applied biosystems
by life technologies

# MicroSEQ® Rapid Microbial Identification System

## Giving you complete control over microbial identification using the gold-standard genotypic method

The MicroSEQ® ID microbial identification system, based on comparative rDNA sequencing of the 16S region (for bacteria) or the LSU D2 region (for fungi), is a proven method for rapid and accurate microbial identification. The optimized design, integrated workflow, and superior accuracy of the MicroSEQ® ID Rapid System enable top pharmaceutical companies, medical devices companies, public health laboratories, and government agencies worldwide to reliably perform a wide range of microbial identification tasks, including routine QC microbiology tests.

The MicroSEQ® ID system was designed for workflow flexibility and highly accurate identification calls. Some system highlights include:

- On-site system installation and validation that allows customers to have in-house capability with end-user control of data from facility environmental monitoring and product testing
- Simplified, streamlined workflow and software that allows users to rapidly gain expertise
- Single standardized identification workflow for both bacteria and fungi
- Option of routine bacterial identification using the first 500 bp of the rDNA, or higher-resolution identification based on the full 1,500 bp region
- Data analysis using automated or manual modes
- Outcome predictions using the phylogenetic tree tool
- Ability to build user-defined and user-validated custom libraries
- Installation and qualification support executed by certified personnel
- Software designed to enable 21 CFR Part 11 compliance

System features will be described in detail in the following pages.

life technologies

## Own the product and the process

The streamlined workflow for MicroSEQ® ID includes optimized chemistry, instrumentation, and user access to raw data if needed for more in-depth analyses.

With MicroSEQ® ID v.3.0, the various steps of data analysis can be automated for high accuracy and reproducibility. At the same time, the user has full visibility and control of the analysis criteria used in the final identification. With this system, users can:

- Perform data analyses with intrinsic data quality assessment
- Search against a validated library for assignment of identification
- Search against custom libraries to identify and track specific microorganisms that are unique to the site or are not in validated libraries

With MicroSEQ® ID, system installations and qualifications are performed at the user site. On-site installation helps the end user to design SOPs that reflect the entire workflow process according to their requirements and internal policies. Having complete control of the sample chain of custody from isolation to identification allows the end user to:

- Have high confidence in the result obtained
- Rapidly obtain identification of isolates from multiple sources
- Rapidly respond to investigation of contamination events
- Minimize transit delays due to off-site analysis

Bringing the MicroSEQ® ID system in-house enables a quick turnaround time for specimen ID results while providing complete ownership of the process and results.

## Identify bacteria and fungi

Ribosomal genes have been used successfully for over two decades for microbial identification, and this method is considered the gold standard [1]. In the MicroSEQ® ID system, the following gene targets are used for identification:

### Bacteria

The target for bacterial identification is the 16S ribosomal RNA (rRNA) gene sequence (rDNA). As part of the small subunit of a prokaryotic ribosome, the 16S rRNA is ubiquitous and can therefore be used to study phylogenetic relationships among all bacteria [2].

With the MicroSEQ® ID method, users have the option of choosing to sequence the first 500 bp of the 16S rDNA gene or the full 1,500 bp. Kits and the supporting validated library databases are available for both options. For routine identifications, the first 500 bp are sufficient for an identification that encompasses 3 of the 9 hypervariable regions of the 16S gene (Figure 1).

In some cases the 500 bp region is not sufficient to discriminate among very closely related bacteria and therefore requires a more informative full-gene read (taking into consideration all the hypervariable regions) (Figure 2). Furthermore, sequencing of the entire 1,500 bp sequence is required when describing a new species.

### Fungi and yeast

Comparative sequence analysis of both the expansion region D2 of the larger rRNA molecule in the large subunit of the eukaryotic ribosome (LSU-D2) and ITS regions have been successfully used for identification and classification of fungi down to the species level [1]. The ITS region exhibits more variability, which can give rise to many more "sequence types" than the D2 region when used for sequence-based fungal identifications. Fungal genomes may contain more than 100 copies of the rDNA cluster— it therefore becomes critical to understand and differentiate between simple sequence variability and sequence variability that reflects actual relatedness. The MicroSEQ® ID fungal database is derived from sequences from LSU-D2 rDNA, which helps deliver highly reliable fungal identification and classification in routine identification tests [3].

