Gradual evolution in bacteria: Evidence from Bacillus systematics

M. Feldgarden  
*Broad Institute*

N. Byrd  
*United Therapeutics*

Frederick M. Cohan  
*Wesleyan University, fcohan@wesleyan.edu*

Follow this and additional works at: [https://wesscholar.wesleyan.edu/div3facpubs](https://wesscholar.wesleyan.edu/div3facpubs)

Part of the Biodiversity Commons, Ecology and Evolutionary Biology Commons, and the Environmental Microbiology and Microbial Ecology Commons

Recommended Citation

*Division III Faculty Publications*. 345.  
[https://wesscholar.wesleyan.edu/div3fcpubs/345](https://wesscholar.wesleyan.edu/div3fcpubs/345)

This Article is brought to you for free and open access by the Natural Sciences and Mathematics at WesScholar. It has been accepted for inclusion in Division III Faculty Publications by an authorized administrator of WesScholar. For more information, please contact nmealey@wesleyan.edu, jmlozanowski@wesleyan.edu.
Gradual evolution in bacteria: evidence from *Bacillus* systematics

Michael Feldgarden,† Noah Byrd‡ and Frederick M. Cohan

Department of Biology, Wesleyan University, Middletown, CT 06459-0170, USA

The bacterial genome projects have suggested a central role for horizontal transfer in bacterial adaptation, but it is difficult to rule out an adaptive role for ordinary genetic change in existing genes. The bacterial systematics literature can readily address the importance of gene acquisition in adaptive evolution, since phenotypic characterization typically assesses presence versus absence of metabolic capabilities, and metabolic gains and losses are most likely due to horizontal transfer and/or gene loss. Bacterial systematists have not geared their studies toward quantitative differences in metabolic capabilities, which are more likely to involve adjustments of existing genes. Here, quantitative variation in metabolism within and between three closely related *Bacillus* taxa has been assayed. While these taxa show no qualitative (i.e. presence versus absence) differences in resource utilization, they are quantitatively different in utilization of 8% of 95 resources tested. Moreover, 93% of the resources tested showed significant quantitative variation among strains within a single taxon. These results suggest that ordinary genetic changes in existing genes may play an important role in adaptation. If these results are typical, future genomically based assays of quantitative variation in phenotype (e.g. microarray analysis of mRNA concentrations) may identify hundreds of genes whose expression has been modified. A protocol is presented for identifying those modifications of gene expression and those gene acquisitions that are most likely to have played a role in adaptive evolution.

INTRODUCTION

The peculiar nature of sex in bacteria opens each bacterium to a world of adaptations. Owing to the promiscuity of bacterial sex, bacteria can take up existing adaptations from widely divergent donor species (Ochman *et al.*, 2000). Moreover, because bacterial recombination is infrequent, sexual promiscuity incurs no significant cost to fitness; uptake of incompatible genes from other species is much too rare to disrupt a species' existing adaptations (Cohan, 1994a).

The bacterial genome projects have suggested a central role for horizontal transfer in bacterial adaptation (Lawrence, 2002; Ochman *et al.*, 2000). For example, much of the capacity for virulence in *Salmonella* stems from horizontally acquired 'pathogenicity islands' (Baumler *et al.*, 1997; Ochman & Groisman, 1996). The potential importance of horizontal transfer is suggested by the observation that a typical bacterial genome has received at least 5–10% of its genes from extremely divergent donors (Ochman *et al.*, 2000). Lawrence (1997, 1999) has argued that bacteria evolve into new niches principally by acquiring new gene loci and operons from other species.

It is difficult, however, to rule out a role for ordinary genetic change in bacterial adaptation. Like the eukaryotes, bacteria can evolve adaptations by modifying the genes they already have. In some cases, bacteria have evolved adaptations by substituting, through recombination, an existing allele with a homologous allele from another species. Penicillin resistance has been acquired in this way in *Streptococcus* and *Neisseria* (Maynard Smith *et al.*, 1991). Mutations have also played a role in bacterial adaptation. For example, synonymous substitutions have optimized mRNA secondary structure (Katz & Burge, 2003) and amino acid substitutions have minimized the use of energetically costly amino acids (Akashi & Gojobori, 2002; Seligmann, 2003).

In other cases, mutations have allowed a strain to step into a new ecological niche. For example, Sokurenko *et al.* (1998) have shown that single point mutations in the FimH adhesin can enable pathogenesis in *Escherichia coli*. Mutations have also been shown to routinely allow invasion of new niches...
in laboratory microcosms of evolution. Several laboratories have found that a clone and its descendants, when allowed to evolve on their own (and without benefit of recombination), inevitably undergo an adaptive radiation: the original clone evolves into a diversity of populations, each with its own ecological niche (Rainey & Travisano, 1998; Rozen & Lenski, 2000; Treves et al., 1998). Thus, some niche invasions are clearly accessible through mutation. Here we demonstrate a role for systematics in assessing the relative importance of heterologous horizontal transfer versus modification of existing genes in bacterial evolution.

