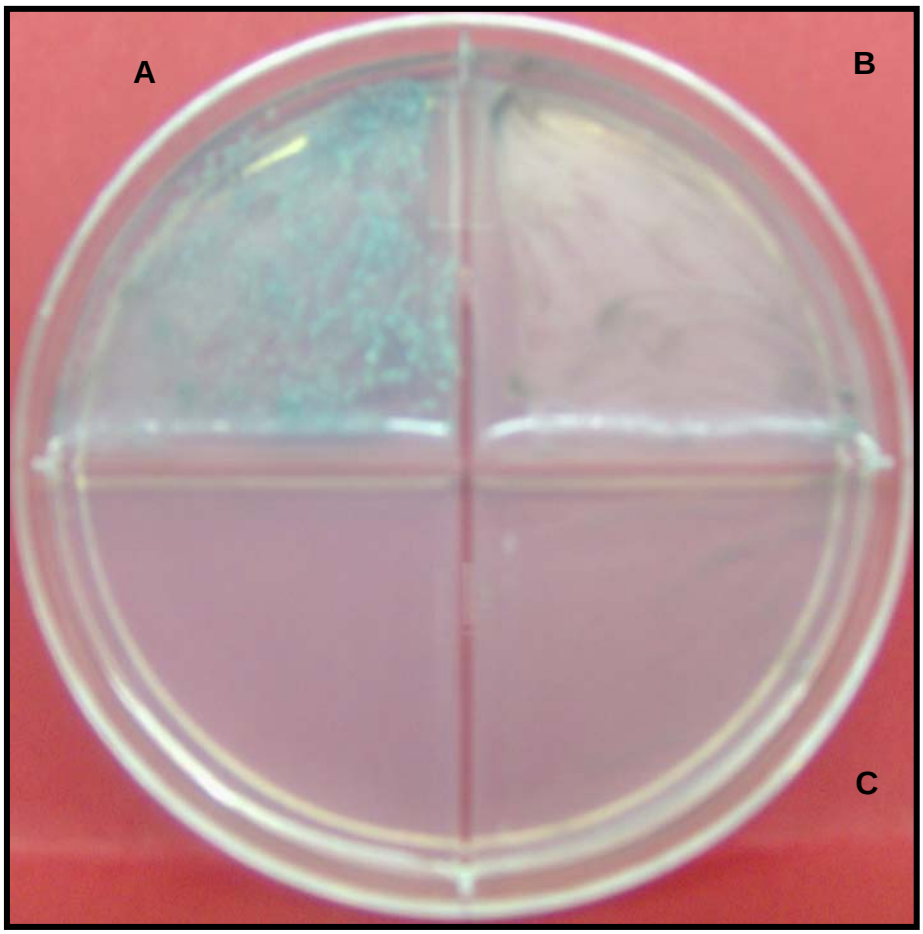


Figure 7

HK1 interacts with VarA in yeast. Matα yeast strain transformed with HK1 mated with MatA yeast transformed with VarA-W and plated on high-stringency selective indicator medium (SD –Ade –His –Leu –Trp + X α-gal) (A). Auto-activation controls Matα (HK1) mated with MatA-pGADT7 (B) and MatA–VarA-W mated with Mat α -pLAM5'-1 (C).

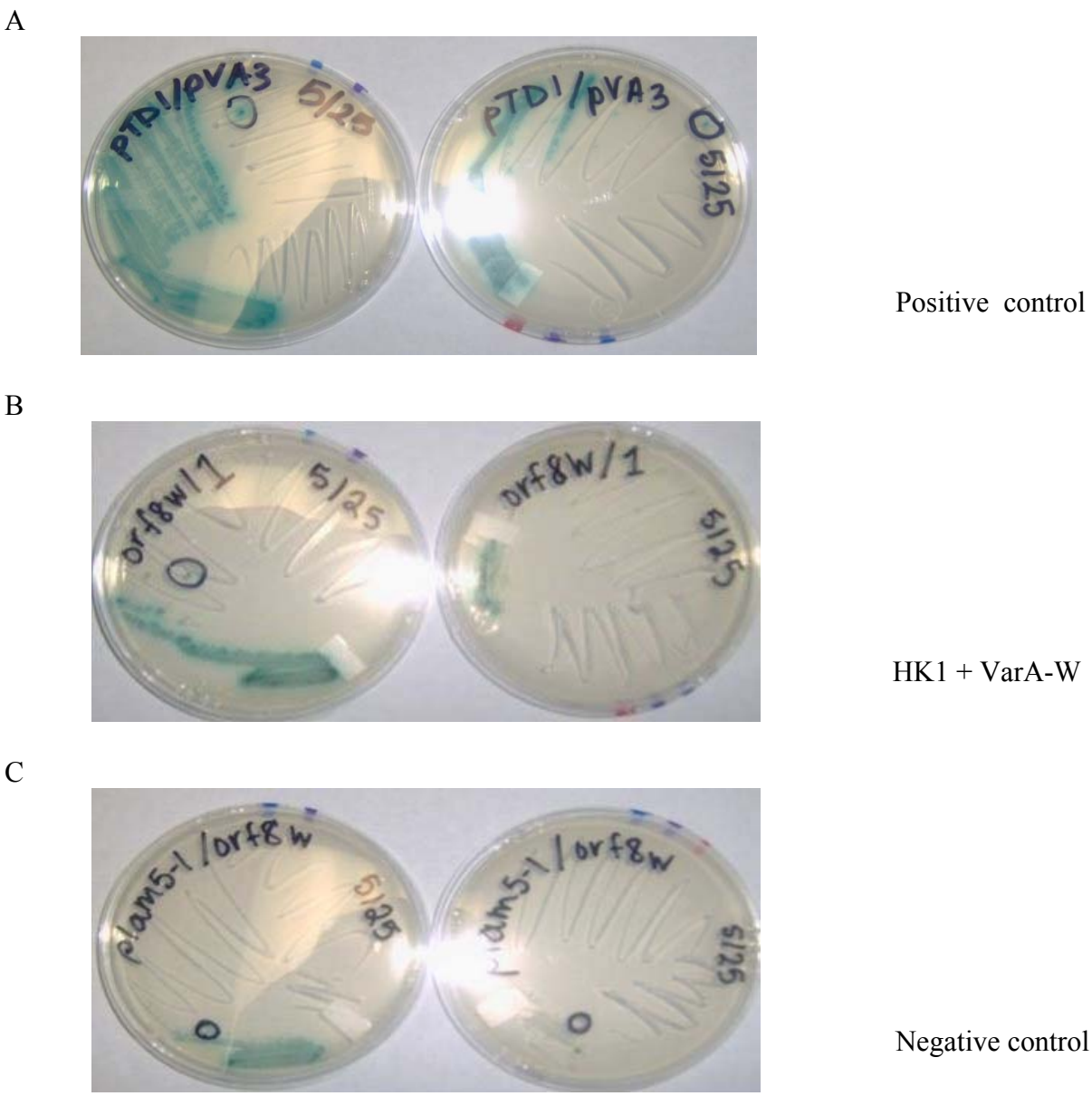


Figure 8

Analysis of HK1/VarA-W interaction under higher stringency. pTD1/pVA3 positive control mating (A) VarA-W/HK1 mating (B) and pLAM5'-1/VarA-W mating (C) was plated on high-stringency selective indicator media (SD-ADE-HIS-LEU-TRP + X α-gal in the absence (-) or presence (+) of 3 mM 3-aminotriazole (3-AT). Growth and blue color represents Gal4 reporter gene activity resulting from an interaction between mating pairs. The addition of 3-AT (toxic *HIS3* gene inhibitor) was used to eliminate nonspecific activation of Gal4-responsive genes as seen in the pLAM5'-1/VarA-W negative control mating. Like the positive control mating pTD1/pVA3, HK1/VarA-W can overcome histidine starvation induced by 3-AT which decreases expression of the *HIS3* gene which is known to be leaky.

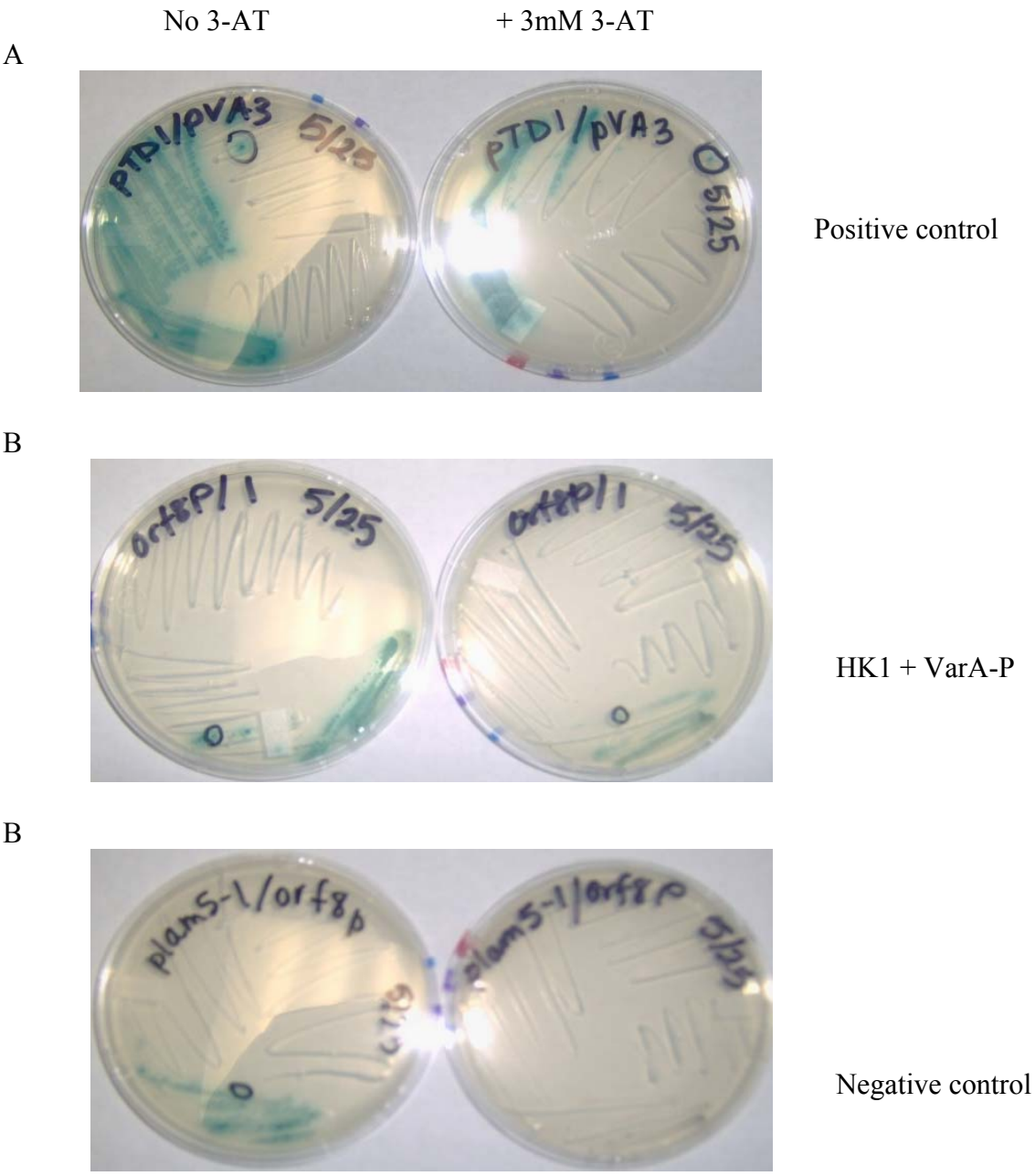


Figure 9

Analysis of HK1/VarA-P interaction under higher stringency. pTD1/pVA3 positive control mating (A) VarA-P/HK1 mating (B) and pLAM5'-1/VarA-P mating (C) was plated on high-stringency selective indicator media (SD-ADE-HIS-LEU-TRP + X α-gal in the absence (-) or presence (+) of 3 mM 3-aminotriazole (3-AT). Growth and blue color represents Gal4 reporter gene activity resulting from an interaction between mating pairs. The addition of 3-AT (toxic *HIS3* gene inhibitor) was used to eliminate nonspecific activation of Gal4-responsive genes as seen in the pLAM5'-1/VarA-P negative control mating. Like the positive control mating pTD1/pVA3, HK1/VarA-P can overcome histidine starvation induced by 3-AT which decreases expression of the *HIS3* gene which is known to be leaky.

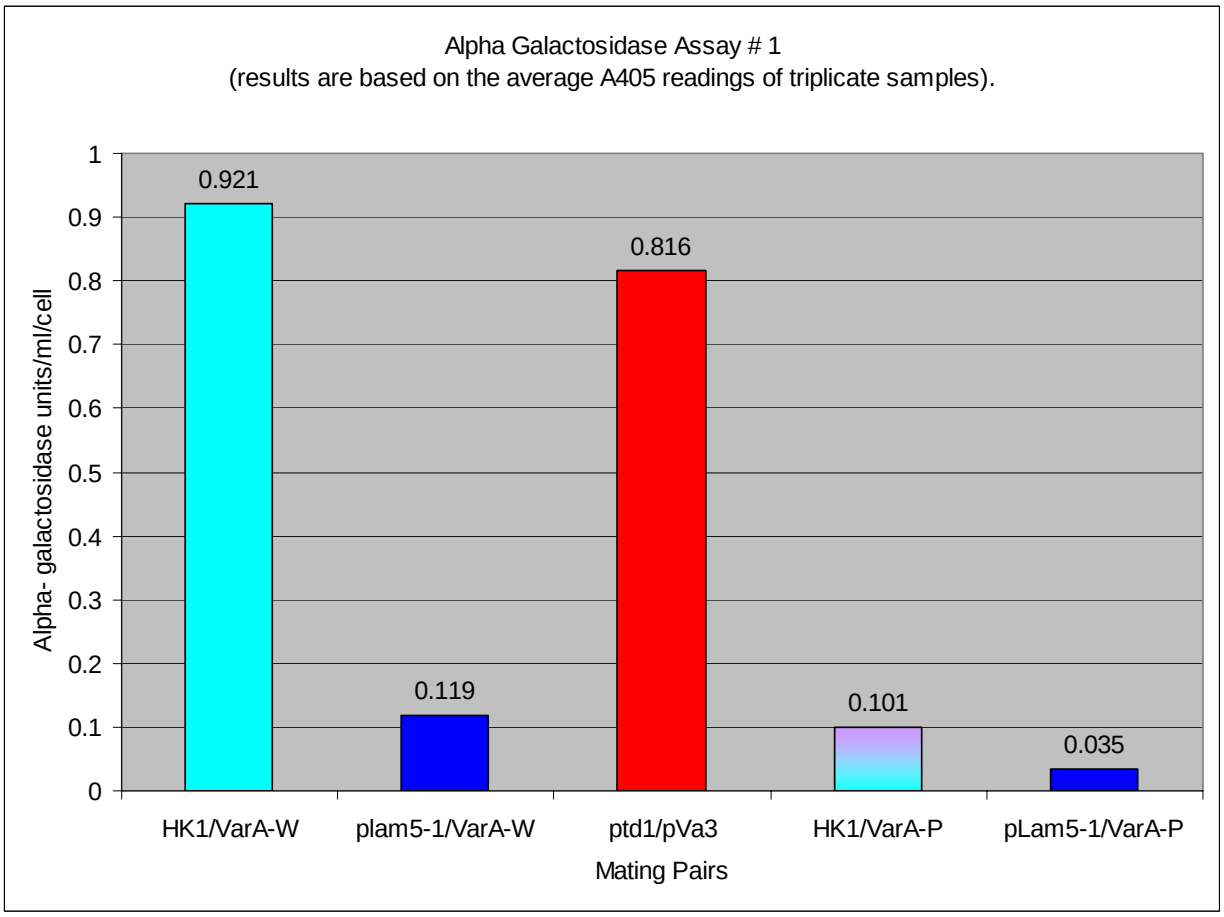


Figure 10

Quantification of HK1/VarA reporter gene activity. The HK1/VarA-W, HK1/VarA-P mating pairs, the negative control mating pairs, pLAM5-1/VarA-W and pLAM5-1/VarA-P, and the positive control mating, pTD1/pVA3 were grown on high-stringency agar medium (SD-Ade-His-Leu-Trp). Overnight broth cultures of each mating pair was used in the generation log phase cultures (OD_{600} 0.5-1). Assays were perfomed with 3 independent colonies of each mating pair. The addition of assay buffer to cell-free supernatant in a 96-well format was used the monitor the catalytic activity of the Gal4-responsive, α-galactosidase, colorimetrically by measuring the rate of hydrolysis of the chromogenic substrate, *p*-nitrophenyl-α-d-galactoside (PNP-α-Gal). One of the products of this reaction, *p*-nitrophenol, displays a strong absorption band at 405 nm.

