

# Ion Torrent™

## Targeted Resequencing Guide

### SUMMARY

- Ion PGM™ system delivers the fastest sequencing runs and overall workflow for targeted resequencing
- Ion PGM™ system offers flexible solutions regardless of the length of the amplicon, or size of target region
- Ion PGM™ system is scalable for resequencing project needs, whether running single samples or multiplexing, or detecting germline or somatic mutations

### Introduction

The identification of common and rare variants in candidate regions of the human genome is essential to better understanding the etiology of complex human diseases. Leveraging the speed and throughput of next-generation sequencing platforms, targeted resequencing allows researchers to focus efficiently and cost-effectively on candidate genomic regions of interest. Whether resequencing amplicons or enriching genomic regions through target capture methods, the Ion Personal Genome Machine™ system provides the fastest sequencing run times and flexible and scalable solutions to fit your dynamic research needs. Multiple options for library preparation allow you to study amplicons of various lengths, Ion Barcoding kits enable sample multiplexing, and multiple Ion sequencing chips support sequencing target regions of variable sizes. With simple workflows and scalable solutions, Ion semiconductor sequencing brings a variety of options to complete targeted resequencing projects accurately, quickly, and affordably.

### Amplicon Resequencing

Whatever the length of your amplicons, the Ion PGM™ sequencer provides the fastest sequencing runs for amplicon resequencing across multiple genes and multiple samples. Various options for library preparation and the ability to scale with chips of varying sequencing capacity allow you to meet diverse project requirements. Here we demonstrate several methods of library preparation on the Ion PGM™ system for amplicon resequencing based on amplicon length.

#### Simple, fast workflow for amplicons with *Ion Xpress™* Fragment Library Kits

Ion has developed a simple, single-day workflow for sequencing standard Sanger (500 bp – 1 kb) or long-range PCR-derived (1 – 10 kb) amplicons utilizing enzymatic shearing. For amplicons that are longer than 150 nucleotides, Ion libraries can be prepared using the new *Ion Xpress™* Fragment Library Kit. This kit reduces library preparation from 6 hours to as little as 3.5 hours (Figure 2). Once amplicons are generated, the amplicons are sheared by enzymatic digestion, followed by ligation of Ion adaptors to the sheared fragments. After size selection and amplification, the adaptor-ligated fragments undergo clonal amplification and are ready to be sequenced on the Ion PGM™ sequencer (see *Ion Xpress™* Template Kit User Guide). This results in a 9-hour workflow for 100 bp reads from amplicons to sequence.

Figure 2. Ion Xpress™ Fragment Library Kit ~3.5-hour workflow.



Figure 1. Overview of Targeted Resequencing Methods and Multiplexing Schemas (one gene is approximately 1 kb).

TARGET REGION SIZE	1 kb	10 kb	100 kb	1 Mb	10 Mb
<b>AMPLICON RESEQUENCING</b> Ion Xpress™ Fragment Library Kits Amplicon Length: ~400 bp – >8 kb  Ion Fragment Library Kits Ligation Amplicon Length: 75 bp – 150 bp Fusion Primers Amplicon Length: 75 bp – 150 bp	<b>TARGET ENRICHMENT</b> TargetSeq™ Custom Enrichment Kits Target Size: 100 kb – >10 Mb  Halo Genomics Target Size: 10 kb – 400 kb SureSelect Target Size: <200 kb – >10 Mb				
NUMBER OF GENES	1	10	100	1 K	10 K
NUMBER OF SAMPLES PER CHIP (100x COVERAGE)					
Ion 314 (10 Mb)	100	10	1		
Ion 316 (100 Mb)		100	10	1	
Ion 318* (1 Gb)			100	10	1

This workflow is also compatible with long-range PCR products, such as those prepared with the *Invitrogen SequelPrep™ Long PCR Kit*. While the current settings of this protocol produce fragments of around 100 bp in length, conditions can be easily modified to produce longer fragments to leverage the readily increasing readlengths for the Ion PGM™ system without the need to redesign amplicon primers. As with physical shearing, enzymatic fragmentation produces multiple random start and end points for improved compatibility of current data analysis tools and subsequent data quality.

