

Development of a resequencing workflow for variant analysis in the *APC* gene

Applied Biosystems® 3730xl and 3500xL Series Genetic Analyzers

The Molecular Diagnostic Unit of the Hereditary Cancer Program at the Catalan Institute of Oncology performs research to better understand hereditary cancers. The director of the Program is Dr. Gabriel Capellá, and the head of the Molecular Diagnostic Unit is Dr. Conxi Lázaro. The laboratory participates in different EMQN (European Molecular Genetics Quality Network) quality control schemes. Currently it participates in the Familial Breast Cancer, Hereditary Non-Polyposis Colon Cancer, and Familial Adenomatous Polyposis external quality assessment (EQA) tests. The Molecular Diagnostic Unit is involved in various research projects. As a part of these research projects, the Capellá laboratory is currently exploring the use of next-generation sequencing platforms, with several ongoing research projects aimed at investigating hereditary cancer syndromes. The familial adenomatous polyposis testing team is led by Sara Gonzalez (third from left), with more than 10 years of experience in the field.



Introduction

The *APC* (adenomatous polyposis coli) tumor suppressor gene is located at 5q21-q22. The most abundant transcript is 8,532 bp long (divided into 15 exons), with exon 15 accounting for 77% of the coding region. *APC* encodes a multifunctional protein with distinct domains [2], that regulates several cellular processes, including transcription, cell cycle control, migration, differentiation, and apoptosis. The APC protein plays an integral role in the Wnt signaling pathway, as it binds and down-regulates β -catenin. More than 300 different mutation of the *APC*

gene are known [3]; in more than 80% of individuals with typical Familial adenomatous polyposis (FAP), a mutation can be identified in the *APC* gene that constitutively activates the Wnt pathway. Pathogenic mutations are scattered throughout the whole coding region. The frequency of *APC* gene mutations is much lower in individuals with attenuated FAP (AFAP) (~25%). There is great interest in correlating individual *APC* mutations with observed phenotypes, and this research goal has highlighted the need for a simple and robust method to identify *APC* gene mutations in research samples.

Automated Sanger sequencing of DNA is widely acknowledged as the gold standard for mutation detection and/or characterization. It is a highly referenced and robust technique that delivers long read lengths and the ability to sequence anywhere from a few to several hundred samples in a single day. Applied Biosystems® Genetic Analyzers are widely used in life science research laboratories for DNA sequence analysis as well as for numerous DNA fragment analysis assays. Here we highlight the workflow developed in a preclinical

research setting at the Institut Català d'Oncologia (ICO), Barcelona. This publication is aimed at:

- Detailing a customer's approach to developing a resequencing protocol for the analysis of the *APC* gene in a research setting
- Demonstrating data concordance between the Applied Biosystems® 3730xl and the 3500xL Genetic Analyzers
- Highlighting key benefits of capillary electrophoresis sequences as delivered by the 3500 Series instrument platform

Materials and methods

This *APC* mutation sequencing protocol was developed as part of a collaborative research effort between Applied Biosystems and the Molecular Diagnostics Unit (Hereditary

Cancer Program) at Institut Català d'Oncologia in Barcelona. The ICO team obtained blood samples identified as potentially having Familial Adenomatous Polyposis (FAP) or attenuated FAP and analyzed them in an anonymous and blinded study

(Table 1). Genomic DNA was isolated from blood using the FlexiGene® DNA Kit (QIAGEN) according to the manufacturer's instructions.

Making use of the Human Genome build (NM_000038.5), M13-tailed PCR

Table 1. Phenotype and mutation identification results for 10 samples.

Sample	Phenotype	Results
1	FAP	c.1958+1G>A
2	AFAP	No mutation identified
3	FAP	c.2468C>A; p.Ser823X
4	FAP	c.3916G>T; p.Glu1306X
5	FAP	c.637C>T; p.Arg213X
6	AFAP	No mutation identified
7	AFAP	No mutation identified
8	AFAP	No mutation identified
9	AFAP	No mutation identified
10	AFAP	No mutation identified
FAP = familial adenomatous polyposis; AFAP = attenuated FAP.		

Table 2. PCR primers and annealing temperatures for the amplification of each amplicon.

