Development of an Interlaboratory-Verified Sequencing Workflow for KRAS Variant Identification

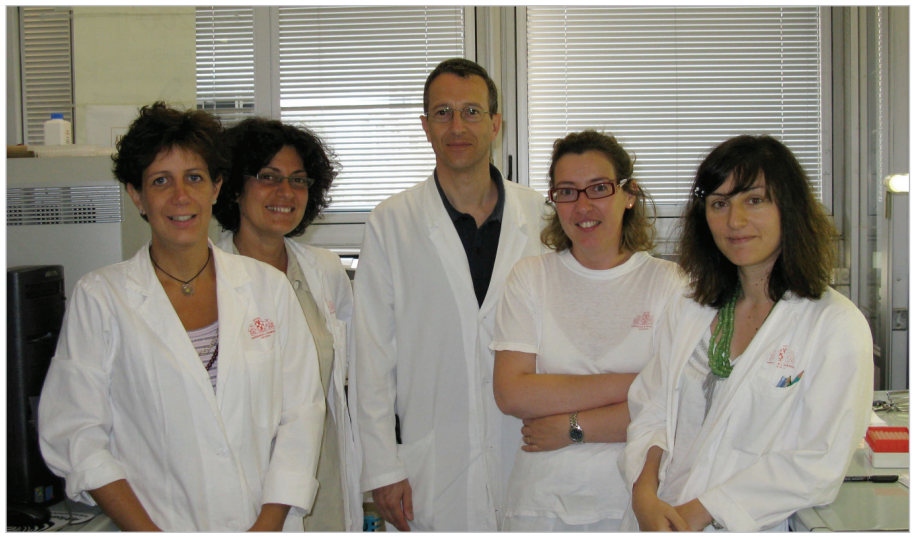
Abstract

KRAS (V-Ki-ras2 Kirsten rat sarcoma viral oncogene) encodes one of the proteins in the epidermal growth factor receptor (EGFR) signaling pathway [1]. This signaling pathway is important in the development and progression of several aggressive cancers. KRAS mutations have been described in codons 12, 13 (exon 2), and, in rare case, at codon 61 (exon 3). These mutations result in EGFR signaling pathway activation in the absence of EGFR itself, making anti-EGFR therapeutic agents ineffective. Ongoing KRAS mutation research has identified the need for a simple, accurate mutation detection protocol. In this publication we describe a collaborative, research-use-only study involving two independent laboratories and Applied Biosystems in the development of a KRAS gene sequencing protocol.

Introduction

Mutations acquired during an individual's lifetime that do not affect germ line cells are called somatic mutations. Certain somatic mutations in the KRAS gene lead to a form of the protein that can direct cells to grow and divide without control, implicating KRAS in the development of several types of cancer. Research suggests that KRAS gene mutations are common in pancreatic, lung, colorectal, and other types of cancer [1,2]. Moreover, KRAS mutants are capable of activating the EGFR signaling pathway independently of EGFR activation, which makes the investigation of KRAS mutations in discovery efforts to find potential anti-EGFR drug candidates. In addition, recent

The Molecular Biology laboratory at Santa Maria della Misericordia Hospital in Perugia, Italy, is directed by Prof. Lucio Crino (not pictured) and coordinated by (left to right) Dr. Vienna Ludovini and Dr. Lorenza Pistola with the collaboration of Dr. Francesca Romana Tofanetti, Dr. Annamaria Siggillino, and Dr. Antonella Flacco (not pictured). It is a highly specialized research laboratory investigating molecular techniques for the study of neoplastic disease.



The Laboratory of Cellular Therapies at the Internal Medicine Department of University of Genova is directed by Prof. Franco Patrone and coordinated by Prof. Alberto Ballestrero. In this picture are some of his collaborators. Left to right, Dr. Anna Garuti, Dr. Claudia Palermo, Prof. Alberto Ballestrero, Dr. Ilaria Rocco, and Dr. Gabriella Cirmena. The research interests of this group include translational genomics and the molecular mechanisms of potential anti-cancer therapeutics.

studies have reported that mutations in the BRAF gene affect the EGFR signaling pathway downstream and may also affect the action of drugs aimed at this pathway [3,4]. KRAS mutations occur most commonly in codon 12 or 13 (exon 2) and, in more rare cases, in codon 61 (exon 3) (Table 1).