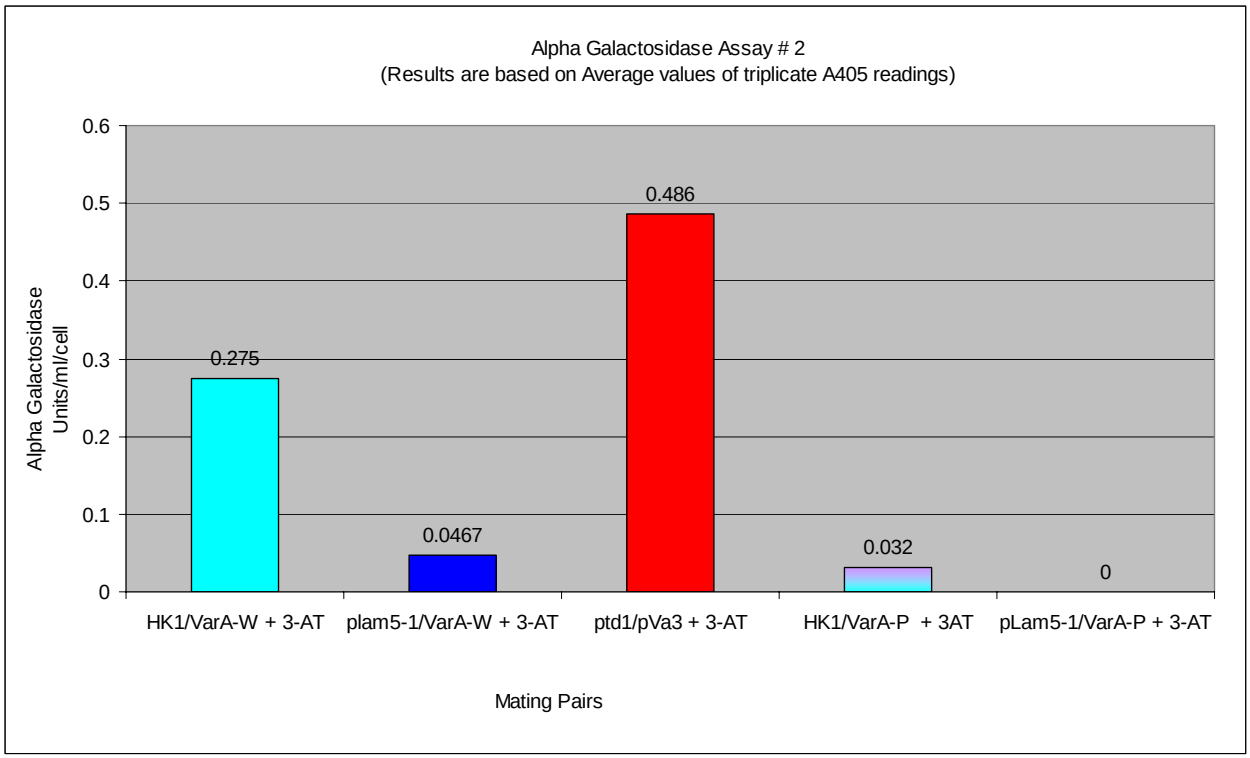


Figure 11

Quantification of HK1/VarA reporter gene activity in the presence of 3-AT. The HK1/VarA-W, HK1/VarA-P mating pairs, the negative control mating pairs, pLAM5-1/VarA-W and pLAM5-1/VarA-P, and the positive control mating, pTD1/pVA3 were grown on high-stringency agar medium (SD-Ade-His-Leu-Trp) supplemented with 3 mM 3-AT to reduce nonspecific reporter gene activity. An overnight broth culture of each mating pair was used to generate log-phase cultures (OD_{600} 0.5-1). Triplicate colonies of each mating pair were analyzed. The addition of assay buffer to cell-free supernatant in a 96-well format was used the monitor the catalytic activity of the Gal4-responsive, α-galactosidase, colorimetrically by measuring the rate of hydrolysis of the chromogenic substrate, *p*-nitrophenyl-α-d-galactoside (PNP-α-Gal). One of the products of this reaction, *p*-nitrophenol, displays a strong absorption band at 405 nm.

```
CLUSTAL W (1.83) multiple sequence alignment

Mycobacterium avium k-12 (PhoR)  ------------MVVTRPRRGLPLRVGLVAATLALVACGLAVSGIAVTSI 38
Mycobacterium avium 104 (PhoR)   ------------MVVTRPRRGLPLRVGLVAATLALVACGLAVSGIAVTSI 38
Mycobacterium ulcerans (PhoR)    -------------MAERLPRAVQLRVGLVAATLALVLCGLLASGVAVTSI 37
Mycobacterium vanbaalenii (PhoR) MDGGDGPTLNNGGVPGRIRRGVPLRVGLVAAMLLLVACGLLASGIAVTTI 50
Rhodococcus equi (HK1)           ---------------MSRLRALPLRVTLVVALLLLAAAGLLASGVAVTSA 35
RHA1 (PhoR)                      ----------------------MPLRITLVGAVLLLSALGLLASGVAVTSA 29
                                                      : **: ** * * *   ** .**:***:

Mycobacterium avium k-12 (PhoR)  LRHSLVSRIDSTLLDASRGWAQAPRRQSSSAYEGPDPGRPPSKFYVRGVS 88
Mycobacterium avium 104 (PhoR)   LRHSLVSRIDSTLLDASRGWAQAPRRQSSSAYEGPDPGRPPSKFYVRGVS 88
Mycobacterium ulcerans (PhoR)    LRHSLVSRIDSTLLEASRTWAQASWRHSMPSVEGPDPARPPSKFYVRSIS 87
Mycobacterium vanbaalenii (PhoR) MQHSLMNRVDDTLLDASRGWAQVPRRVLTVPMEGPNPARPPSDFYVRAVD 10
Rhodococcus equi (HK1)           MQSSLIERTDRQLFDAAHGWAKPRDPWPLPPVSAPGPSRPPSAFYVRTTE 85
RHA1 (PhoR)                      LENSLTNRVDHQLHDAAQGWAKPGKSRPLPPPDEPNPGRPPSPFYVRTTL 79
                                 :. ** .* *  * :*:: **:        . . *.*.**** ****

Mycobacterium avium k-12 (PhoR)  TDGTPFTAINDRNAEPALPANNDVGPNPTTLPSVNGSDIQWRAVSVRGPH 13
Mycobacterium avium 104 (PhoR))  TDGTPFTAINDRNAEPALPADNDVGPNPTTLPSVNGSDIQWRAVSVRGPH 13
Mycobacterium ulcerans (PhoR)    PDGRAMTAINDRNAEPALPADNDVGSDPTTLPSVNGSNIEWRAVSVHGPQ 13
Mycobacterium vanbaalenii (PhoR) PDGHVWMEVNDREAEPALPEDNDVGSEPVTVASVDHSHVQWRAMTVRGLS 15
Rhodococcus equi (HK1))          PDGTVRLEVNDESTNATPEISDDAG--PVTVGSVGGEDIRWRELTVATPS 13
RHA1 (PhoR)                      ANGTVLREINDESVSPALPENPASG--PETVGSTDNSDVRWRVLTTTAPD 12
                                 .:*     :**.....:   .   *  * *: *.. ..:.** ::.

Mycobacterium avium k-12 (PhoR)  G-LTTVAIDLSDVQHTVSSLVWLQIGIGVAVLAVVGIASFAVVQRSLRPL 18
Mycobacterium avium 104 (PhoR)   G-LTTVAIDLSDVQHTVSSLVWLQIGIGVAVLAVVGIASFAVVQRSLRPL 18
Mycobacterium ulcerans (PhoR)    GGLTTVAIDLSDVQHTVRSLVWLQIGIGAGVLVVVGLAGFAVVQRSLRPL 18
Mycobacterium vanbaalenii (PhoR) GELTTVAIDLSDVKSTVRSLVYTQAGFGVAVLVVLGIAGYAVVHRSLRPL 20
Rhodococcus equi (HK1)           G-TTTVAVKLTDIDATVERLVVLQIGIGAVVLVALGFATYFVVRRSLRPL 18
RHA1 (PhoR)                      GTETTVAMTLTENDETVNRLIVLQLIIGVVVLIALGVVAYFVVRRSLRPL 17
                                 *  ****: *:: . **  *:  *  :*. ** .:*.. : **:******

Mycobacterium avium k-12 (PhoR)  VEVEQTAAAIAAGQLDRRVPERDPRTEVGRLSLALNGMLAQIQQALASSE 23
Mycobacterium avium 104 (PhoR)   VEVEQTAAAIAAGQLDRRVPERDPRTEVGRLSLALNGMLAQIQQALASSE 23
Mycobacterium ulcerans (PhoR)    SEVEQTAAAIASGQLDRRVPERDPRTEVGQLSLAMNGMLSQIQVALAASE 23
Mycobacterium vanbaalenii (PhoR) SEVEQTAAAIAAGQLDRRVPERDPRTEVGRLSLALNGMLAQIQRAMASSE 25
Rhodococcus equi (HK1)           REVEEIAADIAAGDLHRRVPEWDQRTEVGRLSSALNGMLAQIQTAFAATA 23
RHA1 (PhoR)                      VEVEETAAAIASGDLHRRVPERDDRTEIGRLSVALNGMLAQIQHAFAATE 22
                                  ***: ** **:*:*.***** * ***:*:** *:****:*** *:*::

Mycobacterium avium k-12 (PhoR)  SSAEKARGSEDRMRRFITDASHELRTPLTTIRGFAELYRQGAARDVAMLL 28
Mycobacterium avium 104 (PhoR)   SSAEKARGSEDRMRRFITDASHELRTPLTTIRGFAELYRQGAARDVAMLL 28
Mycobacterium ulcerans (PhoR)    DSAEKARSSEERMRRFITDASHELRTPLTTIRGFAELYRQGAARDVAMLL 28
Mycobacterium vanbaalenii (PhoR) ASAEQARTSEDRMRRFITDASHELRTPLTTIRGFAELYRQGAARDVEMLM 30
Rhodococcus equi (HK1)           ASEESARRSEERMRRFVADASHELRTPLATIRGFAELYRQGAAPDPEMVM 28
RHA1 (PhoR)                      ASEESARRSEEKMRRFVADASHELRTPLTTIRGFAELYRQGAMTDTETLM 27
                                  * *.** **::****::**********:*************  *   ::
```

Continued on the following page

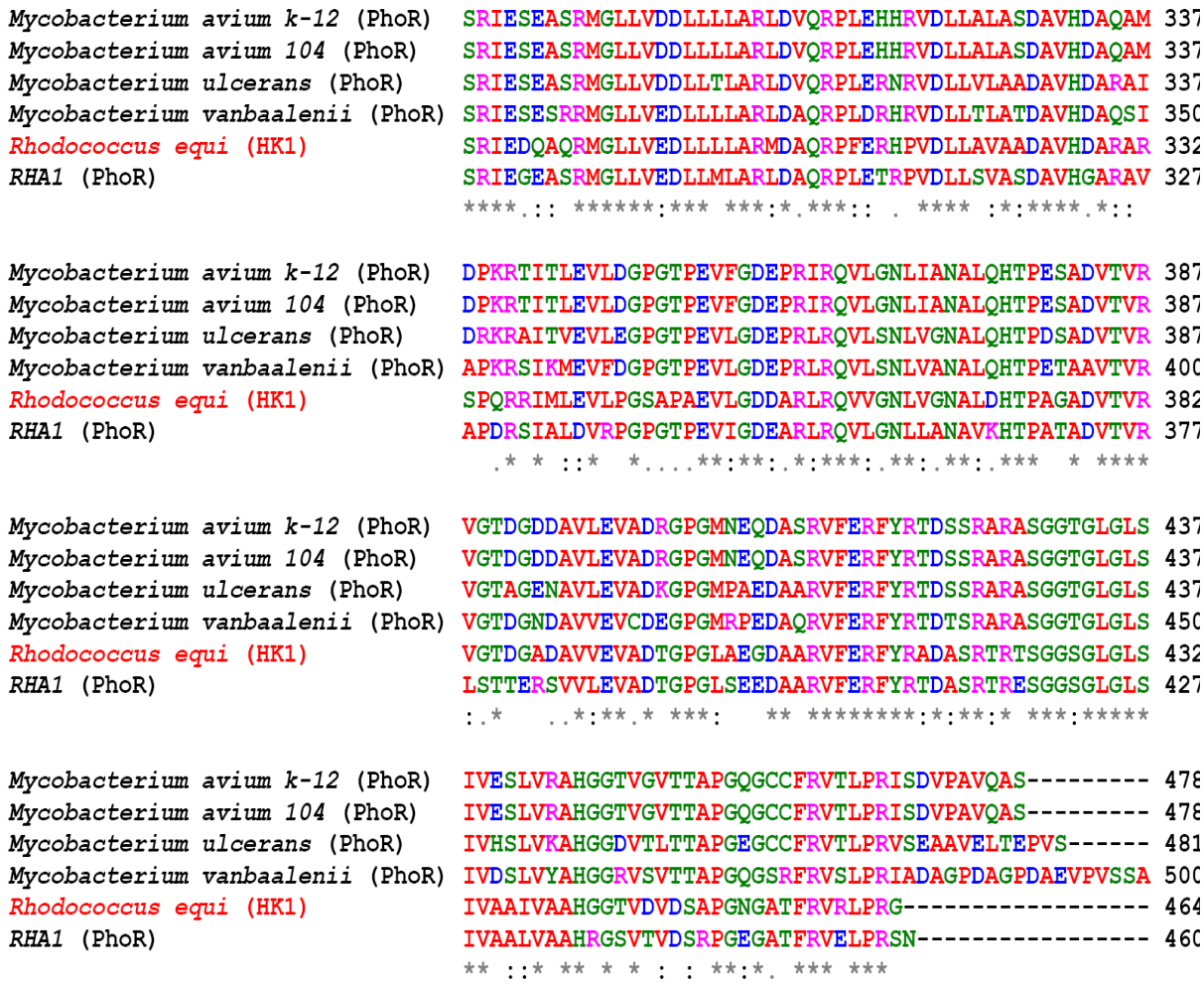


Figure 12

A clustalw multiple sequence alignment of HK1 top matches in Genbank. HK1 alignment with top five highest scoring matches (*Mycobacterium avium k-12, Mycobacterium avium 104, Mycobacterium ulcerans, Mycobacterium vanbaalenii, Rhodococcus RHA1*) shows highly significant global sequence similarity and identity scores to PhoR histidine kinase proteins. Expect (E) values range from $5e^{-131}$ to $3e^{-118}$, affirming the integrity of these matches.

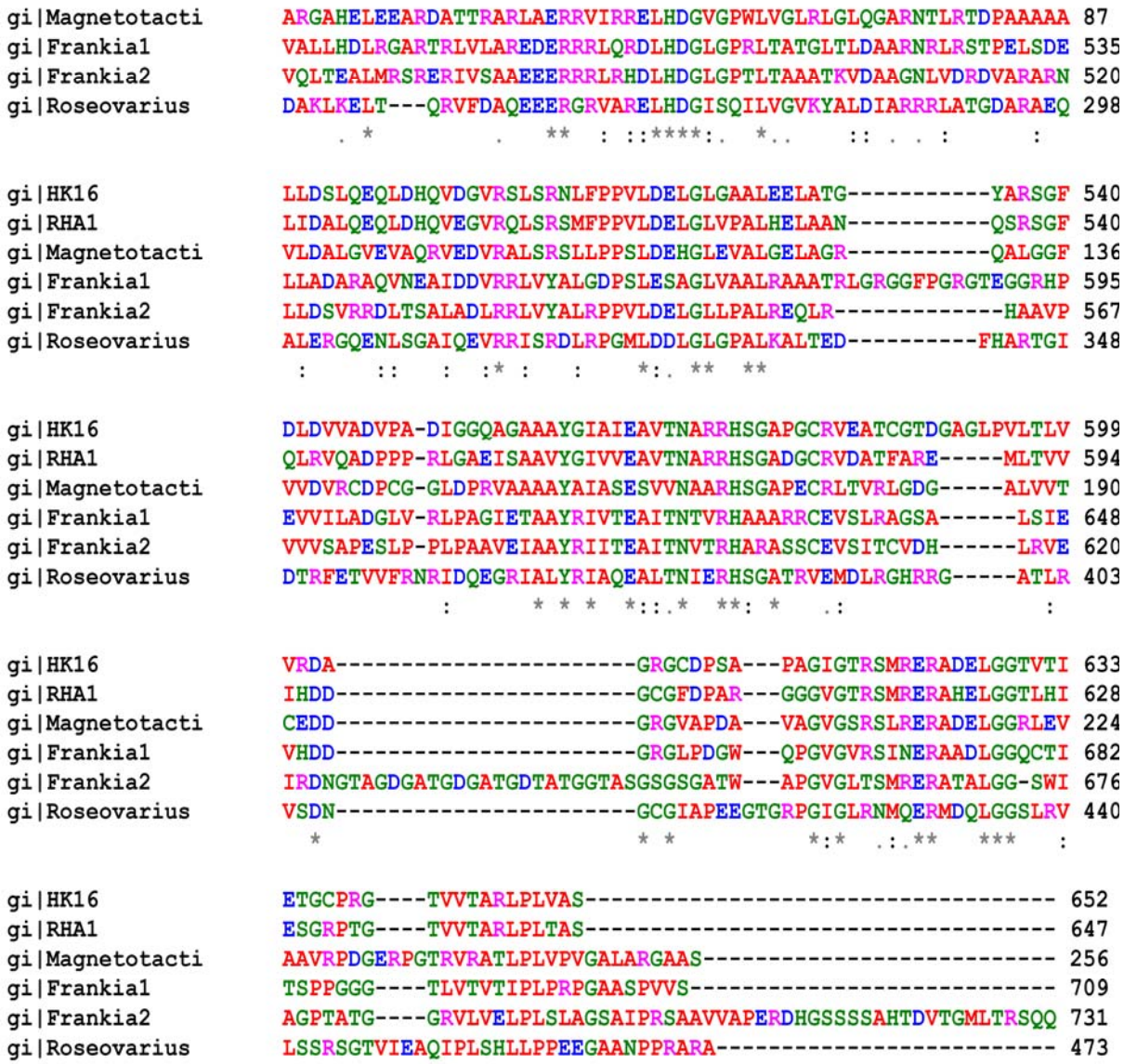
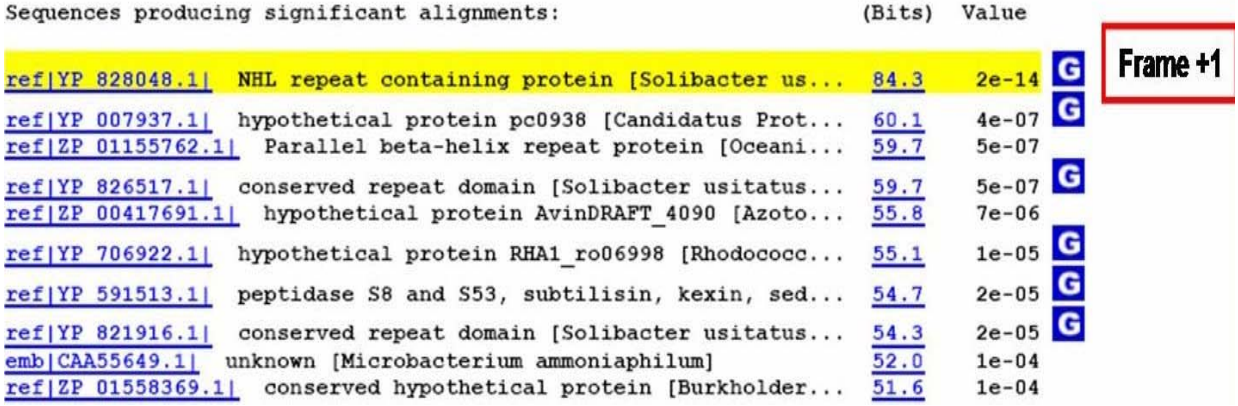


Figure 13

A clustalw multiple sequence alignment of HK16 top five matches in Genbank. Shown is the alignment of *R. equi* HK16 with the top five highest scoring matches in Genbank (*Rhodococcus* sp. RHA1, *Magnetotactic spirillum, Frankia CpI1*, *Frankia* EAN1pec, *Roseovarious nubinhibens*). The alignment demonstrates modest but significant c-terminal sequence similarity with other putative two-component system histidine kinases and weaker similarities at the N-terminal region. Expect (E) values range from $5e^{-41}$ to $3e^{-12}$, affirming the integrity of these matches.

BLASTp search results of A24 (full ORF analysis).



Sequences producing significant alignments:		(Bits)	Value	
ref|YP_828048.1|	NHL repeat containing protein [Solibacter us...	84.3	2e-14	G
ref|YP_007937.1|	hypothetical protein pc0938 [Candidatus Prot...	60.1	4e-07	G
ref|ZP_01155762.1|	Parallel beta-helix repeat protein [Oceani...	59.7	5e-07	
ref|YP_826517.1|	conserved repeat domain [Solibacter usitatus...	59.7	5e-07	G
ref|ZP_00417691.1|	hypothetical protein AvinDRAFT_4090 [Azoto...	55.8	7e-06	
ref|YP_706922.1|	hypothetical protein RHA1_ro06998 [Rhodococc...	55.1	1e-05	G
ref|YP_591513.1|	peptidase S8 and S53, subtilisin, kexin, sed...	54.7	2e-05	G
ref|YP_821916.1|	conserved repeat domain [Solibacter usitatus...	54.3	2e-05	G
emb|CAA55649.1|	unknown [Microbacterium ammoniaphilum]	52.0	1e-04	
ref|ZP_01558369.1|	conserved hypothetical protein [Burkholder...	51.6	1e-04	

Figure 14:

BLASTp analysis of Clone A24. The A24 sequence is a 251 bp *R. equi* genomic library fragment represented on the chromosome at the gene coordinates 195413 - 195663 that reproducibly interacted with VarA in yeast in the reading frame that is being analyzed. The BLAST (Basic Local Alignment Search Tool was used in searching for similar protein sequences. The closest match was a *Solibacter usitatus* NHL repeat containing protein. At $2e^{-14}$, the E-value demonstrates that this is a significant match despite the modest amino acid identity score of 37%. Approximately 51% of the A24 amino acids are similar to those of the NHL repeat containing protein. The NHL repeat containing protein has high sequence similarity to serine threonine kinases.

CHAPTER 4

DISCUSSION AND CONCLUSIONS

4.1 DNA binding studies

VarA is required for vapA expression

R. equi is exposed to various environmental and/or physiological stresses during growth in the host. Within a granulomatous lesion, where *R*. equi is apt to reside during chronic infection, the bacterium encounters stresses that likely include hypoxia, nutrient limitation, reactive oxygen and nitrogen intermediates, low pH, and toxic peptides (4, 19, 20, 43, 44). In order to adapt to these types of conditions, *R. equi* encodes a number of virulence and stress-response genes. It is well-documented that VapA is expressed in response to host intracellular stressors (19, 20, 28, 68, 109). *vir*R, encoding a lysine type transcription factor, is one of the only two pathogenicity island encoded regulatory genes, and was the first described gene to regulate *vap*A expression (183). While *virR* is required for optimal *vapA* expression, alone, it is not sufficient to allow expression of wild type levels of VapA. The *virR* operon consists of five genes, four of which are located downstream of *virR*, including ORF5, *vapH* (a *vapA* homologue), ORF7, and the second regulator located at ORF8, called *varA*. Prior studies have linked this two-component response regulator homologue to *vapA* expression by indirect means. For instance, using a strain deleted for the virulence plasmid, Russell and colleagues were able to show minimal levels of VapA protein production merely by providing *vapA* with its promoter

along with *virR* (183). Wildtype levels of VapA production was achieved only when the entire *virR* operon, inclusive of *varA* was present (183). Because it posed the most relevant of the operon genes to transcriptional regulation, the reconstitution of *vapA* expression was tentatively ascribed to transcriptional activation by the protein encoded by *varA* (183). Consistent with such speculation, work in the Hondalus lab, including my own herein, presents clear evidence that VarA, a two-component response regulator homologue, indeed affects expression of *vap*A at the the RNA level (Hondalus data not shown) and is able to effect this control by binding to the *vapA* promoter region. As judged by the results of a western blot analysis, deletion of *varA* almost completely abrogates production of VapA protein. Residual VapA produced in the absence of *varA* can be attributed to minimal transcriptional activation by VirR. However, of the two plasmid encoded regulators, it is quite clear that VarA is chiefly responsible for maintaining wild type levels of VapA. The *varA* deletion mutant was unable to survive in a SCID mouse model of *R. equi* disease likely because of the negligible amounts of VapA produced in the mutant. However, virulence as measured by *in vivo* growth and organ burden was restored when a wild type copy of *varA* was reintroduced into the *varA* mutant confirming that the attenuated phenotype was due to the *varA* deletion and establishing *varA* as a determinant of *R. equi* virulence.

VarA is a direct regulator of vapA expression

Here, we also report that purified VarA binds to a DNA fragment containing the *vapA* promoter which strongly suggests that it activates *vapA* transcription by a direct interaction with RNA polymerase bound to the *vapA* promoter. With the use of an irrelevant control protein

(which did not bind to the *vapA* promoter probe) and an irrelevant DNA control (to which rVarA does not bind), the binding of recombinant VarA to the *vapA* promoter probe was shown to be a specific interaction. While it was quite clear that rVarA formed complexes specifically with the *vapA* promoter, the binding was suboptimal. Despite titrations of protein:DNA ratios, qualitatively, we observed a modest range of 20-50 percent of VarA binding to the *vapA* promoter region. Two possibilities clarify this observation: Establishing the optimal experimental conditions for gel shift assays is an empirical process that varies between proteins (192). Despite careful consideration of common technical causes of suboptimal binding reactions, it is quite possible that our experimental conditions are not entirely conducive to normal VarA activity. A second, more likely reason for less than optimal binding may be due to the fact that two-component system regulators are typically activated by a phosphorylation event which triggers a conformation change and subsequent binding to a target DNA sequence. Without such an event, it is assumed that the response regulator is essentially in poor binding form and would manifest this in protein interaction assays. Such is the case with regards to DNA-binding studies of OmpR, a response regulator, which, when phosphorylated, is shifted into active conformation that binds DNA with high affinity. In unphosphorylated form, OmpR binds to DNA with low affinity (58). It remains to be determined through phosphorylation assays whether VarA binding of the *vapA* promoter region can be optimized by phosphorylation either by synthetic phosphate donor molecules such as carbamyl phosphate or acetyl phosphate or by an associated kinase molecule. Nevertheless, while suboptimal, these results demonstrate the first account of VarA direct binding to *vapA* promoter sequences.

At lower *varA* concentrations, a single band shift was observed in mobility shift assays, whereas a second band shift was visible at higher VarA concentrations. This has been observed in

DNA binding studies of other regulators, including that of the *Xanthobacter flavus* carbon fixation regulator, CbbR. CbbR binds as a dimer to a high-affinity binding site that is further from the promoter region, giving rise to a band shift with a high mobility (229, 230). A second dimer is recruited by cooperative binding to a promoter-proximal low-affinity binding site, leading to the formation of a second, slower migrating, DNA-protein complex (230). Whether such a scenario exists with respect to *varA* is a likely topic for future studies.

It is notable that the observation of a LysR type transcriptional regulator (LTTR) and a two-component system response regulator that exert control over the same gene is not a unique occurrence. In *Ralstonia solanacearum*, studies of the LTTR PhcA and the response regulator VsrD revealed that both regulators directly but independently regulate the *xpsR* signal integrator gene (102). Meanwhile, LhrA, a homologue of LTTR was found to associate with the response regulator SprE in promotion of RpoS degradation (65). In this study, we have demonstrated that VarA is a chief regulator of *vapA* expression leading to the production of the key virulence determinant and virulence. The significance of the shared regulatory control of *vapA* that is exhibited between *varA* and *virR* remains unclear as is the effect of VarA over other *R. equi* genes. These are topics for future study in the Hondalus lab.

4.2 Yeast two hybrid studies

VarA interacts with two putative HKs and a serine threonine kinase in yeast.

The *R. equi* virulence plasmid does not encode a sensor kinase, suggesting that if VarA is indeed a two-component system response regulator (RR), then it interacts with a sensor histidine

kinase that is chromosomally-encoded. Two-component systems control a wide range of biological processes including virulence. For example, the SpvR protein in *Salmonella enterica* serovar Dublin induces the expression of the *spvABCD* operon during the stationary phase of growth (80, 89). The plasmid-encoded *spv* locus is essential for growth in the liver and the spleen (89). The LTTR, AphB, of *Vibrio cholerae* is required for the activation of the ToxR virulence cascade by transcriptionally activating the two-component system, *tcpPH* (125). This type of signaling system typically consists of a membrane anchored sensor histidine kinase (HK) molecule which auto-phosphorylates at a conserved histidine residue in response to an environmental stimulus that is detected via the extracellularly exposed N-terminal region. The phosphate is transferred to a conserved aspartate residue at the N-terminal receiver domain of a RR. In response, the RR undergoes a conformational change that enables its binding to DNA promoter sequences to control the expression of an appropriate response protein (205). While key to intracellular survival and virulence, *vapA*, remains a poorly understood gene that represents a response to an unknown environmental stimulus. Yet, it is well-documented that VapA is produced under conditions of magnesium limitation, low pH, and high temperature. In fact, it is entirely possible that multiple stress signals may cooperatively induce *vapA* expression (68, 174). With this in mind, our search for the histidine kinase partner to the VarA was focused on a chromosomally-encoded kinase partner. The fact that the available *R. equi* genome sequence is unannotated complicated matters. Yet, with exhaustive bioinformatics methods (detailed elsewhere in this work) 20 independent putative histidine kinase homologues, each of which neighbored one of 20 independent putative response regulator homologues, were identified. No unpaired histidine kinase homologues were located. These findings created a basis for suspicion of crosstalk between VarA and one or more of the putative paired histidine kinases.

The theme of utilization of two sensor kinases by one response regulator protein is not uncommon; the related *Mycobacterium tuberculosis* encodes a regulator DevR which is typically activated by both sensor kinases, DevS and DevT, in response to nitric oxide stress (184, 199, 202). Along this line, two independent nitrate-detecting sensor kinases, NarQ and NarX communicate with response regulator NarL and modulate NarL-dependent gene expression in *E. coli* (34). The reverse is also possible where multiple regulators may interact with a single sensor kinase as is the case with *E. coli* regulatory proteins PhoR and PhoM which both respond to phosphorylation activation by kinase PhoB in *E. coli* (3). These examples exemplify that shared signaling entities can form a multicomponent criss-crossing of pathways which helps the organism to integrate diverse signals into a nuanced response, leading to adaptation to an environment that is a product of multiple signals. With this thinking in mind, in this study, we strategically approached the search for the VarA partner using yeast two hybrid technology from two perspectives: One, a direct mating strategy that allowed for a pair-wise analysis of VarA interactions with the 20 putative chromosomally-encoded histidine kinases; and second, a traditional library screen that facilitated a genome-wide analysis for proteins that form strong interactions with the VarA protein. With the use of both strategies, in yeast, we have discovered two potential VarA interactors from the first direct interaction approach (HK1, a PhoR-homologue and HK16), and it is possible that we have also identified a novel type of interaction between VarA, a two-component response regulator, and a serine threonine kinase molecule by using the second approach.

VarA interacts with a PhoR-histidine kinase homologue and another histidine kinase in yeast.

The putative *phoR* gene of *R. equi* (also annotated as histidine kinase 1 (HK1) in this work) is neighbored by a gene encoding a PhoP response regulator homologue (RR1). These proteins are typically produced within the same operon (40, 205). The PhoPR operon has been previously described in various bacteria as a two-component signal transduction system that is responsible for the positive regulation of adaptation genes in response to an extracellular phosphate deficiency or magnesium limitation (2, 76, 146, 167, 197, 234), both conditions encountered by intracellular pathogens within the host environment. In *Bacillus subtilis*, PhoR is known to sense phosphate limitation and transduce that signal by reversible phosphorylation activation of PhoP (197). As would be expected of a positive regulator of gene expression, a *Mycobacterium tuberculosis phoPR* disruption caused a significant attenuation of growth in macrophages and mice, hampered growth in low magnesium medium and resulted in a failure of the strain to synthesize complex lipids that compose the highly protective mycolic acid rich envelope of the bacterium. Incidentally, the PhoPR system of *R. equi* has been previously described by Ren and colleagues (179). In sharp contrast to the trends demonstrated by *phoPR* deletions within other bacterial systems (including that of the closely related *Mycobacterium tuberculosis* system), an *R. equi* strain lacking the *phoPR* genes was cited as being enhanced in virulence properties (179), although that study was flawed. The hypervirulent phenotype of the *phoPR* mutant was not significantly reduced by an in trans complementation of the mutant with wild-type *phoPR* which leaves open the possibility that a polar mutation is the cause for the unexpected phenotype. Additionally, the researchers did not assess the initial post-infection organ burden and recorded only one time-point four days post infection. Nevertheless, similarly

to genes are known to be under control of the PhoPR operon, earlier studies of *R. equi* have demonstrated that VapA, a key virulence determinant which we showed is regulated by VarA, is also expressed under conditions of low magnesium (178). In the context of our observed interaction between VarA and the *R. equi* PhoP homologue in yeast, Ren and colleagues' observation of a hypervirulent phenotype in the PhoPR mutant would suggest that, together, the *R. equi* PhoPR operon negatively regulates virulence gene expression within *R. equi*. This observation does not follow the established trend of attenuation observed in PhoPR deletion strains of other bacteria which suggest positive regulation of virulence associated genes including in the case of the related intracellular actinomycete, *Mycobacterium tuberculosis* (2, 76, 197, 234). A second, more obscure histidine kinase homologue, HK16 interacted with VarA in yeast. Since the sensory N-terminal region of HK16 contrasts heavily with that of the putative *R. equi phoR*, it is possible that HK16 phosphorylates VarA as a means for transducing an environmental signal other than low phosphate or magnesium which are signals transduced by typical PhoR genes (2, 3, 76, 146, 197). The interactions observed between VarA and the *R. equi* PhoR homologue as well as with HK16 remain to be confirmed at the biochemical level in pull-down assays. Once confirmed, it would be interesting to observe the effect of PhoP and PhoR on virulence independently of one another to distinguish the role that each response regulator homologue (PhoP, VarA) plays in sharing a common PhoR-like histidine kinase. The same consideration can be applied to HK16 as well which is another histidine kinase whose function could be shared by its neighboring RR16 as well as VarA. Also, it would be necessary to demonstrate phosphotransfer from the PhoR homologue or from HK16 to the VarA regulator to confirm their predicted functions.

VarA interacts with a putative serine threonine kinase in yeast.

An interesting observation of reproducible VarA interaction with a serine threonine kinase in yeast arose from a library screen approach to yeast two hybrid methodology. Considering that it is typically response regulators that are activated by histidine kinases which transduce external stimuli to effect the production of adaptive proteins, this finding, if confirmed outside of the yeast system, represents a novel type of interaction with a type of signaling kinase that was originally considered non-existent in bacterial species (83, 246). Protein phosphorylation is the principal mechanism by which extracellular signals are translated into cellular responses. In bacteria, the molecular system that is chiefly responsible for translating extracellular signals into cellular responses are two-component systems which traditionally consist of histidine kinase sensors and their associated response regulators (41, 61, 149, 205, 238). In contrast, in eukaryotes, phosphorylation typically occurs on phosphoester (serine, threonine, or tyrosine) residues conducted by serine threonine kinases (STKs) (84). In the past, STKs were thought to exist only within eukaryotic cells (84). However, within the last decade, large scale bacterial genome sequencing projects have paved the way for the identification of several bacteria encoded STKs (62, 63, 85, 168, 235, 246). For example, the related species, *Mycobacterium tuberculosis* has been found to encode 11 serine threonine kinases (14). These STKs have been largely identified in connection with either developmental processes within the cell or in functions directly related to pathogenicity. Bacteria that have the capacity to sporulate or produce secondary metabolites, including *Streptomyces* (158, 168, 198, 228) *Anabaena* (70, 246, 247) spp., and *Myxococcus xanthus* (85, 156, 160, 226, 247, 248), encode a significant number of STK genes. In addition, STKs have been implicated directly in the survival of human

pathogens within the host and in some cases have been shown to be required for full virulence in mouse models as demonstrated by the *Pseudomonas aeruginosa* STK (235), and the *Yersinia pseudotuberculosis* plasmid-encoded protein kinase YopO also referred to as *ypkA* (18, 62, 63, 246). Secreted YopO is targeted to the inner surface of the host cell plasma membrane and is proposed to interfere with the host response by the disruption of cell signaling events through phosphorylation of eukaryotic substrates. More interestingly, challenging the old paradigm, recent reports have also corroborated that STKs can phosphorylate the response regulator of two-component systems quite independently from their cognate histidine kinase molecules. For example, a *Streptomyces* global response regulator AfsR was shown to be phophorylated by two STKs independently of its cognate AfsK histidine kinase partner (142, 186, 228). In *Streptococcus agalactiae*, a STK regulated expression of cytoxin proteins by phosphorylating a response regulator, CovR independently of its cognate phophokinase, CovS (176). *Myxobacteria* have also been shown to contain transmembrane STK proteins that may also serve as receptors for external signals as well as phosphotransfer proteins for the activation of response regulators of two-component systems (85, 90). In this context, the observed interaction between VarA and the putative STK of *R. equi* may implicate a novel type of convergence between a serine threonine kinase and a response regulator molecule that is directly involved in the regulation of key virulence determinants. Further analysis and confirmation of the interaction observed in yeast is underway.

Implications of these results

Within the yeast system, through yeast two hybrid analyses, we have established that the plasmid-encoded virulence response regulator homologue, VarA, has reproducibly interacted with a PhoR-like histidine kinase as well as another histidine kinase homologue (HK16) and a putative serine threonine kinase (STK1) which are all chromosomally-encoded. Each of the two putative kinases appears to be independently-partnered with a cognate response regulator that is encoded within 29 bp distance and in the same orientation. Although each interaction proved reproducible even under higher stringency selection, it is likely that two approaches of the yeast two hybrid methodology did not yield overlapping results due to the application of highly stringent selection criteria used throughout the library screen method. In eliminating a large number of nonspecific interactions, it is quite likely that these conditions abrogated weaker interactions. Additionally, the *R. equi* library insert sequences of four positive VarA interacting plasmids remain to be analyzed for the presence of recognizable domains and motifs by bioinformatics methods. Nevertheless, a foundation for studying the mechanisms that lead to *vapA* gene regulation by VarA has been established. With this, a model is proposed (Figure 15) wherein the three transmembrane phophokinases (the *R. equi* STK, HK16 and the PhoR-like histisine kinase) either sense independent or overlapping environmental stimuli (low iron, high temperature, low magnesium, oxygen stress) from within the host cell, leading to autophosphorylation and phosphor-transfer directly to VarA. An activated VarA undergoes a conformation change that enables its direct binding to the *vapA* promoter region, thereby, inducing the production of VapA. With its signal sequence, VapA is extruded from the cell and

tethered to the extracellular surface where it can intimately interact with host cells in order to promote intracellular persistence and virulence of *R. equi* in susceptible hosts.

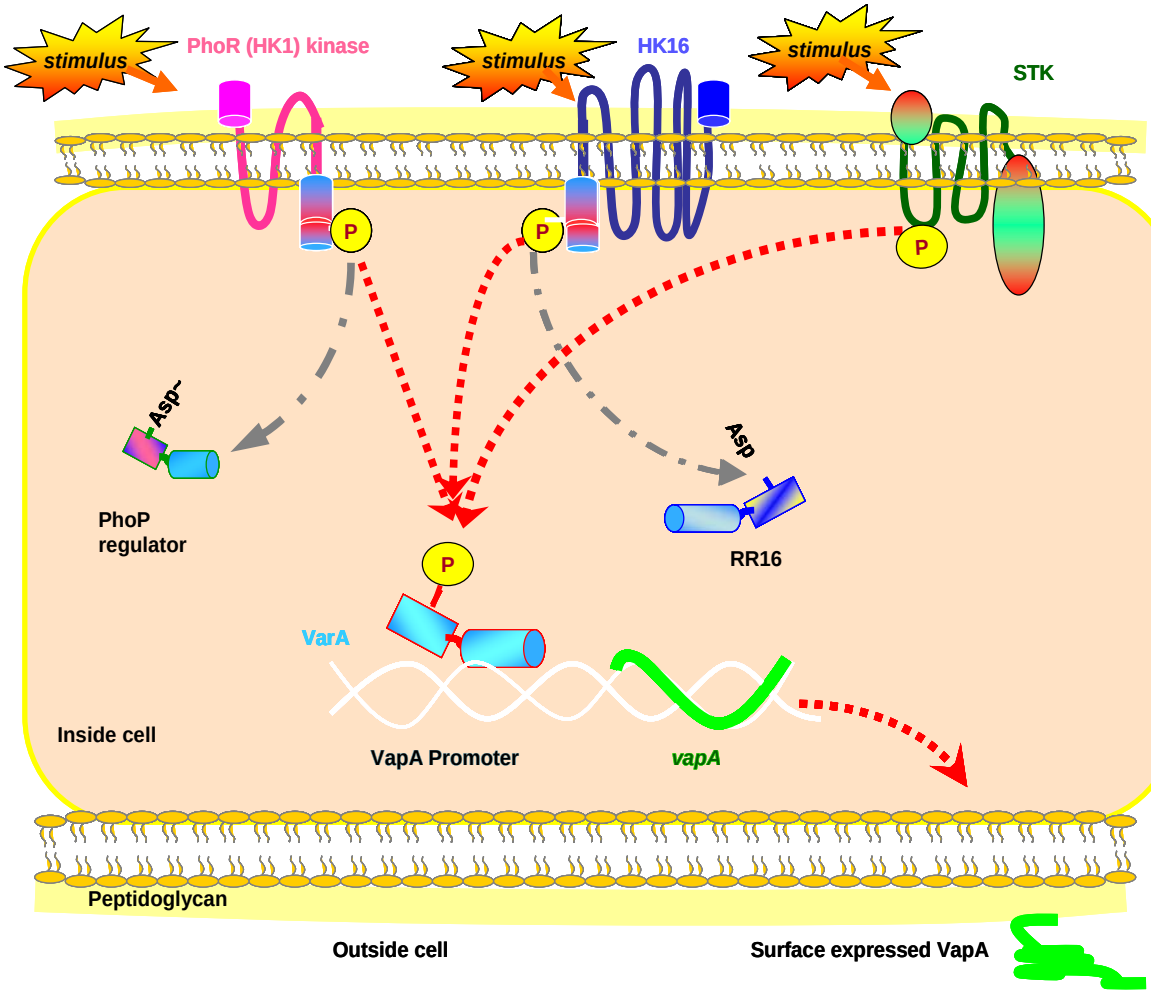


Figure 15

Proposed model for VarA-mediated regulation of VapA production.

REFERENCES

1. **Addo, P. B., and S. M. Dennis.** 1977. *Corynebacteria* associated with diseases of cattle, sheep and goats in Northern Nigeria. Br Vet J **133:**334-9.
2. **Aguirre, A., M. L. Cabeza, S. V. Spinelli, M. McClelland, E. Garcia Vescovi, and F. C. Soncini.** 2006. PhoP-induced genes within *Salmonella* pathogenicity island 1. J Bacteriol **188:**6889-98.
3. **Amemura, M., K. Makino, H. Shinagawa, and A. Nakata.** 1990. Cross talk to the phosphate regulon of *Escherichia coli* by PhoM protein: PhoM is a histidine protein kinase and catalyzes phosphorylation of PhoB and PhoM-open reading frame 2. J Bacteriol **172:**6300-7.
4. **Amouyal, M.** 2005. Gene regulation at-a-distance in *E. coli*: new insights. C R Biol **328:**1-9.
5. **Anantharaman, V., and L. Aravind.** 2001. The CHASE domain: a predicted ligand-binding module in plant cytokinin receptors and other eukaryotic and bacterial receptors. Trends Biochem Sci **26:**579-82.
6. **Appleman, J. A., and V. Stewart.** 2003. Mutational analysis of a conserved signal-transducing element: the HAMP linker of the *Escherichia coli* nitrate sensor NarX. J Bacteriol **185:**89-97.
7. **Aravind, L.** 2007. Opening Pandora's Box: making biological discoveries through computational data exploration. Biol Direct **2:**29.
8. **Aravind, L., V. Anantharaman, S. Balaji, M. M. Babu, and L. M. Iyer.** 2005. The many faces of the helix-turn-helix domain: transcription regulation and beyond. FEMS Microbiol Rev **29:**231-62.
9. **Aravind, L., and C. P. Ponting.** 1997. The GAF domain: an evolutionary link between diverse phototransducing proteins. Trends Biochem Sci **22:**458-9.
10. **Ardans, A. A., S. K. Hietala, M. S. Spensley, and A. Sansome.** 1986. Studies of naturally occurring and experimental *Rhodococcus equi* (*Corynebacterium equi*) pneumonia in foals. Proc. Am. Assoc. Equine Pract. **:**129-144.
11. **Arya, B., S. Hussian, and S. Hariharan.** 2004. *Rhodococcus equi* pneumonia in a renal transplant patient: a case report and review of literature. Clin Transplant **18:**748-52.
12. **Asha, H., and J. Gowrishankar.** 1993. Regulation of kdp operon expression in *Escherichia coli*: evidence against turgor as signal for transcriptional control. J Bacteriol **175:**4528-37.
13. **Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K.** 1994. Current Protocols in Molecular Biology (John Wiley & Sons, Inc.) **1**.
14. **Av-Gay, Y., S. Jamil, and S. J. Drews.** 1999. Expression and characterization of the *Mycobacterium tuberculo*sis serine/threonine protein kinase PknB. Infect Immun **67:**5676-82.
15. **Baker, M. D., and J. B. Stock.** 2007. Signal transduction: networks and integrated circuits in bacterial cognition. Curr Biol **17:**R1021-4.

