SeqStudio Genetic Analyzer

The latest innovation for Sanger sequencing and fragment analysis

In this application guide, we introduce the Applied Biosystems™ SeqStudio™ Genetic Analyzer and demonstrate:

- Sequencing and fragment analysis applications using the all-new SeqStudio Genetic Analyzer
- Intuitive run setup and real-time monitoring of quality control results
- Instrument connectivity that allows seamless integration into the existing sequencing application portfolio in Thermo Fisher Cloud
- Highlights of several powerful Sanger sequencing and fragment analysis applications, including:
  - **Plasmid sequencing**
  - **Oncology research, next-generation confirmation**
  - **Species identification**
  - **CRISPR-Cas9 genome editing analysis**
  - **Human cell line authentication**
  - **Applied Biosystems™ SNaPshot™ genotyping**
  - **Multiplex ligation–dependent probe amplification (MLPA™) analysis of human copy number variation (CNV)**

Introduction

The foundations of Applied Biosystems™ genetic analyzers are built upon reliability and trusted results. The SeqStudio Genetic Analyzer, the newest member of the Applied Biosystems™ line of capillary electrophoresis (CE) genetic analysis instruments, is built on this foundation, but designed from the ground up to be easy to use, easy to maintain, and easy to access and share data. In this application guide, we introduce the state-of-the-art features found nowhere else—features which combine the reliability of Sanger sequencing and fragment analysis with the convenience of remote monitoring and cloud-based applications.

ThermoFisher SCIENTIFIC

All-in-one cartridge

The easy-to-use functional core of the instrument is the result of a new cartridge design that helps maximize efficiency and convenience. The SeqStudio Genetic Analyzer utilizes an all-in-one cartridge that contains the capillary array, polymer reservoir, and anode buffer. The cartridge is removable and can be stored on the instrument for up to four months. Each cartridge contains a new polymer unique to the SeqStudio system that allows Sanger sequencing and fragment analysis to be performed with no reconfiguration. The cartridge has four capillaries, and can process samples from either standard 96-well plates or 8-well strip tubes. The cartridge and cathode buffer container includes radio-frequency identification (RFID) tags that track the number of injections (cartridge) and length of time on the instrument (cathode buffer container). This allows scientists, using the same instrument, to maintain custody of their own cartridges and use them only when needed, providing another level of flexibility. Finally, maintenance of the SeqStudio Genetic Analyzer has also been simplified. Instrument calibrations are all handled automatically by leveraging advancements in imaging and algorithm tools.

Compact instrument design

The SeqStudio instrument has a small footprint, with an onboard computer and integrated touch screen that makes run setup quick, intuitive, and flexible (Figure 1). The running software was configured to allow sequencing and fragment analysis experiments to be set up on the same plate and run on the instrument without interruption between the methods. This opens up new opportunities for streamlining analyses—for example, combining a sequence-based, locus-screening mutation test with a fragment analysis–based CNV test on the same CE run.

Figure 1. Real-time monitoring of runs on the SeqStudio Genetic Analyzer. (A) The SeqStudio instrument displays results for each capillary in real time. **(B)** Once an injection is finished, a number of quality checks are calculated and displayed. If an injection produces poor traces or poor QC values, those samples can be re-injected, with altered injection parameters, if desired. **(C)** Screenshot from an off-site computer monitoring shows the progress of a run. **(D)** Runs set up in PlateManager can be uploaded directly to the instrument. PlateManager allows investigators to assign sequencing and fragment analysis runs on the same plate, taking advantage of the universal polymer in the cartridge.

Connected experience

Another level of convenience was added by integrating wireless connectivity into the instrument (Figure 2). This makes the SeqStudio instrument accessible via the onboard interface, a remote computer, or a mobile device app. Runs can be set up using either the onboard computer or by using PlateManager, the stand-alone software that operates within Thermo Fisher Connect or on a separate computer (Figure 2). By using web browser–based software, access to run setup, plate maps, run conditions, and analysis settings are all immediately available from anywhere you have Internet access. Injection conditions, reinjections, and reordering of injections can all be modified during the run, maximizing the ability to collect quality data from each plate. After data collection, the web browser–based suite of applications (Sanger Quality Check, Sanger Variant Analysis and Next-Generation Confirmation (NGC)) makes analysis painless and accessible. Determination of DNA sequence variants, alignments, and fragment analysis are all available immediately upon completion of a run. Finally, the cloud connectivity enables collaborators in different locations to monitor, access, share data information, and rapidly analyze the same data sets.

In this applications guide, we demonstrate how the new SeqStudio Genetic Analyzer integrates seamlessly into some of the most commonly used Sanger sequencing and fragment analysis applications. The ease of use and ease of data analysis provided by the SeqStudio Genetic Analyzer give investigators maximum flexibility and make it an ideal system to add to the genetic analysis toolbox.