The bacterial systematics literature can readily address the importance of gene acquisition in bacterial evolution, but it is of little help in addressing the importance of mutation and other modifications of existing genes. This is because phenotypic characterization of strains and species has been biased toward assessing presence versus absence of metabolic capabilities. Gains and losses of metabolic capabilities are generally due to heterologous horizontal transfer (as well as deletions of genes) (Lawrence, 1997, 1999; Ochman & Groisman, 1996), although some capabilities can be acquired through point mutation (Hall, 1999; Hall & Malik, 1998; Spiers et al., 2002).

Bacterial systematists have generally not geared their studies toward quantitative comparisons of capabilities – for example, by testing whether strains and species differ in the rate of fermentation or growth on various sugars. Such quantitative variation, when it exists, is likely due to changes in existing genes (e.g. by increasing the transcription of an existing operon) (Cooper et al., 2003; Ferea et al., 1999). To our knowledge, assessment of quantitative variation among and within bacterial species has been limited to comparing the percentage of different fatty acids in whole cells (Nakamura, 1998; Roberts et al., 1994, 1996); we know of no extensive systematic studies of quantitative variation in metabolic capabilities in bacteria.

We will presently demonstrate significant quantitative variation in metabolism both within and between three closely related Bacillus taxa, all formerly identified as Bacillus subtilis. These taxa, B. subtilis subsp. subtilis, B. subtilis subsp. spizizenii and Bacillus mojavensis, are sympatric in some but not all habitats (Roberts & Cohan, 1995). Their identification as separate taxa resulted from our discovery that they form distinct sequence clusters for three protein-coding genes (Nakamura et al., 1999; Roberts et al., 1994). As with all bacterial sequence clusters that retain their distinctness in sympathy, these taxa are likely to represent ecologically distinct populations (Dykhuizen, 1998; Palys et al., 1997). The geographical distributions of these taxa also suggest that these groups are ecologically distinct: although these taxa are capable of extremely high rates of dispersal (Roberts & Cohan, 1995), B. subtilis subsp. spizizenii is the only one of these three taxa to be isolated from Death Valley, CA, USA; also, B. mojavensis is the only one of these taxa to be found at high frequencies near our Hohat collection site in the Gobi Desert (Roberts & Cohan, 1995).

While these taxa appear to fill different ecological niches, the nature of the ecological differences is unknown. These taxa are not distinguishable by known differences in any metabolic capability at the level of presence versus absence. Here we compare these taxa, as well as the outgroup Bacillus licheniformis, for quantitative differences in the ability to utilize 95 carbon sources; we also analyse these taxa for growth rates in different salt concentrations. We will show that, while the three closely related taxa of our ingroup show no qualitative differences in resource utilization, they are quantitatively different in utilization of 8% of the resources tested, and they are different in their levels of salt tolerance.

**METHODS**

**Strains.** The strains used in this study are listed in Table 1, and they have been described previously (Duncan et al., 1994; Istock et al., 1996; Roberts & Cohan, 1995).

**Metabolic measurements.** Biolog GP MicroPlates were used to determine the ability of strains to utilize 95 different carbon sources, including two alcohols, four amides, 11 amino acids, five aromatic compounds, 44 carbohydrates, 13 carboxylic acids, three esters, seven phosphorylated compounds and six polymers (Bochner, 1989). Cells were grown in Luria Broth for 12 h at 30°C, at which point they were washed twice in saline (0.8% NaCl). Then, 150 μl of the saline suspension was inoculated into each of the 96 wells, and the plates were incubated at 30°C for 24 h. Each strain was assayed three times on separate dates. Analysis of variance (ANOVA) indicated that there were no date effects. The OD595 was determined at the time of inoculation and 24 h later using a Bio-Rad 3550 UV microtitre plate reader. The metabolic capability of each strain with each metabolite was quantified as ΔOD595, the starting OD595 subtracted from the final OD595.

We took into account the possibility that apparent differences among strains in their metabolic capabilities might be an artefact of strain differences in their adaptation to the base medium used in all Biolog tests. That is, if some strains were better adapted to Biolog's base medium, these strains might show greater change in OD595 on all metabolites. We therefore adjusted the raw metabolic capability score (ΔOD595) for each strain (i) on each metabolite (j) by the strain's mean raw score over all metabolites τj.

\[
\hat{x}_{ij} = x_{ij} - \bar{x}_j,
\]

where \(x_{ij}\) and \(x_j\) are strain i's adjusted and raw ΔOD595 scores, respectively, on metabolite j. This adjusted score quantifies a strain's ability to metabolize a particular substrate relative to its ability to metabolize all of the substrates tested. We have reported analyses of adjusted metabolic data, but have also indicated wherever results based on raw and adjusted scores are qualitatively different.