16. **Balaji, S., M. M. Babu, and L. Aravind.** 2007. Interplay between network structures, regulatory modes and sensing mechanisms of transcription factors in the transcriptional regulatory network of *E. coli.* J Mol Biol **372:**1108-22.
17. **Barton, M. D., and K. L. Hughes.** 1984. Ecology of *Rhodococcus equi.* Vet Microbiol **9:**65-76.
18. **Barz, C., T. N. Abahji, K. Trulzsch, and J. Heesemann.** 2000. The *Yersinia* Ser/Thr protein kinase YpkA/YopO directly interacts with the small GTPases RhoA and Rac-1. FEBS Lett **482:**139-43.
19. **Benoit, S., A. Benachour, S. Taouji, Y. Auffray, and A. Hartke.** 2002. H(2)O(2), which causes macrophage-related stress, triggers induction of expression of virulence-associated plasmid determinants in *Rhodococcus equi*. Infect Immun **70:**3768-76.
20. **Benoit, S., A. Benachour, S. Taouji, Y. Auffray, and A. Hartke.** 2001. Induction of vap genes encoded by the virulence plasmid of *Rhodococcus equi* during acid tolerance response. Res Microbiol **152:**439-49.
21. **Bohm, A., and W. Boos.** 2004. Gene regulation in prokaryotes by subcellular relocalization of transcription factors. Curr Opin Microbiol **7:**151-6.
22. **Bourret, R. B., K. A. Borkovich, and M. I. Simon.** 1991. Signal transduction pathways involving protein phosphorylation in prokaryotes. Annu Rev Biochem **60:**401-41.
23. **Breathnach, C. C., T. Sturgill-Wright, J. L. Stiltner, A. A. Adams, D. P. Lunn, and D. W. Horohov.** 2006. Foals are interferon gamma-deficient at birth. Vet Immunol Immunopathol **112:**199-209.
24. **Brent, R., and M. Ptashne.** 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. Cell **43:**729-36.
25. **Bruner, D. W., and P. R. Edwards.** 1941. Classification of *Corynebacterium equi*. Ky. Agric. Exp. Stn. Bull. **:**91-107.
26. **Burke, J. D., and K. L. Gould.** 1994. Molecular cloning and characterization of the *Schizosaccharomyces pombe* his3 gene for use as a selectable marker. Mol Gen Genet **242:**169-76.
27. **Buron-Barral, M. C., K. K. Gosink, and J. S. Parkinson.** 2006. Loss- and gain-of-function mutations in the F1-HAMP region of the *Escherichia coli* aerotaxis transducer Aer. J Bacteriol **188:**3477-86.
28. **Byrne, B. A., J. F. Prescott, G. H. Palmer, S. Takai, V. M. Nicholson, D. C. Alperin, and S. A. Hines.** 2001. Virulence plasmid of *Rhodococcus equi* contains inducible gene family encoding secreted proteins. Infect Immun **69:**650-6.
29. **Byrne, G. A., C. A. Boland, E. P. O'Connell, and W. G. Meijer.** 2008. Differential mRNA stability of the vapAICD operon of the facultative intracellular pathogen *Rhodococcus equi*. FEMS Microbiol Lett **280:**89-94.
30. **Byrne, G. A., D. A. Russell, X. Chen, and W. G. Meijer.** 2007. Transcriptional regulation of the virR operon of the intracellular pathogen *Rhodococcus equi*. J Bacteriol **189:**5082-9.
31. **Cann, J. R.** 1993. Theoretical studies on the mobility-shift behavior of binary protein-DNA complexes. Electrophoresis **14:**669-79.
32. **Carrigan, M. J., I. J. Links, and A. G. Morton.** 1988. *Rhodococcus equi* infection in goats. Aust Vet J **65:**331-2.

33. **Cauchard, J., C. Sevin, J. J. Ballet, and S. Taouji.** 2004. Foal IgG and opsonizing anti-*Rhodococcus equi* antibodies after immunization of pregnant mares with a protective VapA candidate vaccine. Vet Microbiol **104:**73-81.
34. **Cavicchioli, R., I. Schroder, M. Constanti, and R. P. Gunsalus.** 1995. The NarX and NarQ sensor-transmitter proteins of *Escherichia coli* each require two conserved histidines for nitrate-dependent signal transduction to NarL. J Bacteriol **177:**2416-24.
35. **Ceglarek, J. A., and A. Revzin.** 1989. Studies of DNA-protein interactions by gel electrophoresis. Electrophoresis **10:**360-5.
36. **Chatterjee, A., Y. Cui, H. Hasegawa, and A. K. Chatterjee.** 2007. PsrA, the *Pseudomonas* sigma regulator, controls regulators of epiphytic fitness, quorum-sensing signals, and plant interactions in *Pseudomonas syringae pv. tomato strain DC3000*. Appl Environ Microbiol **73:**3684-94.
37. **Chien, C. T., P. L. Bartel, R. Sternglanz, and S. Fields.** 1991. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. Proc Natl Acad Sci U S A **88:**9578-82.
38. **Cho, S., S. G. Park, D. H. Lee, and B. C. Park.** 2004. Protein-protein interaction networks: from interactions to networks. J Biochem Mol Biol **37:**45-52.
39. **Cimprich, R. E., and J. R. Rooney.** 1977. *Corynebacterium equi* enteritis in foals. Vet Pathol **14:**95-102.
40. **Cock, P. J., and D. E. Whitworth.** 2007. Evolution of gene overlaps: relative reading frame bias in prokaryotic two-component system genes. J Mol Evol **64:**457-62.
41. **Cock, P. J., and D. E. Whitworth.** 2007. Evolution of prokaryotic two-component system signaling pathways: gene fusions and fissions. Mol Biol Evol **24:**2355-7.
42. **Daniel, J. A., J. Yoo, B. T. Bettinger, D. C. Amberg, and D. J. Burke.** 2006. Eliminating gene conversion improves high-throughput genetics in *Saccharomyces cerevisiae.* Genetics **172:**709-11.
43. **Darrah, P. A., M. K. Hondalus, Q. Chen, H. Ischiropoulos, and D. M. Mosser.** 2000. Cooperation between reactive oxygen and nitrogen intermediates in killing of *Rhodococcus equi* by activated macrophages. Infect Immun **68:**3587-93.
44. **Darrah, P. A., M. C. Monaco, S. Jain, M. K. Hondalus, D. T. Golenbock, and D. M. Mosser.** 2004. Innate immune responses to *Rhodococcus equi*. J Immunol **173:**1914-24.
45. **Davis, W. P., B. A. Steficek, G. L. Watson, B. Yamini, H. Madarame, S. Takai, and J. A. Render.** 1999. Disseminated *Rhodococcus equi* infection in two goats. Vet Pathol **36:**336-9.
46. **De La Pena-Moctezuma, A., J. F. Prescott, and M. Goodfellow.** 1996. Attempts to find phenotypic markers of the virulence plasmid of *Rhodococcus equi*. Can J Vet Res **60:**29-33.
47. **Demple, B.** 1996. Redox signaling and gene control in the *Escherichia coli* soxRS oxidative stress regulon--a review. Gene **179:**53-7.
48. **Dmitriev, A., Y. H. Yang, A. D. Shen, and A. Totolian.** 2006. Adjacent location of the bac gene and two-component regulatory system genes within the putative *Streptococcus agalactiae* pathogenicity island. Folia Microbiol (Praha) **51:**229-35.
49. **Durfee, T., K. Becherer, P. L. Chen, S. H. Yeh, Y. Yang, A. E. Kilburn, W. H. Lee, and S. J. Elledge.** 1993. The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. Genes Dev **7:**555-69.

50. **Dziembowski, A., and B. Seraphin.** 2004. Recent developments in the analysis of protein complexes. FEBS Lett **556:**1-6.
51. **Elienberger, M. A., and R. P. Genetzky.** 1986. *Rhodococcus equi* infections: literature review. . Comp. Cont. Educ. Pract.Vet. **8:**S414-S424.
52. **Epstein, W.** 1992. Kdp, a bacterial P-type ATPase whose expression and activity are regulated by turgor pressure. Acta Physiol Scand Suppl **607:**193-9.
53. **Erdman, S., L. Lin, M. Malczynski, and M. Snyder.** 1998. Pheromone-regulated genes required for yeast mating differentiation. J Cell Biol **140:**461-83.
54. **Esteves, P., A. Mineiro, M. Serrado, and A. Diniz.** 2007. [*Rhodococcus equi* pneumonia in an HIV+ patient: An uncommon association]. Rev Port Pneumol **13:**703-9.
55. **Fernandez-Mora, E., M. Polidori, A. Luhrmann, U. E. Schaible, and A. Haas.** 2005. Maturation of *Rhodococcus equi*-containing vacuoles is arrested after completion of the early endosome stage. Traffic **6:**635-53.
56. **Fields, S.** 2005. High-throughput two-hybrid analysis. The promise and the peril. FEBS J **272:**5391-9.
57. **Fields, S., and O. Song.** 1989. A novel genetic system to detect protein-protein interactions. Nature **340:**245-6.
58. **Forst, S. A., J. Delgado, and M. Inouye.** 1989. DNA-binding properties of the transcription activator (OmpR) for the upstream sequences of ompF in *Escherichia coli* are altered by envZ mutations and medium osmolarity. J Bacteriol **171:**2949-55.
59. **Fried, M. G.** 1989. Measurement of protein-DNA interaction parameters by electrophoresis mobility shift assay. Electrophoresis **10:**366-76.
60. **Fried, M. G., and J. L. Bromberg.** 1997. Factors that affect the stability of protein-DNA complexes during gel electrophoresis. Electrophoresis **18:**6-11.
61. **Galperin, M. Y.** 2004. Bacterial signal transduction network in a genomic perspective. Environ Microbiol **6:**552-67.
62. **Galyov, E. E., S. Hakansson, A. Forsberg, and H. Wolf-Watz.** 1993. A secreted protein kinase of *Yersinia pseudotuberculosis* is an indispensable virulence determinant. Nature **361:**730-2.
63. **Galyov, E. E., S. Hakansson, and H. Wolf-Watz.** 1994. Characterization of the operon encoding the YpkA Ser/Thr protein kinase and the YopJ protein of *Yersinia pseudotuberculosis*. J Bacteriol **176:**4543-8.
64. **Genetsky, R. M., M. P. Bettcher, L. H. Arp, and P. L. White.** 1982. *Corynebacterium equi* infection in a mare. . Mod. Vet. Pract. **63:**876-879.
65. **Gibson, K. E., and T. J. Silhavy.** 1999. The LysR homolog LrhA promotes RpoS degradation by modulating activity of the response regulator sprE. J Bacteriol **181:**563-71.
66. **Gietz, R. D.** 2006. Yeast two-hybrid system screening. Methods Mol Biol **313:**345-71.
67. **Gietz, R. D., and R. H. Schiestl.** 1991. Applications of high efficiency lithium acetate transformation of intact yeast cells using single-stranded nucleic acids as carrier. Yeast **7:**253-63.
68. **Giguere, S., M. K. Hondalus, J. A. Yager, P. Darrah, D. M. Mosser, and J. F. Prescott.** 1999. Role of the 85-kilobase plasmid and plasmid-encoded virulence-associated protein A in intracellular survival and virulence of *Rhodococcus equi*. Infect Immun **67:**3548-57.

69. **Golemis, E. A., I. Serebriiskii, and S. F. Law.** 1999. The yeast two-hybrid system: criteria for detecting physiologically significant protein-protein interactions. Curr Issues Mol Biol **1:**31-45.
70. **Gonzalez, L., V. Phalip, and C. C. Zhang.** 2001. Characterization of PknC, a Ser/Thr kinase with broad substrate specificity from the *cyanobacterium Anabaena* sp. strain PCC 7120. Eur J Biochem **268:**1869-75.
71. **Goodfellow, M.** 1986. Genus *Rhodococcus*, p. 1472. In P. H. A. Sneath, N. S. Mair, M. E. Sharpe, and J. G. Holt (ed.). Bergey's manual of systematic bacteriology **Vol. 2.** .
72. **Goodfellow, M., and G. Alderson.** 1977. The actinomycete-genus *Rhodococcus*: a home for the "rhodochrous" complex. J Gen Microbiol **100:**99-122.
73. **Goodfellow, M., and T. Cross.** 1984. Classification, p. 7. In M. Goodfellow, M. Mordaski, and S. T. Williams (ed.), The biology of the actinomycetes. Academic Press, Inc. (London), Ltd., London.
74. **Goodfellow, M., A. R. Beckham, and M. D. Barton.** 1982. Numerical classification of *Rhodococcus equi* and related actinomycetes. J Appl Bacteriol **53:**199-207.
75. **Goudreau, P. N., and A. M. Stock.** 1998. Signal transduction in bacteria: molecular mechanisms of stimulus-response coupling. Curr Opin Microbiol **1:**160-9.
76. **Grabenstein, J. P., M. Marceau, C. Pujol, M. Simonet, and J. B. Bliska.** 2004. The response regulator PhoP of *Yersinia pseudotuberculosis* is important for replication in macrophages and for virulence. Infect Immun **72:**4973-84.
77. **Gray, K. J., N. French, E. Lugada, C. Watera, and C. F. Gilks.** 2000. *Rhodococcus equi* and HIV-1 infection in Uganda. J Infect **41:**227-31.
78. **Grebe, T. W., and J. B. Stock.** 1999. The histidine protein kinase superfamily. Adv Microb Physiol **41:**139-227.
79. **Grimm, M. B., N. D. Cohen, N. M. Slovis, G. D. Mundy, J. R. Harrington, M. C. Libal, S. Takai, and R. J. Martens.** 2007. Evaluation of fecal samples from mares as a source of *Rhodococcus equi* for their foals by use of quantitative bacteriologic culture and colony immunoblot analyses. Am J Vet Res **68:**63-71.
80. **Grob, P., and D. G. Guiney.** 1996. In vitro binding of the *Salmonella dublin* virulence plasmid regulatory protein SpvR to the promoter regions of spvA and spvR. J Bacteriol **178:**1813-20.
81. **Guthrie, C. F., G. R.** 1991. Guide to yeast genetics and molecular biology. Methods in Enzymology (Academic Press, San Diego) **194:**1–932.
82. **Hakenbeck, R., and J. B. Stock.** 1996. Analysis of two-component signal transduction systems involved in transcriptional regulation. Methods Enzymol **273:**281-300.
83. **Hanks, S. K.** 2003. Genomic analysis of the eukaryotic protein kinase superfamily: a perspective. Genome Biol **4:**111.
84. **Hanks, S. K., A. M. Quinn, and T. Hunter.** 1988. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science **241:**42-52.
85. **Hanlon, W. A., M. Inouye, and S. Inouye.** 1997. Pkn9, a Ser/Thr protein kinase involved in the development of *Myxococcus xanthus*. Mol Microbiol **23:**459-71.
86. **Harvey RL, S. J.** 1991 *Rhodoccus equi* infection in patients with and without human immunodeficiency virus infection. . Rev Infect Dis **13:**139–145.
87. **Haydel, S. E., N. E. Dunlap, and W. H. Benjamin, Jr.** 1999. In vitro evidence of two-component system phosphorylation between the *Mycobacterium tuberculosis* TrcR/TrcS proteins. Microb Pathog **26:**195-206.

