Sanger sequencing using Ion AmpliSeq primers and libraries

Exploiting advantages of Ion AmpliSeq technology for limited-sample genomic analyses

FIndings of this study:

- Sanger sequencing results can be obtained using Ion AmpliSeq™ library primer sequences or from existing Ion AmpliSeq™ library pools
- Robust genotyping results using both Ion AmpliSeq™ next-generation sequencing (NGS) and confirmatory Sanger sequencing can be generated from less than 1 ng of FFPE DNA
- Previously sequenced NGS libraries can be used as a direct input for confirmatory Sanger sequencing

Introduction

Next-generation sequencing (NGS) has revolutionized biology and clinical research, yielding vast amounts of genomic data at an affordable price. In molecular oncology research, NGS panels are now used to analyze genomes for mutations in the search for relevant targets. A major advantage of the Ion AmpliSeq™ library preparation method is that a very low amount of gDNA is required for the process. Typically, as little as 10 ng of gDNA is needed for Ion AmpliSeq library preparation, which allows the use of minute amounts of formalin-fixed, paraffin-embedded (FFPE) tissue or retrospective samples from cells and fine-needle aspiration biopsy (FNAB) material. However, sequence information is occasionally needed from samples containing much less than 10 ng. It is therefore crucial to develop tools that allow extracting maximum sequence information from very small amounts of sample.

In the analysis of tissue samples, germline sequence variants can present as homozygous or, more frequently, as heterozygous sequence differences. For heterozygous single nucleotide polymorphism (SNP) variants, a typical 50:50 allele ratio can easily be observed in NGS results and in Sanger sequencing peak traces. However, somatic variants may occur in unusual ratios, such as 5–10% minor allele frequencies, in tumor samples with heterogeneous mutational histories. These low allele frequencies raise the issue of whether the minor allele is a true variant or an experimental aberration. In these cases, reflex testing using an orthogonal sequencing technology is needed.

This study demonstrates how Ion AmpliSeq panels analyzed by Sanger sequencing can be used to verify the presence of unusual or low-frequency alleles in a sample. To illustrate this, we use the Ion AmpliSeq™ Cancer Hotspot Panel v2 to re-amplify specific oncogene or tumor suppressor gene targets, and subsequently resequence using the Applied Biosystems™ BigDye™ Direct Cycle Sequencing Kit and the Applied Biosystems™ 3500 Series Genetic Analyzer (Figure 1). Further, we present a workflow for extremely limited gDNA samples that uses amplification material from Ion AmpliSeq library preparation as a reservoir for reflex testing of individual targets by Sanger sequencing.