The collaborators for this research study included scientists from the Medical Oncology Laboratory at Santa Maria della Misericordia Hospital (Perugia, Italy), the Laboratory of Cellular Therapy at the Internal Medicine Department (DIMI) of the University of Genova (Genova, Italy), and Applied Biosystems. The scientists at the laboratories in Perugia and Genova chose automated capillary electrophoresis (CE) DNA sequencing for this work because DNA sequencing is widely acknowledged as the industry standard for direct detection of sequence variants. It is a highly referenced and robust technique that delivers long read lengths (800–1,000 bp) and provides scalable throughput from a few to several hundred samples in a single day. The protocols involved are generally straightforward and cost effective. CE sequencing is an attractive option for researching mutation detection because it is the only method that enables researchers to identify unknown mutations, in addition to confirming known or expected variants. Since CE analysis can produce accurate sequence data on longer amplicons, investigators are able to uncover novel SNPs (that may flank the known SNP) and novel insertions or deletions. The platform is easily automated and also allows flexibility for future studies, as the same instrument can be used for both sequencing and for fragment analysis studies (such as microsatellite analysis). When experiments are designed that generate amplicons using M13 tailed primers, a single automated sequencing protocol can be implemented for all amplicons in the set.

Here we present a KRAS exon 2 and exon 3 sequencing protocol. This protocol was developed via a collaborative research study on 50 colorectal cancer samples at two separate laboratories, both using the Applied Biosystems® 3130 Genetic Analyzer. The same sequencing analysis was also performed at Applied Biosystems (Foster City, CA, USA) using the newest capillary electrophoresis platform, the Applied Biosystems® 3500xL Genetic Analyzer.*

This application note is aimed at:

- Defining the sequencing protocol specifically developed for the analysis of these KRAS exons
- Reporting the data that demonstrate concordance between the results obtained using the 3130 and the 3500 systems
- Introducing the benefits of the 3500 Series Genetic Analyzers when used in conjunction with the BigDye® Terminator Cycle Sequencing Kit and the BigDye XTerminator® Purification Kit, and SeqScape® Software, which ensure an efficient, easy-to-use workflow from sequencing to clean-up and data analysis—critical to investigators in clinical research environments

Table 1. Locations of KRAS Mutations Found in Samples Obtained From a Colorectal Cancer Study.

KRAS mutation	Number of mutations found in the study (50 colorectal cancer cases)
Gly12Asp (GGT>GAT)	4
Gly12Val (GGT>GTT)	3
Gly12Cys (GGT>TGT)	4
Gly12Ser (GGT>AGT)	6
Gly12Ala (GGT>GCT)	4
Gly12Arg (GGT>CGT)	None detected
Gly13Asp (GGC>GAC)	7
Gln61Leu (CAA>CTA)	None detected

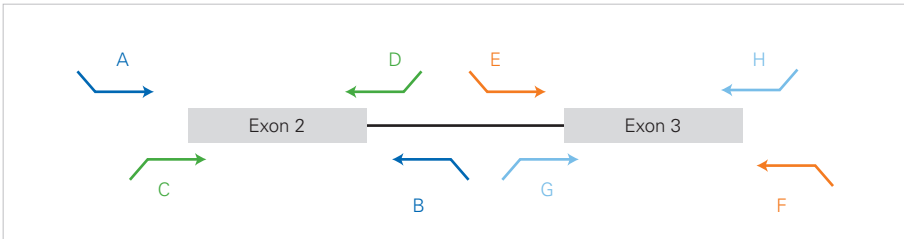


Figure 1. Schematic Diagram Illustrating the Primer Design Strategy for the Interrogation of the KRAS Target Exons.

Materials and Methods

Paraffin-embedded samples from 50 colorectal cancer cases were collected by the medical oncology laboratories in both Perugia and Genova, Italy, and analyzed in the KRAS mutation study anonymously. In order to perform this research, informed consent was obtained in accordance with the Institutional Ethics Committee from Santa Maria della Misericordia Hospital and the Helsinki declaration. Genomic DNA was isolated from five 20 µm paraffin-embedded slices. Researchers from both laboratories extracted gDNA from all 50 samples using either the QIAmp DNA extraction kit (QIAGEN, GmbH) or the RecoverAll™ kit (Ambion, USA).