88. **Heermann, R., K. Altendorf, and K. Jung.** 1998. The turgor sensor KdpD of *Escherichia coli* is a homodimer. Biochim Biophys Acta **1415:**114-24.
89. **Heffernan, E. J., J. Fierer, G. Chikami, and D. Guiney.** 1987. Natural history of oral *Salmonella dublin* infection in BALB/c mice: effect of an 80-kilobase-pair plasmid on virulence. J Infect Dis **155:**1254-9.
90. **Higgs, P. I., S. Jagadeesan, P. Mann, and D. R. Zusman.** 2008. EspA, an orphan hybrid histidine protein kinase, regulates the timing of expression of key developmental proteins of *Myxococcus xanthus*. J Bacteriol.
91. **Hilton, J. L., P. C. Kearney, and B. N. Ames.** 1965. Mode of action of the herbicide, 3-amino-1,2,4-triazole(amitrole): inhibition of an enzyme of histidine biosynthesis. Arch Biochem Biophys **112:**544-7.
92. **Hines, M. T., K. M. Paasch, D. C. Alperin, G. H. Palmer, N. C. Westhoff, and S. A. Hines.** 2001. Immunity to *Rhodococcus equi*: antigen-specific recall responses in the lungs of adult horses. Vet Immunol Immunopathol **79:**101-14.
93. **Hines, S. A., S. T. Kanaly, B. A. Byrne, and G. H. Palmer.** 1997. Immunity to *Rhodococcus equi*. Vet Microbiol **56:**177-85.
94. **Hines, S. A., D. M. Stone, M. T. Hines, D. C. Alperin, D. P. Knowles, L. K. Norton, M. J. Hamilton, W. C. Davis, and T. C. McGuire.** 2003. Clearance of virulent but not avirulent *Rhodococcus equi* from the lungs of adult horses is associated with intracytoplasmic gamma interferon production by CD4+ and CD8+ T lymphocytes. Clin Diagn Lab Immunol **10:**208-15.
95. **Ho, Y. S., L. M. Burden, and J. H. Hurley.** 2000. Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor. EMBO J **19:**5288-99.
96. **Hochschild, A., and S. L. Dove.** 1998. Protein-protein contacts that activate and repress prokaryotic transcription. Cell **92:**597-600.
97. **Hondalus, M. K.** 1997. Pathogenesis and virulence of *Rhodococcus equi*. Vet Microbiol **56:**257-68.
98. **Hondalus, M. K., and D. M. Mosser.** 1994. Survival and replication of *Rhodococcus equi* in macrophages. Infect Immun **62:**4167-75.
99. **Hong, C. B., and J. M. Donahue.** 1995. *Rhodococcus equi*-associated necrotizing lymphadenitis in a llama. J Comp Pathol **113:**85-8.
100. **Hooper-McGrevy, K. E., B. N. Wilkie, and J. F. Prescott.** 2003. Immunoglobulin G subisotype responses of pneumonic and healthy, exposed foals and adult horses to *Rhodococcus equi* virulence-associated proteins. Clin Diagn Lab Immunol **10:**345-51.
101. **Hooper-McGrevy, K. E., B. N. Wilkie, and J. F. Prescott.** 2005. Virulence-associated protein-specific serum immunoglobulin G-isotype expression in young foals protected against *Rhodococcus equi* pneumonia by oral immunization with virulent R. equi. Vaccine **23:**5760-7.
102. **Huang, J., W. Yindeeyoungyeon, R. P. Garg, T. P. Denny, and M. A. Schell.** 1998. Joint transcriptional control of xpsR, the unusual signal integrator of the *Ralstonia solanacearum* virulence gene regulatory network, by a response regulator and a LysR-type transcriptional activator. J Bacteriol **180:**2736-43.
103. **Huffman, J. L., and R. G. Brennan.** 2002. Prokaryotic transcription regulators: more than just the helix-turn-helix motif. Curr Opin Struct Biol **12:**98-106.
104. **Hughes, K. L., and I. Sulaiman.** 1987. The ecology of *Rhodococcus equi* and physicochemical influences on growth. Vet Microbiol **14:**241-50.

105. **Hulko, M., F. Berndt, M. Gruber, J. U. Linder, V. Truffault, A. Schultz, J. Martin, J. E. Schultz, A. N. Lupas, and M. Coles.** 2006. The HAMP domain structure implies helix rotation in transmembrane signaling. Cell **126:**929-40.
106. **Jacks, S., S. Giguere, P. C. Crawford, and W. L. Castleman.** 2007. Experimental infection of neonatal foals with *Rhodococcus equi* triggers adult-like gamma interferon induction. Clin Vaccine Immunol **14:**669-77.
107. **Jacks, S., S. Giguere, and J. F. Prescott.** 2007. In vivo expression of and cell-mediated immune responses to the plasmid-encoded virulence-associated proteins of *Rhodococcus equi* in foals. Clin Vaccine Immunol **14:**369-74.
108. **Jaeger, S., G. Eriani, and F. Martin.** 2004. Results and prospects of the yeast three-hybrid system. FEBS Lett **556:**7-12.
109. **Jain, S., B. R. Bloom, and M. K. Hondalus.** 2003. Deletion of vapA encoding Virulence Associated Protein A attenuates the intracellular actinomycete *Rhodococcus equi*. Mol Microbiol **50:**115-28.
110. **James, P., J. Halladay, and E. A. Craig.** 1996. Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. Genetics **144:**1425-36.
111. **Janga, S. C., and J. Collado-Vides.** 2007. Structure and evolution of gene regulatory networks in microbial genomes. Res Microbiol **158:**787-94.
112. **Jenny, R. J., K. G. Mann, and R. L. Lundblad.** 2003. A critical review of the methods for cleavage of fusion proteins with thrombin and factor Xa. Protein Expr Purif **31:**1-11.
113. **John R. Cann, O. P., David E. Pettijohn.** 1995. Theory of the mobility-shift assay of nonspecific protein-DNA complexes governed by conditional probabilities: The HU:DNA complex. Electrophoresis **Volume 16:**881-887.
114. **Jung, K., and K. Altendorf.** 2002. Towards an understanding of the molecular mechanisms of stimulus perception and signal transduction by the KdpD/KdpE system of *Escherichia coli*. J Mol Microbiol Biotechnol **4:**223-8.
115. **K Wilson, F. A., R Brent, RE Kingston** 1990. Current Protocols in Molecular Biology. Greene Publishing and Wiley-Interscience.
116. **Kanacher, T., A. Schultz, J. U. Linder, and J. E. Schultz.** 2002. A GAF-domain-regulated adenylyl cyclase from Anabaena is a self-activating cAMP switch. EMBO J **21:**3672-80.
117. **Kanaly, S. T., S. A. Hines, and G. H. Palmer.** 1995. Cytokine modulation alters pulmonary clearance of *Rhodococcus equi* and development of granulomatous pneumonia. Infect Immun **63:**3037-41.
118. **Kasuga-Aoki, H., S. Takai, Y. Sasaki, S. Tsubaki, H. Madarame, and A. Nakane.** 1999. Tumour necrosis factor and interferon-gamma are required in host resistance against virulent *Rhodococcus equi* infection in mice: cytokine production depends on the virulence levels of R. equi. Immunology **96:**122-7.
119. **Keegan, L., G. Gill, and M. Ptashne.** 1986. Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. Science **231:**699-704.
120. **Khorchid, A., and M. Ikura.** 2006. Bacterial histidine kinase as signal sensor and transducer. Int J Biochem Cell Biol **38:**307-12.
121. **Kishii, R., L. Falzon, T. Yoshida, H. Kobayashi, and M. Inouye.** 2007. Structural and functional studies of the HAMP domain of EnvZ, an osmosensing transmembrane histidine kinase in *Escherichia coli*. J Biol Chem **282:**26401-8.

122. **Klopotowski, T., and A. Wiater.** 1965. Synergism of aminotriazole and phosphate on the inhibition of yeast imidazole glycerol phosphate dehydratase. Arch Biochem Biophys **112:**562-6.
123. **Klug, A.** 1999. Zinc finger peptides for the regulation of gene expression. J Mol Biol **293:**215-8.
124. **Komijn, R. E., H. J. Wisselink, V. M. Rijsman, N. Stockhofe-Zurwieden, D. Bakker, F. G. van Zijderveld, T. Eger, J. A. Wagenaar, F. F. Putirulan, and B. A. Urlings.** 2007. Granulomatous lesions in lymph nodes of slaughter pigs bacteriologically negative for Mycobacterium avium subsp. avium and positive for *Rhodococcus equi*. Vet Microbiol **120:**352-7.
125. **Kovacikova, G., and K. Skorupski.** 1999. A *Vibrio cholerae* LysR homolog, AphB, cooperates with AphA at the tcpPH promoter to activate expression of the ToxR virulence cascade. J Bacteriol **181:**4250-6.
126. **Kreikemeyer, B., K. S. McIver, and A. Podbielski.** 2003. Virulence factor regulation and regulatory networks in *Streptococcus pyogenes* and their impact on pathogen-host interactions. Trends Microbiol **11:**224-32.
127. **Kvint, K., L. Nachin, A. Diez, and T. Nystrom.** 2003. The bacterial universal stress protein: function and regulation. Curr Opin Microbiol **6:**140-5.
128. **Landsman, D.** 1992. RNP-1, an RNA-binding motif is conserved in the DNA-binding cold shock domain. Nucleic Acids Res **20:**2861-4.
129. **Latchman, D. S.** 1997. Transcription factors: an overview. Int J Biochem Cell Biol **29:**1305-12.
130. **Laughon, A., and R. F. Gesteland.** 1982. Isolation and preliminary characterization of the GAL4 gene, a positive regulator of transcription in yeast. Proc Natl Acad Sci U S A **79:**6827-31.
131. **Lee-Chiong, T., M. Sadigh, M. Simms, and G. Buller.** 1995. Case reports: pericarditis and lymphadenitis due to *Rhodococcus equi*. Am J Med Sci **310:**31-3.
132. **Lopez, A. M., M. T. Hines, G. H. Palmer, D. C. Alperin, and S. A. Hines.** 2002. Identification of pulmonary T-lymphocyte and serum antibody isotype responses associated with protection against *Rhodococcus equi*. Clin Diagn Lab Immunol **9:**1270-6.
133. **Lopez, A. M., H. G. Townsend, A. L. Allen, and M. K. Hondalus.** 2008. Safety and immunogenicity of a live-attenuated auxotrophic candidate vaccine against the intracellular pathogen *Rhodococcus equi*. Vaccine **26:**998-1009.
134. **Lozada-Chavez, I., S. C. Janga, and J. Collado-Vides.** 2006. Bacterial regulatory networks are extremely flexible in evolution. Nucleic Acids Res **34:**3434-45.
135. **Machang'u, R. S., and J. F. Prescott.** 1991. Role of antibody to extracellular proteins of *Rhodococcus equi* in protection against R. equi pneumonia in foals. Vet Microbiol **26:**323-33.
136. **Madan Babu, M., S. A. Teichmann, and L. Aravind.** 2006. Evolutionary dynamics of prokaryotic transcriptional regulatory networks. J Mol Biol **358:**614-33.
137. **Magnusson, H.** 1938. Pyaemia in foals caused by *Corynebacterium equi. Vet. Rec.***:**1459-1468.
138. **Marchiori, E., N. L. Muller, R. G. de Mendonca, D. Capone, A. S. Souza, Jr., D. L. Escuissato, E. L. Gasparetto, and E. M. de Cerqueira.** 2005. *Rhodococcus equi* pneumonia in AIDS: high-resolution CT findings in five patients. Br J Radiol **78:**783-6.

139. **Maris, C., C. Dominguez, and F. H. Allain.** 2005. The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. FEBS J **272:**2118-31.
140. **Marodi, L.** 2002. Down-regulation of Th1 responses in human neonates. Clin Exp Immunol **128:**1-2.
141. **Mascher, T., J. D. Helmann, and G. Unden.** 2006. Stimulus perception in bacterial signal-transducing histidine kinases. Microbiol Mol Biol Rev **70:**910-38.
142. **Matsumoto, A., S. K. Hong, H. Ishizuka, S. Horinouchi, and T. Beppu.** 1994. Phosphorylation of the AfsR protein involved in secondary metabolism in *Streptomyces* species by a eukaryotic-type protein kinase. Gene **146:**47-56.
143. **Medhurst, A. L., P. A. Huber, Q. Waisfisz, J. P. de Winter, and C. G. Mathew.** 2001. Direct interactions of the five known Fanconi anaemia proteins suggest a common functional pathway. Hum Mol Genet **10:**423-9.
144. **Meeuse, J. J., H. G. Sprenger, S. van Assen, D. Leduc, S. M. Daenen, J. P. Arends, and T. S. van der Werf.** 2007. *Rhodococcus equi* infection after alemtuzumab therapy for T-cell prolymphocytic leukemia. Emerg Infect Dis **13:**1942-3.
145. **Meijer, W. G., and J. F. Prescott.** 2004. *Rhodococcus equi*. Vet Res **35:**383-96.
146. **Mendes, M. V., S. Tunca, N. Anton, E. Recio, A. Sola-Landa, J. F. Aparicio, and J. F. Martin.** 2007. The two-component phoR-phoP system of *Streptomyces natalensis*: Inactivation or deletion of phoP reduces the negative phosphate regulation of pimaricin biosynthesis. Metab Eng **9:**217-27.
147. **Mikic, D.** 2006. [*Rhodococcus equi* infection]. Vojnosanit Pregl **63:**957-62.
148. **Miller, J., and I. Stagljar.** 2004. Using the yeast two-hybrid system to identify interacting proteins. Methods Mol Biol **261:**247-62.
149. **Mizuno, T.** 1998. His-Asp phosphotransfer signal transduction. J Biochem **123:**555-63.
150. **Mizuno, T., and T. Suzuki.** 2002. [His-Asp phosphorelay signal transduction]. Tanpakushitsu Kakusan Koso **47:**113-24.
151. **Mizuno, Y., F. Sato, M. Sakamoto, K. Yoshikawa, M. Yoshida, K. Shiba, S. Onodera, R. Matsuura, and S. Takai.** 2005. VapB-positive *Rhodococcus equi* infection in an HIV-infected patient in Japan. J Infect Chemother **11:**37-40.
152. **Mohapatra, N. P., S. Soni, B. L. Bell, R. Warren, R. K. Ernst, A. Muszynski, R. W. Carlson, and J. S. Gunn.** 2007. Identification of an orphan response regulator required for the virulence of *Francisella spp*. and transcription of pathogenicity island genes. Infect Immun **75:**3305-14.
153. **Mosser, D. M., and M. K. Hondalus.** 1996. *Rhodococcus equi*: an emerging opportunistic pathogen. Trends Microbiol **4:**29-33.
154. **Mougel, C., and I. B. Zhulin.** 2001. CHASE: an extracellular sensing domain common to transmembrane receptors from prokaryotes, lower eukaryotes and plants. Trends Biochem Sci **26:**582-4.
155. **Mulder, N. J., R. Apweiler, T. K. Attwood, A. Bairoch, D. Barrell, A. Bateman, D. Binns, M. Biswas, P. Bradley, P. Bork, P. Bucher, R. R. Copley, E. Courcelle, U. Das, R. Durbin, L. Falquet, W. Fleischmann, S. Griffiths-Jones, D. Haft, N. Harte, N. Hulo, D. Kahn, A. Kanapin, M. Krestyaninova, R. Lopez, I. Letunic, D. Lonsdale, V. Silventoinen, S. E. Orchard, M. Pagni, D. Peyruc, C. P. Ponting, J. D. Selengut, F. Servant, C. J. Sigrist, R. Vaughan, and E. M. Zdobnov.** 2003. The Interpro Database, 2003 brings increased coverage and new features. Nucleic Acids Res **31:**315-8.