**Salinity tolerance measurements.** We assessed salinity tolerance by measuring growth rates in Luria Broth adjusted to a final salt concentration of 1°C, 2°C, 5°C and 10°C. In each well, 4 x 10^6 cells were added to 0.2 ml of growth medium. The optical density of cultures was read by using a Bio-Rad 3550 UV microtitre plate reader.

**Statistical analyses.** Analyses were performed using JMP 3.1.5 STATISTICAL DISCOVERY SOFTWARE (SAS Institute, 1996).
RESULTS

Metabolism of individual resources

Based on raw, unadjusted metabolic scores, *B. subtilis* subsp. *spizizenii* showed greater (but not necessarily significantly greater) metabolic capability on 79 of 95 metabolites than both *B. subtilis* subsp. *subtilis* and *B. mojavensis* (see Table 2 for data on 22 metabolites; other data not shown). This result suggested the possibility that *B. subtilis* subsp. *spizizenii* might simply be better adapted to Biolog base medium than the other taxa. This also raised the possibility that metabolic differences among other taxa and among strains might be artefacts of differences in level of adaptation to Biolog base medium. To control for differences in adaptation to the base medium, we adjusted each metabolic OD\textsubscript{595} score for a strain by its average performance on all metabolites (see Methods).

Metabolic measurements for each resource were analysed separately by ANOVA using a nested model: strains were nested within taxa, with taxon and strain both treated as random effects. Because 95 ANOVAs were performed, we reduced the significance values using the sequential Bonferroni adjustment (Rice, 1989). Most carbon sources (88/95) revealed significant differences among strains of the same taxon (data not shown). Using adjusted ΔOD\textsubscript{595} scores, 12 of 95 carbon sources (13 %) showed significant variation among taxa; using raw ΔOD\textsubscript{595} scores, 19 of 95 carbon sources (20 %) showed significant variation among taxa (Table 2). Nine of the 95 carbon sources (9 %) showed significant variation among taxa by both raw and adjusted analysis – this is our most conservative estimate for the fraction of metabolites for which the taxa are significantly different.

The nine carbon sources with inter-taxon differences were widely distributed across categories of metabolites: taxa were significantly different in metabolism of three of the 11 amino acids, two of the 44 carbohydrates, three of the 13 carboxylic acids and one of the seven phosphorylated compounds.

We next addressed whether the differences among taxa indicated by the ANOVA were due primarily to inclusion of the outgroup, *B. licheniformis*. Of the 22 carbon sources that demonstrated taxon-specific differences (by analysis of either raw or adjusted ΔOD\textsubscript{595} scores), only two such