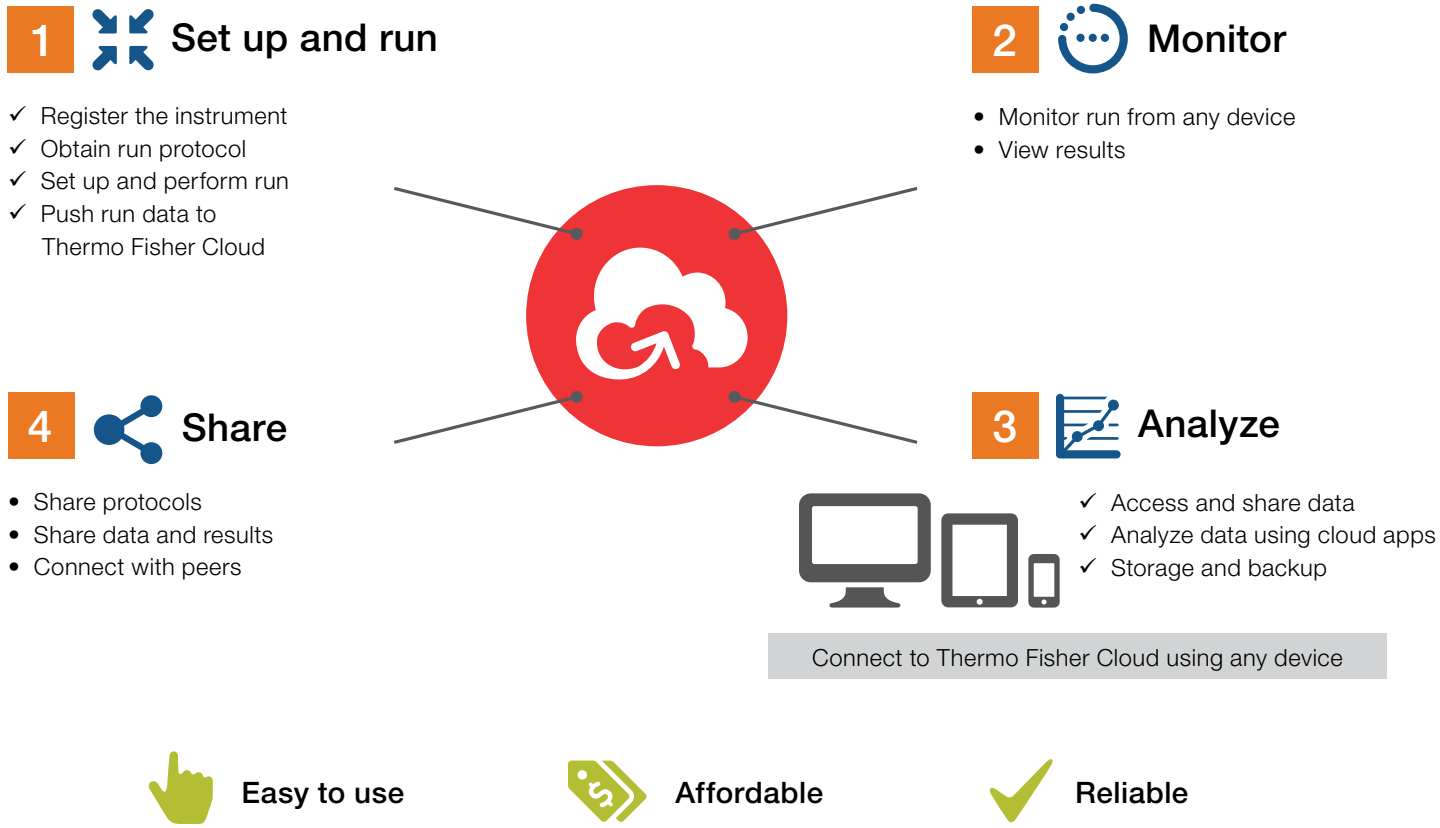


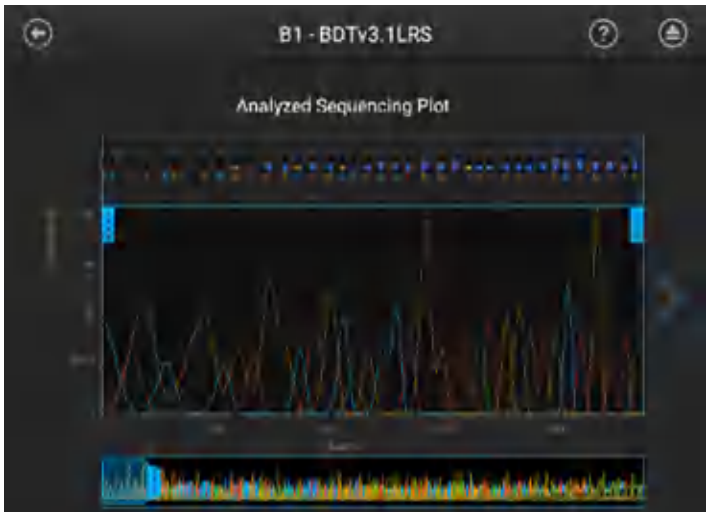
Figure 2. SeqStudio Genetic Analyzer integrates seamlessly into the Thermo Fisher Cloud. By registering your instrument and logging into your cloud account, remote features are accessible, including instrument monitoring, data analysis using cloud apps, and data sharing with your colleagues.

Plasmid sequencing

One of the most common applications of Sanger sequencing is the analysis of inserts subcloned into plasmids. Applied Biosystems™ BigDye™ chemistries are widely used for Sanger sequencing and an integral part of plasmid sequencing workflows. Several of the new features on the SeqStudio platform offer benefits to researchers performing basic plasmid sequencing methods. The instrument is preloaded with sequencing modules optimized for short (<300 bp), medium (500 bp), and long (>600 bp) read lengths, and can also be customized on the instrument to meet specific needs. The swappable cartridges can be associated with individual projects and users. The cloud-based Sanger **Quality Check** application provides an intuitive set of tools to analyze sequencing traces. Finally, the cloud connectivity for remote monitoring, accessing, and sharing sequencing information can help collaborators rapidly analyze the same data sets.