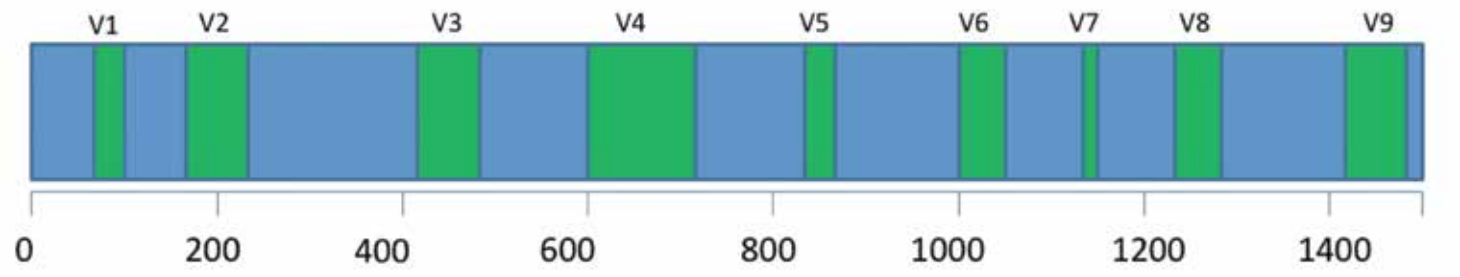


**Figure 1. Hypervariable regions within the 16S rRNA gene.** There are 9 hypervariable regions (V1–V9) within the bacterial 16S gene [4], indicated in green. The conserved regions are indicated in blue.

## Obtain quality result with validated libraries

Results from a microbial identification system are only as good as the quality of the reference libraries. Rigorous development of the reference library is critical for two reasons:

1. High confidence in the accuracy of the identification results
2. Ability to trace back to an identified source or type strain

Although the number of sequences in public databases is quite high, the taxonomy (naming) and sequence information that is deposited may not always be correct. The libraries provided with the MicroSEQ® ID system are verified at multiple levels The sequences in the MicroSEQ® ID library are derived almost exclusively from type strains or from strains that can reliably be traced back. Users can follow a similar process to create a custom library containing sequences from organisms that are unique to their sites. Such custom libraries can be used in combination with the MicroSEQ® ID validated libraries for highly accurate identification. Having well-validated libraries is particularly important in cGMP/cGLP environments during audits by regulatory agencies, as this provides complete and traceable documentation of the source and the sequence.


**Figure 2. Sequence data taken from the first 500 bp (lower tree) is not sufficient to differentiate between *Mycobacterium abscessus* (ATCC 19977) and *Mycobacterium chelonae chelonae* (ATCC 35752).** Extending the sequence to the full 1,500 bp allows for differentiation between these two organisms (upper tree). Similarly, the differences between *Mycobacterium farcinogenes* (ATCC 35753) and *Mycobacterium senegalense* (ATCC 35796) as well as *Mycobacterium africanum* (ATCC 25420) and *Mycobacterium microti* (ATCC 19422) become more apparent when analyzed using the full 1,500 bp sequence.

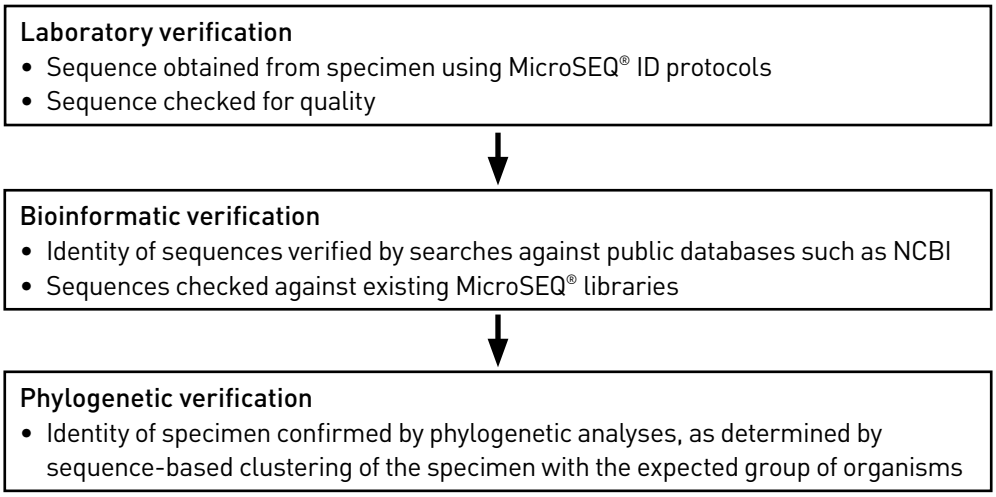


**Figure 3. MicroSEQ® ID libraries are derived almost exclusively from type strains or from strains that can reliably be traced back, and are validated using a very stringent process.**