### Table 1. Strains used in this study

All strains were collected from desert soil. Strains from Tucson were isolated by Conrad Istock; all other strains were isolated by F. M. Cohan.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>RO-A-4</td>
<td>Rosamond, CA, USA</td>
</tr>
<tr>
<td>RO-J-2</td>
<td>Rosamond, CA, USA</td>
</tr>
<tr>
<td>RO-NN-1</td>
<td>Rosamond, CA, USA</td>
</tr>
<tr>
<td>RS-C-4</td>
<td>Rosamond, CA, USA</td>
</tr>
<tr>
<td>RS-D-2</td>
<td>Rosamond, CA, USA</td>
</tr>
<tr>
<td>T-89-14</td>
<td>Tucson, AZ, USA</td>
</tr>
<tr>
<td>T-89-18</td>
<td>Tucson, AZ, USA</td>
</tr>
<tr>
<td>T-89-48</td>
<td>Tucson, AZ, USA</td>
</tr>
<tr>
<td>T-F-32</td>
<td>Tucson, AZ, USA</td>
</tr>
<tr>
<td>T-G-2-5</td>
<td>Tucson, AZ, USA</td>
</tr>
<tr>
<td>TU-C-7</td>
<td>Nefta, Tunisia</td>
</tr>
<tr>
<td>DV1-B-1</td>
<td>Death Valley, CA, USA</td>
</tr>
<tr>
<td>DV2-A-2</td>
<td>Death Valley, CA, USA</td>
</tr>
<tr>
<td>DV3-E-3</td>
<td>Death Valley, CA, USA</td>
</tr>
<tr>
<td>DV4-D-5</td>
<td>Death Valley, CA, USA</td>
</tr>
<tr>
<td>DV6-A-2-3</td>
<td>Death Valley, CA, USA</td>
</tr>
<tr>
<td>IM-C-216</td>
<td>Hohat, Inner Mongolia</td>
</tr>
<tr>
<td>RO-E-2</td>
<td>Rosamond, CA, USA</td>
</tr>
<tr>
<td>RO-G-4</td>
<td>Rosamond, CA, USA</td>
</tr>
<tr>
<td>RO-PP-2</td>
<td>Rosamond, CA, USA</td>
</tr>
<tr>
<td>RO-SS-1</td>
<td>Rosamond, CA, USA</td>
</tr>
<tr>
<td>RS-B-1</td>
<td>Rosamond, CA, USA</td>
</tr>
<tr>
<td>T-88-10</td>
<td>Tucson, AZ, USA</td>
</tr>
<tr>
<td>T-88-8</td>
<td>Tucson, AZ, USA</td>
</tr>
<tr>
<td>T-89-3</td>
<td>Tucson, AZ, USA</td>
</tr>
<tr>
<td>B. subtilis subsp. spizizenii</td>
<td></td>
</tr>
<tr>
<td>T-G-1-16</td>
<td>Tucson, AZ, USA</td>
</tr>
<tr>
<td>T-T-1-23</td>
<td>Tucson, AZ, USA</td>
</tr>
<tr>
<td>TU-B-8</td>
<td>Nefta, Tunisia</td>
</tr>
<tr>
<td>TU-D-6</td>
<td>Nefta, Tunisia</td>
</tr>
<tr>
<td>TU-D-8</td>
<td>Nefta, Tunisia</td>
</tr>
<tr>
<td>TU-E-6</td>
<td>Nefta, Tunisia</td>
</tr>
<tr>
<td>TU-F-6</td>
<td>Nefta, Tunisia</td>
</tr>
<tr>
<td>B. mojavensis</td>
<td></td>
</tr>
<tr>
<td>RO-B-2</td>
<td>Rosamond, CA, USA</td>
</tr>
<tr>
<td>RO-C-2</td>
<td>Rosamond, CA, USA</td>
</tr>
<tr>
<td>RO-H-1</td>
<td>Rosamond, CA, USA</td>
</tr>
<tr>
<td>RO-QQ-2</td>
<td>Rosamond, CA, USA</td>
</tr>
<tr>
<td>RS-A-2</td>
<td>Rosamond, CA, USA</td>
</tr>
<tr>
<td>T-G-2-42</td>
<td>Tucson, AZ, USA</td>
</tr>
<tr>
<td>T-G-3-41</td>
<td>Tucson, AZ, USA</td>
</tr>
<tr>
<td>B. licheniformis</td>
<td></td>
</tr>
<tr>
<td>DV5-A-1-4</td>
<td>Death Valley, CA, USA</td>
</tr>
<tr>
<td>DV5-A-2-2</td>
<td>Death Valley, CA, USA</td>
</tr>
<tr>
<td>DV7-B-2</td>
<td>Death Valley, CA, USA</td>
</tr>
<tr>
<td>DV8-3-1</td>
<td>Death Valley, CA, USA</td>
</tr>
<tr>
<td>Rr1</td>
<td>Tucson, AZ, USA</td>
</tr>
<tr>
<td>T-88-14</td>
<td>Tucson, AZ, USA</td>
</tr>
<tr>
<td>T-89-27</td>
<td>Tucson, AZ, USA</td>
</tr>
<tr>
<td>T-89-36</td>
<td>Tucson, AZ, USA</td>
</tr>
<tr>
<td>T-89-37</td>
<td>Tucson, AZ, USA</td>
</tr>
</tbody>
</table>
Table 2. Mean metabolic capability (ΔOD₉₅₅) of each taxon, for the 22 metabolites showing significant variation among taxa, based on either raw or adjusted ΔOD₉₅₅ scores

The raw metabolic score is the upper number and the adjusted score is the lower number. The standard errors were calculated based on the variance among strains.