Although fragmentation for long amplicons can also be done physically, as described in the *Ion Fragment Library Kit User Guide*, enzymatic fragmentation has several advantages:

- Easy transfer of existing Sanger amplicons to semiconductor sequencing.
- No time-consuming concatemerization required (greater than ~3kb required for physical shearing).
- No costly sonicators or other physical shearing devices required.

### Germline variant detection in CFTR amplicons with *Ion Xpress™ Fragment Library Kits*

In this study, multiple amplicons from the CFTR gene (cystic fibrosis transmembrane conductance regulator) were sequenced, spanning ~12 kb (average amplicon 531 nucleotides in length). Amplicons were generated from 3 samples, HapMap sample NA12878 and Coriell samples CD00009 and CD00010. Amplicons were pooled, sheared, and adapted using the *Ion Xpress™ Fragment Library Kit*.

The amplicon library from each sample was sequenced on the Ion PGM™ sequencer following template prep with the *Ion Xpress™ Template Kit*. Using the same set of amplicons, the data was confirmed by Sanger sequencing on the ABI 3500 Genetic Analyzer using the *BigDye® Direct Cycle Sequencing Kit*.

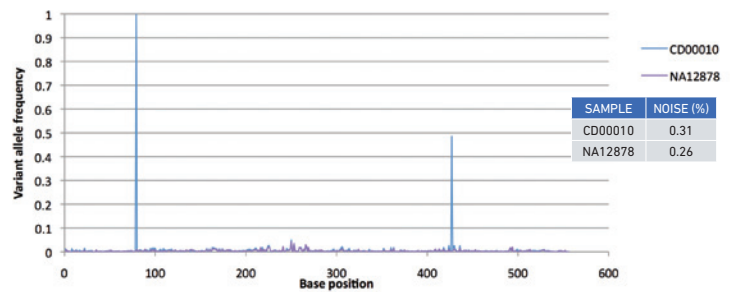
Table 1. Comparison of CFTR variants detected using Sanger and PGM sequencing.

rsID	Primer name	NA12878 Reference allele	CD00010		CD00009		Amplicon length (bp)
			Sanger genotype	PGM genotype	Sanger genotype	PGM genotype	
rs34855237	CFTR.10.80s.1	A	G/G	G/G	G/G	G/G	527
rs213950		G	A/A	A/A	A/A	A/A	
rs4148711	CFTR.12.570s.1	T	A/T	A/T	reference	reference	518
rs11978434	CFTR.13.120s.1	T	C/T	C/T	reference	reference	499
rs1042077	CFTR.14.550s.1	T	G/T	G/T	reference	reference	573
rs4148712		AT	AT/-	AT/-	reference	reference	
hg19 chr7:117251899	CFTR.19.560s.1	G	A/G	A/G	reference	reference	604
rs121908769	CFTR.2.520s.1	TT	reference	reference	TT/-	TT/-	506
rs213989	CFTR.21.420s.1	C	A/C	A/C	reference	reference	503
rs214164	CFTR.23.80s.1	G	A/G	A/G	reference	reference	500
rs4727855	CFTR.25.190s.1	G	A/G	A/G	reference	reference	517
rs35516286	CFTR.3.70s.1	T	C/T	C/T	reference	reference	556
rs1429566		G	A/A	A/A	A/G	A/G	
rs34159932	CFTR.5.520s.1	G	A/A	A/A	A/G	A/G	513
rs1800503	CFTR.6.210s.1	C	T/T	T/T	C/T	C/T	512
rs67140043		GATT	-/-	-/-	GATT/-	GATT/-	

Table 1 shows the correlation of germline variants between PGM sequencing and Sanger sequencing. For those amplicons that were amplified successfully, we observed 100% concordance between semiconductor sequencing and Sanger sequencing. The average depth of coverage achieved with PGM sequencing was 919x and 1235x for CD00010 and CD00009, respectively. PGM variants in Table 1 were identified using the Variant Caller Plugin deployed with Torrent Suite v.1.4.\* The Variant Caller Plugin allows automated identification of both SNPs and indels using the community standard Samtools package, and provides streamlined single-click access via a Web browser interface to the Integrative Genome Browser (IGV) from the Broad Institute.

Figure 3 shows the noise level across the whole length of amplicon CFTR.3.70s.1 on the Ion PGM™ sequencer. The average depth of coverage of this amplicon was 835x. This low error rate can enable detection of somatic mutations at very low frequencies in heterogeneous cancer samples.

Figure 3. Amplicon CFTR3.70s.1 showing homozygous and heterozygous variants.



## Shorter amplicons for FFPE samples and disease research applications

With highly degraded samples and certain disease research applications, the design of longer amplicons is not always feasible. When sequencing short amplicons that are under 150 nucleotides in length, either a fusion primer or ligation approach is recommended. Rapidly increasing readlengths up to 200 bp on the Ion PGM™ system in 2011 will allow for the sequencing of larger amplicons of non-degraded samples.

With fusion primers, amplicon libraries are created via PCR with a pair of amplicon-specific primers that include the Ion adaptor sequences. Briefly, one fusion primer will have the A adaptor region followed by the proximal end of the target amplicon, and the other fusion primer will have the P1 adaptor region followed by the distal end of the target amplicon. For optimal results with fusion primers, we recommend using four fusion primers for bidirectional sequencing to produce high-quality reads from both ends and across the full length of the amplicon (see the *Ion Fragment Library Kit User Guide*). Fusion primers are often the fastest approach when routinely sequencing a small number of short amplicons.

An adaptor-ligation method is also applicable for short amplicons, and may be an easier and more affordable method when sequencing a large number of amplicons per sample (see the *Ion Fragment Library Kit User Guide*). Briefly, Ion adaptors are ligated to existing amplicons, and then the A-P1 ligated amplicons are amplified by PCR. As Ion Torrent works rapidly towards longer readlengths in the second half of 2011, there is no need for primer redesign with the ligation method, thus being more economical than fusion primers. Additionally, bidirectional sequencing is achieved without the need for amplicon-specific primers, as is the case with fusion primers.

## Confirmation of somatic mutations in pancreatic cancer research samples

The Ion PGM™ system was used to validate somatic mutations that were discovered using SOLiD™ System exome sequencing in a primary pancreatic adenocarcinoma research sample with fusion primers. First, fusion primers were designed to generate amplicons that contain somatic mutations, followed by whole genome amplification of tumor, normal, and adjacent tissue DNA. After PCR with fusion primers, amplicons were then pooled and sequenced using Ion 314™ Chips. As recommended in the *Ion Torrent Amplicon Sequencing Application Note*, greater than 1,000x depth of coverage was achieved for detection of somatic mutations. Table 2 shows 4 variants that were 100% concordant and confirmed on the Ion PGM™ sequencer. Allele frequency is also provided, giving high coverage and accuracy with detection of the somatic variants with greater than 1,000x coverage.

Table 2. Confirmed somatic mutations on the Ion PGM™ system.

Gene	Chromosome: position	Variant	Sample	Allele frequency		Total coverage
				Normal allele	Somatic allele	
GPR137B	chr1: 236343437	G>C	Tumor	81.8%	18.2%	6,359
			Normal	100.0%	0.0%	4,135
			Adjacent	100.0%	0.0%	1,514
IQCH	chr15: 67687770	G>C	Tumor	76.3%	23.7%	4,477
			Normal	99.9%	0.1%	3,770
			Adjacent	100.0%	0.0%	981
KLKB1	chr4: 187159516	A>G	Tumor	68.3%	31.7%	1,918
			Normal	99.8%	0.2%	3,011
			Adjacent	100.0%	0.0%	399
KRAS	chr12: 25398284	C>A	Tumor	76.1%	23.9%	3,866
			Normal	99.9%	0.1%	4,201
			Adjacent	100.0%	0.0%	2,136

Data courtesy of Sean Grimmond, Queensland Centre of Medical Genomics

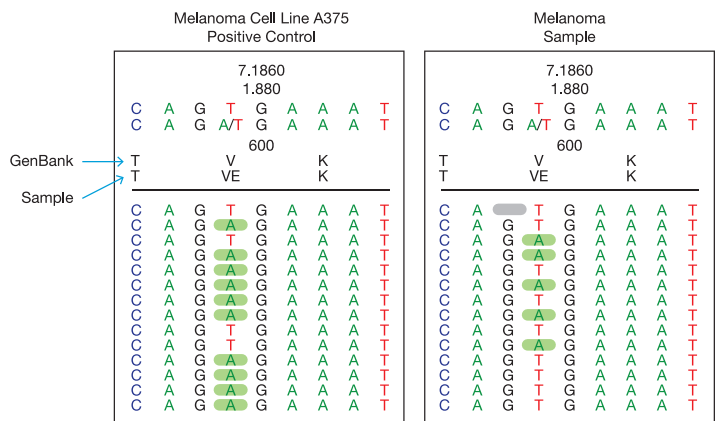
## Multiplexing with barcodes

Multiplexing allows multiple samples to be processed simultaneously, for instance on the same chip. Barcodes act as indexing tags to identify the data from each sample when multiplexing. Depending on throughput requirements, the Ion PGM™ system offers the Ion 314™, Ion 316™, and Ion 318™ Chips with 10 Mb, 100 Mb, and 1 Gb of sequencing capacity, respectively. Literally, the chip is the machine. This scalability allows for multiplexing numerous samples on a single Ion chip, either with the *Ion DNA Barcoding 1-16 Kit\** or with customer-designed barcodes. Both the amplicon size and desired depth of coverage need to be considered when determining how many samples can be multiplexed on the same chip. Figure 2 shows an overview of targeted resequencing methods and the number of samples that can be multiplexed on each Ion chip at 100x coverage with varying target region sizes.

## Barcoded amplicon resequencing of FFPE samples for somatic mutation detection

Barcodes were used to multiplex samples from formalin-fixed, paraffin-embedded (FFPE) melanoma research samples to detect BRAF and NRAS mutations. Fifty-three samples were analyzed from 6-year old FFPE tissue blocks (47 unique samples). After performing two rounds of PCR using barcoded primer sets, 16 samples per chip were sequenced on 4 Ion 314™ chips (14 samples and 2 controls, one A375 melanoma cell line and one normal FFPE tissue sample; 5 amplicons per sample). Mutations were found in 47% of samples: 10 BRAF (V600E,K) and 12 NRAS codon 61 (Q61R,K,H). Also, BRAF mutations were found in replicate tissue blocks, illustrating the reproducibility of the platform. The normal control sample showed 100% accuracy with no mutations in any amplicon across the four chip runs. The positive control sample A375 cell line confirmed the BRAF V600E mutation in all four runs (Figure 4).

Figure 4. Presence of V600E BRAF mutation in both positive control cell line and melanoma research samples.



Data courtesy of George S. Watts, PhD, Arizona Cancer Center

## Target Enrichment

When the genomic region of interest approaches 100 kb or greater, it is typically most efficient to enrich those target regions using a method different from amplicon resequencing. Whether capturing target regions by hybridization, or through parallelized PCR amplification of multiple regions, the scaling throughput of the Ion PGM™ chips can be leveraged based on the size of the target region and the coverage required for variant detection.

## TargetSeq™ Custom Enrichment Kits\*

TargetSeq™ Custom Enrichment Kits\* deliver an in-solution method for capturing target regions via probe hybridization. Just select your regions of interest by submitting a gene list or genomic coordinates (total target region 0.1–10 Mb), and specific probe pools will be designed using optimized algorithms for your targets. Biotinylated probe pools are then shipped for hybridization capture. After genomic DNA is fragmented and Ion adaptors ligated, biotinylated probes are hybridized (Figure 5). These target region fragments are then captured using streptavidin-coated magnetic beads. Finally, the captured DNA is eluted from the beads and is ready for subsequent template prep and sequencing on the Ion PGM™ system. The TargetSeq™ Custom Enrichment Kit workflow will support the enrichment of an indexed library of multiple samples.

Figure 5. Overview of TargetSeq™ Custom Enrichment workflow.



### Compatibility with SureSelect and Halo Genomics

Other in-solution hybridization capture technologies are also compatible with Ion Torrent, including SureSelect and Halo Genomics.

SureSelect is another hybridization-based target enrichment solution for target regions ranging from under 200 kb to the whole exome, with a similar workflow as described in Figure 5. The SureSelect product user guide currently only supports indexing or barcoding after the library has been enriched.

For regions totaling under 400 kb, Halo Genomics offers a highly specific method for targeting regions of interest by combining probe capture hybridization with highly specific PCR amplification (see *Halo Genomics Selector Technology™ Target Enrichment Protocol*). This method provides a high percentage of completeness of target and high specificity for capture of smaller target region panels, particularly important for disease research applications.

## Conclusions

No matter what the combination of parameters, the Ion PGM™ system delivers the optimal solution for fast and affordable targeted resequencing. Ion offers:

- The fastest workflow, with 2-hour sequencing run times and a single-day workflow for amplicon resequencing.
- The ability to sequence amplicons of any size for germline or somatic mutation detection with 10 Mb chips to 1 Gb\* chips.
- Barcodes for multiplexing samples at the depth of coverage needed.
- TargetSeq™ Custom Enrichment Kits, a high quality custom target enrichment solution, and full compatibility with other target enrichment vendors, such as Halo Genomics and SureSelect.
- Scalable current protocols to accommodate upcoming improvements in longer readlengths.

## Ordering Information

Description	Item	Part No.
<b>Ion Torrent™ Systems</b>	Ion PGM™ System—Includes PGM™ Sequencer [4462917] and Torrent Server [4462918]	4462921
	Ion OneTouch™ System	4470001
<b>Service Contracts</b>	Rapid Exchange Program	ZGEXSCIONPGMSYS
	AB Assurance Plan	ZG11SCIONPGMSYS
<b>Semiconductor Sequencing Chips</b>	Ion 314™ Chip Kit (8 Chips)	4462923
	Ion 316™ Chip Kit (4 Chips)	4466616
	Ion 316™ Chip Kit (8 Chips)	4469496
<b>Reagent Kits</b>	Ion Xpress™ Fragment Library Kit (up to 20 reactions)	4468987
	Ion Fragment Library Kit (up to 20 reactions)	4466464
	Ion DNA Barcoding 1-16 Kit* (10 sets of 16 libraries)	4468654
	Ion Xpress™ Template Kit (10 reactions)	4469001
	Ion Sequencing Kit (8 reactions)	4468997
<b>Ion Control Kits</b>	Ion Control Materials Kit (3 reactions each per control)	4466465
	Ion Sphere™ Quality Control Kit (20 measurements)	4468656
	Ion Library Quantitation Kit (250 reactions)	4468802
<b>Invitrogen PCR Kit</b>	SequalPrep™ Long PCR Kit (1000 units)	A10498
<b>AB Cycle Sequencing Kit</b>	BigDye® Direct Cycle Sequencing Kit (100 reactions)	4458687

\* The content provided herein may relate to products that have not been officially released and is subject to change without notice.

**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.

Ion Torrent by Life Technologies | 7000 Shoreline Court | Suite 201 | South San Francisco, CA 94080 USA  
Phone +1-203-458-8552 | Toll Free in North America 1-877-SEQUENCE (1-877-378-3623)

www.lifetechnologies.com | www.iontorrent.com | http://ioncommunity.iontorrent.com

CO19872 0711