

## Suggested project topics

These project topics are broadly stated, leaving you scope to define a specific phylogenetic hypotheses that you will then explore for the remainder of the course. Some of these are broad enough to encompass several different phylogenetic questions. Projects on the same topic should not overlap extensively. Phylogenetics projects will be carried out in teams of 1 or 2 people. Let me know once you choose a topic to make sure it is still available. I may group people up who are interested in the same topic if there are sufficient phylogenetic questions and gene families (otherwise you may be asked to pick a different topic). I can also help you find a project partner, if you identify a phylogenetic question that requires analysis of two or more gene families.

Once you have a project topic, it may be useful to read the papers linked on each topic to prepare to write your project plan. Your project plan should include a brief discussion of what is known about the gene family (or families) you intend to study, your phylogenetic question(s), and the trees you intend to build. If you are working with a partner, your plan should include at least two gene families and also include a coordination plan. Partners need only turn in one project plan and it can be up to two pages long.

- Email me your proposed project topic and team members by **November 7th**.
- Your one page project plan is due **November 14<sup>th</sup>**.

## Distribution and Phylogenetics of Stage 0 Sporulation genes in Firmicutes

Signal transduction pathways allow bacteria to sense and respond to environmental challenges. A common type of bacterial signaling pathway, the two-component signaling system, consists of a sensor kinase and a response regulator which transduce a specific signal by passing a phosphoryl group. Phosphorylation activates the response regulator causing it to initiate an appropriate response to the incoming signal.

The signaling pathway that initiates spore formation in *Bacillus subtilis* is unusual in that it consists of four proteins, instead of just two. In this so-called phosphorelay, the phosphoryl group passes through two intermediate proteins on its path from the sensor to response regulator. This unusual pathway architecture is found throughout the Bacilli and has recently been discovered in Clostridia, as well<sup>1</sup>.

It has been hypothesized that these intermediate proteins provide additional control points supporting more nuanced regulation of sporulation. A number of potential regulators have

---

<sup>1</sup> Evolutionary remodeling of the sporulation initiation pathway. Davidson P. 2017 Jul. PhD thesis, Carnegie Mellon University.

been identified<sup>2,3</sup>, though mostly in the phosphorelay type-species, *B. subtilis*. This includes multiple families of known regulatory proteins that interact directly with the phosphorelay proteins, including at least one family that directly inhibits kinase activity and two different types of aspartyl phosphatases (which draw phosphoryl groups out of the phosphorelay). However, little is known about the regulators of the Spo0 phosphorelay outside of *Bacillus subtilis*. Is this system of regulation at intermediate control points a conserved feature of the system? Or have specialized regulatory programs evolved in different species, allowing for lineage specific adaptation to different ecological niches?

In addition to those regulators that act directly on the Spo0 phosphorelay, there are more general regulatory proteins that, when inactivated, arrest sporulation in its earliest stages, including proteins involved in chromosome segregation, GTP/GDP sensors, regulators of alternative differentiation pathways, and others with unknown function. What is the distribution of these alternative regulatory proteins in this diverse phylum? Are they vertically inherited with the Spo0 pathway or are they lineage specific regulators?

### **Comparative evolution: methylamine methyltransferases with the pyrrolysine synthesis system**

Some methylotrophic organisms exploit methylamines, a non-standard carbon source. Known methylotrophs include bacterial organisms and methanogenic archaea. These organisms encode at least one of three known methylamine methyltransferase pathway. The first step in each of these pathways is encoded by a protein that includes a non-standard amino acid (pyrrolysine) in its active site. The special chemical properties of pyrrolysine are required for the methyltransferase function to proceed. In order to include the non-standard amino acid into these proteins, these genomes must also encode the Pyrrolysine biosynthesis pathway and a Pyl-tRNA synthetase.