<table>
<thead>
<tr>
<th>Substrate</th>
<th>B. subtilis subsp. subtilis</th>
<th>B. subtilis subsp. spizizenii</th>
<th>B. mojavensis</th>
<th>B. licheniformis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Amide</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Succinamic acid*</td>
<td>0.083 ± 0.039</td>
<td>0.111 ± 0.041</td>
<td>0.104 ± 0.042</td>
<td>0.137 ± 0.035</td>
</tr>
<tr>
<td>Amino acid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D-Alanine*†</td>
<td>0.052 ± 0.024</td>
<td>0.066 ± 0.023</td>
<td>0.054 ± 0.019</td>
<td>0.112 ± 0.087</td>
</tr>
<tr>
<td>Glycyl-L-glutamic acid**†</td>
<td>0.063 ± 0.024</td>
<td>0.098 ± 0.027</td>
<td>0.078 ± 0.034</td>
<td>0.132 ± 0.040</td>
</tr>
<tr>
<td>L-Alanyl-glycine*†</td>
<td>0.062 ± 0.024</td>
<td>0.083 ± 0.021</td>
<td>0.067 ± 0.025</td>
<td>0.111 ± 0.049</td>
</tr>
<tr>
<td>L-Glutamic acid*</td>
<td>0.079 ± 0.029</td>
<td>0.117 ± 0.032</td>
<td>0.102 ± 0.033</td>
<td>0.142 ± 0.064</td>
</tr>
<tr>
<td><strong>Carbohydrate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arbutin†</td>
<td>0.126 ± 0.017</td>
<td>0.209 ± 0.012</td>
<td>0.136 ± 0.022</td>
<td>0.143 ± 0.02</td>
</tr>
<tr>
<td>D-Mannitol*†</td>
<td>0.112 ± 0.041</td>
<td>0.155 ± 0.054</td>
<td>0.178 ± 0.07</td>
<td>0.093 ± 0.022</td>
</tr>
<tr>
<td>D-Melibiose</td>
<td>0.089 ± 0.036</td>
<td>0.127 ± 0.036</td>
<td>0.110 ± 0.035</td>
<td>0.12 ± 0.021</td>
</tr>
<tr>
<td>D-Ribose*</td>
<td>0.082 ± 0.031</td>
<td>0.130 ± 0.036</td>
<td>0.130 ± 0.048</td>
<td>0.139 ± 0.032</td>
</tr>
<tr>
<td>D-Sorbitol†</td>
<td>0.144 ± 0.012</td>
<td>0.193 ± 0.009</td>
<td>0.203 ± 0.016</td>
<td>0.149 ± 0.014</td>
</tr>
<tr>
<td>Lactulose*†</td>
<td>0.038 ± 0.019</td>
<td>0.043 ± 0.019</td>
<td>0.033 ± 0.01</td>
<td>0.125 ± 0.069</td>
</tr>
<tr>
<td>N-Acetyl-D-glucosamine*</td>
<td>0.039 ± 0.024</td>
<td>0.067 ± 0.041</td>
<td>0.064 ± 0.025</td>
<td>0.093 ± 0.032</td>
</tr>
<tr>
<td>N-Acetyl-D-mannosamine*</td>
<td>0.034 ± 0.022</td>
<td>0.062 ± 0.038</td>
<td>0.044 ± 0.023</td>
<td>0.091 ± 0.031</td>
</tr>
<tr>
<td>Sedoheptulosan*</td>
<td>0.065 ± 0.029</td>
<td>0.090 ± 0.035</td>
<td>0.074 ± 0.027</td>
<td>0.105 ± 0.025</td>
</tr>
<tr>
<td>Stachyrose*</td>
<td>0.072 ± 0.034</td>
<td>0.094 ± 0.036</td>
<td>0.084 ± 0.032</td>
<td>0.115 ± 0.026</td>
</tr>
<tr>
<td>Xylitol†</td>
<td>0.057 ± 0.025</td>
<td>0.078 ± 0.035</td>
<td>0.065 ± 0.026</td>
<td>0.045 ± 0.016</td>
</tr>
<tr>
<td><strong>Carboxylic acid</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D-Malic acid*</td>
<td>0.045 ± 0.019</td>
<td>0.071 ± 0.025</td>
<td>0.060 ± 0.023</td>
<td>0.054 ± 0.022</td>
</tr>
<tr>
<td>2-Hydroxybutyric acid*</td>
<td>0.055 ± 0.024</td>
<td>0.090 ± 0.033</td>
<td>0.071 ± 0.021</td>
<td>0.075 ± 0.02</td>
</tr>
<tr>
<td>L-Lactic acid*†</td>
<td>0.088 ± 0.043</td>
<td>0.139 ± 0.052</td>
<td>0.124 ± 0.064</td>
<td>0.067 ± 0.025</td>
</tr>
<tr>
<td>L-Malic acid*†</td>
<td>0.084 ± 0.035</td>
<td>0.120 ± 0.033</td>
<td>0.112 ± 0.045</td>
<td>0.145 ± 0.051</td>
</tr>
<tr>
<td>Pyruvic acid*†</td>
<td>0.153 ± 0.065</td>
<td>0.230 ± 0.068</td>
<td>0.229 ± 0.083</td>
<td>0.173 ± 0.047</td>
</tr>
<tr>
<td><strong>Phosphorylated compound</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DL-2-Glycerol phosphate*†</td>
<td>0.033 ± 0.019</td>
<td>0.054 ± 0.026</td>
<td>0.047 ± 0.026</td>
<td>0.026 ± 0.016</td>
</tr>
</tbody>
</table>
| *Variance among taxa significant for raw ΔOD₉₅₅ scores, after adjusting for 95 tests, using the sequential Bonferroni method.