Amplicon	Primer sequence
	Forward
APC-E1	5' - CCTCTTAGATGCTGCTACTTGAACA -3'
APC-E2 (with M13 tail)	5' - TGTAACACGACGGCCAGTTGCGTGCTTTGAGAGTGATCTGA -3'
APC-E3	5' - AAGCAATTGTTGTATAAAACTT -3'
APC-E4	5' - CAACTGATGTAAGTATTGCTC -3'
APC-E5	5' - CCATGACTGACCTATTTGCTTATTC -3'
APC-E6 (with M13 tail)	5' - TGTAACACGACGGCCAGTTGAGAATGATTTGACATAACCCTGAGC -3'
APC-E7 (with M13 tail)	5' - TGTAACACGACGGCCAGTCGACCGCAATCGTACTGGA -3'
APC-E8 (with M13 tail)	5' - TGTAACACGACGGCCAGTTTCAGTCTTTGGTTAAGTCCATTCTGC -3'
APC-E9A (with M13 tail)	5' - TGTAACACGACGGCCAGTGCACCAACTTATCTAGGCAAACAGCA -3'
APC-E9B (with M13 tail)	5' - TGTAACACGACGGCCAGTTGTTTCTAACTCATTGCCCCACAG -3'
APC-E10 (with M13 tail)	5' - TGTAACACGACGGCCAGTTCAAGGGCAGATGAGTGGTAAACA -3'
APC-E11 (with M13 tail)	5' - TGTAACACGACGGCCAGTTCACTGTTGGCAAGGTGCAGTG -3'
APC-E12 (with M13 tail)	5' - TGTAACACGACGGCCAGTGACCAAGGCAAGTGTACACACACA -3'
APC-E13	5' - TACTGCTAGCATTAAAAACAAA -3'
APC-E14	5' - GGGACGGGCAATAGGATAGA -3'
APC-E15.1	5' - GTTACTGCATACACATTGTGAC -3'
APC-E15.2	5' - AGTCTTAAATATTCAGATGAGCAG -3'
APCE-15.3	5' - AGTAAATGCTGCAGTTCAGAGG -3'
APC-E15.4 (with M13 tail)	5' -TGTAACACGACGGCCAGTTGTTGAAGATACCCAGTTTGTT -3'
APC-E15.5 (with M13 tail)	5' -TGTAACACGACGGCCAGTCAGGGGAGAAAAGTACATTGGA -3'
APC-E15.6 (with M13 tail)	5' - TGTAACACGACGGCCAGTCAGATGAGCCAAACAGAACCTT -3'
APC-E15.7 (with M13 tail)	5' - TGTAACACGACGGCCAGTCGTGAGCACAGCAAACATTC -3'
APC-E15.8 (with M13 tail)	5' - TGTAACACGACGGCCAGTCCCCTGACCAAAAAGGAACT -3'

primers were chosen from the NCBI Probe database, and primers were optimized for 100% coverage of the APC gene's coding sequences (Table 2). The APC gene comprises 15 exons, with exon 15 accounting for 77% of the coding sequence. In order to scan the exons for mutations, the gene was divided into 23 amplicons using specific primer pairs, with 8 amplicons covering exon 15.

Amplification reactions and conditions

To enable a fast sequencing approach, the amplifications were performed using 100 ng of the extracted DNA with the AmpliTaq Gold® Fast PCR Master Mix on a Veriti® 96-well, 0.2 mL Thermal Cycler.

The structure of the APC gene sequence did not permit the use of a single

optimal thermal profile for all amplification primers, so different amplification conditions were designed to fit the primer sequences and amplicon lengths (Tables 2 and 3). For each of the amplified products, 2 µL was analyzed by agarose gel electrophoresis to check the amplification quality and quantity.

For more information on the amplification protocol and primers, please contact the researchers directly at sgonzalez@iconcologia.net.

Purification of PCR products and cycle sequencing conditions

The PCR products (8 µL each) were purified with Exonuclease I (New England Biolabs) and Antarctic phosphatase (New England Biolabs) following the protocol described in Table 4. Cycle sequencing was performed using the BigDye® Terminator

Cycle Sequencing Kit V3.1 with modifications to the manufacturer's recommended protocol (Table 5). Each amplicon was sequenced in both forward and reverse directions using M13 primers (Table 6) (exon 2, 6, 7, 8, 9, 10, 12, 15.4, 15.5, 15.6, 15.7, 15.8), the same primers used for amplification (exon 1, 3, 5, 13, 14), or internal primers (Table 6). Sequencing reactions were purified using the Applied Biosystems® BigDye XTerminator® Purification Kit according to the manufacturer's protocol.