The performance of the SeqStudio instrument for plasmid sequencing was determined by sequencing the pGEM7zf+ plasmid with M13 primers and Applied Biosystems™ BigDye™ Terminator v3.1 chemistry. Results were obtained by analyzing the sequencing traces using the Sanger Quality Check module on the **Thermo Fisher Cloud** (Figure 3). In the example shown, the same plasmid was sequenced in 16 wells and analyzed on the SeqStudio Genetic Analyzer in 4 different injections. Note that the trace score, peak under peak (PUP) values, contiguous read length (CRL), and QV20+ (length with quality values >20) are similar for each sample. Similar results were obtained in traces on the other strand, and in other experiments by using Applied Biosystems™ BigDye™ Terminator v1.1 chemistry. These data demonstrate that the SeqStudio platform can generate plasmid sequencing results of very high quality.

A



B

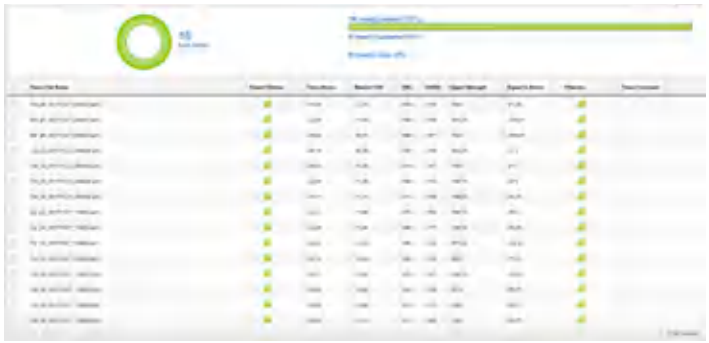


Figure 3. Analysis of sequencing quality using the Sanger Quality Check Cloud app. (A) Once a run is completed, the SeqStudio instrument displays the resulting sequence file as well as the quality scores for each base. **(B)** Sixteen separate pGEM7zf+ sequencing reactions were run on the SeqStudio instrument and the .ab1 files were uploaded to the cloud and analyzed. Note that the sequencing metrics were very similar in the sixteen different reactions. CRL = contiguous read length, QV20+ = number of nucleotides with a quality value >20.

Oncology research, next-generation confirmation

The SeqStudio Genetic Analyzer can be used by clinical researchers to maintain the gold-standard quality for detecting and verifying the presence of mutant alleles in tumor tissue. The SeqStudio system integrates with the following tools to simplify Sanger sequencing workflows:

- The SeqStudio Genetic Analyzer comes preloaded with running modules optimized for fragmented DNA extracted from formalin-fixed, paraffin-embedded tissue.
- The cloud-based **NGC module** allows investigators to rapidly verify variants identified in next-generation sequencing (NGS) .vcf files using Sanger sequencing traces.
- Allelic variants at frequencies down to 5% can be detected using the Applied Biosystems™ Minor Variant Finder (MVF) Software and Sanger traces generated by the SeqStudio instrument.
- Applied Biosystems™ BigDye™ Direct and BigDye XTerminator™ chemistries simplify the Sanger sequencing workflow by providing one-tube sequencing and clean-up.

The performance of the SeqStudio Genetic Analyzer for detecting mutant alleles in tumor samples was determined by analyzing genomic DNA extracted from 10 different FFPE tumor samples, and determining variant frequencies at 4 different hotspot regions. The frequency of mutant alleles was determined by NGS using the Ion Torrent™ Oncomine™ Oncology Focus Panel, and Sanger sequencing using BigDye Direct/BigDye XTerminator chemistries and MVF Software. The correlation between the frequencies measured by the SeqStudio Genetic Analyzer was excellent when compared to NGS at allele frequencies—from about 9% to about 70% (Figure 4).

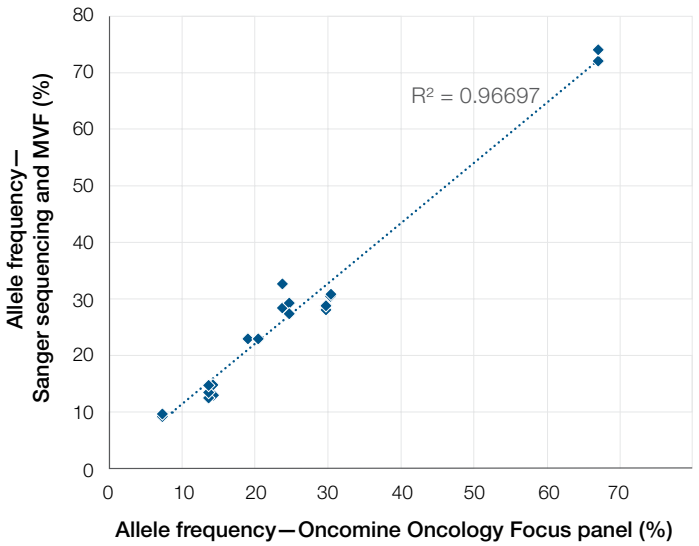
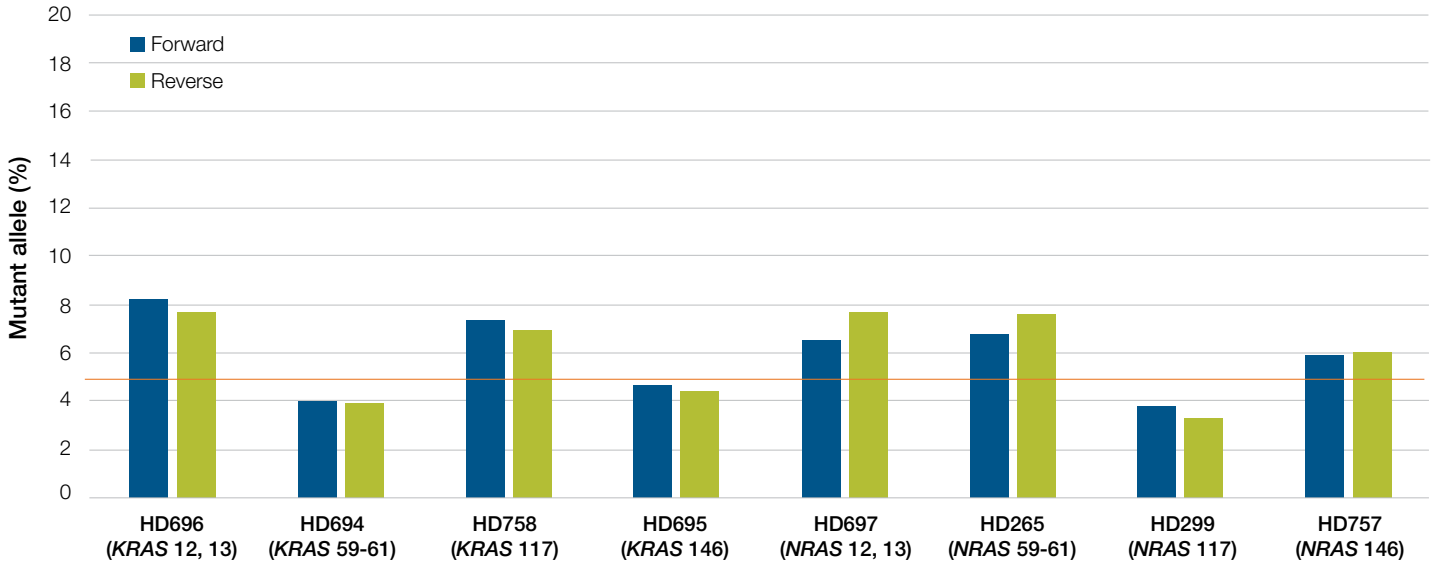


Figure 4. Analysis of FFPE samples using the SeqStudio instrument at allele frequency of 9–70%. Ten different tumor samples were analyzed for mutations at 4 different hotspots by Ion Torrent™ Sequencing and Sanger sequencing on the SeqStudio instrument. The mutant allele frequency correlated extremely well between the two methods across a wide range of allele frequencies.

The ability of the SeqStudio Genetic Analyzer to analyze variant frequencies was also determined using a 96-well plate containing Sanger sequencing primers that query the most common tumorigenic mutations in *KRAS* and *NRAS*. The minor allele frequency analysis of SeqStudio instrument traces accurately measured the allele frequencies in 1 ng of diluted FFPE-extracted DNA (Figure 5A). Therefore, researchers needing to detect rare alleles can be confident that the SeqStudio Genetic Analyzer will produce accurate results on FFPE tissues. For more details, see [1].

Finally, the cloud-based NGC application simplifies the confirmation of variants identified by NGS by organizing Sanger sequencing traces by amplicons and specimens, and aligning them in the proper orientation to the candidate variant sequences in a .vcf file. To show the utility of the NGC app in an oncology workflow, we confirmed the presence of an *NRAS* mutation identified using the Oncomine Oncology Focus panel by Sanger sequencing (Figure 5B). The SeqStudio results verified that the mutation in *NRAS* (p.Ala59Thr) was present. Therefore, focused and rapid examination of the most meaningful portions of sequencing traces by the NGC app facilitates NGS variant confirmation.