†Variance among taxa significant for adjusted ΔOD₉₅₅ scores, after adjusting for 95 tests, using the sequential Bonferroni method.
differences (xylitol and glycy1-L-glutamate) could be attributed entirely to *B. licheniformis* being distinct from the three other taxa. The latter was included among metabolites yielding significant inter-taxon variation for both raw and adjusted ΔOD595 scores. Therefore, our conservative estimate for the fraction of metabolites with significant inter-taxon variation within our ingroup is 8% (8/95).

**Overall species differences in metabolism**

Through canonical analysis (Dillon, 1984), differences among ingroup taxa in 95 metabolic measurements were projected onto a two-dimensional phenotype space. The canonical analysis shows *B. mojavensis* and the two subspecies of *B. subtilis* to be distinct in their average metabolic measurements (Fig. 1).

A multiple analysis of variance (MANOVA) of the pool of 95 metabolic characters showed the taxa to be significantly distinct in overall metabolic phenotype (Pillai’s Trace = 2.69, *F* = 16.11, d.f. = 285, *P* < 0.0001). The MANOVA also shows striking variation among strains within each taxon (Pillai’s Trace = 35.25, *F* = 1.91, d.f. = 4180, *P* < 0.0001). A principal components analysis (Fig. 2) shows that the inter-taxon variation leads to considerable overlap of phenotypes among species. Therefore, quantitative differences among taxa are not diagnostic, even while the modal values of the taxa are significantly different.

**Salinity tolerance**

As salinity increased, growth rates decreased (Table 3), and no strains grew at 10% salinity. Taxon effects were significant, with *B. mojavensis* growing approximately 10% faster than either *B. subtilis* subspecies at both 1 and 2.5% NaCl. There were no significant differences at 5% NaCl. The *B. subtilis* subspecies were not significantly different in their growth rates at any salinity.

**DISCUSSION**

**Quantitative variation in diagnostic systematics**

This study presents the first extensive survey, to our knowledge, of quantitative variation for metabolic capabilities within and between closely related bacterial taxa. The two subspecies of *B. subtilis* and the closely related *B. mojavensis*, as well as the outgroup *B. licheniformis*, all showed wide-ranging quantitative variation, both within and between taxa. Quantitative variation was significant among strains of the same taxon for 93% of the 95 carbon sources tested, and the ingroup taxa were significantly different in their average rates of metabolism for 8% of the carbon sources.

Nevertheless, these quantitatively varying metabolic characters do not contribute to diagnostic systematics for the *Bacillus* taxa studied here. Each of the metabolic traits investigated showed considerable overlap among taxa, even for traits whose mean values differed significantly among taxa (Fig. 2). Likewise, *qualitative* variation in metabolic
Nevertheless, the present study suggests that acquisition of
in adaptive divergence.

differences are responsible for adaptive differences, but
know at this point what fraction of these genomic
(Nakamura et al, 1994). We do not
a larger part in adaptive
evolution. We cannot be certain exactly how many genetic
changes were responsible for variation in the metabolic
characters studied, since a single genetic change can affect
the expression of multiple genes (Cavalieri et al., 2000;
Hamilton, 2002). However, the observation that metabolic
capability varied over so many substrates, and over several
categories of substrates, suggests that quantitative metabolic
changes have occurred frequently in Bacillus evolution. Even
so, it is unknown whether the metabolic variants we have
observed reflect adaptations in nature.

What is the likelihood that incremental change in metabolic
capabilities has led to invasion of new ecological niches in
bacterial evolution? Experiments in evolutionary micro-
cosms have demonstrated that evolution by mutation alone
can yield new, ecologically distinct populations, which
can co-exist indefinitely with the parental population. For
example, consider the case of the clonal descendants of one
E. coli cell, evolving in medium with glucose as the only
carbon source. Here mutational increases in acetate meta-
bolism founded a new population that could survive by
feeding on the acetate secreted as a waste product by the
parental glucose-specialized population (Treves et al., 1998).
Mutational adjustment of a large number of metabolic
capabilities could potentially find an entire ensemble of
stably co-existing, ecologically distinct populations.

However, co-existence between populations that are only
incrementally different in their metabolic capabilities is
severely challenged by temporal variation in nutrient
levels (Feldgarden et al., 2003; Lunzer et al., 2002). With
fluctuating nutrient levels, one population will almost
inevitably out-compete to extinction any other population
that uses the same set of resources, albeit in different
proportions. In contrast, acquiring a metabolic capability by
horizontal transfer can give a population a novel resource
not utilized by the parental population, thus preventing
extinction by the parental population (Cohan, 2004; Holt,
1987). Populations distinguished by horizontal transfer
events are much more likely to co-exist over ecological time.

### Table 3. Mean growth rates (doublings per hour) in Luria broth with different salt concentrations

Asterisks indicate those taxa that grew significantly faster or slower than others at a given salt concentration, as determined by the Tukey–Kramer test (*, P<0.05; **, P<0.005). Standard errors are based on the variance among strains within the taxon.

<table>
<thead>
<tr>
<th>Salt concn</th>
<th>Taxon</th>
<th>B. subtilis subsp. subtilis</th>
<th>B. subtilis subsp. spizizenii</th>
<th>B. mojavensis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1% NaCl</td>
<td></td>
<td>1.09 ± 0.04</td>
<td>1.03 ± 0.14</td>
<td>1.17 ± 0.04*</td>
</tr>
<tr>
<td>2.5% NaCl</td>
<td></td>
<td>0.95 ± 0.15</td>
<td>1.01 ± 0.15</td>
<td>1.16 ± 0.11**</td>
</tr>
<tr>
<td>5% NaCl</td>
<td></td>
<td>0.46 ± 0.12</td>
<td>0.51 ± 0.02</td>
<td>0.43 ± 0.03</td>
</tr>
</tbody>
</table>

capabilities has also failed to provide diagnostic pheno-
typic characters for taxa of the B. subtilis complex. As yet,
the only diagnostic phenotypic characters distinguishing
B. mojavensis and the B. subtilis subspecies are the quanti-
tatively varying levels of fatty acids (Nakamura et al., 1999;
Roberts et al., 1994, 1996).

### Gradual evolution versus abrupt evolution by horizontal transfer

Comparative genomics has led to increased interest in
adaptation via abrupt evolutionary change effected by
horizontal transfer, at the expense of interest in incremental
change effected by mutations in existing genes (Lawrence,
hypothesized that nearly all invasions of new niches, and
perhaps much of adaptation within a niche, is effected by
horizontal transfer. That bacteria have acquired hetero-
logous genes by horizontal transfer is demonstrated by
large differences in genomic content among close relatives,
as assayed in three decades of DNA–DNA hybridization
studies (Johnson, 1973; Lan & Reeves, 1996), and more
recently by sequencing multiple genomes from the same
species (Alm et al., 1999; Baba et al., 2002; Beres et al., 2002;
Perna et al., 2001; Shirai et al., 2000).

In the case of the Bacillus taxa we have investigated, only
B. subtilis strain 168 has been fully sequenced (Kunst et al.,
1997), but DNA–DNA hybridization indicates considerable
variation in genome content within and between named
taxa (Nakamura et al., 1999; Roberts et al., 1994). Within
each of the taxa studied here, pairs of strains are on average
6.3 ± 0.5 % different in the genes they contain (Nakamura
et al., 1999; Roberts et al., 1994), and pairs of strains from
different taxa differ on average by 37.0 ± 0.4 % of their
genomes for the two subspecies of B. subtilis, by 59.1 ± 1.2 %
for B. mojavensis and B. subtilis subsp. subtilis, and by
58.5 ± 1.4 % for B. mojavensis and B. subtilis subsp. spizizenii
(Nakamura et al., 1999; Roberts et al., 1994). We do not
know at this point what fraction of these genomic
differences are responsible for adaptive differences, but
there is clearly room for horizontal transfer to play a role
in adaptive divergence.

Nevertheless, the present study suggests that acquisition of
genes is not the whole story of phenotypic evolution in
bacteria. A sizeable fraction of diverse metabolic traits were
shown to vary quantitatively within (93 %) and between
(8 %) ingroup taxa, so modest changes in the expression of
many genes could have played a large part in adaptive
evolution. We cannot be certain exactly how many genetic
changes were responsible for variation in the metabolic
characters studied, since a single genetic change can affect
the expression of multiple genes (Cavalieri et al., 2000;
Hamilton, 2002). However, the observation that metabolic
capability varied over so many substrates, and over several
categories of substrates, suggests that quantitative metabolic
changes have occurred frequently in Bacillus evolution. Even
so, it is unknown whether the metabolic variants we have
observed reflect adaptations in nature.
Over evolutionary time, the co-existence of quantitatively different populations is also challenged by adaptive mutations appearing in each population. As long as two populations use entirely the same set of resources (but in different proportions), an adaptive mutation in one population threatens to out-compete the other to extinction (Cohan, 2004). However, a population that has diverged by acquiring a new metabolic operon instantaneously escapes this risk of extinction. Thus, while there are likely many more opportunities for ecological divergence to originate by mutation than by horizontal transfer, persistence of these populations is more likely when they have diverged through horizontal transfer.

Gradual evolution may have its most important role in improving the level of adaptation in a newly divergent, ecologically distinct population. Many new adaptations have somewhat deleterious side effects, and incremental change may play a role in compensating for novel adaptations that bring a lineage into a new niche (Bouma & Lenski, 1988; Cohan et al., 1994; Schrag et al., 1997). Furthermore, when a population enters a new niche, there may be new opportunities for adaptation by incremental changes. For example, when birds first evolved flight, this created natural selection to take advantage of new feeding opportunities (e.g. involving changes in behaviour and feeding morphology). Likewise, a bacterial population that can live in new microhabitats owing to a horizontally transferred gene may find its established resources available in different proportions. This would precipitate gradual evolution in metabolic capabilities.

**Metabolic and genomic diversity within taxa**

Each taxon has shown considerable quantitative variation in metabolic rates (Fig. 2) and salt tolerance (Table 3), as well as in genomic content (as revealed by DNA–DNA hybridization) (Nakamura et al., 1999; Roberts et al., 1994, 1996). What are the dynamics that allow such a high degree of variation within a bacterial species?

The level of genetic diversity within an ecologically homogeneous population of bacteria (i.e. an ecotype) is most likely regulated by periodic selection (Cohan, 1994a, 2002b, 2004; Koch, 1974; Levin, 1981). Owing to the extremely low rates of recombination within bacterial populations (Cohan, 2002a, c; Maynard Smith et al., 1993), selection favouring an adaptive mutation is expected to bring the entire genome of the adaptive mutant cell to nearly 100% frequency (Cohan, 1994b, 2004). This extinguishes variation at all loci within the genome, including coding and regulatory sequences, as well as variation in genomic content (i.e. the set of genes borne by a genome). If all members of a named taxon are actually members of the same ecotype (i.e. the strains are ecologically interchangeable), then the intra-taxon variation would represent only ecologically meaningless variation that appears randomly between periodic selection events, only to be extinguished with the next periodic selection.

Most bacterial species, however, are unlikely to represent a single ecotype. Most species investigated show considerable variation in protein-coding sequences, and this variation typically falls into multiple clusters of sequence similarity (Cohan, 2002b; Feil et al., 1999, 2000), although there are some notable exceptions (Yersinia pestis and Bacillus anthracis) (Keim & Smith, 2002; Klevetska et al., 2001), owing to systematists’ practice of splitting lethal human pathogens (nomen periculosum) from extremely close relatives that are less harmful (Stackebrandt et al., 2002). Any long-divergent clusters appearing within a species are unlikely to be members of the same ecotype, since the first periodic selection event would have collapsed the species diversity to a single cluster (Cohan, 2002b; Dykuizen, 1998; Palys et al., 1997). Moreover, the clusters observed within a named species are typically ecologically distinct (Achtman et al., 2001; Chan et al., 2001; Maiden et al., 1998; Zhu et al., 2001) and their phylogenies are consistent with single ecotypes (Cohan, 2002b).

Thus, a more likely interpretation of the great genomic and phenotypic diversity within a named species is that most of this diversity represents adaptive divergence among ecotypes. For example, variation within B. mojavensis in the ability to metabolize sorbitol may reflect differences between ecotypes that are adapted to environments with different sorbitol levels. Perhaps only a small fraction of the genomic and phenotypic diversity within a named species represents transient, random divergence within ecotypes.

We suggest a two-pronged approach to investigating the adaptive and non-adaptive origins of genomic and phenotypic variation within a named species. First, strains should be classified into putative ecotypes using one of several sequence-based approaches (Cohan, 2002b; Maiden et al., 1998). The principle is that multilocus sequence clusters are likely to correspond to ecotypes, since periodic selection recurrently purges the diversity within but not between ecotypes. The putative ecotypes then establish a framework for mapping the metabolic and genomic diversity within a species.

Second, we should apply genomic technology to expand greatly the breadth of phenotypic surveys we can accomplish. Recently, microarray assays of whole-genome mRNA concentrations have shown adaptive differences between natural, conspecific populations of fish (Oleksiak et al., 2002). While this approach has not yet been applied to natural populations of bacteria, the magnitude of variation found in the present study suggests that taxa may vary in the expression patterns of hundreds of genes.

We can then attempt to map phenotypic variation in gene expression and metabolism (and perhaps differences in genomic content as well) onto the putative ecotypes inferred from sequence clusters. Variation appearing within a putative ecotype is likely to represent merely the random changes occurring within a population between periodic selection events. However, any differences that correspond
to the putative ecotypes will demand further study. To the extent that differences in gene expression appear to determine niche differences, these differences will provide evidence that the putative ecotypes are really ecologically distinct and will suggest the nature of ecological differences among ecotypes.

ACKNOWLEDGEMENTS

This research was supported by National Science Foundation grant DEB-9818576 and by research funds from Wesleyan University.

REFERENCES