Capillary electrophoresis and analysis

Purified sequencing reactions were analyzed independently at both ICO and Applied Biosystems laboratories. At ICO the reactions were carried out on the Applied Biosystems® 3730xl Genetic Analyzer using the BigDye XTerminator® Ultraseq 36_POP7

Reverse	Amplicon (exon)	Size (bp)	Annealing temp. (°C)	PCR condition (see Table 3)
5' - GAGGGAATACTGAATAAAATGGA -3'	1 (1)	380	60	A
5' - CAGGAAACAGCTATGACCCCAACACCCAAATCGAGAGAAGC -3'	2 (2)	339	60	A
5' - ACAATAAACTGGAGTACACAAGG -3'	3 (3)	293	50	B
5' - TTTAATGGATTACCTAGGTACT -3'	4 (4)	264	55	C
5' - TCCAAGGCAGAACAGACAGT -3'	5 (5)	458	60	D
5' - CAGGAAACAGCTATGACCGGGTTTCTGGAGTAAACACAGGTAAA -3'	6 (6)	591	60	A
5' - CAGGAAACAGCTATGACCTGCTTCTGGAATATGCATTGAGGA -3'	7 (7)	440	60	A
5' - CAGGAAACAGCTATGACCTGTGTGCCACCACACTGGCT -3'	8 (8)	604	60	A
5' - CAGGAAACAGCTATGACCGCCCGAGCCTCTTTACTGCC -3'	9A (9)	559	60	A
5' - CAGGAAACAGCTATGACCTGAGTAGCACAAATGGCTGATATGAAT -3'	9B (9)	577	60	A
5' - CAGGAAACAGCTATGACCTCCTTTCTATGCTGGAACCAGGG -3'	10 (10)	550	60	A
5' - CAGGAAACAGCTATGACCAAGCGAATGTGAAGCACAGGTTT -3'	11 (11)	620	60	A
5' - CAGGAAACAGCTATGACCGCACCTGTAGGCCACGCCTC -3'	12 (12)	600	60	A
5' - TGAACATGGGAGACGGAGG -3'	13 (13)	460	60	A
5' - ACATTGCTTACAATTAGGTCTTTTT -3'	14 (14)	405	60	A
5' - TGTTTGGGTCTTGCCCATCTT -3'	15.1 (15)	1,500	59	E
5' - TATCAGCATCTGGAAGAACCT -3'	15.2 (15)	1,650	59	E
5' - CTTTTTGGCATTGCGGAGCT -3'	15.3 (15)	2,000	63	F
5' - CAGGAAACAGCTATGACCTTGTCTGCTCGAGAGATT -3'	15.4 (15)	650	60	A
5' -CAGGAAACAGCTATGACCTCCTTTGGAGGCAGACTCAC -3'	15.5 (15)	692	60	A
5' - CAGGAAACAGCTATGACCTCACTGGATTCTGATGAAGCA -3'	15.6 (15)	632	60	A
5' - CAGGAAACAGCTATGACCTTTGCTTGAGCTGCTAGAAGT -3'	15.7 (15)	663	60	A
5' - CAGGAAACAGCTATGACCGAAGTTGGGATGGGATGCTA -3'	15.8 (15)	623	60	A

Table 3. PCR mixes and thermal cycling conditions for PCR using the Veriti® Thermal Cycler.

PCR condition	PCR mix					Thermal cycling conditions			
	AmpliTaq Gold® Fast PCR Master Mix	Forward primer (1 µM)	Reverse primer (1 µM)	DNA (100 ng total)	Final volume	Denaturation	Annealing	Extension	Cycle number
A	10 µL	4 µL	4 µL	2 µL	20 µL	96°C 15 sec	60°C 15 sec	68°C 15 sec	35
B	10 µL	4 µL	4 µL	2 µL	20 µL	96°C 15 sec	50°C 15 sec	68°C 15 sec	35
C	10 µL	4 µL	4 µL	2 µL	20 µL	96°C 15 sec	55°C 15 sec	68°C 15 sec	35
D	10 µL	3 µL	3 µL	2 µL + 2 µL H ₂ O	20 µL	96°C 15 sec	60°C 15 sec	68°C 15 sec	35
E	15 µL	6 µL	6 µL	3 µL	30 µL	96°C 30 sec	59°C 30 sec	74°C 30 sec	40
F	15 µL	4 µL	4 µL	3 µL + 4 µL H ₂ O	30 µL	96°C 30 sec	63°C 30 sec	74°C 30 sec	40

Table 4. PCR purification protocol.