156. **Munoz-Dorado, J., S. Inouye, and M. Inouye.** 1991. A gene encoding a protein serine/threonine kinase is required for normal development of *M. xanthus*, a gram-negative bacterium. Cell **67:**995-1006.
157. **Munoz, P., J. Palomo, J. Guinea, J. Yanez, M. Giannella, and E. Bouza.** 2008. Relapsing *Rhodococcus equi* infection in a heart transplant recipient successfully treated with long-term linezolid. Diagn Microbiol Infect Dis **60:**197-9.
158. **Nadvornik, R., T. Vomastek, J. Janecek, Z. Technikova, and P. Branny.** 1999. Pkg2, a novel transmembrane protein Ser/Thr kinase of *Streptomyces granaticolor*. J Bacteriol **181:**15-23.
159. **Nagy, G., U. Dobrindt, G. Schneider, A. S. Khan, J. Hacker, and L. Emody.** 2002. Loss of regulatory protein RfaH attenuates virulence of uropathogenic Escherichia coli. Infect Immun **70:**4406-13.
160. **Nariya, H., and S. Inouye.** 2005. Identification of a protein Ser/Thr kinase cascade that regulates essential transcriptional activators in *Myxococcus xanthus* development. Mol Microbiol **58:**367-79.
161. **O'Toole, R., and H. D. Williams.** 2003. Universal stress proteins and Mycobacterium tuberculosis. Res Microbiol **154:**387-92.
162. **Ohta, N., and A. Newton.** 2003. The core dimerization domains of histidine kinases contain recognition specificity for the cognate response regulator. J Bacteriol **185:**4424-31.
163. **Park, S. H., and R. T. Raines.** 2004. Fluorescence gel retardation assay to detect protein-protein interactions. Methods Mol Biol **261:**155-60.
164. **Parrish, J. R., K. D. Gulyas, and R. L. Finley, Jr.** 2006. Yeast two-hybrid contributions to interactome mapping. Curr Opin Biotechnol **17:**387-93.
165. **Pate, M., T. Pirs, I. Zdovc, B. Krt, and M. Ocepek.** 2004. Haemolytic *Rhodococcus equi* Isolated from a Swine Lymph Node with Granulomatous Lesions. J Vet Med B Infect Dis Vet Public Health **51:**249-50.
166. **Perkins, G. A., A. Yeager, H. N. Erb, D. V. Nydam, T. J. Divers, and J. L. Bowman.** 2002. Survival of foals with experimentally induced *Rhodococcus equi* infection given either hyperimmune plasma containing R. equi antibody or normal equine plasma. Vet Ther **3:**334-46.
167. **Perron-Savard, P., G. De Crescenzo, and H. Le Moual.** 2005. Dimerization and DNA binding of the Salmonella enterica PhoP response regulator are phosphorylation independent. Microbiology **151:**3979-87.
168. **Petrickova, K., and M. Petricek.** 2003. Eukaryotic-type protein kinases in *Streptomyces coelicolor*: variations on a common theme. Microbiology **149:**1609-21.
169. **Phumoonna, T., M. D. Barton, and M. W. Heuzenroeder.** 2005. Recognition of a B-cell epitope of the VapA protein of *Rhodococcus equi* in newborn and experimentally infected foals. J Vet Med B Infect Dis Vet Public Health **52:**291-5.
170. **Polidori, M., and A. Haas.** 2006. VapI, a new member of the *Rhodococcus equi* Vap family. Antonie Van Leeuwenhoek **90:**299-304.
171. **Ponting, C. P., J. Schultz, F. Milpetz, and P. Bork.** 1999. SMART: identification and annotation of domains from signalling and extracellular protein sequences. Nucleic Acids Res **27:**229-32.
172. **Pradip, I. S., A. D. Larson, and C. S. McClesky. 1966., and** 1966. Nutritional factors affecting growth and pigmentation of *Corynebacteriumequi*. Bacteriol. Proc. **1:**20.

173. **Prescott, J. F.** 1987. Epidemiology of *Rhodococcus equi* infection in horses. Vet Microbiol **14:**211-4.
174. **Prescott, J. F.** 1991. *Rhodococcus equi*: an animal and human pathogen. Clin Microbiol Rev **4:**20-34.
175. **Puthucheary, S. D., V. Sangkar, A. Hafeez, R. Karunakaran, N. S. Raja, and H. H. Hassan.** 2006. *Rhodococcus equi*--an emerging human pathogen in immunocompromized hosts: a report of four cases from Malaysia. Southeast Asian J Trop Med Public Health **37:**157-61.
176. **Rajagopal, L., A. Vo, A. Silvestroni, and C. E. Rubens.** 2006. Regulation of cytotoxin expression by converging eukaryotic-type and two-component signalling mechanisms in *Streptococcus agalactiae*. Mol Microbiol **62:**941-57.
177. **Ramos, J. L., M. Martinez-Bueno, A. J. Molina-Henares, W. Teran, K. Watanabe, X. Zhang, M. T. Gallegos, R. Brennan, and R. Tobes.** 2005. The TetR family of transcriptional repressors. Microbiol Mol Biol Rev **69:**326-56.
178. **Ren, J., and J. F. Prescott.** 2003. Analysis of virulence plasmid gene expression of intra-macrophage and in vitro grown *Rhodococcus equi* ATCC 33701. Vet Microbiol **94:**167-82.
179. **Ren, J., and J. F. Prescott.** 2004. The effect of mutation on *Rhodococcus equi* virulence plasmid gene expression and mouse virulence. Vet Microbiol **103:**219-30.
180. **Roder, K., and M. Schweizer.** 2001. Running-buffer composition influences DNA-protein and protein-protein complexes detected by electrophoretic mobility-shift assay (EMSA). Biotechnol Appl Biochem **33:**209-14.
181. **Rozsypal, H., V. Aster, M. Stahikova, and B. Horova.** 2007. [*Rhodococcus equi* infection in subjects infected with human immunodeficiencv virus (HIV)]. Cas Lek Cesk **146:**163-7.
182. **Ruscher, K., M. Reuter, D. Kupper, G. Trendelenburg, U. Dirnagl, and A. Meisel.** 2000. A fluorescence based non-radioactive electrophoretic mobility shift assay. J Biotechnol **78:**163-70.
183. **Russell, D. A., G. A. Byrne, E. P. O'Connell, C. A. Boland, and W. G. Meijer.** 2004. The LysR-type transcriptional regulator VirR is required for expression of the virulence gene vapA of *Rhodococcus equi* ATCC 33701. J Bacteriol **186:**5576-84.
184. **Saini, D. K., V. Malhotra, and J. S. Tyagi.** 2004. Cross talk between DevS sensor kinase homologue, Rv2027c, and DevR response regulator of *Mycobacterium tuberculosis*. FEBS Lett **565:**75-80.
185. **Sauer, R. T.** 1996. Lac repressor at last. Structure **4:**219-22.
186. **Sawai, R., A. Suzuki, Y. Takano, P. C. Lee, and S. Horinouchi.** 2004. Phosphorylation of AfsR by multiple serine/threonine kinases in *Streptomyces coelicolor* A3(2). Gene **334:**53-61.
187. **Schell, M. A.** 1993. Molecular biology of the LysR family of transcriptional regulators. Annu Rev Microbiol **47:**597-626.
188. **Schiestl, R. H., and R. D. Gietz.** 1989. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. Curr Genet **16:**339-46.
189. **Schmitt, C. K., S. C. Darnell, V. L. Tesh, B. A. Stocker, and A. D. O'Brien.** 1994. Mutation of flgM attenuates virulence of *Salmonella typhimurium*, and mutation of fliA represses the attenuated phenotype. J Bacteriol **176:**368-77.

190. **Schneider, M. M.** 2000. Change in natural history of opportunistic infections in HIV-infected patients. Neth J Med **57:**1-3.
191. **Schultz, J., F. Milpetz, P. Bork, and C. P. Ponting.** 1998. SMART, a simple modular architecture research tool: identification of signaling domains. Proc Natl Acad Sci U S A **95:**5857-64.
192. **Scott, V., A. R. Clark, and K. Docherty.** 1994. The gel retardation assay. Methods Mol Biol **31:**339-47.
193. **Secchi, C., F. Pereira, L. R. Perez, P. A. d'Azevedo, and S. da Silva Rios.** 2006. [Bacteremia due to *Rhodococcus equi* in a patient with acquired immunodeficiency syndrome: case report]. Rev Soc Bras Med Trop **39:**570-2.
194. **Senear, D. F., and M. Brenowitz.** 1991. Determination of binding constants for cooperative site-specific protein-DNA interactions using the gel mobility-shift assay. J Biol Chem **266:**13661-71.
195. **Shah, R. M., A. V. Kaji, B. J. Ostrum, and A. C. Friedman.** 1997. Interpretation of chest radiographs in AIDS patients: usefulness of CD4 lymphocyte counts. Radiographics **17:**47-58; discussion 59-61.
196. **Shenoy, A. R.** 2000. His kinase or mine? histidine kinases through evolution. J Biosci **25:**317-22.
197. **Shi, L., and F. M. Hulett.** 1999. The cytoplasmic kinase domain of PhoR is sufficient for the low phosphate-inducible expression of pho regulon genes in *Bacillus subtilis*. Mol Microbiol **31:**211-22.
198. **Shi, L., and W. Zhang.** 2004. Comparative analysis of eukaryotic-type protein phosphatases in two streptomycete genomes. Microbiology **150:**2247-56.
199. **Singh, A., D. Mai, A. Kumar, and A. J. Steyn.** 2006. Dissecting virulence pathways of *Mycobacterium tuberculosis* through protein-protein association. Proc Natl Acad Sci U S A **103:**11346-51.
200. **Smith, B. P., and R. C. Robinson.** 1981. Studies of an outbreak of *Corynebacterium equi* pneumonia in foals. Equine Vet J **13:**223-8.
201. **Soler-Palacin, P., C. Margareto, P. Llobet, O. Asensio, M. Hernandez, I. Caragol, and T. Espanol.** 2007. Chronic granulomatous disease in pediatric patients: 25 years of experience. Allergol Immunopathol (Madr) **35:**83-9.
202. **Spiro, S.** 2007. Regulators of bacterial responses to nitric oxide. FEMS Microbiol Rev **31:**193-211.
203. **Stagljar, I.** 2003. Finding partners: emerging protein interaction technologies applied to signaling networks. Sci STKE **2003:**pe56.
204. **Stekel, D. J., and D. J. Jenkins.** 2008. Strong negative self regulation of prokaryotic transcription factors increases the intrinsic noise of protein expression. BMC Syst Biol **2:**6.
205. **Stock, A. M., V. L. Robinson, and P. N. Goudreau.** 2000. Two-component signal transduction. Annu Rev Biochem **69:**183-215.
206. **Strunk, T., K. Gardiner, K. Simmer, D. Atlas, and A. D. Keil.** 2007. *Rhodococcus equi* meningitis after ventriculoperitoneal shunt insertion in a preterm infant. Pediatr Infect Dis J **26:**1076-7.
207. **Sumner-Smith, M., R. P. Bozzato, N. Skipper, R. W. Davies, and J. E. Hopper.** 1985. Analysis of the inducible MEL1 gene of *Saccharomyces carlsbergensis* and its secreted product, alpha-galactosidase (melibiase). Gene **36:**333-40.

208. **Takai, S.** 1996. [Current concepts and implications for future research on *Rhodococcus equi*]. Nippon Saikingaku Zasshi **51:**485-96.
209. **Takai, S.** 1997. Epidemiology of *Rhodococcus equi* infections: a review. Vet Microbiol **56:**167-76.
210. **Takai, S., T. Anzai, Y. Fujita, O. Akita, M. Shoda, S. Tsubaki, and R. Wada.** 2000. Pathogenicity of *Rhodococcus equi* expressing a virulence-associated 20 kDa protein (VapB) in foals. Vet Microbiol **76:**71-80.
211. **Takai, S., T. Fujimori, K. Katsuzaki, and S. Tsubaki.** 1987. Ecology of *Rhodococcus equi* in horses and their environment on horse-breeding farms. Vet Microbiol **14:**233-9.
212. **Takai, S., N. Fukunaga, K. Kamisawa, Y. Imai, Y. Sasaki, and S. Tsubaki.** 1996. Expression of virulence-associated antigens of *Rhodococcus equi* is regulated by temperature and pH. Microbiol Immunol **40:**591-4.
213. **Takai, S., S. A. Hines, T. Sekizaki, V. M. Nicholson, D. A. Alperin, M. Osaki, D. Takamatsu, M. Nakamura, K. Suzuki, N. Ogino, T. Kakuda, H. Dan, and J. F. Prescott.** 2000. DNA sequence and comparison of virulence plasmids from *Rhodococcus equi* ATCC 33701 and 103. Infect Immun **68:**6840-7.
214. **Takai, S., M. Iie, Y. Watanabe, S. Tsubaki, and T. Sekizaki.** 1992. Virulence-associated 15- to 17-kilodalton antigens in *Rhodococcus equi*: temperature-dependent expression and location of the antigens. Infect Immun **60:**2995-7.
215. **Takai, S., S. Iimori, and S. Tsubaki.** 1986. Quantitative fecal culture for early diagnosis of *Corynebacterium* (*Rhodococcus*) *equi* enteritis in foals. Can J Vet Res **50:**479-84.
216. **Takai, S., H. Madarame, C. Matsumoto, M. Inoue, Y. Sasaki, Y. Hasegawa, S. Tsubaki, and A. Nakane.** 1995. Pathogenesis of *Rhodococcus equi* infection in mice: roles of virulence plasmids and granulomagenic activity of bacteria. FEMS Immunol Med Microbiol **11:**181-90.
217. **Takai, S., T. Sekizaki, T. Ozawa, T. Sugawara, Y. Watanabe, and S. Tsubaki.** 1991. Association between a large plasmid and 15- to 17-kilodalton antigens in virulent *Rhodococcus equi*. Infect Immun **59:**4056-60.
218. **Takeda, S.** 2004. [High quality and high throughput yeast two-hybrid system]. Tanpakushitsu Kakusan Koso **49:**2751-7.
219. **Temme, K., H. Salis, D. Tullman-Ercek, A. Levskaya, S. H. Hong, and C. A. Voigt.** 2008. Induction and relaxation dynamics of the regulatory network controlling the type III secretion system encoded within *Salmonella* pathogenicity island 1. J Mol Biol **377:**47-61.
220. **Terpe, K.** 2003. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. Appl Microbiol Biotechnol **60:**523-33.
221. **Tkachuk-Saad, O., P. Lusis, R. D. Welsh, and J. F. Prescott.** 1998. *Rhodococcus equi* infections in goats. Vet Rec **143:**311-2.
222. **Tkachuk-Saad, O., and J. Prescott.** 1991. *Rhodococcus equi* plasmids: isolation and partial characterization. J Clin Microbiol **29:**2696-700.
223. **Tomenius, H., A. K. Pernestig, C. F. Mendez-Catala, D. Georgellis, S. Normark, and O. Melefors.** 2005. Genetic and functional characterization of the *Escherichia coli* BarA-UvrY two-component system: point mutations in the HAMP linker of the BarA sensor give a dominant-negative phenotype. J Bacteriol **187:**7317-24.