A



B

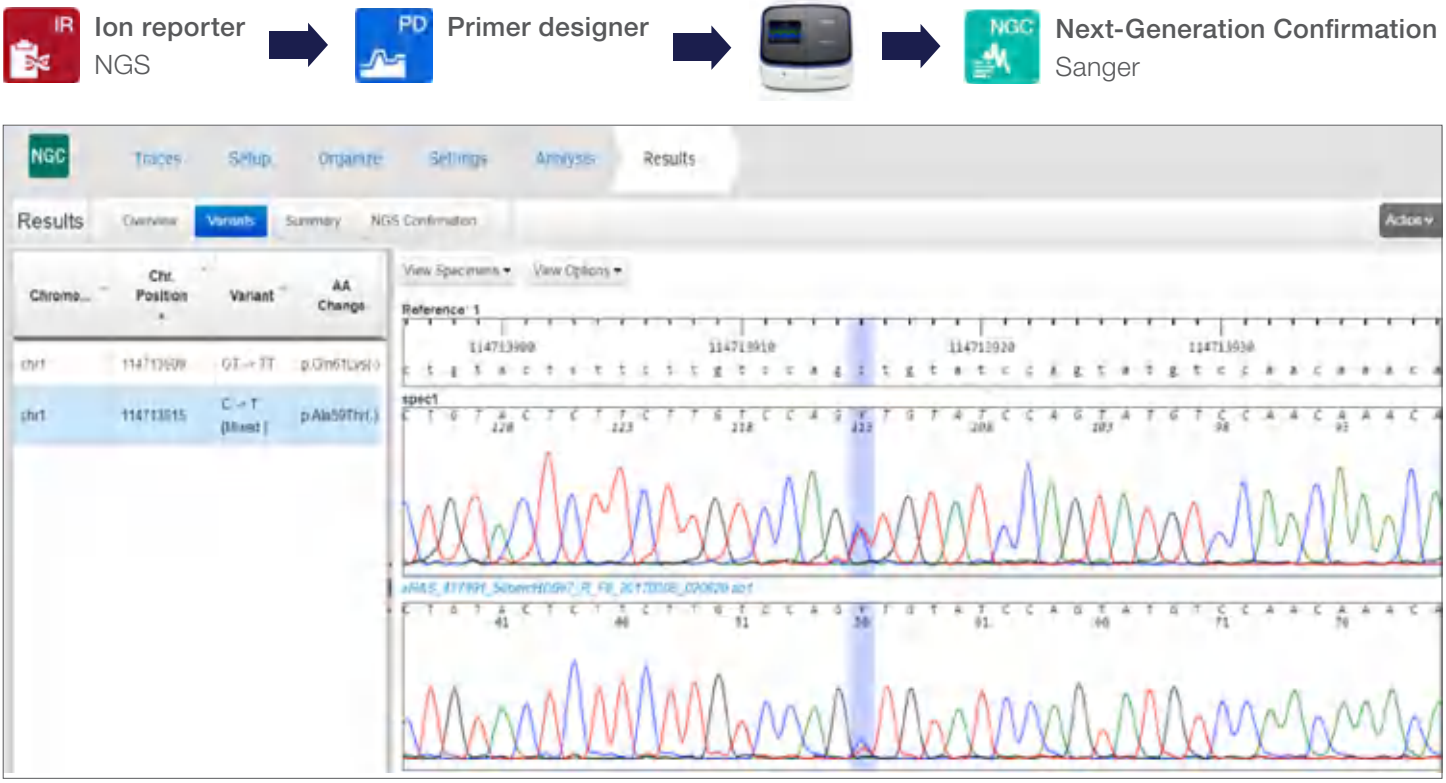


Figure 5. Analysis and confirmation of variants by SeqStudio Genetic Analyzer and the NGC application, respectively. (A) Eight different FFPE samples with mutations at known RAS hotspots were diluted to 5% allele frequency, then analyzed using a 96-well plate containing Sanger sequencing primers that query the most common tumorigenic mutations in *KRAS* and *NRAS*, and using the SeqStudio Genetic Analyzer. Each of the allele queries accurately measured the allele frequencies; deviations from 5% reflected slight inconsistencies in starting concentration of the samples. Yellow line is 5% frequency. Similar results were seen with 10% and 50% dilutions. **(B)** Confirmation of variants identified by NGS. From a .vcf file generated using Ion Reporter™ Software, Sanger sequencing primers targeting loci of interest were ordered from Primer Designer, samples were sequenced on the SeqStudio instrument, and variants common to the .vcf file and the Sanger sequencing traces were highlighted using the NGC cloud app.

Species identification

The ready availability of genomic data opens the opportunity to identify species in an unknown sample by sequencing DNA of "fingerprint" loci. The Applied Biosystems™ family of kits, for example, the MicroSEQ™ kit, has simplified the identification of prokaryotes and fungi by Sanger sequencing ribosomal DNA (rDNA) sequences [2]. Similarly, eukaryotic organisms can be identified using the mitochondrial-specific loci as the identifying locus. This strategy has been exploited in the Barcode of Life project (**barcodeoflife.org**, for review see [3]), providing a means for rapidly establishing the identity of unknown eukaryotic samples.

To illustrate the performance of the SeqStudio Genetic Analyzer for microbial identification, we obtained genomic DNA samples from ATCC for a variety of microorganisms, and sequenced them using the Applied Biosystems™ MicroSEQ™ 500 PCR kit and the SeqStudio instrument. The resulting sequences were queried against the BLAST database. For each sequencing reaction, the correct organism was identified with the highest BLAST confidence. Similarly, using primers for fish mitochondrial sequences (*CO1* gene) and fish samples, the fish species was correctly identified as the top BLAST hit. The accurate identification of the species queried with BLAST illustrates how well the SeqStudio platform can be used for species identification.

Table 1. Analysis of species ID using the SeqStudio Genetic Analyzer. Samples of microorganism DNA or genomic DNA extracted from fish were sequenced using primers for 16s rDNA and the MicroSEQ kit (BigDye Terminator v1.1 chemistry), or using primers for fish mitochondrial *CO1* sequences and BigDye Terminator v3.1 chemistry.

	Number of organisms	Number of queries	Percent correct
Microorganisms	24	48	100
Piscine organisms	12	24	100

Genome editing

Genome editing technologies, including CRISPR-Cas9–mediated editing events, are rapidly becoming accessible to a majority of biological science researchers, and are poised to revolutionize all fields of biology and health care. Thermo Fisher Scientific offers all the tools necessary for a genome editing project. As an integral part of such a project, the features of the SeqStudio Genetic Analyzer facilitate Sanger sequencing analyses and fit well within a genome editing workflow. In particular, the data generated are compatible with Tracking of Indels by Decomposition (TIDE) software [4], a widely available tool for analyzing the efficiency of genome editing events.

The utility of the SeqStudio Genetic Analyzer in a genome editing project was shown by obtaining whole-cell lysates from HEK293 cells that were edited to introduce random deletions around a targeted site in the *HPRT* or the *relA* locus. To confirm the position of the edit, the Sanger sequencing traces were uploaded to the cloud and analyzed using the **Sanger Variant Analysis module** (Figure 6). Note that the position of the edit is clearly indicated and can be visualized by the abundant mixed-base peaks downstream of the break. The efficiency of the edits in this mixed primary culture was determined by analyzing these trace files using the TIDE software. In each case, the spectrum and frequencies of deletions at each locus was nearly identical using the data generated in the forward and reverse directions (Figure 7). These frequencies confirm results obtained using Invitrogen™ TOPO™ cloning and followed by Sanger sequencing results of the same edited cell populations [5].