Two PCR primer pairs were designed for the amplification of exon 2 (primer pair A/B and primer pair C/D), and two pairs were designed for the amplification of exon 3 (primer pair E/F and primer pair

*The 3130 and 3500 Series Genetic Analyzers are FOR RESEARCH USE ONLY. NOT FOR USE IN DIAGNOSTIC PROCEDURES.

G/H). Primers A and B (exon 2) and E and F (exon 3) were designed to anneal outside the exon region (Figure 1). Primer pair C/D and primer pair G/H were designed to amplify an amplicon shorter than 200 bp. This strategy was employed to ensure that good amplification efficiency could still be achieved even when using degraded DNA. The use of pairs of independent primers in this study for the optimization of primer design provided a replicate data set. Following this verification, the teams at the collaborating laboratories made the decision to use the primer pairs that generated smaller amplicons (primer pairs C/D and G/H) for subsequent KRAS mutation research, since those gave more efficient and robust amplification.

PCR primers, reactions, and cycling conditions

The two collaborating laboratories both performed the following cycling and clean-up protocol on the 50 DNA extracted samples to independently verify the results. The amplifications were performed using 80 ng of the extracted DNA and standard reaction conditions with AmpliTaq Gold® Master Mix on an Applied Biosystems Veriti® Thermal Cycler (conditions shown below). The primer sequences used to produce both KRAS amplicons from the 50 samples are shown in Table 2. In addition, primers for another EGFR pathway gene, BRAF, were synthesized (Table 3) and used to amplify sections of exons 11 and 15. The collaborators were able to use the KRAS sequencing workflow and protocol without modification to successfully amplify the BRAF gene amplicons from exons 11 and 15.

Sequencing primers, reactions, and cycling conditions

Each amplicon was sequenced in both forward and reverse directions using M13 primers.

Electrophoresis

Purified sequencing reactions were analyzed independently in the two collaborating sites using the Applied

Table 1. Sequences of the Primers Used in the KRAS Mutation Study.

EXON 2	Primer A	5' tgtaaaacgacggccagtTATTTGATAGTGTATTAACCTTATGTGTG 3'
	Primer B	5' caggaaacagctatgaccGAAACCTTTATCTGTATCAAAGAATG 3'
	Primer C	5' tgtaaaacgacggccagt TGTGACATGTTCTAATATAGTCACATT 3'
	Primer D	5' caggaaacagctatgacc ACCAGTAATATGCATATTAAAACAAGA 3'
EXON 3	Primer E	5' tgtaaaacgacggccagtAGGTGCACTGTAATAATCCAGA 3'
	Primer F	5' caggaaacagctatgaccCTATAATTACTCCTTAATGTCAGCTTATT 3'
	Primer G	5' tgtaaaacgacggccagt GACTGTGTTTCTCCCTTCTCA 3'
	Primer H	5' caggaaacagctatgacc AGCTTATTATATTCAATTTAAACCCAC 3'

Each primer includes the M13 tail in order to facilitate subsequent sequencing using universal M13 primers.

Table 3. Sequences of the Primers Used in the BRAF Mutation Study.

Primer Exon 11 forward	5' tgtaaaacgacggccagtGCATAAGGTAATGTACTTAGGGTG 3'
Primer Exon 11 reverse	5' caggaaacagctatgaccCCTATTATGACTTGTCACAATGTC 3'
Primer Exon 15 forward	5' tgtaaaacgacggccagtCTAAACTCTTCATAATGCTTGCTC 3'
Primer Exon 15 reverse	5' caggaaacagctatgaccTCTAGTAACTCAGCAGCATCTCA 3'

Each primer includes the M13 tail to facilitate subsequent sequencing using universal M13 primers. Sequences of the M13 primers are: M13 forward **5'** TGTAAAACGACGGCCAGT **3'**, M13 reverse **5'** CAGGAAACAGCTATGACC **3'**.

Amplification Reactions

AmpliTaq Gold® PCR Master Mix	25 µL
Reverse primer (5 pM)	5 µL
Forward primer (5 pM)	5 µL
Purified genomic DNA	(80 ng, volume varies)
H_2O	to 50 µL final volume

Amplification Cycling Protocol

Initial denaturation		96°C for 5 min
35 cycles	Denaturation	94°C for 30 sec
	Annealing	58°C for 45 sec
	Extension	72°C for 45 sec
Final extension		72°C for 10 min

The amplified products were treated with ExoSAP-IT® enzyme (USB Corporation) for PCR primer and dNTP removal, following the manufacturer's protocol.