224. **Toyooka, K., S. Takai, and T. Kirikae.** 2005. *Rhodococcus equi* can survive a phagolysosomal environment in macrophages by suppressing acidification of the phagolysosome. J Med Microbiol **54:**1007-15.
225. **Traven, A., B. Jelicic, and M. Sopta.** 2006. Yeast Gal4: a transcriptional paradigm revisited. EMBO Rep **7:**496-9.
226. **Udo, H., J. Munoz-Dorado, M. Inouye, and S. Inouye.** 1995. *Myxococcus xanthus*, a gram-negative bacterium, contains a transmembrane protein serine/threonine kinase that blocks the secretion of beta-lactamase by phosphorylation. Genes Dev **9:**972-83.
227. **Ulivieri, S., and G. Oliveri.** 2006. Cerebellar abscess due to *Rhodococcus equi* in an immunocompetent patient: case report and literature review. J Neurosurg Sci **50:**127-9.
228. **Umeyama, T., and S. Horinouchi.** 2001. Autophosphorylation of a bacterial serine/threonine kinase, AfsK, is inhibited by KbpA, an AfsK-binding protein. J Bacteriol **183:**5506-12.
229. **van Keulen, G., L. Girbal, E. R. van den Bergh, L. Dijkhuizen, and W. G. Meijer.** 1998. The LysR-type transcriptional regulator CbbR controlling autotrophic CO2 fixation by *Xanthobacter flavus* is an NADPH sensor. J Bacteriol **180:**1411-7.
230. **van Keulen, G., A. N. Ridder, L. Dijkhuizen, and W. G. Meijer.** 2003. Analysis of DNA binding and transcriptional activation by the LysR-type transcriptional regulator CbbR of *Xanthobacter flavus*. J Bacteriol **185:**1245-52.
231. **Venner, M., B. Meyer-Hamme, J. Verspohl, F. Hatori, N. Shimizu, Y. Sasaki, T. Kakuda, S. Tsubaki, and S. Takai.** 2007. Genotypic characterization of VapA positive *Rhodococcus equi* in foals with pulmonary affection and their soil environment on a warmblood horse breeding farm in Germany. Res Vet Sci **83:**311-7.
232. **Verville TD, H. M., GreenWeld RA, Fine DP, Kuhls TL, Slater LN** 1994. *Rhodococcus equi* infections of humans Medicine **73:**119–131.
233. **Voelkner, P., W. Puppe, and K. Altendorf.** 1993. Characterization of the KdpD protein, the sensor kinase of the K(+)-translocating Kdp system of *Escherichia coli*. Eur J Biochem **217:**1019-26.
234. **Walters, S. B., E. Dubnau, I. Kolesnikova, F. Laval, M. Daffe, and I. Smith.** 2006. The *Mycobacterium tuberculosis* PhoPR two-component system regulates genes essential for virulence and complex lipid biosynthesis. Mol Microbiol **60:**312-30.
235. **Wang, J., C. Li, H. Yang, A. Mushegian, and S. Jin.** 1998. A novel serine/threonine protein kinase homologue of *Pseudomonas aeruginosa* is specifically inducible within the host infection site and is required for full virulence in neutropenic mice. J Bacteriol **180:**6764-8.
236. **Watts, K. J., Q. Ma, M. S. Johnson, and B. L. Taylor.** 2004. Interactions between the PAS and HAMP domains of the *Escherichia coli* aerotaxis receptor Aer. J Bacteriol **186:**7440-9.
237. **Weiss, V., G. Kramer, T. Dunnebier, and A. Flotho.** 2002. Mechanism of regulation of the bifunctional histidine kinase NtrB in Escherichia coli. J Mol Microbiol Biotechnol **4:**229-33.
238. **West, A. H., and A. M. Stock.** 2001. Histidine kinases and response regulator proteins in two-component signaling systems. Trends Biochem Sci **26:**369-76.
239. **Wiater, A., D. Hulanicka, and T. Klopotowski.** 1971. Structural requirements for inhibition of yeast imidazoleglycerol phosphate dehydratase by triazole and anion inhibitors. Acta Biochim Pol **18:**289-97.

240. **Wolanin, P. M., P. A. Thomason, and J. B. Stock.** 2002. Histidine protein kinases: key signal transducers outside the animal kingdom. Genome Biol **3:**REVIEWS3013.
241. **Woolcock, J. B., M. D. Mutimer, and A.-M. T. Farmer.** 1980. Epidemiology of *Corynebacterium equi* in horses. . Res. Vet. Sci. **:**87-90.
242. **Yang, Y., D. P. Harris, F. Luo, L. Wu, A. B. Parsons, A. V. Palumbo, and J. Zhou.** 2008. Characterization of the *Shewanella oneidensis* Fur gene: roles in iron and acid tolerance response. BMC Genomics **9 Suppl 1:**S11.
243. **Yarmolinsky, M.** 2000. Transcriptional silencing in bacteria. Curr Opin Microbiol **3:**138-43.
244. **Young, K. H.** 1998. Yeast two-hybrid: so many interactions, (in) so little time. Biol Reprod **58:**302-11.
245. **Zaim, J., and A. M. Kierzek.** 2003. The structure of full-length LysR-type transcriptional regulators. Modeling of the full-length OxyR transcription factor dimer. Nucleic Acids Res **31:**1444-54.
246. **Zhang, C. C.** 1996. Bacterial signalling involving eukaryotic-type protein kinases. Mol Microbiol **20:**9-15.
247. **Zhang, C. C.** 1993. A gene encoding a protein related to eukaryotic protein kinases from the filamentous heterocystous cyanobacterium *Anabaena* PCC 7120. Proc Natl Acad Sci U S A **90:**11840-4.
248. **Zhang, W., M. Inouye, and S. Inouye.** 1996. Reciprocal regulation of the differentiation of *Myxococcus xanthus* by Pkn5 and Pkn6, eukaryotic-like Ser/Thr protein kinases. Mol Microbiol **20:**435-47.
249. **Zhu, Y., and M. Inouye.** 2002. The role of the G2 box, a conserved motif in the histidine kinase superfamily, in modulating the function of EnvZ. Mol Microbiol **45:**653-63.
250. **Zhulin, I. B., A. N. Nikolskaya, and M. Y. Galperin.** 2003. Common extracellular sensory domains in transmembrane receptors for diverse signal transduction pathways in bacteria and archaea. J Bacteriol **185:**285-94.
251. **Zink, M. C., J. A. Yager, and N. L. Smart.** 1986. *Corynebacterium equi* infections in horses,1958-1984: a review of 131 cases. . Can. Vet. J. **34** 213-217.

Appendix

Appendix A: Plasmids used in this study.

Plasmid	Fusions	Epitope	Yeast selection marker	Bacterial Selection marker	Source/reference
pGBKT7	GAL4 (1-147) DNA-BD	c-myc	TRP1	Kan^R	Chien *et al.* (1991)
pGADT7	GAL4 (768-881) AD	HA	LEU2	Amp^R	Louret *et al.* (1997)
p*Requi*	GAL4AD:400-2000bp *R.equi* genomic library fragments.	HA	LEU2	Amp^R	Shruti Jain
pIA13	GAL4BD:Orf8W	c-myc	TRP1	Kan^R	This study
pIA14	GAL4BD:Orf8P	c-myc	TRP1	Kan^R	This study
pIA15	GAL4BD:HK1	c-myc	TRP1	Kan^R	This study
pIA16	GAL4BD:HK2	c-myc	TRP1	Kan^R	This study
pIA17	GAL4BD:HK3	c-myc	TRP1	Kan^R	This study
pIA18	GAL4BD:HK4	c-myc	TRP1	Kan^R	This study
pIA19	GAL4BD:HK5	c-myc	TRP1	Kan^R	This study
pIA20	GAL4BD:HK6	c-myc	TRP1	Kan^R	This study
pIA21	GAL4BD:HK7	c-myc	TRP1	Kan^R	This study
pIA22	GAL4BD:HK8	c-myc	TRP1	Kan^R	This study
pIA23	GAL4BD:HK9	c-myc	TRP1	Kan^R	This study
pIA24	GAL4BD:HK10	c-myc	TRP1	Kan^R	This study
pIA25	GAL4BD:HK11	c-myc	TRP1	Kan^R	This study
pIA26	GAL4BD:HK12	c-myc	TRP1	Kan^R	This study
pIA27	GAL4BD:HK13	c-myc	TRP1	Kan^R	This study
pIA28	GAL4BD:HK14	c-myc	TRP1	Kan^R	This study
pIA29	GAL4BD:HK15	c-myc	TRP1	Kan^R	This study
pIA30	GAL4BD:HK16	c-myc	TRP1	Kan^R	This study
pIA31	GAL4BD:HK17	c-myc	TRP1	Kan^R	This study
pIA32	GAL4BD:HK18	c-myc	TRP1	Kan^R	This study
pIA33	GAL4BD:HK19	c-myc	TRP1	Kan^R	This study
pIA34	GAL4BD:HK20	c-myc	TRP1	Kan^R	This study
pIA35	GAL4AD:Orf8P	HA	LEU2	Amp^R	This study
pIA36	GAL4AD:Orf8W	HA	LEU2	Amp^R	This study
pLAM5'-1	GAL4BD::human Lamin C protein	c-myc	TRP1	Kan^R	Clontech
pVA3	Gal4 DNA-BD::murine p53 protein	c-myc	TRP1	Amp^R	Clontech
pTD1	GAL4 AD::SV40 large T-antigen	HA	LEU2	Amp^R	Clontech

Appendix B: Primers and *R. equi* gene coordinates of cloned genes.

PLASMID	FUSION PROTEIN		5' PRIMER / gene coordinates synthesized by celtek genes		3'PRIMER
pIA13	GAL4BD:Orf8W	fVarA1	ATG CCA TGG TTA TGG AGT CCA TTC GCG GTA T	rVarA1	ATA GTC GAC ATG CCG TGG CCC TGC CAA
pIA14	GAL4BD:Orf8P	fVarA1	ATG CCA TGG TTA TGG AGT CCA TTC GCG GTA T	rVarArec	ATA GTC GAC GAT ACG CAA TGC CAG CTC AC
pIA15	GAL4BD:HK1	fHK1	TAT CCA TGG GTG AGG AGT CGG CGC GCC GGT C	rHK1	ATT GAA TTC CCC TCT CGG CAG ACG CAC
pIA16*	GAL4BD:HK2		celtek genes (gene coordinates 983739 - 984719)		n/a
pIA18	GAL4BD:HK4	fHK4	ATTCCATGGGTGGCACCACCGTGGGAC	rHK4	ATTGAATTCGCGTCCCTCGGCCTCGGGCT
pIA19	GAL4BD:HK5	fHK5	ATT CCA TGG GTG TGT ACC AGT TGC TGG TGC GGG AG	rHK5	ATT GAA TTC CAG CGG CCT CCC CAC CGG AAG T
pIA20*	GAL4BD:HK6		celtek genes (gene coordinates* 1277283 - 1277939)		n/a
pIA22*	GAL4BD:HK8		celtek genes (gene coordinates* 3304990 - 3306045)		n/a
pIA23	GAL4BD:HK9	fHK9	ATT CCA TGG GTG CGC AGA CCG AAC GTC TC	rHK9	ATT GAA TTC TGC CTG GGG CAA TGC AAC
pIA24	GAL4BD:HK10	fHK10	ATT CCA TGG GTC GTC AGG TCG TGT TG	rHK10	ATT GAA TTC CGA TTC CCT CTC CCG CAC CCT
pIA25*	GAL4BD:HK11		celtek genes (gene coordinates* 35992506 - 3592026)		n/a
pIA26*	GAL4BD:HK12	fHK12	ATT CCA TGG GTC CCC GAC TGA GCA GGC GTC AC	rHK12	ATT GAA TTC ACG GTT GGC CTC CAG
pIA28*	GAL4BD:HK14		celtek genes (gene coordinates 1158511 - 1159188)		n/a
pIA29*	GAL4BD:HK15		celtek genes (gene coordinates 143585 - 142994)		n/a
pIA30*	GAL4BD:HK16		celtek genes (gene coordinates 395565 - 396566)		n/a
pIA31	GAL4BD:HK17	fHK17	ATT CCA TGG GTG AAG AGG AGG CCG GTG AG	rHK17	ATT GAA TTC GGC GAG CGG GGC GTT CGG GAT
pIA32	GAL4BD:HK18	fHK18	ATT CCA TGG GTA CCC TCG GAC TCG ACC CCG ACC A	rHK18	ATT GAA TTC TCG ATC CTC GAC CTC CCC GGA ATC GT
pIA33*	GAL4BD:HK19		celtek genes (gene coordinates 111949 - 112599)		n/a
pIA34*	GAL4BD:HK20		celtek genes (gene coordinates 643347 - 642805)		n/a
pIA35	GAL4AD:Orf8P		celtek genes (gene coordinates 111,949 - 112599)		n/a
pIA36	GAL4AD:Orf8W		celtek genes (gene coordinates 111,949 - 112599)		n/a

* gene is encoded on the anti-sense strand.

Appendix C: Bacterial and yeast strains used in this study.

Strains	**Relevant phenotype/genotype**
E.coli TAM1	*mcr* A D(*mrr -hsd* RMS-*mcr* BC) F80*lac* ZDM15 D*lac* X74 *rec* A1 *araD139 D(ara-leu)7697 galU galK rpsL endA1 nupG*
S. cerevisiae pJ69-4A	MATA trp1-901 leu2-3,112 ura3-52 his3-200 gal4(deleted) gal80(deleted) LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ
S. cerevisiae pJ69-4alpha	MATalpha trp1-901 leu2-3,112 ura3-52 his3-200 gal4(deleted) gal80(deleted) LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ

Appendix D: Media formulations for yeast two hybrid selection.

SD-medium (Synthetic Dropout) - 1 Liter

6.7 g/L Yeast nitrogen base (Difco laboratories, Detroit, MI)
15 g/L Bacto agar (Difco laboratories, Detroit, MI) for plates only
20 g/L Dextrose (Difco laboratories)
variable Complete Supplement Mixture (CSM) of appropriate amino acids for selection (Bio101, Vista, Calif.).

CSM-Leu (0.69 g/L)
CSM-Trp (0.74 g/L)
CSM-Leu-trp (0.64 g/L)
CSM-Ade-His-Leu-Trp (0.61 g /L)