

## The 16S Direct workflow

### Microbial identification using 16S gene sequencing on the SeqStudio Genetic Analyzer and analysis with the SmartGene web application

#### In this application note we present:

- A fast and economical workflow called “16S Direct” for bacterial identification at the species level by PCR and Sanger sequencing of the 16S rRNA gene
- The use of the Applied Biosystems™ BigDye™ Direct kit for high-resolution Sanger sequencing
- The use of the innovative Applied Biosystems™ SeqStudio™ Genetic Analyzer for capillary electrophoresis
- The use of the SmartGene™ web application service for easy sequence data management and organism identification
- The option to use the Applied Biosystems™ MicroSEQ™ Full Gene 16S rDNA PCR Kit and Sequencing Kit for customers who prefer quality-controlled reagents

#### Introduction

Rapid and accurate identification of infectious, fastidious, or noncultivable bacteria is a major challenge for microbiology laboratories and for public health surveillance. Based on the groundbreaking phylogenetic research work of Woese and others [1,2], sequencing of the 16S ribosomal RNA (rRNA) gene has emerged as the preferred method for taxonomic classification and identification of bacteria [3]. Today, Sanger sequencing of the 16S gene is recognized as the gold standard for identification at the species level.



The Applied Biosystems™ MicroSEQ™ 16S gene PCR sequencing system has long been trusted for microbial identification. In addition, several databases from public institutions and commercial companies are available for comparing 16S sequences with typed or annotated entries. Until recently, automated Sanger sequencing required the use of high-throughput instruments such as the Applied Biosystems™ 3500 or 3730 Genetic Analyzer. However, many investigators or laboratories may not need high-throughput solutions.

The SeqStudio Genetic Analyzer provides a rapid, low-cost alternative to high-throughput instruments. This instrument is ideal for microbiology laboratories that need to sequence individual samples or small batches in a fast, easy, and economical manner. The SmartGene web application service offers a module for bacteria that is the perfect complement to achieve rapid, reliable results. With the SmartGene Bacteria Module, the laboratory can deliver sequence-based identification of bacteria efficiently and confidently, while avoiding the complexity and expense of building a bioinformatics team and managing software and servers in-house. The SmartGene Bacteria Module provides users with constantly updated, proprietary reference databases, representing more than 15,400 bacterial species and reflecting up-to-date taxonomy.

Here we present a novel, fast, and economical 16S gene sequencing workflow called “16S Direct” that is optimized for use on the SeqStudio Genetic Analyzer, with downstream data analysis using the SmartGene Bacteria Module for bacterial identification.

### SmartGene Bacteria Module

The SmartGene Bacteria Module is an integrated suite of reference databases and bioinformatics tools, a workflow solution that guides the user from .ab1 files through to final report. Delivered securely via the internet as a multiuser software service, the SmartGene Bacteria Module requires neither implementation of hardware or specific software, nor local bioinformatics staff. The module provides a fully searchable sample sequence archive for its users, enabling easy comparisons with earlier cases and the tracking of organisms of interest over time. The SmartGene Bacteria Module is the ideal infrastructure to expand a laboratory’s expertise in bacterial sequencing and to keep up to date with emerging pathogens, advances in bacteriology, and changes in nomenclature and taxonomy.

### SeqStudio Genetic Analyzer and 16S Direct workflow

The SeqStudio Genetic Analyzer is an affordable capillary electrophoresis (CE) system that can perform both automated Sanger DNA sequencing and high-precision sizing and analysis of multicolor fluorescent DNA fragments. The instrument can accommodate a 96-well plate loaded with sequencing reactions that are sequentially electrophoresed in batches of four samples per run. The sequencing samples can be processed in a 30 min “short” run cycle, resulting in DNA sequences of about 500 bases each. This is sufficient to reconstruct the almost complete sequence (~1.5 kb) of the 16S gene and thus identify the species of a particular sample in only 4 sequencing reactions and 1 run.

The procedure described here can be adapted to individual needs, such as different primer designs or running the sequencing reactions on CE instruments (Research Use Only versions) such as the 3500 Series Genetic Analyzers. Other established or commercial 16S sequencing methods and protocols using Applied Biosystems™ BigDye™ sequencing chemistry will also work on the SeqStudio instrument. However, these methods may require more steps and time to prepare additional CE runs since more sequencing reactions are needed for assembling the complete gene sequence (Table 1).

**Table 1. Number of PCR reactions, sequencing reactions, and runs on the SeqStudio Genetic Analyzer (30 min each) required to identify a number of bacteria after 4 hr CE time (arbitrary time point).**

	No. of PCR reactions	No. of sequencing reactions	No. of CE runs needed on SeqStudio instrument	No. of bacteria identified after 4 hr (i.e., after 8 CE runs)
16S Direct workflow	2	4	1	8
MicroSEQ Full Gene 16S rDNA Sequencing Kit	3	6	2	5
Traditional lab-developed protocol	1	8	2	4

The basic idea behind the 16S Direct workflow is to reduce the number of sequencing reactions to 4 in order to fully use the running capacity of 4 capillaries on the SeqStudio instrument. This maximizes the throughput of samples within a given time window, while taking full advantage of the discriminatory power of the full 16S gene sequence.