Cycle Sequencing Reactions

BigDye® Terminator Cycle Sequencing Kit v1.1 Master Mix	4 µL
M13 primer (0.8 pmol/µL)	4 µL
5X buffer	2 µL
Purified PCR product	(10 ng, volume varies)
H_2O	to 20 µL final volume

Cycle Sequencing Protocol

Initial denaturation		96°C for 1 min
25 cycles	Denaturation	96°C for 10 sec
	Annealing	50°C for 5 sec
	Extension	60°C for 2 min

Sequencing reactions were purified using the Applied Biosystems BigDye XTerminator® Purification Kit according to the manufacturer's protocol.

Biosystems® 3130 Genetic Analyzer, and at Applied Biosystems using the Applied Biosystems® 3500 Genetic Analyzer (see sidebar). The sequence results for each of the 50 samples were analyzed independently at the three laboratories (Applied Biosystems and the two collaborating laboratories) using SeqScape® Software v2.7 to verify the sequencing results and to identify the putative mutations for each sample (Figures 2 and 3). The collaborators found SeqScape® software to be invaluable in the automatic identification of mutations, which then are confirmed by the researcher. Previously, the electropherograms had to be read manually for the presence of mutations, which took extensive time.

Quality Control

The 3500 Series Data Collection Software allows for real-time data quality evaluation of sequencing results. The Monitor Run functionality allows scientists to review the progress of a run in real time, and provides access to key quality control parameters (Figure 4). Scientists can immediately assess the quality of data as they are generated.

3500 Data Collection Software includes primary analysis of sequence data to allow direct viewing of base-called files for sequencing experiments (size-called files for fragment analysis runs), and combines powerful quality control parameters predefined by the user, for direct analysis of data generated by the platform. This allows scientists to assess and make decisions about the quality of data as they are produced on the instrument. By providing immediate access to base-called or size-called data, the software allows scientists to make decisions about the quality of data as they are generated, without first transferring output files to secondary analysis software packages.

After completion of the sequencing run, secondary software packages such as Applied Biosystems SeqScape® Software provide additional features to review and view the sequencing data produced.

Figure 2. Mutation Analysis Using SeqScape® Software. Through automated sequence analysis, superior basecalling, and sophisticated data visualization tools, SeqScape® Software v2.7 produces definitive mutation analysis, as exemplified in this study of KRAS mutation 35G>A(Gly12Asp). The region showing the sequence variation is boxed in red.

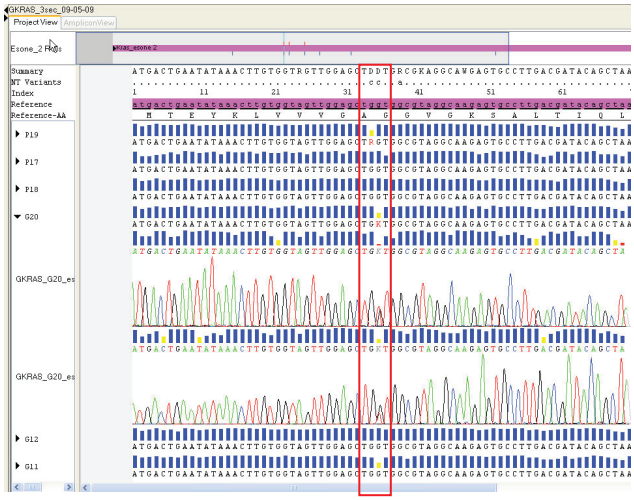


Figure 3. SeqScape® Software Readout of the KRAS Mutation 35G>T(Gly12Val). The region showing the sequence variation is boxed in red.