## Choose method of data analysis

MicroSEQ® ID users can be trained by qualified and certified Life Technologies Field Application Specialists to analyze and interpret their own data based on user-determined acceptance criteria. Support and expertise are provided along with a full implementation training program. Additionally, a detailed guideline for data analysis and interpretation is provided during training so that facility staff can fully understand the scientific rationale behind MicroSEQ® ID data analysis. In this way, users can accurately analyze the sequence data they gather and can confidently report microbial identifications based on that analysis.

MicroSEQ® ID software version 3.0 provides the user with the option to use the automated quality assessments configured in the software or to review the data manually.

### Sequence quality

Sequence quality is a key parameter for accurate identification. Using the MicroSEQ® ID analysis software, sequence quality assessments can be made for each run by examining specific criteria such as the length of the consensus sequence generated from sample processing and the quality of the signal. For example, the final consensus sequence length generated must be at least 90% of the library entry length to ensure that the identification is based on a reliable sequence length (i.e., one that encompasses a sufficient number of the variable regions shown in Figure 1). The quality assessment performed by the MicroSEQ® ID analysis software provides an output to the user in the form of a "specimen score". Life Technologies Field Application Specialists provide users with specific guidelines during training to establish specimen score ranges as part of the acceptance criteria. In addition to the numeric specimen score output, the software also displays easy-to-spot, color-coded indicators to point out the specific criteria that may be contributing to a low specimen score. Based on the resulting specimen score, the user has the option to reject sequences that are of poor overall quality (Figure 4).

### Consensus sequence comparisons give more reliable data

With MicroSEQ® ID analysis, identification of an organism is based on comparison of a consensus sequence generated from the unknown sample against validated reference strains in the MicroSEQ® libraries. A consensus sequence is generated using the sequence from the forward strand aligned to the sequence from the reverse strand. While the sequence from a single strand may be sufficient to generate a match, use of a consensus sequence allows for internal confirmation of a sequence in both directions (Figure 5), resulting in higher confidence in the sequence data generated.

### Default setting for automated analysis

With the MicroSEQ® ID software, the user can set up an automated analysis program based on a set of pre-determined, user-defined, and validated criteria. Guidance on how to set up appropriate acceptance criteria can be provided by experienced and certified Life Technologies Field Application Specialists. Once the acceptance criteria are configured they can be included as an approved procedure in the SOP, and the automated analysis can be set as a default program. This helps to ensure that analysis is performed using the same set of acceptance criteria each time, without the requirement to manually re-enter parameters for each run. Any deviations from the default analysis program will be captured in an audit trail, allowing for better management of the data and results.

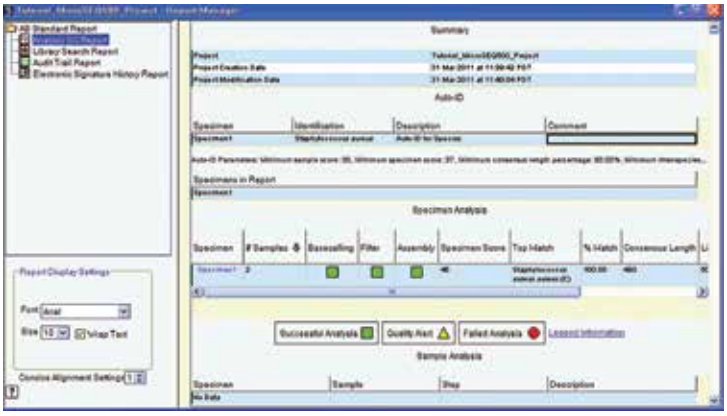

**Figure 4. The MicroSEQ® ID analysis software QC report gives consensus length and percent match for each run and assigns a specimen score.** Based on these quality parameters, users have the option of rejecting low-quality sequences. Color-coded indicators allow users to quickly spot specific criteria that may be affecting quality.

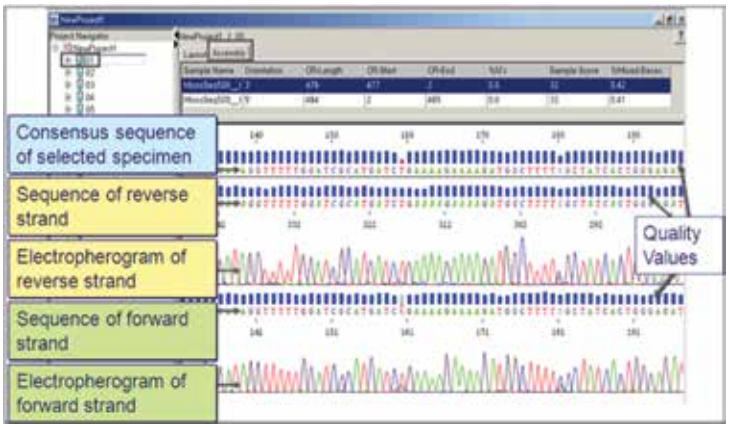


**Figure 5. The electropherograms, sequence, and associated base quality values for each strand can be viewed using the MicroSEQ® ID analysis software.**

### Auto-ID feature

For routine use, the Auto-ID feature can be enabled. Auto-ID evaluates each consensus sequence in a project against the specified Auto-ID requirements and thresholds, then assigns a specimen identification from a library (Figure 6).

### Semiautomated analysis

Although the MicroSEQ® ID system analysis software can be configured to perform automated analysis as the default setting, the user still has the option to perform semi-automated analysis of the generated sequence data. A MicroSEQ® ID system user may decide to perform a semi-automated analysis based on individually defined rules for accepting, rejecting, or reassessing a specimen.

Through both automated and semiautomated data analysis processes, the MicroSEQ® ID system provides a powerful and flexible tool for accurate data interpretation. The software also includes features to enable 21 CFR Part 11 compliance, providing the necessary security, audit trail, and electronic signature features.

Finally, the MicroSEQ® ID report itself captures the key criteria used to determine the final identification, as shown in Figure 7, as well as information about the phylogenetic relationship of the identified specimen.


Figure 6. The Auto-ID feature of MicroSEQ® ID analysis software allows rapid identification of multiple specimens based on a predetermined set of acceptance criteria.

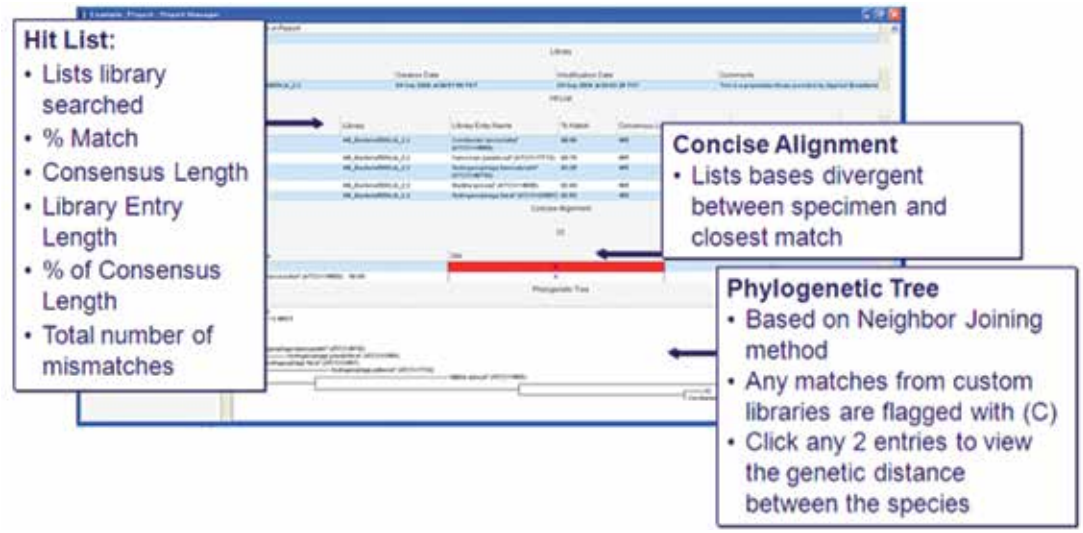


Figure 7. The MicroSEQ® ID report shows the criteria used to make the identification, as well as information about the phylogenetic relationship of the specimen.

## Predict and verify results with enhanced analysis

Phylogenetic tree analysis is one of the most powerful tools to examine, assess, or study relatedness among microorganisms (Figure 8). The creation of such a tree is one of the key elements of the MicroSEQ® ID software and is automatically generated when analyzing data. The accuracy of the sequence itself and the authenticity of the organism used for comparison are prerequisites for a sound analysis that can only be provided by a well-developed sequencing technology combined with truly validated libraries, both of which are advantages of the MicroSEQ® ID system.

The phylogenetic tree tool within the MicroSEQ® ID software also allows the user to quickly predict the outcome of MicroSEQ® ID without actually completing real identification tests. As an example, the trees in Figure 2 were created using this phylogenetic tree tool, with no actual wet chemistry performed. This *in silico* option helps reduce the need to employ a costly trial-and-error approach (e.g., when users have to determine if two closely related organisms can be adequately differentiated by analysis of 500 bp, or if correct identification will require full-gene analyses). It also recommends the next step—in the case of the analysis depicted in Figure 2, the use of the full-gene analysis for the organism in question is the preferred option.

## Customize libraries for unique environments

The MicroSEQ® ID software has a feature for building custom libraries, which gives end users flexibility for including sequences from unique or commonly found microorganisms in their environment or from published and verifiable rDNA sequences. Using the custom library feature, users can upload sequences into a custom library. These could be a set of sequences that are unique to a particular site or to a particular specimen (type). New specimens can then be matched against the validated MicroSEQ® ID libraries as well as the custom library to allow for comparison of specimen identity from different MicroSEQ® ID runs. This feature is also a powerful tool for maintaining a database of sequences from environmental isolates that may differ from the type strains obtained from culture collections [5]. Users may choose to validate a custom library based on their own set of criteria. Additionally, Life Technologies can provide support and training for creating and validating custom libraries. Custom libraries can be screened simultaneously with MicroSEQ® validated libraries by selecting more than one library.