1. PCR purification master mix					2. Reaction setup	3. Incubation steps (37°C)	
Exonuclease I (New England Biolabs, 20,000 U/mL)	Antarctic phosphatase (New England Biolabs, 5,000 U/mL)	Antarctic phosphatase buffer 10X	H ₂ O		8 µL PCR reaction + 2 µL PCR Purification Master Mix	Incubation 1	Incubation 2
50 µL	40 µL	26 µL	144 µL			15 min	15 min

Table 5. Sequencing mix and thermal cycling conditions for sequencing reactions using the Veriti® Thermal Cycler.

1. PCR mix					2. Thermal cycling conditions				
BigDye® Terminator v3.1	Buffer 3.1	Sequencing primer (3.2 µM)	PCR product	Final volume	Initial Denaturation (°C)	Denaturation (°C)	Annealing	Extension	Cycle number
1 µL	1.5 µL	1 µL	1 µL + 5.5 µL H ₂ O	10 µL	96°C 1 min	96°C 10 sec	50°C 5 sec	60°C 4 min	25

Table 6. Internal sequencing primers used for sequencing reactions.

Amplicon	Primer name	Sequence
APC-E4	4INT forward	5' - ATAGGTCATTGCTTCTTGCTGA -3'
	APC-E4 reverse	5' - TTTAATGGATTACCTAGGTACT -3'
APC-E9A	M13 reverse	5' - CAGGAAACAGCTATGACC -3'
APC-E11	E11INT forward	5' - GATGATTGTCTTTTCCTCTTGC -3'
	M13 reverse	5' - CAGGAAACAGCTATGACC -3'
15.1	15 C forward	5' - ATTTGA ATACTACAGTGTTACCC -3'
	15 B reverse	5' - ACTTCTATCTTTTCAGAACGAG -3'
15.2	15 G forward	5' - AAGAAACAATACAGACTTATTGT -3'
	15 F2 reverse	5' - AGCTGATGACAAAGATGATAATG -3'
15.3	15 K forward	5' - CCCTCCAAATGAGTTAGCTGC -3'
	15 J reverse	5' - GAGCCTCATCTGTACTTCTGC -3'
	15 M forward	5' - ATGATGTTGACCTTTCCAGGG -3'
	15 L reverse	5' - TTTTGCCTTTCTTAATTCAGCC -3'
APC-E2, APC-E6, APC-E7, APC-E8, APC-E9B, APC-E10, APC-E11, APC-E12, APC-E15.4, APC-E15.5, APC-E15.6, APC-E15.7, APC-E15.8	M13 forward	5' - TGTA AACGACGGCCAGT -3'
	M13 reverse	5' - CAGGAAACAGCTATGACC -3'

run module. A concordance analysis of the samples was performed at Applied Biosystems using the 3500xL Genetic Analyzer (see Highlight) using the ShortReadSeq50_POP7xl run module. The sequence results for each sample were analyzed independently at the two laboratories using Mutation Surveyor® (SoftGenetics) and SeqScape® Software (Figure 1) to verify the results and to identify putative mutations in each sample.

Results

The ICO team developed the optimized sequencing protocol presented in this publication. The protocol has been subsequently verified on 10 human DNA samples. DNA samples were amplified with AmpliTaq Gold® Fast PCR Master Mix in 23 amplicons to cover the *APC* gene. After PCR clean-up, amplicons were sequenced using BigDye® Terminator v3.1 chemistry. Following sequence analysis by both teams, the mutations observed using the Applied Biosystems® 3730xl Genetic Analyzer were found to be identical to the results obtained by Applied Biosystems scientists analyzing the same samples on the 3500xL Genetic Analyzer. Table 1 shows the results of the 10-sample study; we identified the mutation in all the classical FAP cases but not in the attenuated FAP cases.

Conclusion

In comparing sequence data from the 3730xl and 3500xL Series Genetic Analyzers, mutation detection was demonstrated with 100% concordance. The ICO team reported the following advantages when using this Sanger sequencing-based method for *APC* mutation studies:

- The optimized workflow presented in this research study is easy to implement in research laboratories, and enables reliable and reproducible results.