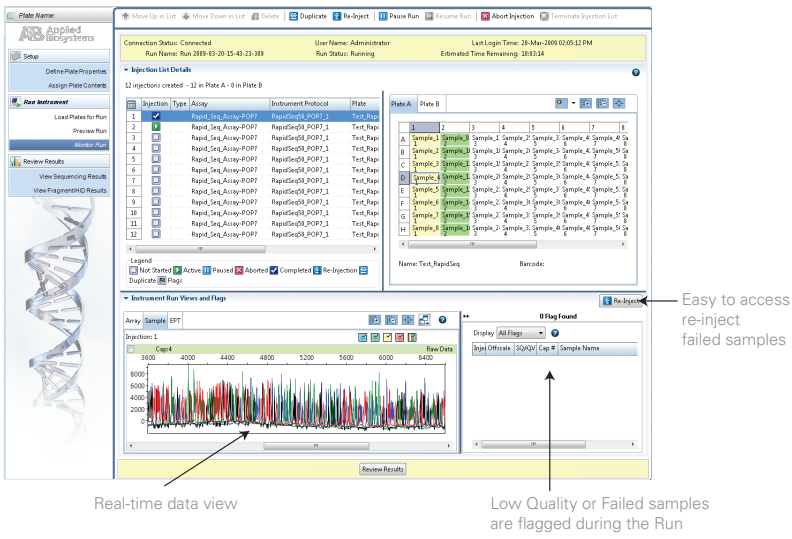


Figure 4. The 3500 Series Genetic Analyzer Data Collection Software. The software offers quality control features critical for accurate data analysis.

SeqScape® Software detects all variations between the reference sequence and the subject sample sequence, including deletions, insertions, mismatches, heterozygous bases, and heterozygous insertion/deletions. The variations between the consensus and the reference sequence are reported as mutations in the mutations report (Figure 5). For researchers working in process-controlled environments, SeqScape® Software offers a full suite of features to provide data security, an audit trail of data changes, and electronic signatures. These features are designed to assist customers with 21 CFR Part 11 requirements.

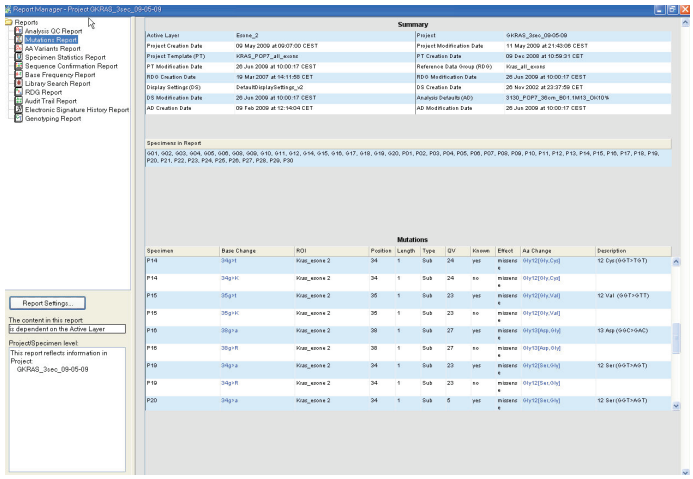


Figure 5. Mutation Report View From SeqScape® Software. At a glance, the user can see the name of the specimen, the base change, and the amino acid change. Clicking on the base change in the Mutation Report View takes the user directly to electropherogram associated with this result.

Results

The medical oncology research laboratories in both Perugia and Genova, Italy, determined that the optimized sequencing protocol presented in this application note identified the same mutations in all 50 colorectal cancer samples, regardless of which amplicon size was queried. Moreover, the mutations observed by our collaborators were identical to the mutations identified by Applied Biosystems scientists when the samples were run on the 3500xL Genetic Analyzer.

Conclusion

In the current research study designed to detect mutations in the KRAS gene, the Applied Biosystems 3500 Genetic Analyzer with an optimized workflow delivers reliable and reproducible results.

The same workflow was used successfully by collaborators, without modification, to sequence the BRAF gene (exon 11 and exon 15). Our collaborators at the laboratories in Perugia and Genova report the following advantages when using this method for KRAS and BRAF mutation studies:

- The method is based on industry-standard Sanger sequencing
- CE sequencing allows the sequencing of longer amplicons (>500 bp), which gives confirmation of predicted mutations as well as permits the analysis of the regions flanking the suspected SNP; thus, unanticipated or new variants, including SNPs or sequence insertions or deletions can be easily identified to provide complete and accurate variant detection
- The platform is flexible and scalable, allowing more applications on the same instrument (for example, fragment analysis studies)
- When experiments are designed that generate amplicons using tailed primers, a single automated sequencing protocol can be implemented for all amplicons in the set, allowing researchers to get their results on the same day