## The 16S Direct workflow

An overview and an example timeline of the 16S Direct workflow are shown in Figure 1 and in more detail in Figure 2. Starting in the morning, individual bacterial colonies are picked from agar culture plates and DNA is extracted using any preferred manual or automated nucleic acid extraction method. Here we use the Applied Biosystems™ PrepMan™ Ultra Sample Preparation Reagent. PCR reactions “A” and “B” are then set up, followed by a quick check for quantity and expected sizes of the two amplicons by standard agarose or Invitrogen™ E-Gel™ agarose gels.

A total of 4 sequencing reactions are then set up using the BigDye Direct sequencing reagent and cycled. The reactions are then purified with Applied Biosystems™ BigDye™ XTerminator™ reagent, and the samples are electrophoresed on the SeqStudio Genetic Analyzer in the early afternoon. The instrument continuously generates a data set of 4 sequences every 30 min when operating in short-run mode. The sequence files are then immediately available for secondary data analysis using the SmartGene Bacteria Module to align the .ab1 files, create a consensus sequence for the sample, and interrogate the SmartGene reference databases to identify the organism.

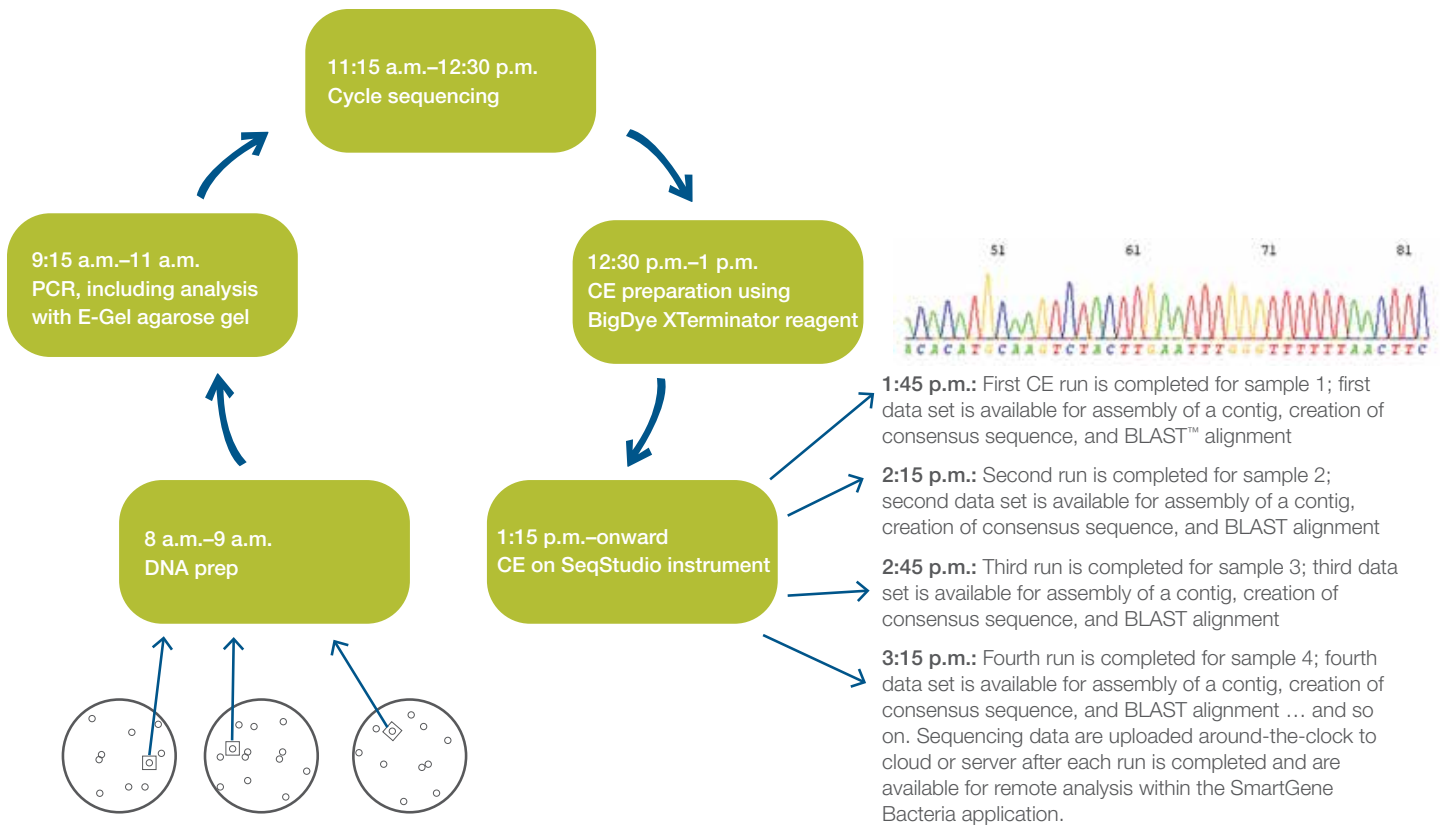


Figure 1. The 16S Direct workflow—an example timeline.

<b>DNA</b>	<ul style="list-style-type: none"> <li>• Prepare gDNA from bacterial isolate</li> <li>• Use your preferred or verified method to prepare PCR-ready DNA</li> </ul>	
<b>PCR</b>	<ul style="list-style-type: none"> <li>• Set up PCR reactions "A" and "B" with BigDye Direct PCR kit</li> </ul>	Time (manual): 2–10 min Time (on PCR cycler): 60 min
<b>E-Gel agarose gel</b>	<ul style="list-style-type: none"> <li>• Check PCR success with an E-Gel or other agarose gel</li> </ul>	Time (manual): 2–5 min Time (on device): 10–15 min
<b>Sequencing</b>	<ul style="list-style-type: none"> <li>• Set up forward and reverse sequencing reactions with amplicons "A" and "B" using the BigDye Direct sequencing kit</li> </ul>	Time (manual): 2–10 min Time (on PCR cycler): 70 min
<b>BigDye XTerminator reagent</b>	<ul style="list-style-type: none"> <li>• Purify sequencing reaction with BigDye XTerminator reagent</li> <li>• Transfer supernatant to CE plate</li> </ul>	Time (manual): 2–5 min Time (on device): 20–30 min
<b>CE</b>	<ul style="list-style-type: none"> <li>• Set up CE run on the SeqStudio Genetic Analyzer</li> <li>• Short run (each run of 4 sequencing samples = 1 specimen) or long run (each run of 4 sequencing samples = 1 specimen)</li> </ul>	Time (manual): 1–5 min Time (on instrument): 30 min Time (on instrument): 100 min
<b>SmartGene proofreader</b>	<ul style="list-style-type: none"> <li>• Assemble sequencing files with SmartGene target-specific proofreader</li> <li>• Save consensus sequence in SmartGene Bacteria Module</li> </ul>	Time (manual): 1–10 min
<b>SmartGene database</b>	<ul style="list-style-type: none"> <li>• Search sample sequence against SmartGene reference databases (e.g., 16S Centroids)</li> <li>• Review and verify top hit to identify organism</li> </ul>	Time (analytical data processing): 1–2 min typical, longer if complex

**Figure 2. Steps of the 16S Direct workflow.**

## Materials and methods

### DNA

The genomic DNA (gDNA) used for this research was from known bacterial strains obtained from ATCC and diluted to 10 ng/μL in Invitrogen™ TE buffer (10 mM Tris, 0.1 mM EDTA). To prepare purified DNA from bacteria growing on an agar plate or other biological or environmental matrix, the PrepMan Ultra Sample Preparation Reagent is a suitable product that is fast, economical, and easy to use. Other options are described at [thermofisher.com/gdna](http://thermofisher.com/gdna).

### PCR

For PCR of the 16S rRNA gene, the following primers were custom ordered from [thermofisher.com/oligo](http://thermofisher.com/oligo):

#### 16S\_A\_FWD (F8):

5'-tgtaaacgacggccagtAGAGTTTGATCMTGGCTCAG-3'

#### 16S\_A\_REV (R802):

5'-caggaacagctatgaccTACCAGGGTATCTAATCC-3'

#### 16S\_B\_FWD (F785):

5'-tgtaaacgacggccagtGGATTAGATACCCTGGTA-3'

#### 16S\_B\_REV (R1511):

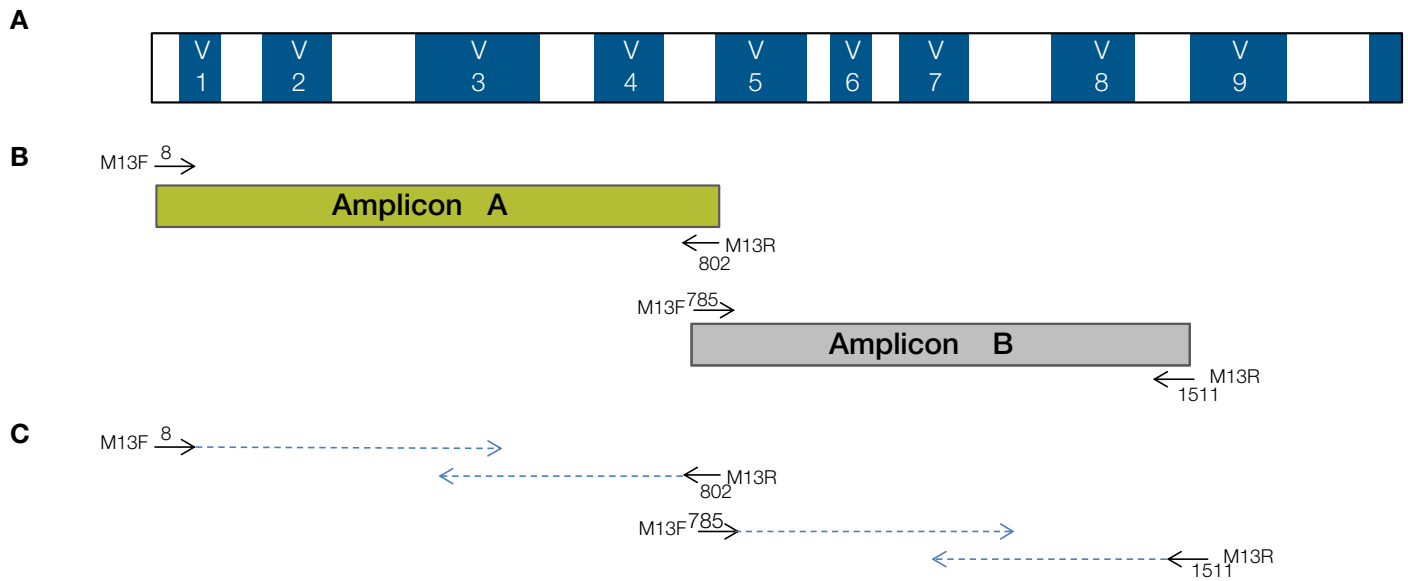
5'-caggaacagctatgaccCGGTTACCTTGTTACGACTT-3'

PCR set “A” amplifies the 5’ half of the 16S rRNA gene from position 8 to 802, and PCR set “B” amplifies the 3’ half of the 16S rRNA gene from position 785 to 1,511 (Figure 3).

Forward PCR primers are modified at the 5’ end with the M13 forward primer sequence, and likewise, reverse PCR primers with the M13 reverse primer sequence. M13 primer sequences are shown in lowercase letters. Note that the internal 16S gene-specific primers are complementary (Figure 3) and thus form a short overlapping segment of 18 bp that can be used to link segments A and B for full gene sequence assembly using the SmartGene Bacteria Module. The advantage of this primer design is that only 4 sequencing reactions are required for near full-length coverage of the 16S gene, which matches the 4-capillary capacity of the SeqStudio instrument for a single run.

For PCR, 5  $\mu$ L of the BigDye Direct PCR reagent was combined with 3.5  $\mu$ L of water and 0.5  $\mu$ L of primer pair A or B (each primer at 10  $\mu$ M) and 1  $\mu$ L of DNA (typically 1–10 ng) and amplified in an Applied Biosystems™ ProFlex™ PCR System using these cycling conditions: initial hot start at 95°C for 5 min, followed by 8 cycles at 95°C for 15 sec, 50°C for 15 sec, and 68°C for 90 sec, followed by 27 cycles at 95°C for 15 sec, 65°C for 15 sec, and 72°C for 90 sec.

After PCR, 15  $\mu$ L of low-EDTA TE buffer was added to the reaction and 2.5  $\mu$ L (~10%) was used for electrophoresis on a 2% E-Gel agarose gel to verify the presence of amplicon A (~830 bp) or amplicon B (~750 bp) as a clearly visible single band in an estimated amount of 20–50 ng.



**Figure 3. Sequencing primer-binding sites.** Schematic representation of the 16S rRNA gene showing (A) variable regions (V1–V9) with high sequence diversity between species, and general and conserved regions (white segments) that are targets for the M13-tagged primers. The segment between nucleotides 785–802 where the sequences overlap serves as a primer-binding site for the (B) PCR amplicons A and B. (C) The M13 tags are used as sequencing primer-binding sites.

## High-resolution Sanger sequencing with the BigDye Direct sequencing reagent

The BigDye Direct kit is comprehensive and economical, containing PCR reagent, BigDye Direct sequencing reagent, and unique M13 sequencing primers that bind to the M13 forward and reverse sequences present in the PCR amplicon. The BigDye Direct reagent provides high resolution of the nucleotide sequences close to the 5' end and does not require a separate PCR primer purification step, thereby saving time and extra reagents. A sample setup of 4 cycle sequencing reactions for sequencing M13-tagged PCR amplicons A and B in both directions is shown in Table 2.

Cycle sequencing on a ProFlex PCR System was performed using this profile: 1 cycle at 37°C for 15 min, 80°C for 2 min, and 96°C for 1 min; 13 cycles at 96°C for 10 sec, 50°C for 5 sec, and 60°C for 75 sec; 6 cycles at 96°C for 10 sec, 50°C for 5 sec, and 60°C for 90 sec; 6 cycles at 96°C for 10 sec, 50°C for 5 sec, and 60°C for 150 sec. After cycle sequencing, 50 µL of BigDye XTerminator suspension was added to each sequencing reaction. The suspension consisted of 1 part BigDye XTerminator beads (10 µL/sample) and 4 parts of SAM

Solution (40 µL/sample) and was prepared in bulk for all samples shortly before use. The sequencing plate with the BigDye XTerminator suspension added to the samples was vigorously vortexed for 30 min and then centrifuged for 2 min to pellet the beads; 20 µL of the supernatant was transferred to a fresh plate, centrifuged for 1 min to remove any residual air bubbles, and then placed on the SeqStudio Genetic Analyzer for CE.

## Sequencing by CE on the SeqStudio Genetic Analyzer

The samples can be run as either short (30 min), medium (45 min), or long runs (105 min) using BigDye Direct chemistry and a regular sequencing run module (note: do not use a BigDye XTerminator run module). Long runs yield nearly complete bidirectional coverage, whereas short runs yield approximately complete, 40% bidirectional coverage. Complete and accurate assembly of nearly full-length 16S gene sequences (~1,450 bases) can be reliably achieved with high-quality (quality value >35) short-run data. After completion of a CE run cycle, the 4 sequencing files for PCR segments A and B are immediately available for data analysis and also are remotely retrievable through internet cloud connectivity. This helps enable urgent identification analysis away from the lab or after lab hours.

**Table 2. Setting up the forward and reverse BigDye Direct sequencing reactions.** Note that the Applied Biosystems™ BigDye™ 5X sequencing buffer is not included in the BigDye Direct kit and must be obtained separately. The 5X buffer is part of the standard BigDye sequencing v1.1 and v3.1 kits and can also be purchased as a stand-alone item.

Reagent	Volume for 2 forward sequencing reactions (1 specimen)	Volume for 2 reverse sequencing reactions (1 specimen)
Molecular biology–grade water (not included in BigDye Direct kit)	12 µL	12 µL
BigDye v1.1 and v3.1 sequencing buffer, 5X (not included in the BigDye Direct kit)	3 µL	3 µL
BigDye Direct M13 forward primer (included in the BigDye Direct kit)	1 µL	–
BigDye Direct M13 reverse primer (included in the BigDye Direct kit)	–	1 µL
BigDye Direct sequencing reagent (included in the BigDye Direct kit)	2 µL	2 µL
Total volume	18 µL (= 2 x 9 µL)	18 µL (= 2 x 9 µL)
	↓	↓
Well or tube 1	9 µL forward sequencing reaction + 1 µL amplicon A	
Well or tube 2	9 µL reverse sequencing reaction + 1 µL amplicon A	
Well or tube 3	9 µL forward sequencing reaction + 1 µL amplicon B	
Well or tube 4	9 µL reverse sequencing reaction + 1 µL amplicon B	



## Data analysis

The output files from the SeqStudio Genetic Analyzer are sequencing files in .ab1 format, which can be viewed in electropherogram trace mode or in text mode. In order to reconstitute the entire sequence of the 16S rRNA gene from both strands of amplicons A and B, the .ab1 files need to be assembled, i.e., an overlap needs to be established and a contiguous consensus sequence (“contig”) needs to be generated. This can be conveniently done using the target-specific 16S proofreader tool within the SmartGene Bacteria Module.

1. Open a web browser, use the secure URL to go to the SmartGene web application, and log in with a username and password to access the SmartGene Bacteria Module (Figure 4).
2. Click on “Add record to Demo Bacteria” in the “Sample Sequences” menu and enter the sample number and metadata.
3. Open the target-specific 16S proofreader tool and upload the four .ab1 files from the SeqStudio Genetic Analyzer for the relevant sample. The SmartGene Bacteria Module automatically selects a close alignment reference for the unknown bacterium, presenting a draft alignment and consensus sequence for review by the user (Figure 5).



Figure 4. Example home page of the SmartGene Bacteria Module.



Figure 5. Display of the draft alignment and consensus sequence for proofreading.

- Use the embedded tools in the SmartGene Bacteria Module to efficiently review the alignment, trim sequences, and correct base-calling errors, if necessary. Conserve true mixed-base positions (dual peaks in the electropherograms) that reflect inter-operon variability of the 16S gene. All editing steps by the user are recorded automatically in the audit and log trail, which is embedded in the SmartGene Bacteria Module.
- Save the consensus 16S sequence within the sample record in the SmartGene Bacteria Module, together with the alignment and the original .ab1 files from the SeqStudio Genetic Analyzer, for full traceability.
- Use the consensus sequence to search SmartGene curated reference databases (under "Reference Sequences"; Figure 4), which power the SmartGene Bacteria Module and enable confident species identification. Start by searching the sample sequence against the proprietary SmartGene 16S Centroids reference database, which contains the best representative sequence for each valid species according to current bacterial taxonomy (Figure 6).

Select	Action	Dataset	Seq. length	Last updated	Last modified by	Sample date	Sample ID	16S creation date	Length 16S	Query sequence - locus 16S bacteria		
<input type="checkbox"/>	more...	Demo Bacteria	1484	13.04.2020	sgdemo	09.05.2018	TF1short	13.04.2020	1484			
Similar sequences found												
Select	Action	Dataset	AC - Accession	OS - Organism	Species group size	Centroid	IDNS valid OS	Seq. length	Identities	Mismatches	Match length	Score
<input type="checkbox"/>	more...	IDNS 16S Centroids	KP326374	Bacteroides fragilis	140	Y	Bacteroides fragilis	1533	1479 (99.73%)	4	1483	2341
<input type="checkbox"/>	more...	IDNS 16S Centroids	AE015928	Bacteroides thetaiotaomicron VPI-5482	78	Y	Bacteroides thetaiotaomicron	1534	1394 (93.87%)	91	1485	2039
<input type="checkbox"/>	more...	IDNS 16S Centroids	AB510704	Bacteroides nordii	13	Y	Bacteroides nordii	1473	1386 (93.78%)	92	1478	2029
<input type="checkbox"/>	more...	IDNS 16S Centroids	AB547646	Bacteroides rodentium	3	Y	Bacteroides rodentium	1488	1382 (93.19%)	101	1483	2023
<input type="checkbox"/>	more...	IDNS 16S Centroids	AB510700	Bacteroides eggerthii	17	Y	Bacteroides eggerthii	1488	1379 (92.99%)	104	1483	2013
<input type="checkbox"/>	more...	IDNS 16S Centroids	LT622246	Bacteroides ovatus V975	55	Y	Bacteroides ovatus	1534	1385 (93.27%)	100	1485	2010
<input type="checkbox"/>	more...	IDNS 16S Centroids	CP015401	Bacteroides caecimuris	6	Y	Bacteroides caecimuris	1534	1385 (93.27%)	100	1485	2010
<input type="checkbox"/>	more...	IDNS 16S Centroids	MK929066	Bacteroides faecichinchillae	2	Y	Bacteroides faecichinchillae	1529	1381 (93.12%)	102	1483	2008
<input type="checkbox"/>	more...	IDNS 16S Centroids	AB547640	Bacteroides faecis	12	Y	Bacteroides faecis	1475	1377 (93.61%)	94	1471	2005

Figure 6. Example of a display of the top matches for a BLAST search using the SmartGene 16S Centroids reference database.



If the sample sequence does not match conclusively to a 16S Centroids sequence, the user can search for a better-matching variant of the species in the SmartGene 16S Eubacteria reference database, which contains more than 950,000 profile-filtered entries (Figure 7).

## Results

### Long- and short-read sequencing runs

To demonstrate the feasibility of the primer design and workflow described above, we have processed DNA derived from 5 representative microbial organisms belonging to 4 different phyla: Bacteroidetes, Proteobacteria, Actinobacteria, Firmicutes, and Verrucomicrobia. Table 3 shows a summary of the long- and short-read sequencing results from the SeqStudio

Genetic Analyzer along with the identification results obtained using the SmartGene 16S Centroids and 16S Eubacteria reference databases.

The quartet of sequencing files was analyzed using the SmartGene Bacteria Module in 3 ways:

- Consensus of forward and reverse reads of amplicon A (sequencing reactions 1 and 2)
- Consensus of forward and reverse reads of amplicon B (sequencing reactions 3 and 4)
- Consensus of forward and reverse reads of segments "A and B" (sequencing reactions 1, 2, 3, and 4)

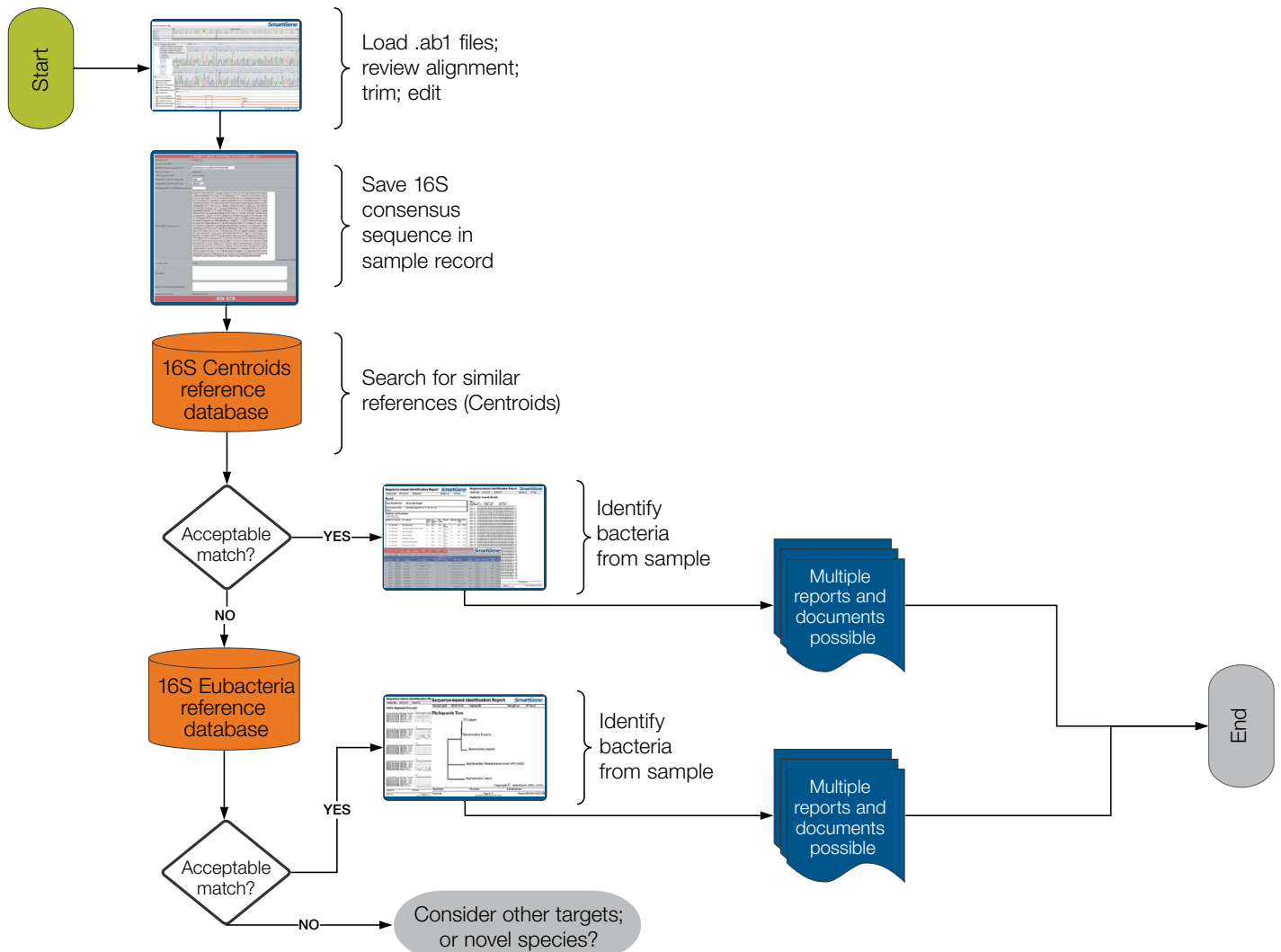


Figure 7. Bacterial identification workflow using reference databases in the SmartGene Bacteria Module.

The short sequencing run only takes about 30 min for the 4 sequencing reactions and yields a total of over 2,000 nucleotides with high quality values (QVs). About 40% of the total sequence is covered in both forward and reverse directions, and about 60% of the sequence is covered in only one direction. The sample score values for the portions with single-strand coverage were usually >40 and thus exhibited very high confidence in

base call accuracy. Table 3 shows that all samples were identified correctly at above 99.5% identity rates using the consensus sequences for PCR segments A or B alone or A + B in combination. This demonstrates that a full 16S gene sequence can be obtained for a given specimen with a short CE run of 30 min (compared to about 105 min with a long CE run).

**Table 3. Summary of long and short CE run sequence data metrics, alignment results, and identification of the microbial DNA test panel representing 4 bacterial phyla.**

DNA input of known organism (sample)	Amplicon	SeqStudio run mode	Length of consensus sequence (nt)	Best match in SmartGene 16S Centroids reference database	Identity	Number of mismatches (IUPAC aware)	Match length
<i>Bacteroides fragilis</i> ATCC 25285D-5	A contig	Long	797	<i>Bacteroides fragilis</i>	100.00%	0	797
	B contig	Long	710	<i>Bacteroides fragilis</i>	99.58%	3	707
	A + B contig	Long	1,496	<i>Bacteroides fragilis</i>	99.73%	4	1,496
				<i>Bacteroides thetaiotaomicron*</i>	92.97%	107	1,522
	A contig	Short	792	<i>Bacteroides fragilis</i>	100.00%	0	792
	B contig	Short	708	<i>Bacteroides fragilis</i>	99.58%	3	708
	A + B contig	Short	1,484	<i>Bacteroides fragilis</i>	99.80%	3	1,483
<i>Acinetobacter baumannii</i> ATCC 17978D-5	A contig	Long	790	<i>Acinetobacter baumannii</i>	99.87%	1	790
	B contig	Long	727	<i>Acinetobacter baumannii</i>	99.86%	1	727
	A + B contig	Long	1,500	<i>Acinetobacter baumannii</i>	99.87%	2	1,499
				<i>Acinetobacter nosocomialis*</i>	97.94%	31	1,504
	A contig	Short	790	<i>Acinetobacter baumannii</i>	99.87%	1	790
	B contig	Short	684	<i>Acinetobacter baumannii</i>	100.00%	0	684
A + B contig	Short	1,499	<i>Acinetobacter baumannii</i>	99.87%	2	1,499	
<i>Actinomyces naeslundii</i> ATCC 12104D-5	A contig	Long	810	<i>Actinomyces naeslundii</i>	99.87%	1	789
	B contig	Long	729	<i>Actinomyces naeslundii</i>	99.86%	1	726
	A + B contig	Long	1,522	<i>Actinomyces naeslundii</i>	99.87%	2	1,498
				<i>Actinomyces viscosus*</i>	97.31%	41	1,526
	A contig	Short	811	<i>Actinomyces naeslundii</i>	99.87%	1	790
	B contig	Short	664	<i>Actinomyces naeslundii</i>	99.85%	1	664
A + B contig	Short	1,522	<i>Actinomyces naeslundii</i>	99.87%	2	1,498	
<i>Akkermansia muciniphila</i> ATCC BAA-835D-5	A contig	Long	764	<i>Akkermansia muciniphila</i>	99.87%	1	758
	B contig	Long	728	<i>Akkermansia muciniphila</i>	99.72%	2	725
	A + B contig	Long	1,425	<i>Akkermansia muciniphila</i>	99.79%	3	1,425
				<i>Akkermansia glycaniphila*</i>	93.53%	93	1,438
	A contig	Short	758	<i>Akkermansia muciniphila</i>	99.87%	1	758
	B contig	Short	725	<i>Akkermansia muciniphila</i>	99.72%	2	725
A + B contig	Short	1,425	<i>Akkermansia muciniphila</i>	99.79%	3	1,425	
<i>Streptococcus agalactiae</i> ATCC BAA-1138D-5	A contig	Long	802	<i>Streptococcus agalactiae</i>	99.75%	2	802
	B contig	Long	728	<i>Streptococcus agalactiae</i>	99.86%	1	728
	A + B contig	Long	1,511	<i>Streptococcus agalactiae</i>	99.80%	3	1,511
				<i>Streptococcus dysgalactiae*</i>	97.55%	37	1,511
	A contig	Short	801	<i>Streptococcus agalactiae</i>	99.75%	2	801
	B contig	Short	708	<i>Streptococcus agalactiae</i>	100.00%	0	708
A + B contig	Short	1,511	<i>Streptococcus agalactiae</i>	99.80%	3	1,511	

\* Next-best match in SmartGene 16S Centroids reference database to A + B contig, long run mode.

## Conclusions

In this application note, we demonstrated the feasibility of sequencing the 16S rRNA gene on the SeqStudio Genetic Analyzer for the purpose of bacterial identification. The 16S Direct workflow that was presented here for use on the SeqStudio instrument could also be applied to other Applied Biosystems™ CE instruments such as the 3130 or 3500 Series instruments with similar run modules.

The 16S Direct workflow includes a novel and streamlined primer design that reduces the number of PCR amplicons to 2 and the number of sequencing reactions to 4 to obtain nearly full-length DNA sequences for querying genomic databases such as the curated 16S reference databases contained within the SmartGene Bacteria Module. To that end, the BigDye Direct PCR and sequencing reagents generate high-resolution and

high-quality sequence reads right from the beginning of the primer. A short CE sequencing run of 30 min loaded with 4 sequencing reactions is sufficient to yield approximately 2,000 high-quality bases (i.e., 4 x ~500 bases) for subsequent 16S gene contig assembly and downstream analysis using the SmartGene web application service.

Complementing the SeqStudio Genetic Analyzer, the SmartGene Bacteria Module enables reliable, quick, and confident sequence-based identification of bacteria. This fast and economical workflow will benefit microbiology laboratories that use DNA sequencing as their ultimate bacterial identification and characterization tool. The affordable, compact, and versatile SeqStudio Genetic Analyzer makes it easy to introduce molecular genetic techniques into the general microbiology laboratory.

## Ordering information

Product	Cat. No.
SeqStudio Genetic Analyzer	A35644
BigDye Direct Cycle Sequencing Kit	4458687
BigDye Terminator v1.1 & v3.1 5X Sequencing Buffer	4336697
BigDye XTerminator Purification Kit	4376486
ProFlex 3 x 32-Well PCR System	4484073
PrepMan Ultra Sample Preparation Reagent	4318930
E-Gel General Purpose Agarose Gels, 2%	G501802
TE Buffer	12090015
Nuclease-Free Water	AM9937

For SmartGene Bacteria license terms, email [SeqStudio@smartgene.com](mailto:SeqStudio@smartgene.com).

## References

1. Woese CR et al. (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci USA* 87:4576–4579.
2. Weisburg WG et al. (1991) 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* 173:697–703.
3. Clarridge JE 3rd (2004) Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. *Clin Microbiol Rev* 17:840–862.

Find out more at [thermofisher.com/seqstudio](http://thermofisher.com/seqstudio)