Niche Adaptation in the Dynamics and Evolution of a Bacterial Stress Response

Amoolya H. Singh[1,3], Peggy Wang[1], Adam P. Arkin[1,2]
[1]Depts of Bioengineering and Computer Science, UC Berkeley, Berkeley CA [2]Physical Biosciences Division, Lawrence Berkeley Lab, Berkeley CA
[3]Present address: Structural & Computational Biology Unit, European Molecular Biology Lab, Heidelberg, Germany

ESPP Environmental Stress Pathway Project

VIMSS Virtual Institute for Microbial Stress and Survival · Miami University · Missouri · University of Washington · DIVERSA · Sandia National Laboratories · Oak Ridge National Laboratory · Berkeley · http://vimss.lbl.gov/

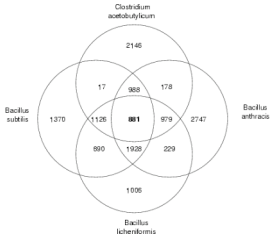
DOE Genomics:GTL — Accelerating Discovery for Energy and Environment

Overview

Bacterial endospore formation is a complex developmental pathway involving differential and staged expression of hundreds of genes in response to overcrowding or nutritional stress. Most of the insights into the genetics and molecular biology of this pathway have come from studies of *Bacillus subtilis*. However, as new genomic and expression data for other spore-formers becomes available, one may ask to what extent this pathway is conserved phylogenetically, over evolutionary time, as well as transcriptionally, over physiological time. Here, we collect and compare genomic and expression data from four spore-formers (*B. subtilis, B. anthracis, C. acetobutylicum, B. licheniformis*) using a combination of statistical and phylogenetic methods. Preliminary results show that genes sharing expression patterns are not always evolutionarily conserved, but that genes sharing phylogenetic patterns are typically co-expressed. Furthermore, the sporulation pathway shows clear evidence of niche modifications in that the free-living but phylogenetically divergent bacteria (*B. subtilis* and *C. acetobutylicum*) have significantly different expression patterns in conserved genes as compared to the more closely related pathogenic bacteria (*B. licheniformis* and *B. anthracis*). These results shed light on the conservation of spore-formation in terms of its temporal and ecological roles, and provide an early insight into the evolution of pathway dynamics.

Methods and Data

Orthologs shared by four spore-formers

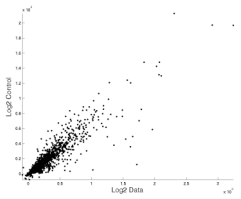


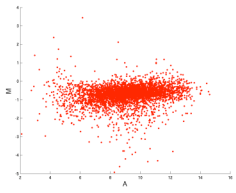
- Expression data normalization (global normalization per slide, geometric average between replicates, arithmetic normalization between conditions)
- Significance assessed by t-statistic when replicates available, z-score otherwise
- Deviation from normal distribution measured using normal probability plot

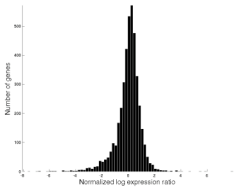
Microarray data sources (52 expts)

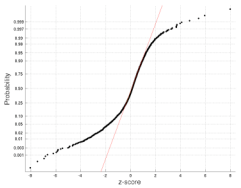
Table 1: Microarray data sources.

No.	Species	Experiment	Ref.
1–6	*B. subtilis*	Wild type and mutant strains (mutant for sigmaE, sigmaK, GerE, GerK, or SpoIIIG) were grown and induced to sporulate. www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE1629, acc=GPL1353	(2, 3)
7–14	*B. subtilis*	Global transcriptional response of wt and perR mutant cells 3, 18, 20, and 40 min after exposure to low (H_2O_2M) and intermediate (50μM) levels of peroxide (H_2O_2) with high (100μM) levels of tert-butyl peroxide (t-bOOH). www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL2740	(1)
15–22	*B. subtilis*	Transcriptional profiling during sporulation of spo0A and sigF mutants and wt cells at timepoints -1.5, 0, 2 hours relative to the onset of sporulation. Transcriptional profiling of cells with a Spo0A mutant (Spo0A Sas267) that does not require phosphorylation, at timepoints 0, 30, 60 minutes relative to the onset of sporulation. mcb.harvard.edu/Losick/Pubs/FawcettPaperMainPage.htm	(4)
23–37	*B. anthracis*	Gene expression of B. anthracis StarnesMP2 strain grown in nutrient-poor medium to promote sporulation. Cells examined at logarithmic growth to sporulation, corresponding to 15 minutes interval time points up to 5 hours. www.ncbi.nlm.nih.gov/geo/gds/gds_browse.cgi?gds=929, gds=930	(9)
38–51	*C. acetobutylicum*	Transcriptome of sporulation for a wt strain grown in a fermenter with controlled pH $\geq$ 5 (9 timepoints spanning 15 hours of exponential-growth, transitional, and early stationary phases), and a non-fermenting, non-sporulating strain (M5) grown in a static flask with no pH control (5 timepoints spanning 20 hours exponential-growth, transitional, and early stationary phases). jb.asm.org/cgi/content/full/187/20/7103/DC1	(1)
52–53	*B. licheniformis*	Increasing pH stress in wt strain (pH 6.0, 7.0, 8.0) and stationary phase development in inocula propagated for 12 and 24 h. dx.doi.org/10.1016/j.femsle.2004.05.035	(6, 7)