The 3500 Genetic Analyzer offers the following features, which are beneficial to basic research labs and also to teams working in process-controlled environments:

- Accurate and reliable data quality
- Easy instrument setup
- Ready-to-use consumables that are easy to load and run
- RFID (Radio Frequency Identification) tracking to assist users with consumables information required during experimental design and setup and to help track consumables usage
- Simplified user interface with easy display of consumables and array usage information, quick-start functionality, and system maintenance reminders
- Performance-check functionality that allows researchers to self-check instrument performance periodically, as required by their operating procedures
- Faster run times compared to previous generations of capillary electrophoresis platforms
- Security, Audit Trail, and Electronic Signature features to help process-controlled labs comply with 21CFR Part 11 requirements

References

1. van Krieken JH, Jung A, Kirchner T et al. (2008) KRAS mutation testing for predicting response to anti-EGFR therapy for colorectal carcinoma: proposal for an European quality assurance program. *Virchows Arch* 453:417–443.
2. Karnoub AE, Weinberg RA (2008) Ras oncogenes: split personalities. *Nat Rev Mol Cell Biol* 9:517–531.
3. Brose MS, Volpe P, Feldman M et al. (2002) BRAF and RAS mutations in human lung cancer and melanoma. *Cancer Res* 62(23):6997–7000.
4. Di Nicolantonio F, Martini M, Molinari F et al. (2008) Wild-type BRAF is required for response to panitumumab or cetuximab in metastatic colorectal cancer. *J Clin Oncol* 26(35):5705–5712.

The 3500 Series Genetic Analyzers

The 3500 Series Genetic Analyzers are automated fluorescence-based capillary electrophoresis systems. Samples are sequenced in less than 40 minutes on the 24-capillary 3500xL system or on the 8-capillary 3500 system, using 3500 POP-7™ Polymer and the 50 cm capillary array. For convenience, the 3500 Series instruments also include a sequencing module that incorporates the BigDye XTerminator® Purification Kit workflow. After the sequence data are collected, the KB™ Basecaller software embedded in the 3500 Data Collection Software automatically processes the sequenced samples and provides read length greater than 850 bp, with average base Quality Values greater than 20 (QV20). Furthermore, the 3500xL system, using the RapidSeq50_ POP7 run module, can efficiently sequence up to 840 samples in a 23-hour period, generating high-quality, high-resolution data with minimal hands-on time.

Fully automated from the moment you place an 8-tube strip, 96-well, or 384-well sample plate on the instrument and start the run, the instrument provides continuous, unattended operation for every phase of the process, including polymer loading, sample injection, separation, detection, and primary data analysis. The 3500 Series Systems feature simplified, easy-to-install consumables. The Anode and Cathode Buffer Containers are supplied with ready-to-use 1X Genetic Analysis Buffer formulations.

Easy-to-use wizards for instrument operation and maintenance ensure predictable, hassle-free performance. The majority of applications can be analyzed on a single configuration of POP-7™ polymer with a 50 cm capillary array. Integrated primary analysis and QC software includes improved Data Collection software, with an intuitive workflow from plate setup to primary analysis that performs base calls and applies quality control flags to alert the user to failed or low-quality samples.

ORDERING INFORMATION

Description	Quantity	P/N
DNA Isolation From Paraffin-Embedded Samples		
RecoverAll™ Total Nucleic Acid Isolation Kit	1 kit	AM1975
PCR Amplification		
AmpliTaq Gold® PCR Master Mix	200 rxn	4318739
Sequence Detection Primer (medium scale)	1 tube	4304971
Sequencing Reactions and Purification		
BigDye® Terminator Cycle Sequencing Kit	100 rxn	4337450
BigDye XTerminator® Purification Kit	20 mL	4376487