Figure 6. Analysis of genome-edited samples using the SeqStudio instrument. A mixed population of cells with a genome editing event at the human *HPRT* locus was analyzed using the cloud-enabled Sanger Variant Analysis app. Note that this app finds single-nucleotide variants common to both forward and reverse strands, but is also able to detect where the genome cleavage event occurred, producing a population of mixed sequences downstream (to the right in this example) of the breakpoint.

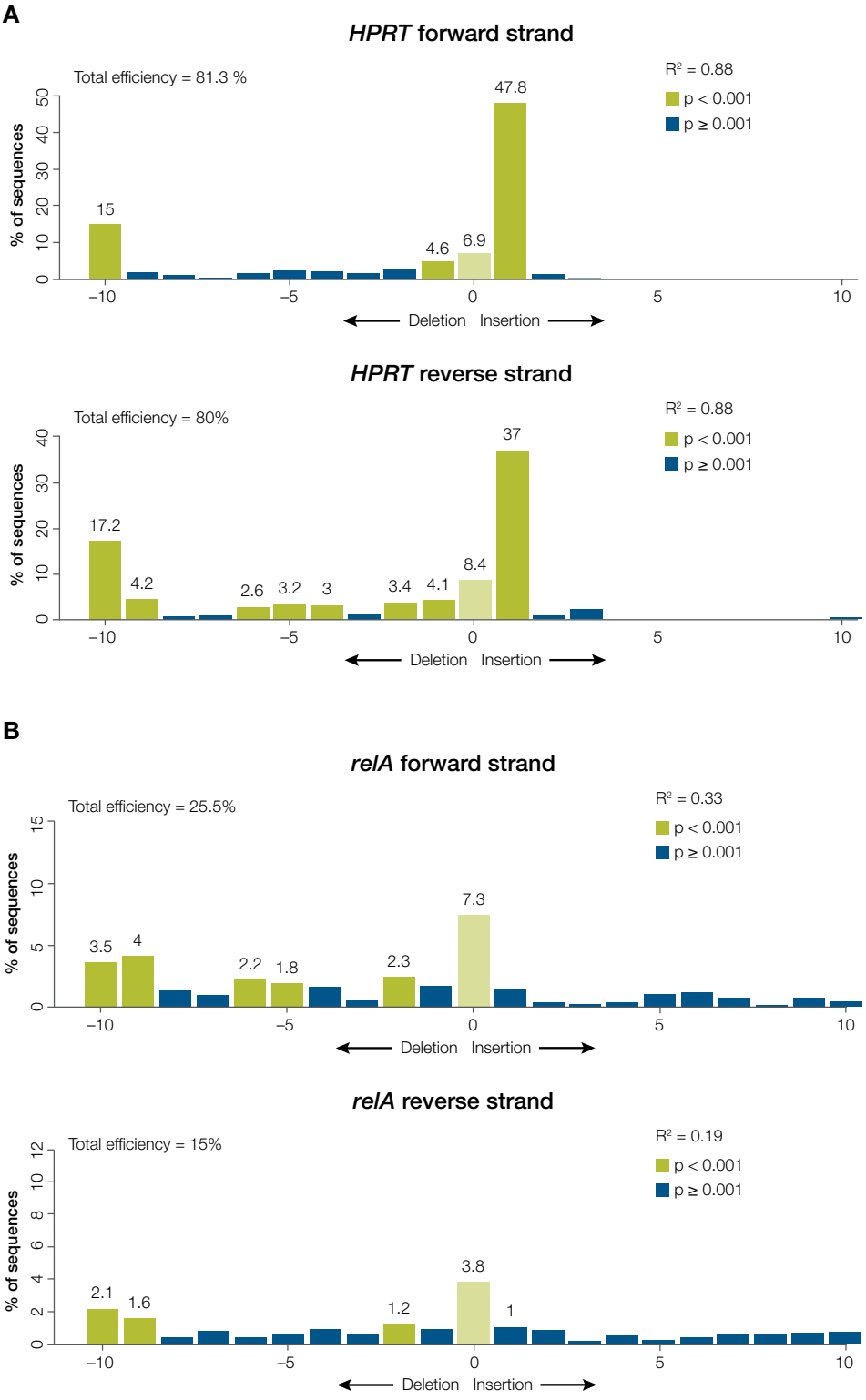


Figure 7. Analysis of two different genome editing events at the *HPRT* and *relA* loci using TIDE software and mixed population sequencing traces generated by the SeqStudio instrument. The bars show the proportion of the population having the indicated number of nucleotides deleted or inserted. For **(A)** *HPRT,* the overall efficiency of the edit was around 80%, whereas the overall efficiency at the **(B)** *relA* locus was around 20%.

Human cell line authentication

The study of development of human diseases relies heavily on the analysis of dissociated human cell lines grown in culture. However, an increasingly acknowledged problem is that cells grown *in vitro* can be misidentified or contaminated with other unrelated cell lines. The identity of cell lines can be verified by analysis of a highly specific genetic "fingerprint" of highly variable short tandem repeats (STRs). The SeqStudio platform integrates well with the Thermo Fisher Scientific cell line authentication solution. The Applied Biosystems™ Identifiler™ Plus and Identifiler™ Direct kits can be used on purified and crude DNA preparations, respectively, for analyzing 16 highly variable human STR loci commonly used for verifying cell line authenticity. The Applied Biosystems™ GeneMapper™ Software, used for analyzing alleles identified by Identifiler kits, is compatible with data produced by the SeqStudio instrument, and the results can be used to query ATCC or other STR databases to verify authenticity [6].