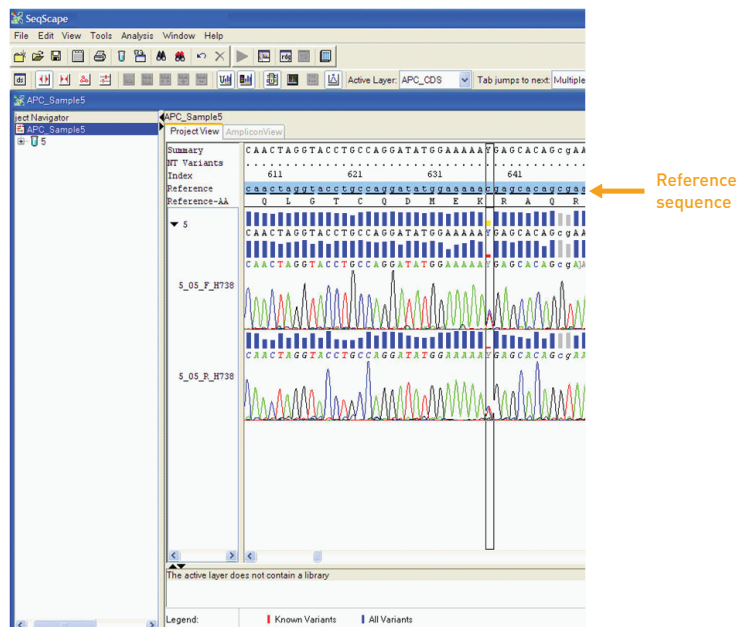


Figure 1. Identifying sequence variation. Through automated sequence analysis, superior basecalling, and sophisticated data visualization tools, SeqScape® Software v2.7 facilitates definitive mutation analysis, as exemplified in this study of *APC* gene mutations. Through the project view in SeqScape® Software, mutations are easily identified—the region showing the sequence variation (exon 5: c.637C>T) is boxed in black.

- The AmpliTaq Gold® Fast Master Mix gives faster, more reproducible and robust results; following this protocol optimization, AmpliTaq Gold® Fast Master Mix is now widely used in the routine workflows of the Institut Català d'Oncologia laboratory.
- The BigDye XTerminator® Purification Kit used for purifying sequencing reactions facilitates automation and streamlines the workflow.
- The workflow was optimized to complete the entire sequencing effort using a single thermal profile, which significantly reduced sequencing time.

References

1. Vasen HF, Möslin G, Alonso A et al. (2007) Guidelines for the clinical management of familial adenomatous polyposis (FAP). *Gut* 57:704–713.
2. Fodde R, Smits R, Clevers H (2001) APC signal transduction and genetic instability in colorectal cancer. *Nat Rev Cancer* 1:55–67.

3. Laurent-Puig P, Bérout C, Soussi T (1998) APC gene: database of germline and somatic mutations in human tumors and cell lines. *Nucleic Acids Res* 26:269–270.

Scientific contributors

Special thanks to Beatriz Cabot, Sr. Scientific Application Specialist, Life Technologies, for her contribution in this collaboration. This collaboration was led by Rosella Petraroli, FALCON Team, Europe, Life Technologies.

3500 Series Genetic Analyzers: built on a legacy of innovation

The 3500 Series is setting a new standard in capillary electrophoresis—integrating a number of platform improvements designed for the optical and thermal subsystems, and pioneering an innovative consumables system approach. Working together, these elements provide the highest level of performance from genetic analyzers to date.

The 3500 Series platform can run a wide variety of research applications—including *de novo* sequencing and resequencing (mutational profiling), as well as microsatellite analysis, MLPA™, AFLP®, LOH, MLST, and SNP validation or screening. The majority of research applications can be run on a single polymer and capillary array, and the 3500 Series Data Collection Software integrates seamlessly with several downstream Applied Biosystems® software packages to provide comprehensive analysis of genetic data:

- Variant Reporter® Software—designed for mutation detection and analysis, SNP discovery and validation, and sequence confirmation
- Sequencing Analysis Software with KB™ Basecaller—designed to analyze, display, edit, save, and print sequencing data
- SeqScape® Software—used for mutation detection and library-based allele identification
- GeneMapper® Software—an ideal tool for genotyping, allele calling, fragment sizing, and SNP analysis

For more information on the 3500 Series Genetic Analyzers, visit us online at www.appliedbiosystems.com/3500



This “Your Innovative Research” article describes work done by researchers using Life Technologies products. Note that the protocols and data described here are those of the contributing laboratory and may include modifications to the recommended protocol from Life Technologies. Optimization may be required for best results in your laboratory setting.

For Research Use Only. Not intended for any animal or human therapeutic or diagnostic use.

© 2011 Life Technologies Corporation. All rights reserved. The trademarks mentioned herein are the property of Life Technologies Corporation or their respective owners. AmpliTaq Gold is a registered trademark of Roche Molecular Systems, Inc. FlexiGene is a registered trademark of Qiagen Corp. Mutation Surveyor is a registered trademark of SoftGenetics, LLC. Printed in the USA. **CO18247 0611**