Genetic Analyzers

3500 Series System Packages		3500	3500xL
Package Name	**Description**	**P/N**	**P/N**
3500 Series Genetic Analyzer for Resequencing & Fragment Analysis	3500 Series System with Data Collection Software, Sequencing Analysis, Variant Reporter®, and GeneMapper® Software packages. System package also includes DNA Sequencing and Fragment Analysis reagent kits for system qualification.	4440462	4440463
3500 Series Genetic Analyzer for Resequencing & Fragment Analysis With SAE	3500 Series System with Data Collection Software (includes additional functionality for Security, Audit Trail, and Electronic Signature (SAE) capabilities), Sequencing Analysis, Variant Reporter®, and GeneMapper® Software packages. System package also includes DNA Sequencing and Fragment Analysis reagent kits for system qualification.	4440464	4440465
3500 Series Genetic Analyzer for Resequencing	3500 Series System with Data Collection Software, Sequencing Analysis, and Variant Reporter® Software packages. System package also includes DNA Sequencing reagent kits for system qualification.	4440466	4440467
3500 Series Genetic Analyzer for Fragment Analysis	3500 Series System with Data Collection Software and GeneMapper® Software. System package also includes DNA Fragment Analysis reagent kits for system qualification.	4440468	4440469
3500 Series Genetic Analyzer for Sequence Typing & Fragment Analysis	3500 Series System with Data Collection Software, Sequencing Analysis, SeqScape®, and GeneMapper® Software packages. System package also includes DNA Sequencing and Fragment Analysis reagent kits for system qualification.	4440470	4440471

System Consumables and Reagents

Description	P/N
3500xL Capillary Array (50 cm)	4404689
3500 Capillary Array (50 cm)	4404685
3500 POP-7™ Polymer (960 samples)	4393714
3500 POP-7™ Polymer (384 samples)	4393708
Anode Buffer Container (ABC) 3500 Series	4393927
Cathode Buffer Container (CBC) 3500 Series	4408256
Septa Cathode Buffer Container 3500 Series	4410715
Conditioning Reagent 3500 Series	4393718
Hi-Di™ Formamide (5 mL) 1 bottle	4401457
BigDye® Terminator v3.1 Cycle Sequencing Kit (1,000 rxns)	4337456
BigDye XTerminator® Purification Kit (1,000 rxns)	4376487

Visit www.appliedbiosystems.com/3500Series or contact your local Sales Representative for more information.

FOR RESEARCH USE ONLY. NOT FOR USE IN DIAGNOSTIC PROCEDURES.

For those who require IVD-marked devices, the 3500 Dx and 3500xL Dx Genetic Analyzers and system accessories meet the requirements of the In Vitro Diagnostic Medical Devices Directive (98/79/EC). The 3500 Dx and 3500xL Dx systems are for distribution and use in specific European countries only. For more information about the 3500 Dx Series Systems, contact your Applied Biosystems representative.

The purchase price of this Instrument includes a grant of a limited, non-transferable license under U.S. patents and method claims of its foreign counterparts, and under U.S. patents and element claims of its foreign counterparts, to use this particular instrument for electrophoresis methods employing fluorescence as a means of detection. No other licenses or rights are hereby conveyed either expressly, by implication, or estoppel including, but not limited to, any claims to a composition. This instrument incorporates technology subject to one or more patents licensed from Hitachi, Ltd. as well as patents and patented technology owned by or under control of Applied Biosystems.

This instrument is Authorized for use in DNA sequencing and fragments analysis only. This Authorization is included in the purchase price of the instrument and corresponds to the up-front fee component of a license under process claims of U.S. patents and under all process claims for DNA sequence and fragment analysis of U.S. patents now or hereafter owned or licensable by Applied Biosystems for which an Authorization is required, and under corresponding process claims in foreign counterparts of the foregoing for which an Authorization is required. The running royalty component of licenses may be purchased from Applied Biosystems or obtained by using Authorized reagents purchased from Authorized suppliers in accordance with the label rights accompanying such reagents. Purchase of this instrument does not itself convey to the purchaser a complete license or right to perform the above processes. This instrument is also licensed under U.S. patents and apparatus and system claims in foreign counterparts thereof. No rights are granted expressly, by implication, or by estoppel under composition claims or under other process or system claims owned licensable by Applied Biosystems. For more information regarding licenses, please contact the Director of Outlicensing at Applied Biosystems, 850 Lincoln Centre Drive, Foster City, California 94404, USA.

Printed in the USA, 8/2010 Publication 106AP30-01

applied biosystems™

Headquarters
850 Lincoln Centre Drive | Foster City, CA 94404 USA
Phone 650.638.5800 | Toll Free 800.327.3002
www.appliedbiosystems.com

International Sales
For our office locations please call the division headquarters or refer to our website at
www.appliedbiosystems.com/about/offices.cfm