Pyrrolysine incorporation machinery and pyrrolysine-methyltransferase enzymes are functionally linked as both are necessary for methyltrophism. In turn, methyltrophism likely provides a strong fitness advantage in the adverse environments in which methanogens are typically identified. Taken together, the linked function and strong fitness advantage suggest that these two sets of proteins are maintained together. Do these sets of proteins share an evolutionary history?

There are three types of methylamine methyltransferase pathways that are specific for the three different moieties of methylamine (i.e. mono-, di-, or tri-methylamine). Some genomes contain more than one of these pathways, but their exact distribution is not well-studied. Did these three proteins share a common ancestor that also required pyrrolysine? Or are they the result of convergent evolution from disparate ancestors that led to 1) convergent incorporation

---

<sup>2</sup> Genomic determinants of sporulation in *Bacilli* and *Clostridia*: towards the minimal set of sporulation-specific genes. Galperin M, Sergei M, Puigbo P, Smirnov S, Wolf Y, Rigden D. *EnvMicroBio*. 2012 14(11), 28-70-2890. doi:10.1111/j.1462-2920.2012.02841.x.

<sup>3</sup> Genome diversity of spore-forming *Firmicutes*. Galperin M. *MicrobiolSpectrum*. 1(2):TBS-0015. doi:10.1128/microbiolspectrum.TBS-0015-2012.

of pyrrolysine and 2) similar function on different substrates? Regardless of their origin, do these proteins co-evolve with each other following their acquisition?

### **Evolution of yeast scaffolding proteins, FAR1 and STE5**

STE5 and FAR1 are scaffolding proteins with similar structural organization and related, but distinct roles in the mating pathway in Baker's yeast, *Saccharomyces cerevisiae*. The human pathogen, *Candida albicans*, has two analogous proteins, as do other closely related yeast species in the Saccharomycotina. More distantly related fungi have only one Ste5/Far1-like scaffolding protein. This raises the intriguing question of whether the mating pathway functions performed by two scaffolding proteins in Baker's yeast are all carried out by a single protein in those species. A recent study by Cote<sup>4</sup> and colleagues provides a thoughtful analysis of the evolution of fungal mating pathway scaffolding proteins, but includes no phylogenetic analysis. Their study is based on simple counting arguments that can lead to incorrect conclusions. Your phylogenetic analysis of FAR1 and STE5 has the potential to either confirm their results or offer a new hypothesis for the evolution of this pathway.

This could be a project for a multiperson team or two separate projects.

### **Evolution of a toxin-antitoxin system in *Streptococcus***

Toxin-antitoxin (TA) and Abortive infection (Abi) systems are mechanisms of bacterial defense against invasion by mobile elements<sup>5</sup>. In the absence of the antitoxin, the toxin causes cell death or inhibits growth. Since the antitoxin is degraded more rapidly than the toxin, upon inhibition of the antitoxin by foreign DNA, the toxin causes the cell to commit altruistic suicide. An Abi system was recently discovered in the genomes of *Streptococcus pneumoniae* strains isolated from conjunctivitis (Antic et al., mSphee, in press). Further, the genomes of all these are genetically isolated from other *S. pneumo* strains.

This raises the question of whether this Abi system contributes to isolation of these strains. Your phylogenetic analysis of the two components of this Abi has the potential to better understand strain isolation. It will also elucidate the source of the Abi system in these strains, whether from vertical descent or horizontal gene transfer, and many provide good candidates for studying kin recognition in these strains.

---

<sup>4</sup> Evolutionary Reshaping of Fungal Mating Pathway Scaffold Proteins. Côte P, Sulea T, Dignard D, Wu C, Whiteway M. *MBio*. 2011 Jan 11;2(1):e00230-10. doi: 10.1128/mBio.00230-10.

<sup>5</sup> A widespread bacteriophage abortive infection system functions through a Type IV toxin-antitoxin mechanism. Dy R, Przybilski R, Semeijn K, Salmond G, Fineran P. *NAR*, 2014 Jan 24;42(7). doi:10.1093/nar/gk1419.