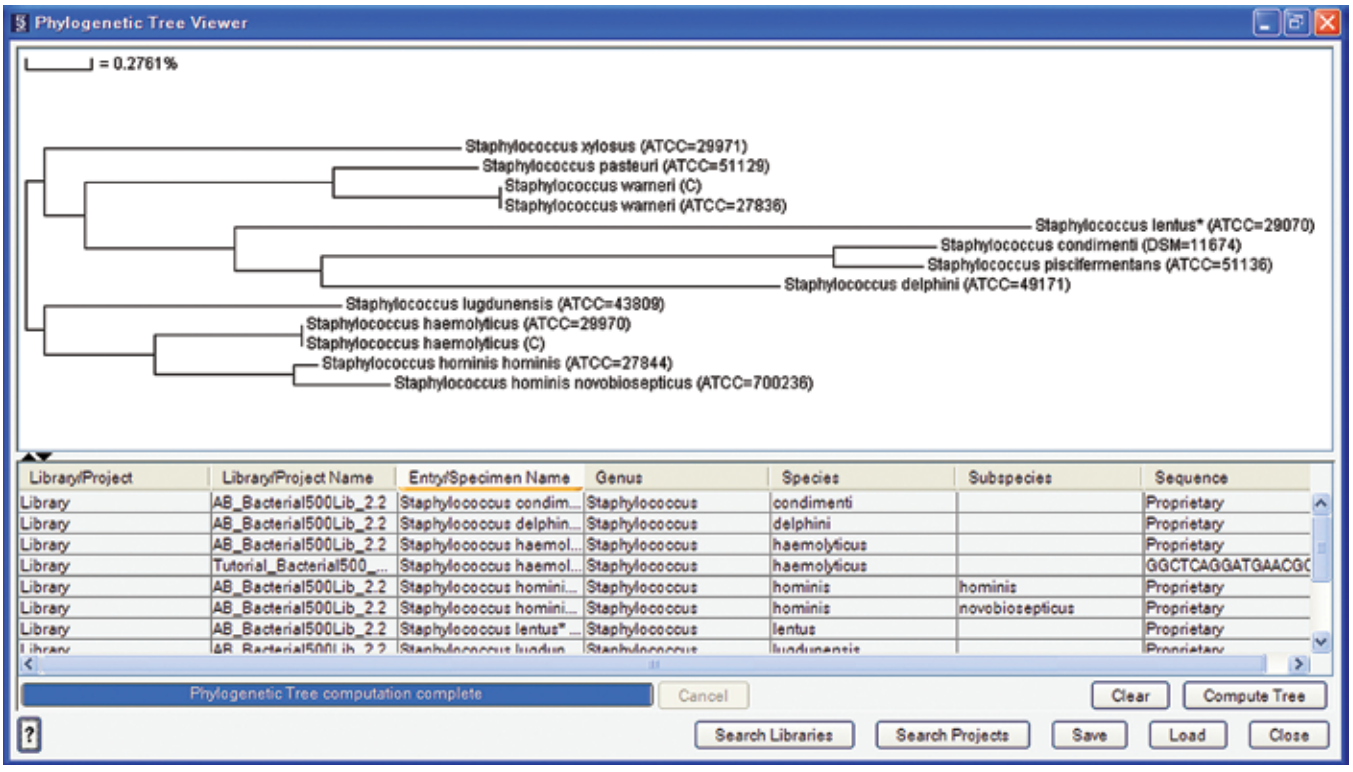


Figure 8. The MicroSEQ® ID analysis software's phylogenetic tree tool helps users study relatedness among organisms.

**Control the process, own the data**

The features of the MicroSEQ® ID Rapid Microbial Identification System offer a gold-standard solution for rapid and accurate microbial identification—rRNA gene sequence–based genotypic identification. The system, process, and data analyses are completely owned by the customer, which allows complete visibility and control. In addition, users have access to domain experts from Life Technologies who can provide training and support for validation and routine testing. MicroSEQ® ID is currently being used in various pharmaceutical, biopharmaceutical, and biotechnology companies, where customers routinely perform microbial identifications in line with their own SOPs and regulatory requirements and in compliance with the strict regulations of the cGMP environment.

"Species databases built on validated sequence data linked to vouchered specimens and cultures are the real future for rapid and accurate identification."

**Director**
Culture collection

"MicroSEQ® provides a very comprehensive coverage of needs in the field of microbial identification...the validated libraries, analysis criteria, and safety of analyzed data comply with our needs in a GMP regulated environment."

**Customer**
cGMP lab

"Public databases for identification purposes with isolate identification have their place and should be utilized in a confirmatory capacity only. Reliance upon unvalidated sequences of uncontrolled organisms would be of real concern when audited by any regulatory agency under any cGMP/cGLP requirements. In our lab, our use of such databases is verification of a validated library ID when the quality of the data and phylogenetic tree may not portray precise criteria as expected with duplicate runs."

**Customer**
Tissue bank agency

## References

1. Petti CA, Bosshard PP, Brandt ME et al. (2008) Interpretive Criteria for Identification of Bacteria and Fungi by DNA Target Sequencing; Approved Guideline. p. 1–73. In: Vol. 28, Clinical and Laboratory Standards Institute.
2. Weisburg WG, Barns SM, Pelletier DA et al. (1991) 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* 173:697–703.
3. Abraham Z (2012) Fungal Identification. p. 107–150. In M. Griffin and D. Reber (ed.), Microbial Identification: The Keys to a Succesful Program PDA and DHI Publishing, LLC, Bethesda, MD, USA.
4. Hartmann M, Howes CG, Abarenkov K et al. (2010) V-Xtractor: an open-source, high-throughput software tool to identify and extract hypervariable regions of small subunit (16S/18S) ribosomal RNA gene sequences. *J Microbiol Methods* 83:250–253.
5. Hellberg RS, Haney CJ, Shen Y et al. (2012) Development of a custom 16S rRNA gene library for the identification and molecular subtyping of *Salmonella enterica*. *J Microbiol Methods* 91(3):448–458.


life technologies™