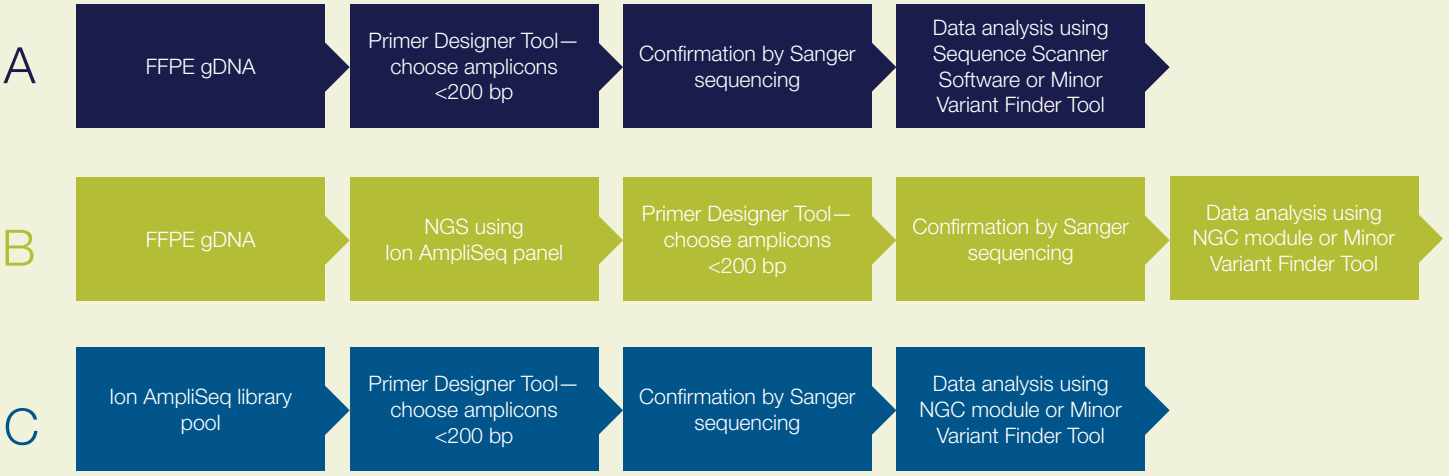


Figure 1. Three scenarios for variant confirmation.

Ion AmpliSeq primer design and conversion to Sanger PCR primers

The Ion AmpliSeq Cancer Hotspot Panel v2 used for this study is a single pool of amplification primers for targeted sequencing of genes frequently mutated in human tumor samples. This panel is designed to amplify 207 amplicons covering approximately 2,800 Catalogue Of Somatic Mutations In Cancer (COSMIC) mutations from 50 oncogenes and tumor suppressor genes in a single multiplex PCR (Figure 2). PCR primer sequences for other commercial Ion AmpliSeq™ panels and Ion AmpliSeq™ Community Panels are available for download at ampliseq.com.

Before using the Ion AmpliSeq primers for Sanger sequencing, universal M13 primer sequences must be added to the 5′ ends of the Ion AmpliSeq forward and reverse primers. In this study, the M13 forward primer sequence 5′-TGTAAAACGACGGCCAGT-3′ was added to the 5′ end of the Ion AmpliSeq forward primer sequence, and the M13 reverse primer sequence 5′-CAGGAAACAGCTATGACC-3′ was added to the 5′ end of the Ion AmpliSeq reverse primer sequence. Alternatively, native or M13-modified primer pairs from the Ion AmpliSeq Cancer Hotspot Panel v2 can be ordered using the Primer Designer™ Tool at **thermofisher.com/primerdesigner**

Streamlined PCR-to-sequencing workflow

In this study, we used the DNA control sample (CEPH-02), which is included in the BigDye Direct Cycle Sequencing Kit, for intact DNA. We also used a mixture of two Coriell Institute DNA samples, referred to as NA8020. For FFPE DNA preparations, we extracted DNA from FFPE tissue sections. We used the BigDye Direct Cycle Sequencing Kit for integrated PCR and sequencing of individual targets. For post-PCR cleanup after cycle sequencing, BigDye™ Xterminator bead suspension was added to the reactions and vigorously vortexed for 20–30 min, followed by centrifugation for 1 min to pellet the beads. The plate was loaded on a 3500 Series Genetic Analyzer, and capillary electrophoresis was performed using a rapid run module. For details, see the Appendix.

	A	B	C	D	E	F	G	H	I	J	K
1	Amplicon_ID	Ion_AmpliSeq_Fwd_Primer	Ion_AmpliSeq_Rev_Primer	LineItem_	Gene_Syn	Genome_	Chr	Amplicon_Start	Amplicon_Stop	Insert_Star	Insert_Stop
2	CHP2_ABL1_1	TCTATGGTGTGTCCCCCAACT	CGTCAGGCTGTATTTCTTCCACA	GENE	ABL1	hg19	chr9	133738274	133738401	133738295	133738378
3	CHP2_ABL1_2	CGCTGAAGCTCCATTTTGCAT	CAGCTTCTTTCAAGAACTCTTCCAC	GENE	ABL1	hg19	chr9	133747422	133747555	133747443	133747530
4	CHP2_ABL1_3	GGGAGCCCCCGTTCTATATCA	AGGCCCCTACCTGTGGA	GENE	ABL1	hg19	chr9	133748259	133748434	133748280	133748417
5	CHP2_ABL1_4	GGAGAACCACTTGGTGAAGGTAG	CGGACTTGATGGAGAACTTGTTGTA	GENE	ABL1	hg19	chr9	133750285	133750430	133750308	133750405
6	CHP2_AKT1_1	GCGCCACAGAGAAGTTGTTGA	GGGTCTGACGGGTAGAGTGT	GENE	AKT1	hg19	chr14	105246425	105246603	105246446	105246583
7	CHP2_AKT1_2	CTTGGCCACGATGACTTCCTT	CCATGAACGAGTTTGAGTACCTGA	GENE	AKT1	hg19	chr14	105241413	105241543	105241434	105241519
8	CHP2_ALK_1	TCTCTCGGAGGAAGGACTTGAG	GCCCAGACTCAGCTCAGTTAAT	GENE	ALK	hg19	chr2	29443586	29443751	29443608	29443729
9	CHP2_ALK_2	ACAGGGTACCAGGAGATGATGTAAG	GGAAGAGTGGCCAAGATTGGA	GENE	ALK	hg19	chr2	29432548	29432701	29432573	29432680
10	CHP2_APC_1	GAGAGAACGCGGAATTGGTCTA	GTATGAATGGCTGACACTTCTTCCA	GENE	APC	hg19	chr5	112173850	112173987	112173872	112173962
11	CHP2_APC_2	AGCACTGATGATAAACACCTCAAGTT	ATCTTCTTGACACAAAGACTGGCT	GENE	APC	hg19	chr5	112174532	112174690	112174558	112174666

Figure 2. Excerpt of the PCR primer data sheet for the Ion AmpliSeq Cancer Hotspot Panel v2. Shown are the first 10 of the 207 amplicons. Genome coordinate annotations and hyperlinks to COSMIC and SNP databases as well as the UCSC Genome Browser are available at ampliseq.com.

Sanger sequencing data analysis tools

There are several options for analyzing Sanger sequencing data, depending on the needs and the complexity of the project:

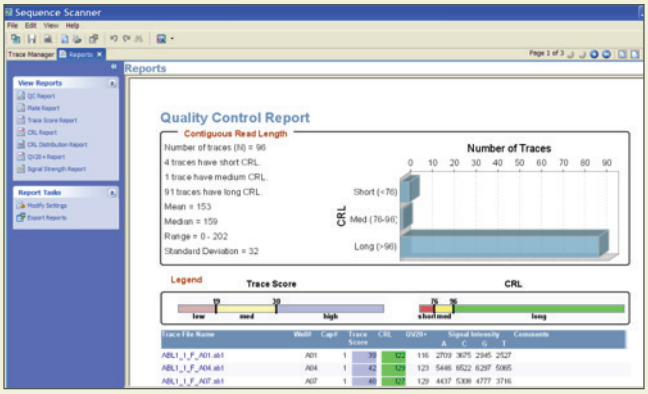
- **Sequence Scanner Software**—a free download of this software, which is useful for visual inspection and base calling, can be obtained at **thermofisher.com/sangersoftware**
- **Applied Biosystems™ Variant Reporter™ Software**—downloadable desktop software specialized for detection and reporting of typical Mendelian inheritance mutations, such as heterozygous or homozygous single nucleotide variations (SNVs) or short insertions or deletions (indels); can be obtained at **thermofisher.com/sangersoftware**
- **Minor Variant Finder Tool**—specialized desktop software for enhanced and automated detection of minor variants in Sanger sequencing traces; can be obtained at **thermofisher.com/mvf**
- **Thermo Fisher Cloud NGC module**—a specialized cloud-based application that enables direct side-by-side comparison of variants from NGS-derived .vcf files and Sanger sequencing files and generates a comprehensive report; available at **apps.thermofisher.com** (requires registration)

For the data analysis described below, sequencing data files (.ab1 files) were analyzed with Sequence Scanner Software, and QC reports were generated and exported as tables readable in Microsoft Excel™ software (Figure 3). The Excel file with the QC data was further analyzed using an Excel software macro, and the sample files were flagged with a numerical penalty flag (+1) according to the following criteria: trace score is <40, contiguous read length is <100, QV20 bases is <100, signal strength RFU for the (A) trace is <300. If the trace score was below 30, the trace file was given a penalty of +4 and called as failed. A combined score is generated by adding the number of flags:

Score result = 0: good data
Score result = 1: mostly good data; may require visual review
Score result = 2: lower-quality data; usable after manual review, but stretches of poor data may exist
Score result = 3 or more: very poor data quality; not usable

1 ng of input DNA can be used for Sanger sequencing of individual Ion AmpliSeq targets

We established the workflow and protocol for resequencing of individual hotspot targets by using 48 amplicons representing a subset of the Ion AmpliSeq Cancer Hotspot Panel v2 targets. In addition, we tested 24 Ion AmpliSeq panel targets covering the exons of the human *TP53* gene. For DNA samples, we used the CEPH-02 control gDNA that is included in the BigDye Direct Cycle Sequencing Kit; a mixture of two Coriell Institute DNA samples, referred to as NA8020; and two DNA preparations from FFPE tissue sections, referred to as FFPE 1 and FFPE 5.



	A	B	C	D	E	F	G
1	Trace File Name	Well#	Cap#	Trace Score	Contiguous Read Length	QV20+	Signal Intensity (A)
2	ABL1_1_F_A01.ab1	A01	1	39	122	116	2709
3	ABL1_1_F_A04.ab1	A04	1	42	129	123	5446
4	ABL1_1_F_A07.ab1	A07	1	40	127	129	4437
5	ABL1_1_F_A10.ab1	A10	1	39	125	126	3058
6	ABL1_2_F_B01.ab1	B01	4	44	134	142	5620
7	ABL1_2_F_B04.ab1	B04	4	46	134	145	4821
8	ABL1_2_F_B07.ab1	B07	4	47	134	150	5534
9	ABL1_2_F_B10.ab1	B10	4	46	134	145	3995
10	ABL1_4_F_C01.ab1	C01	7	41	146	149	1489
11	ABL1_4_F_C04.ab1	C04	7	41	144	156	3268

Figure 3. Generating and exporting results using Sequence Scanner Software. A QC report was generated and exported for further analysis in Excel software.

To determine the minimal amount of gDNA that is needed for robust Sanger sequencing of Ion AmpliSeq panel targets, we diluted CEPH-02 gDNA to 2.5 ng/µL, 1 ng/µL, 0.5 ng/µL, and 0.125 ng/µL for use as input DNA in individual PCR reactions using a subpanel of 24 primer pairs. Robust amplification and sequencing was obtained with as little as 1 ng of input DNA. Lowering the amount to 0.5 ng per reaction still yielded good results for over 80% of assays, but data quality for some amplicons started to deteriorate in this range. Data quality further deteriorated when the amount was lowered to 0.125 ng per reaction, but nearly 80% of assays produced usable sequencing data even with this very low amount of input DNA. Taken together, these results show that 1 ng of gDNA is an acceptable minimal amount of template DNA for both forward and reverse sequencing reactions using the BigDye Direct Cycle Sequencing Kit.

1 ng of FFPE DNA can be used for Sanger sequencing of individual Ion AmpliSeq targets

Although FFPE tissue is a standard sample type in histology and pathology laboratories, the fixation process damages DNA and makes it a challenging sample type for molecular genetic analysis. To determine the compatibility of Ion AmpliSeq panel primer designs with the BigDye Direct Cycle Sequencing Kit workflow (Figure 1, workflow A and B), we used two DNA preparations from FFPE tissue sections, FFPE 1 and FFPE 5. Here we used the Ion AmpliSeq primer designs for 24 coding segments of the human *TP53* gene (available at ampliseq.com). Similar to our findings with intact genomic DNA, we found that all 24 segments were successfully amplified and sequenced at an input amount of 1 ng DNA per PCR reaction. These results demonstrate that DNA extracted from FFPE-preserved samples is suitable for Sanger sequencing using M13-modified Ion AmpliSeq primer sets.

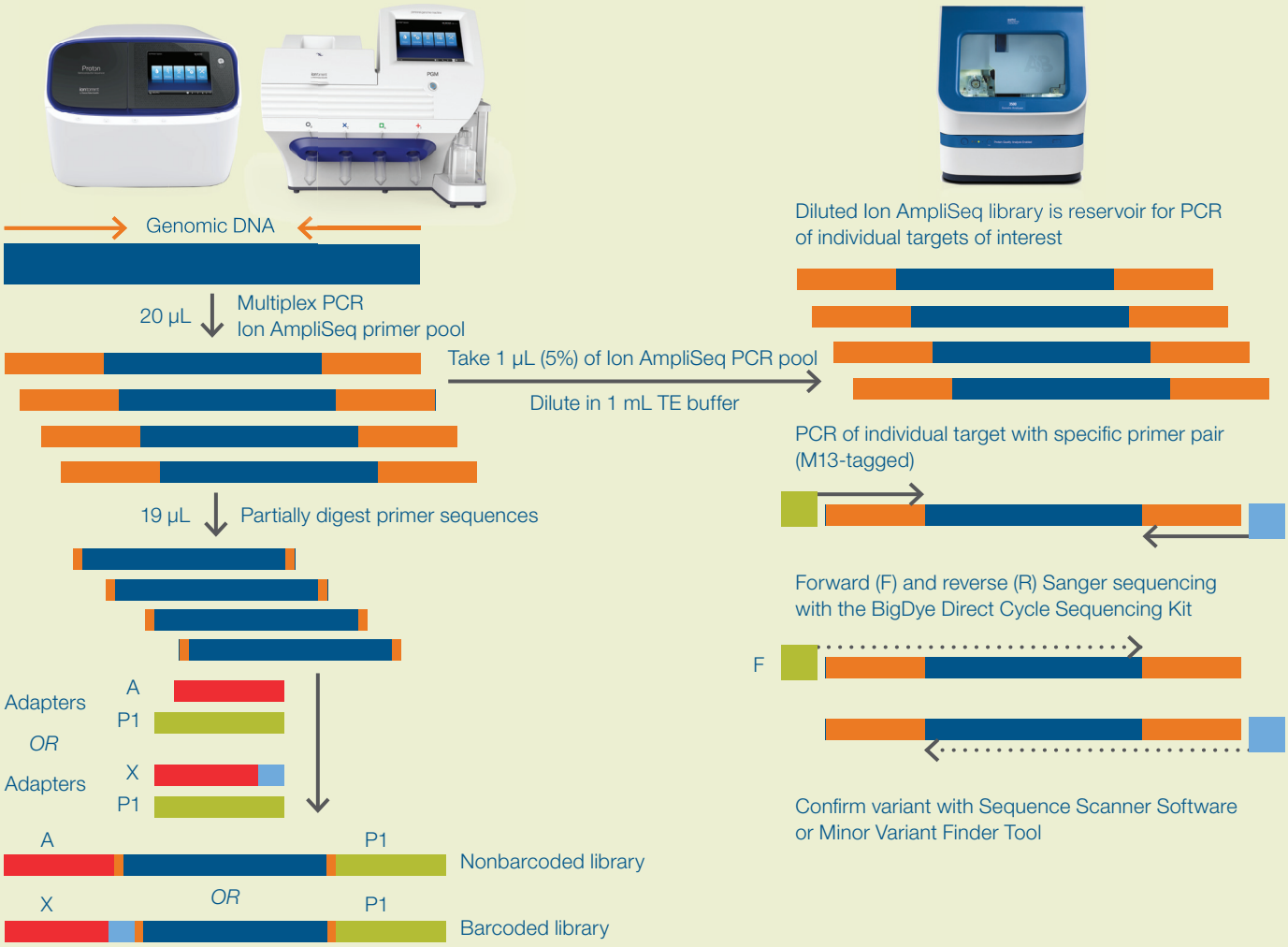


Figure 4. The Ion AmpliSeq library pool as a reservoir for Sanger sequencing of individual amplicons. The Ion AmpliSeq library preparation workflow is shown on the left. The PCR amplification material from the first step (the initial multiplex PCR) is the source of the Sanger sequencing reservoir. A 1 µL aliquot, representing approximately 5% of the amplification reaction, is taken before the library is further processed and is diluted in 1 mL of TE buffer. This material contains an enrichment of the targets of interest, which are typically amplified for 17 cycles. Assuming ideal PCR conditions and efficiency, it is estimated that each target is present at around 1,000–2,000 copies/µL in the 1:1,000 dilution (in 1 mL TE) of the aliquot; 1 µL of this dilution is then used as input DNA template in the PCR reaction using the BigDye Direct Cycle Sequencing Kit.

Table 1. Mutation frequencies observed for the *TP53* gene following amplification using an Ion AmpliSeq Community Panel and NGS on an Ion 318™ Chip.

	TP53_10.1.470 (TP53_02) chr17:7578645 C > T	TP53_11.1324631 (TP53_08) chr17:7579472 G > C	TP53_8.884088 (TP53_23) chr17:7578115 T > C	TP53_10.215568 (TP53_5) chr17:7578406 C > T
FFPE 5	21.8%	20.2 %	17.9 %	60.6%

Minor variants can be verified by Sanger sequencing of an Ion AmpliSeq library pool

In many instances, the small size and limited amount of DNA extracted from a sample limits further confirmatory analyses. Nevertheless, alleles appearing at low frequencies (10–20%) should be confirmed by orthogonal techniques. One other source of DNA for such analyses is the Ion AmpliSeq library pool generated for NGS. Sequences present in these libraries reflect the endogenous alleles in the sample, but because they are amplified, they are less limiting. We therefore determined whether Sanger sequencing of individual Ion AmpliSeq targets directly from a preexisting Ion AmpliSeq library pool would reflect the allele frequencies of the unamplified DNA source (Figure 1, workflow C). We used a retained aliquot (1 µL, approximately 5%) of the original FFPE 5 Ion AmpliSeq library material. A dilution of this aliquot was used as a template for specific, individual PCR and sequencing reactions. This workflow is illustrated in Figure 4. Mutation frequencies observed for the *TP53* gene after NGS are shown in Table 1, and Sanger sequencing results are shown in Figure 5. The frequencies of the alleles observed by NGS and Sanger sequencing

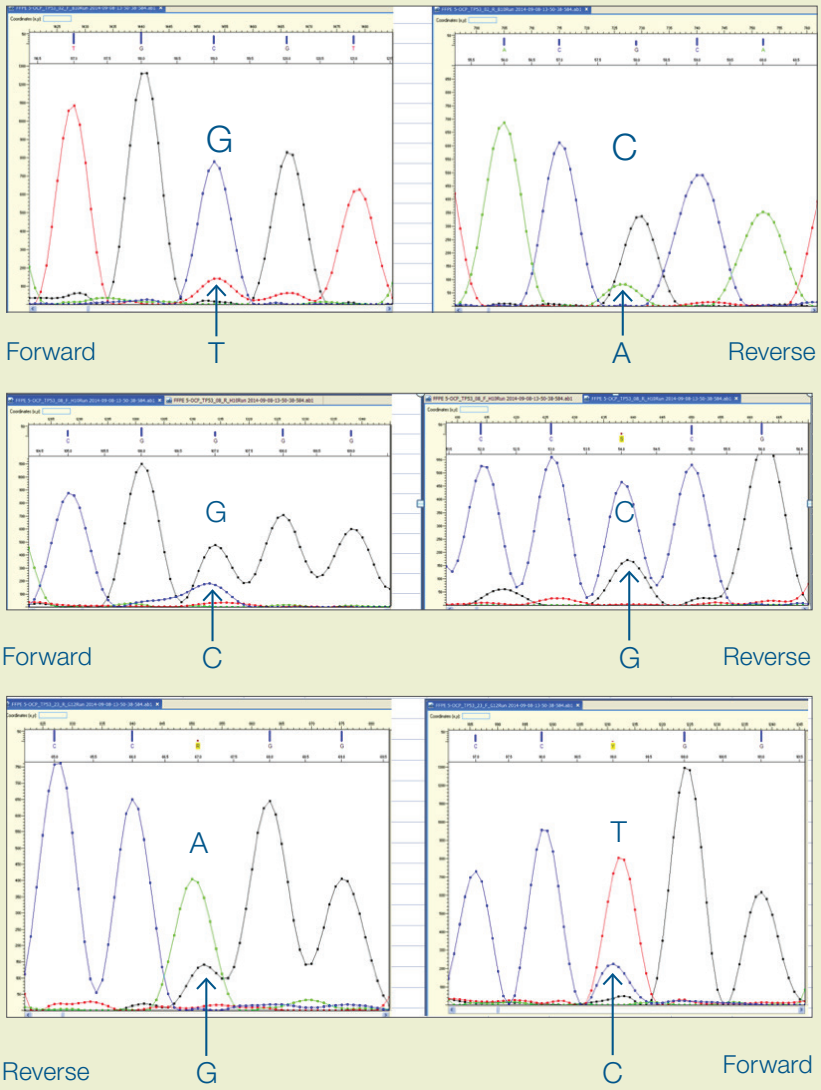


Figure 5. Corresponding Sanger sequencing traces for minor variants previously detected by NGS (Table 1). All minor variants could be visually confirmed using Sequence Scanner Software as discrete “peak under peak” entities in approximate proportion to the NGS results in both forward and reverse strands.

correlated very well. Thus, Sanger sequencing of the AmpliSeq library pool can be used to verify allele frequencies in the original sample DNA.

Conclusions

We have shown that Ion AmpliSeq multiplex PCR primer design can be readily converted to a singleplex PCR primer pair design by attaching M13 tags to forward and reverse primers, allowing robust PCR amplification and Sanger sequencing from limited amounts of genomic DNA (0.5–1 ng per PCR reaction). Ion AmpliSeq primer sequences are available at ampliseq.com or searchable and orderable from the Primer Designer Tool at thermofisher.com/primerdesigner. When DNA availability is extremely low, the amplification product generated as a first step in the Ion AmpliSeq library preparation process can serve as a potential reservoir for reflex testing by PCR and Sanger sequencing. However, the method is not yet optimized for high-complexity Ion AmpliSeq panels, such as the Oncomine™ Comprehensive Assay or human exome panel. These data demonstrate that researchers needing a fast and economical solution for confirmation of uncertain NGS results can rely on the robustness and sensitivity of PCR coupled with Sanger sequencing.

Appendix—protocol

Primer design and ordering summary

1. Add an M13 forward tag (5′-TGTAAAACGACGGCCAGT-3′) to the 5′ end of the Ion AmpliSeq forward primer, and an M13 reverse tag (5′-CAGGAAACAGCTATGACC-3′) to the 5′ end of Ion AmpliSeq reverse primer.
2. Order oligonucleotides using the Invitrogen Custom DNA Oligos service or through the Primer Designer Tool (25 nmol synthesis scale and desalted purification is sufficient for >100 PCR reactions).
3. Resuspend oligonucleotides in TE buffer to 100 µM (primer stock).
4. Generate a working solution of the primer pair (10 µM each) by combining 10 µL of the forward primer stock with 10 µL of the reverse primer stock and diluting them into 80 µL of deionized water or TE buffer.

Set up PCR amplification

A typical PCR reaction for bidirectional sequencing is set up in an Applied Biosystems™ MicroAmp™ Standard or Fast Optical 96-Well Reaction Plate as described below.

Combine:

- 10 µL BigDye™ Direct PCR Master Mix (included in kit)
- 2 µL PCR primer pair with M13 tags (10 µM each)
- 7 µL deionized water
- 1 µL DNA template (typically 1 ng) or 1 µL diluted (i.e., 1:1,000) Ion AmpliSeq post-PCR pool

Cover the plate with Applied Biosystems™ MicroAmp™ Optical Adhesive Film and perform PCR using an Applied Biosystems™ Veriti™ 96-Well Fast Thermal Cycler (or similar) using this profile:

Stage 1: 94°C for 10 min (1x)
Stage 2: 95°C for 3 sec, 60°C for 15 sec, 68°C for 45 sec (8x)
Stage 3: 95°C for 3 sec, 70°C for 50 sec (28x)
Stage 4: Hold at 4°C

Note: Only half of the plate is used for PCR (i.e., 48 samples in columns 1–6 to allow the later use of columns 7–12 for the reverse sequencing reaction).

Preparing the sequencing mix

The BigDye Direct Sequencing Master Mix contains a nuclease that degrades the remaining PCR primers when it comes in contact with the PCR material. This obviates the need for and cost of any further PCR purification. The BigDye Direct Sequencing Master Mix is simply combined in a 2:1 ratio with an M13 sequencing primer (either forward or reverse), which is chemically modified to be nuclease resistant. A sufficient quantity of modified nuclease-resistant M13 forward and reverse sequencing primers is supplied with the kit.

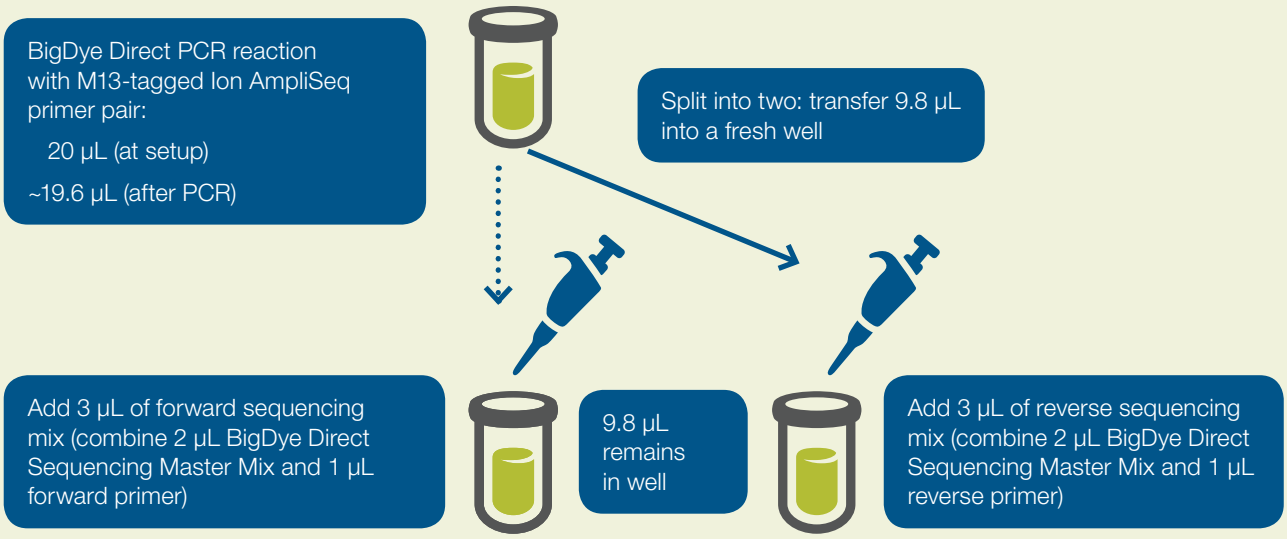


Figure 6. Streamlined PCR-to-sequencing workflow. The PCR reaction is split into 2 wells, and sequencing mix is added to each well. Cycle sequencing is then performed.

Splitting the PCR material for forward and reverse sequencing

Before adding the complete BigDye Direct Sequencing Master Mix plus primer (total of 3 µL), the PCR reaction needs to be split in half: one for the forward and one for the reverse sequencing reaction (Figure 6). This is done by transferring 9.8 µL of PCR material into free wells on the PCR plate with a multichannel pipettor (e.g., wells A–H of column 1 go into A–H of column 7). After splitting the PCR reactions, 3 µL of the appropriate sequencing mix (forward or reverse) is added to each well and the reaction mix is subjected to cycle sequencing.

Cycle sequencing is performed on the Veriti 96-Well Fast Thermal Cycler using this profile:

- Stage 1: 37°C for 20 min (1x)
- Stage 2: 80°C for 2 min (1x)
- Stage 3: 96°C for 1 min (1x)
- Stage 4: 96°C for 3 sec, 50°C for 5 sec, 60°C for 45 sec (27x)
- Stage 5: Hold at 4°C

After cycle sequencing, 50 µL of BigDye Xterminator bead suspension is added to the reactions and vigorously vortexed for 20–30 min followed by centrifugation for 1 min to pellet the beads. The plate is now ready for capillary electrophoresis on the 3500 Series Genetic Analyzer. Capillary electrophoresis is performed using a rapid run module that delivers sequence data files in 45 minutes.

applied
biosystems

Find out more at **thermofisher.com/primerdesigner**

ThermoFisher
SCIENTIFIC