To demonstrate the utility of the SeqStudio instrument in a cell line authentication workflow, allelic information on STRs was obtained from five different, commonly used human cell lines. The identity of the cell lines was confirmed even with as little as 300 pg of gDNA. To show the ability to detect contaminating cells, a population of M4A4GFP cells was spiked with varying amounts of HeLa cells and analyzed using the Identifiler Direct kit. HeLa cell–specific alleles could be detected even if only 10% of the population had HeLa cells (Figure 8). Therefore, when coupled with the Identifiler kits, the SeqStudio instrument can be a central component for a cell line authentication solution.

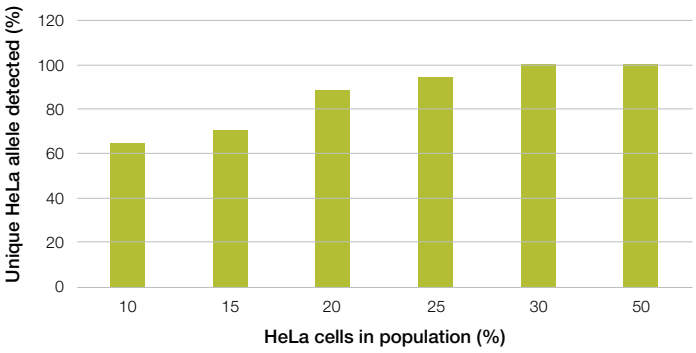


Figure 8. Analysis of cell line contamination on the SeqStudio instrument. HeLa cells and M4A4GFP cell suspensions were diluted to 5 x 10^5 cells/mL, mixed in the indicated proportions, and spotted onto NUCLEIC-CARD™ Sample Collection Device. Contaminating HeLa cells can be detected with high confidence on the SeqStudio instrument if they make up approximately 20% of a population; however, some alleles unique to HeLa can be detected if they make up as little as 10% of a population.

SNaPshot Multipex System

The ability to detect single-nucleotide polymorphisms (SNPs) plays a critical role in understanding how the genome influences biological phenotypes. To analyze SNP variants, the Applied Biosystems™ SNaPshot™ Multiplex System was developed [7]. Customizable, color-coded fragments of differing sizes, corresponding to specific alleles, are analyzed by fragment analysis. The SeqStudio system includes new features that facilitate SNaPshot analysis, including built-in reporting of fragment analysis results of size and peak area. Additionally, the ability to mix fragment analysis and sequencing reactions on one plate enables investigators to perform SNP profiling and Sanger sequencing on a single run.

To illustrate the functional utility of the SeqStudio instrument in SNaPshot workflows, genomic DNA from FFPE-preserved tumor slices was collected and analyzed using probes targeting *KRAS* G12X and G13X alleles using the SNaPshot multiplex reagent kit. The SeqStudio instrument produced results that clearly showed the presence and accurate calls of the different alleles at this position (Figure 9). Note that although the detection of the alleles was accurate on SeqStudio instrument, the absolute migration of all peaks will differ slightly when compared to that in other platforms due to the different chemical nature of the different polymers. Therefore, to associate a peak with an allele without an ambiguity, a calibration with known alleles should be performed before undertaking a large-scale analysis.

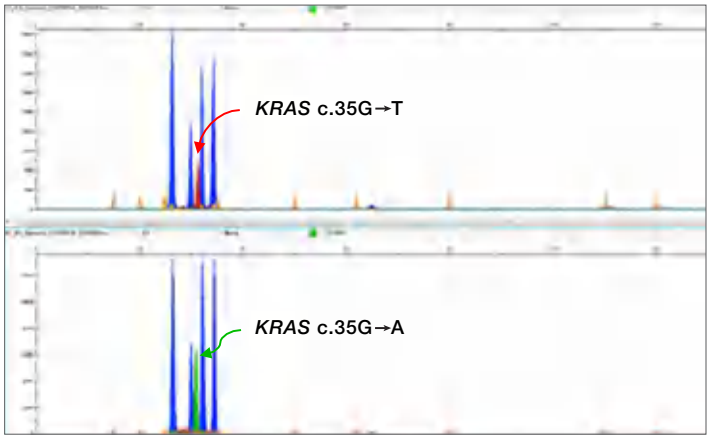


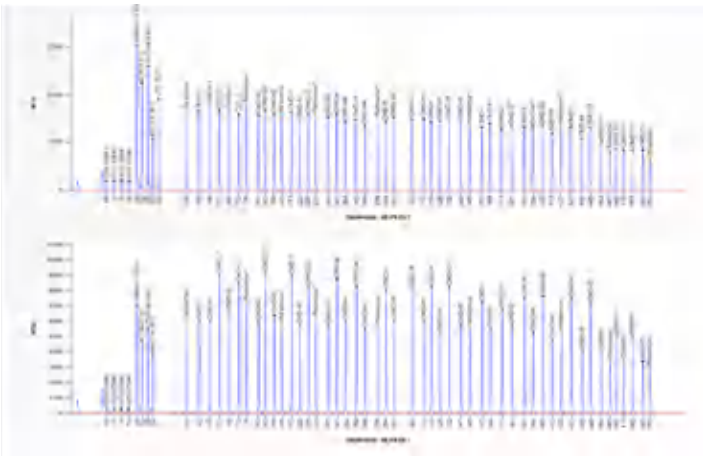
Figure 9. Analysis of SNPs using kit on the SeqStudio instrument. One nanogram of FFPE-extracted DNA from two different tumors was processed using the SNaPshot multiplex kit and *KRAS*-specific primers, followed by fragment analysis on the SeqStudio system. The SNaPshot multiplex kit produces fragments with allele-specific colors and lengths. The four blue peaks represent wild type alleles at *KRAS* c.34, c.35, c.37, and c.38. The red peak (top graph) indicates that the sample also had a *KRAS* c.35G➔T mutation present, whereas the green peak (bottom graph) results from a different allele at the same position in a different sample.