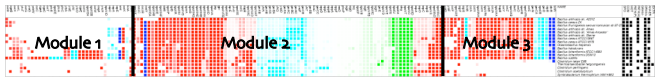


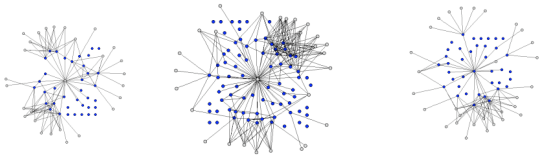


Results

Evolutionary modules are transcriptionally co-expressed

Phylogenetic profile of 153 sporulation genes in 18 fully sequenced spore formers yields 3 major evolutionary modules with distinct pathway functions
For genes in each module, how are their expression profiles correlated? (|r| <= 0.7)
Measure clustering coefficient of resulting graph, compare to graph of same number of randomly chosen genes. Results are significant at P<0.05.

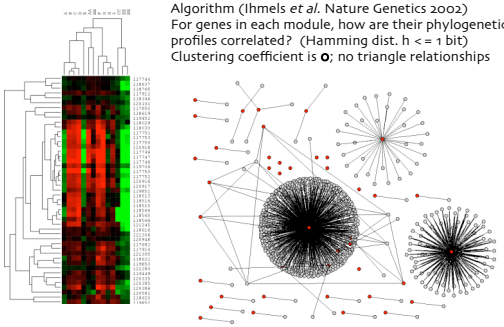




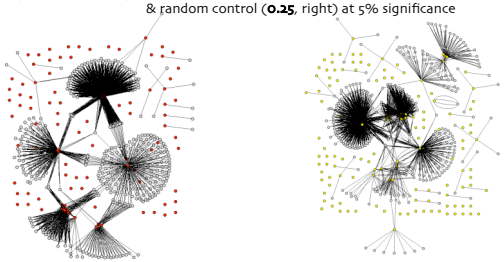
37 genes	77 genes	39 genes
response modulation	regulation, spore coat	cross-talk, sensing
0.35 (0.05)	0.35 (0.19)	0.46 (0.25)

Transcriptional modules are not evolutionarily conserved

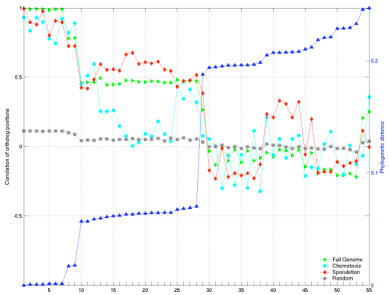
Sample transcriptional module from *C. acetobutylicum* (49 genes) obtained using Iterative Signature Algorithm (Ihmels *et al.* Nature Genetics 2002)
For genes in each module, how are their phylogenetic profiles correlated? (Hamming dist. h <= 1 bit)
Clustering coefficient is **0**; no triangle relationships



Sample transcriptional module from *B. subtilis* (154 genes) obtained as above
For genes in each module, how are their phylogenetic profiles correlated? (Hamming dist. h <= 1 bit)
Clustering coefficient comparable between data (**0.28**, left)
& random control (**0.25**, right) at 5% significance



Genome location is surprisingly conserved



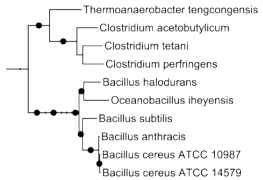
Well known that genomic order not conserved in microbes in general due to chromosomal rearrangements, inversions, insertions, and deletions (Wolf *et al.* Genome Research 2001). However, in subset of genes enriched for cellular function (e.g. sporulation pathway), significant conservation of operon structure, chromosomal location, and gene order.

X-axis: pair of sporulation genomes
Y-axis: correlation coefficient of scatterplot of ortholog positions on chromosome

Differences reflect niche adaptation

Hierarchical clustering of entire expression data set (not shown) indicates that free-living spore-formers cluster separately from pathogenic spore-formers, distinct from taxonomic relationships shown right. Lifestyle and niche adaptation of bacteria thus play a role in shaping its transcriptional profile, even for identical pathways.

ML tree of 31 essential genes, dots signifying bootstrap > 80%, pruned from tree of life (Ciccarelli *et al.* Science 2006).



Discussion

Several areas of future work:
-Include multiple testing analysis in microarray data
-Systematically analyze noise resulting from use of diverse data sets and different kinds of chips
-Incorporate current knowledge from dynamical model of sporulation, to elucidate which aspects of transcriptional correlation are significant versus accidental

Despite these limitations, the major contribution of this work is that it begins to incorporate expression dynamics into the lexicon of evolution. That is, knowing the dynamics of one network, can we predict what it will be in homologous networks? This leads to the interesting and intricate question of how to systematically compare dynamics, given the difficulty of model-building in multiple species, gathering experimental conditions and parameters, and understanding valid parameter combinations and ranges. Finally, it provides a framework for developing effective predictive tools that bridge the gap between physiological and evolutionary timescales.

Acknowledgment

ESPP is part of the Virtual Institute for Microbial Stress and Survival supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomics Program:GTL through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy.