

Abstract

The accurate identification of microorganisms in industrial settings, especially in sterile pharmaceutical manufacturing, is an important part of a comprehensive environmental monitoring program. Thorough characterization of the microbial population of the manufacturing environment allows for rapid and definitive resolution of sterility failures and other excursions. Standard tracking and trending EM programs that use the species names alone for tracking microorganisms do not always provide enough information to make definitive conclusions in an investigation or provide the appropriate information to develop a response to the contamination. In cases where the microorganism cannot be identified to the species level, reliance on the closest species or genus name may be either inappropriate or inconclusive. This study describes the advantages of using 16S rDNA sequences to track isolates of bacteria. The use of 16S rDNA sequences negates the need to characterize the microorganism by phenotypic means, using the DNA sequence as a tool for identification as well as a unique identifier. Additionally, in most instances a single species of bacteria can be further divided into several subgroups based on their unique DNA sequences, which further help the tracking process.

Introduction

Tracking microorganisms found in the manufacturing areas of pharmaceutical and biotech facilities is an important part of the company's environmental monitoring program. This information is used to create trending reports on a regular basis (monthly, quarterly, etc.), which provides information regarding the state of environmental control within the manufacturing area. An increase in the number of microorganisms recovered in certain areas of the facility may indicate breaches in the HVAC system or other sources of microbial contamination. A shift in the types of organisms recovered may be helpful in finding the source of the contamination. Historically, most manufacturing facilities have used information regarding the identity of the microorganisms to help aid in the tracking process. Ideally a genus and species name would be used, but in some cases where a species name is not possible, either a genus name alone is used, or in others, a Gram stain. Even when a species name is available, consistent and accurate names are required for tracking purposes. According to the FDA's Guidance for Industry, "The goal of microbiological monitoring is to reproducibly detect microorganisms for purposes of monitoring the state of environmental control. Consistent methods will yield a database that allows for sound data comparisons and interpretations."¹ Identification methods which provide inconsistent identifications or no identification for the same isolate are not useful for tracking isolates to their source or for generating trending reports. Name-based tracking is also affected by developments in taxonomy that reclassify organisms. Commercially available genotypic systems, such as MicroSEQ® 2.0, may not consistently provide adequate sequence coverage of the 16S gene, which is necessary, as the sequence of the organism is its true identity and provides more information than a name alone. Tracking and trending by sequence is more effective as the DNA sequences will never change.

In this study, we demonstrate the use of 16S rDNA sequences in tracking unnamed bacteria, instead of just relying on the taxonomic name, for the purposes of tracking and trending. This approach is especially useful when only a genus name, or higher level taxa, can be assigned.

Methods

Organisms

The isolates used in this study were all environmental isolates from the pharmaceutical manufacturing environment. All 13 isolates were identified as *Bacillus sp.* (Genus level identifications) with a closest species match of *Bacillus benzoevorans*.

Generation of 16S rDNA sequences

High accuracy, partial consensus sequences were generated using Accugenix procedures. The process included PCR amplification and sequencing using 0005F and 0531R flanking 16S rDNA primers, Applied Biosystems BigDye™ v1.1 terminator chemistry and Applied Biosystems POP-6 polymer, using a 3130 xl Genetic Analyzer.

The raw data was manually reviewed to correct base calling errors, and assembled with a full length reference sequence. This process has been shown to provide a more complete and consistent consensus sequence than the MicroSEQ® 2.0 automated approach.

Phylogenetic Analysis

Pairwise alignments were used to calculate percent difference, and neighbor joining trees were used to represent relatedness. The unknown isolates' DNA sequences were compared to their 10 closest matches from a validated, cGMP compliant database.

Results

Because the raw data was manually assembled, we were able to compare all 13 organisms with assurance that the full 500 base pair region of the 16S gene was used. The MicroSEQ® 2.0 assembly process truncates the sequence, while the manual process does not (Figure 1). MicroSeq® 2.0 was not able to analyze the remaining strains.

Bacillus benzoovorans was the closest match for all 13 unknown organisms. They were called "*Bacillus sp.*," because, when compared to their 10 closest matches, they were not close enough to any one species to make a Species level call (Figure 2). Furthermore, phylogenetic analysis of the 13 unknown *Bacillus* species revealed 4 groups of isolates, with respect to their percent genetic difference from their closest match, *B. benzoovorans*. Group A was 3.09% different, Group B was 3.37% different, Group C was 4.21% different and Group D was 4.40% different (Figure 3).

Phylogenetic analysis of all 13 isolates to one another using a neighbor joining tree, on the other hand, revealed 8 potentially different species of *Bacillus* within 10 distinct sequences (Figure 4). Identical sequences are illustrated with a vertical line, indicating no genetic difference within this region of the 16S gene.

A comparison of the above 2 types of phylogenetic analyses shows that more information can be gathered by comparing isolates to one another than by comparing isolates to a DNA sequence library alone. Isolates which may appear to be similar based on a library search alignment (Figure 5) are actually genetically quite different from one another (Figure 6). *Bacillus sp.* – 8 was not included, as it has no genetic difference from *Bacillus sp.* – 7. There appear to be 3 different potential species of *Bacillus* within Group C, and 5 distinct sequences.

Figure 1. Alignment and Neighbor Joining Tree demonstrating inconsistencies with MicroSEQ® 2.0 data assembly

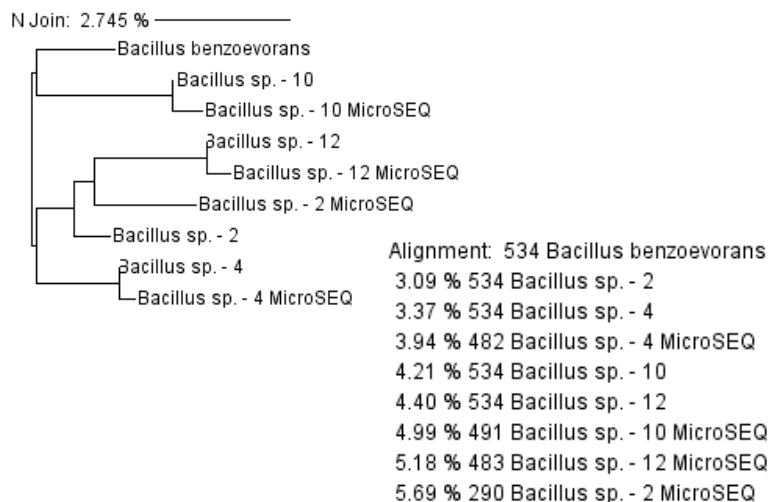


Figure 2. Neighbor Joining Tree with representatives from 4 groups and their 10 closest matches

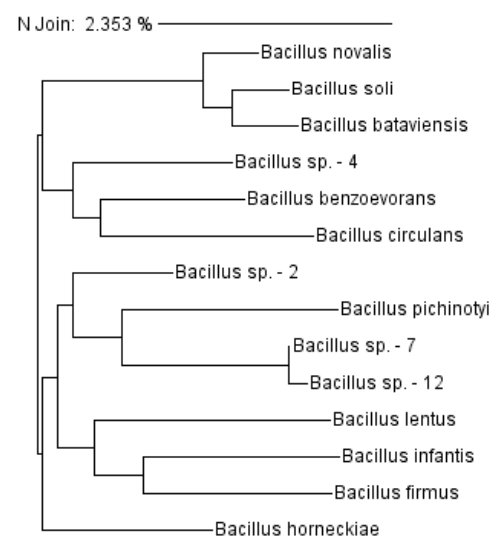


Figure 3. Percent difference of isolates to *Bacillus benzoevorans*

Alignment: 534 *Bacillus benzoevorans*

3.09 % 534 <i>Bacillus sp.</i> - 2	Group A
3.37 % 534 <i>Bacillus sp.</i> - 1	Group B
3.37 % 534 <i>Bacillus sp.</i> - 4	
3.37 % 534 <i>Bacillus sp.</i> - 3	
4.21 % 535 <i>Bacillus sp.</i> - 5	Group C
4.21 % 535 <i>Bacillus sp.</i> - 9	
4.21 % 534 <i>Bacillus sp.</i> - 8	
4.21 % 534 <i>Bacillus sp.</i> - 7	
4.21 % 534 <i>Bacillus sp.</i> - 10	
4.21 % 534 <i>Bacillus sp.</i> - 6	
4.40 % 534 <i>Bacillus sp.</i> - 12	Group D
4.40 % 534 <i>Bacillus sp.</i> - 11	
4.40 % 534 <i>Bacillus sp.</i> - 13	

Figure 4. Phylogenetic tree showing genetic relatedness of isolates

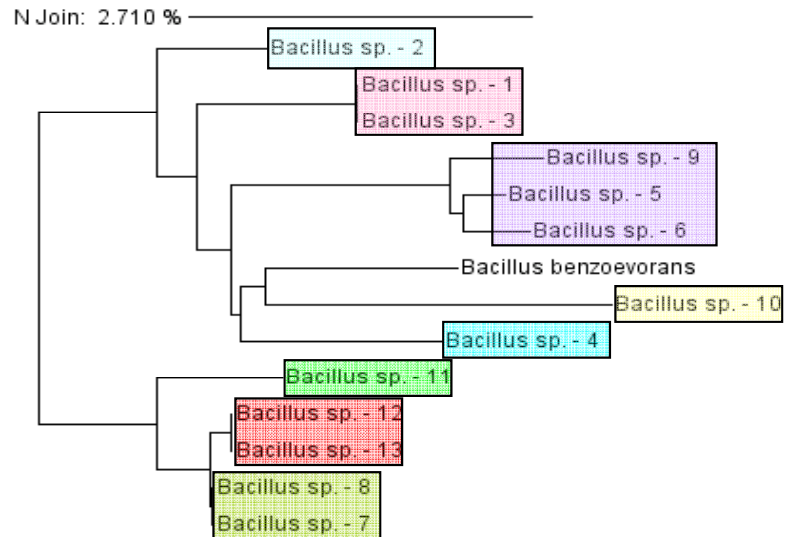
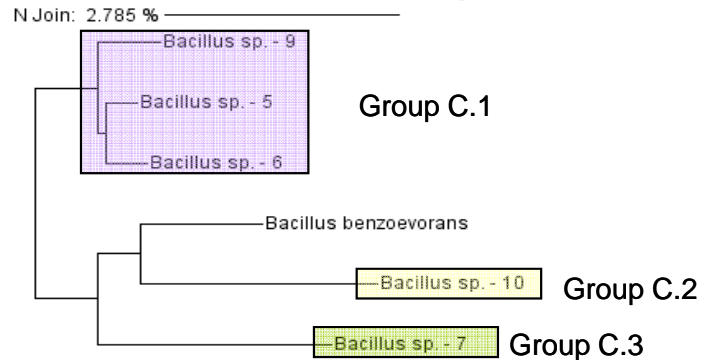


Figure 5. Percent difference of Group C isolates to *Bacillus benzoevorans*

Alignment: 534 *Bacillus benzoevorans*

4.21 % 535 <i>Bacillus sp.</i> - 5	Group C
4.21 % 535 <i>Bacillus sp.</i> - 9	
4.21 % 534 <i>Bacillus sp.</i> - 10	
4.21 % 534 <i>Bacillus sp.</i> - 6	
4.21 % 534 <i>Bacillus sp.</i> - 7	

Figure 6. Phylogenetic tree showing genetic relatedness of Group C isolates



Discussion

Assigning unknown environmental isolates to an appropriate taxa is an important part of a good environmental monitoring program. The names given to the isolates will be used not only for identity, but also for trending and tracking purposes.

The use of the assigned name, however, becomes difficult when a species level designation is not possible, either due to the identification system used, or when the organism has not been described by taxonomists. As the above data illustrates, one cannot assume all isolates identified simply as *Bacillus sp.* are the same, even if their closest match is the same.

The use of DNA sequences in addition to the name of the organism provides much more information as to the relatedness of the recovered organisms than the name alone. The use of this approach, especially in the cases of sterility failures and other investigations, will allow a faster and more complete resolution.

Although this approach will not differentiate isolates to the strain level, differences in the analyzed 500 base pair region of the 16S gene are enough to eliminate some isolates from consideration, and thereby reduce the scope of the investigation and the number of isolates which need to be characterized at the strain level. Automated systems available will not provide a consistent sequence for the 500 base pair region, so a manual assembly process is preferred for comparison purposes.

About Us

Accugenix, Inc. provides leading-edge technology in microbial identification and characterization services. Our FDA-registered lab is cGMP compliant and maintains rigorous standards competitive at the global level. We specialize in testing, analyzing and interpreting data from environmental isolates commonly found in pharmaceutical, biotechnology, medical device, nutraceutical, personal care and other manufacturing industries.

For more than 20 years, Accugenix has provided the fastest, most accurate and reliable microbial identification services to over 400 facilities around the world. Accugenix updates its validated, proprietary DNA sequence libraries annually to reflect current taxonomy and newly described relevant species. We have the industry's first Fungal Library based on the ITS region. Since inception, we have tested more than 400,000 microorganisms – more than any other service laboratory in the industry, while maintaining an on-time delivery of over 99%.

History of Accugenix

1990.

Accugenix, Inc. began as Acculab, Inc., a reference laboratory specializing in microbial identification for industry and research clients. At the time we were one of only a few service laboratories in the world offering cellular fatty acid analysis, beginning a tradition of bringing cutting-edge microbiology methods to full commercial potential and utilization.

1999.

To reflect the addition of comparative DNA sequencing to our menu of validated methods, we created Accugenix, A Division of Acculab, Inc. Since then we have sequenced hundreds of thousands of environmental isolates from over 1000 pharmaceutical and biotechnology production facilities around the world, allowing us to build the largest and most unique industry database for bacteria and fungi that often occur in clean room manufacturing environments.

2005.

Our official name changed to Accugenix, Inc. on February 25, 2005.

2008.

Accugenix GmbH, our European subsidiary, was launched in Spring 2008.

Today.

Dedicated to being the industry leader for providing the most progressive microbiology methods available, Accugenix has invested in the technology, instrumentation and expertise to conquer genetic-based testing methods, their process validation, cGMP compliance, and other rigorous regulatory standards at the global level. Accugenix continues to staff its ranks with scientists and experts to guide and/or fast-forward your transition to genotypic microbial identification.



Visit www.accugenix.com or call +1 302.292.8888/ +49 (0)621 3709 556.