Multiplex ligation–dependent probe amplification (MLPA)

One widely used method for studying inherited human diseases arising from variations in copy number of a locus is multiplex ligation–dependent probe amplification [8]. This method, developed and commercialized by MRC Holland, can analyze up to 50 multiplexed pairs of adjacently located probes hybridizing to the loci of interest. The high dynamic range, sizing precision, and peak-height fidelity necessary for analyzing MLPA probe amplicons make the SeqStudio system an ideal platform for performing MLPA analyses. Results obtained on the SeqStudio instrument are compatible with MRC Holland's Coffalyzer.Net software for analyzing MLPA data.

MLPA on the SeqStudio instrument was used to analyze a DNA sample from a probe that is known to carry a duplication of exons 2–30 in the Duchenne muscular dystrophy (*DMD*) gene and a normal sample using the P034 DMD assay set from MRC Holland. The peak heights and relative sizes of these samples can readily be translated into an accurate detection of the region containing the duplication (Figure 10). Similar results were obtained using probes for large and small deletions. Therefore, the SeqStudio instrument can be an integrated tool for MLPA investigations of regions containing CNVs.

A



B

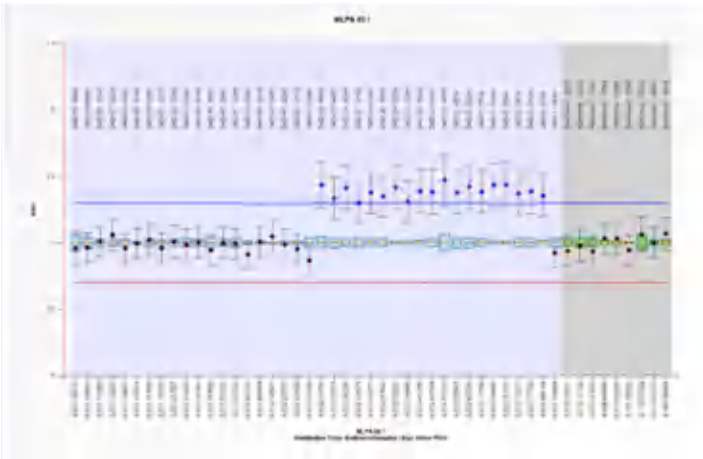


Figure 10. Analysis of CNV on the SeqStudio instrument using MLPA. **(A)** Fragment profile of a sample analyzed for CNVs at the *DMD* gene. Note that peak heights of some of the fragments are larger on the bottom sample versus the top, indicating an overrepresentation of DNA in that region. **(B)** The corresponding ratio chart, aligning the probes on the DMD locus, and clearly displaying increased ratios for *DMD* probes in exons 2–30. Probes that cross the blue threshold are indicative of a gain in copy number from 2 to 3.

Summary

Capillary electrophoresis remains a powerful method for characterizing genetic information. Sanger sequencing and fragment analysis remain the gold standard methods for sequencing and DNA analysis. In this application guide, we have described how the newly introduced SeqStudio Genetic Analyzer produces data completely compatible with the most common CE applications. We also described how the new features of the SeqStudio Genetic Analyzer make running CE experiments easier with minimal hands-on time, facilitate collaboration through Thermo Fisher Cloud–based sharing and applications, and introduce new opportunities to run sequencing and fragment analysis samples at one time. Together, the redesign and enhancements provided by the SeqStudio Genetic Analyzer enable exciting new possibilities for utilizing the gold-standard methods of capillary electrophoresis in an integrated ecosystem.

References

1. Low-level somatic variant detection in tumor FFPE samples by Sanger sequencing COL31176 0616
2. Technical note: MicroSEQ Rapid Microbial Identification System
3. Borisenko AV et al. (2009) The front-end logistics of DNA barcoding: challenges and prospects. *Mol Ecol Resour* s1:27-34.
4. Brinkman EK et al. (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* 42:e168.
5. Application note: Using Sanger sequencing to facilitate CRISPR- and TALEN-mediated genome editing workflows
6. https://www.atcc.org/STR_Database.aspx
7. Product bulletin: SNaPshot Multiplex System for SNP genotyping
8. Schouten JP et al. (2002) Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res* 30:e57.

applied biosystems

Ordering information

Product	Quantity	Cat. No.
BigDye Terminator v1.1 Cycle Sequencing Kit	100 reactions	4337450
BigDye Terminator v3.1 Cycle Sequencing Kit	100 reactions	4337455
BigDye Direct Cycle Sequencing Kit	100 reactions	4458687
BigDye Xterminator Purification Kit	100 preps	4376486
ExoSAP-IT PCR Product Cleanup Reagent	100 reactions	78200.200.UL
MicroSEQ 500 16S rDNA PCR Kit	1 kit	4348228
MicroSEQ 500 16S rDNA Sequencing Kit	1 kit	4346480
Identifiler Plus PCR Amplification Kit	200 reactions	4427368
Identifiler Direct PCR Amplification Kit	200 reactions	4467831
SNaPshot Multiplex kit	200 reactions	4323151
SeqStudio Genetic Analyzer		A34274
SeqStudio Analysis Software		4443764
SeqStudio Cartridge Assay		A33671
SeqStudio Starter Kit		A35000
SeqStudio Full-Day SmartStart Training		A34684

ThermoFisher SCIENTIFIC