Application Note

Simplified, high-performance somatic variant analysis

Authors

Nitya Furtado
Senior Scientist – Reagent and Assay Development

Florence Crary-Dooley
Senior Scientist – Reagent and Assay Development

Liu Xi
Principal Bioinformatics Scientist II

Mingyang Cai
Principal Bioinformatics Scientist II

Junyan Lin
Senior Scientist – Reagent and Assay Development

Jingchuan Li
Principal Scientist II – Reagent and Assay Development

Brian Godwin
Director – Reagent and Assay Development

Markos Mihalatos
Senior International Product Manager

Roche Diagnostics Solutions, Pleasanton, CA, USA

Mariana Fitarelli-Kiehl
Senior Applications Scientist – Support & Applications

Spencer Debenport
Applications Manager

Roche Diagnostics Corp. US, Wilmington, MA, USA

Liesl Joubert
Senior Scientist – Sequencing Assay Development

Marius Snyman
Senior Scientist – Sequencing Assay Development

Roche Diagnostics, Cape Town, South Africa

The KAPA HyperPETE Workflow with pre-capture multiplexing simplifies high-performance somatic variant analysis

The KAPA HyperPETE Workflow offers a novel, comprehensive, end-to-end approach for high-performance somatic variant analysis from plasma cfDNA, as well as tissue (fresh frozen or FFPE) and cell line DNA and RNA. Pre-capture multiplexing simplifies the overall workflow, thereby improving turnaround times, throughput, and cost—without an impact on performance.

Introduction

Research has demonstrated strong correlations between somatic (acquired) genetic variants and cancer progression/response to therapeutic intervention. Every cancer is characterized by a unique combination of variants, most notably single nucleotide variants (SNV) and short insertions/deletions (indels), microsatellite instability (MSI), and large structural rearrangements (gene fusions and/or gene deletions/duplications that cause copy number variation).

Targeted next-generation sequencing (NGS) is an established and powerful tool for studying the genetic basis of cancer, offering the ability to survey large numbers of targets (for potential variants) in parallel, and at high sensitivity (variant allele frequencies, VAF of 5% or less).[1] NGS pipelines for the analysis of somatic variants require three critical components: (i) efficient methods for the preparation of enriched libraries from challenging sample types, including cell free/circulating tumor DNA (cf/ctDNA) and DNA or RNA from formalin-fixed paraffin-embedded (FFPE) tissue samples; (ii) relevant, optimized enrichment panels (content), and (iii) advanced bioinformatic tools.

We have previously introduced the KAPA HyperPETE Workflow,[2,3] which meets all of the above criteria, and offers a streamlined, versatile, end-to-end solution for high-performance somatic variant analysis of all major mutation classes using small (catalog or custom, <300 kb capture target) panels. Based on Roche's novel Primer Extension Target Enrichment (PETE) technology (Figure 1), the KAPA HyperPETE Workflow offers the speed and ease-of-use of amplicon-based target enrichment with the high performance of hybridization-based capture.

Pre-capture multiplexing is an established approach in capture-based target enrichment methods to simplify sample handling, reduce turnaround times, and decrease per-sample reagent usage; but can impact enrichment performance, especially when working with challenging oncology samples and targeting low-abundance variants (VAF <5%).

In this Application Note, we demonstrate efficient pre-capture multiplexing in the KAPA HyperPETE Plasma cfDNA, Tissue DNA, and Tissue RNA enrichment workflows. Sequencing performance metrics generated with oncology samples were comparable for the standard single-plex workflow and the pre-capture 8-plex workflow described here.

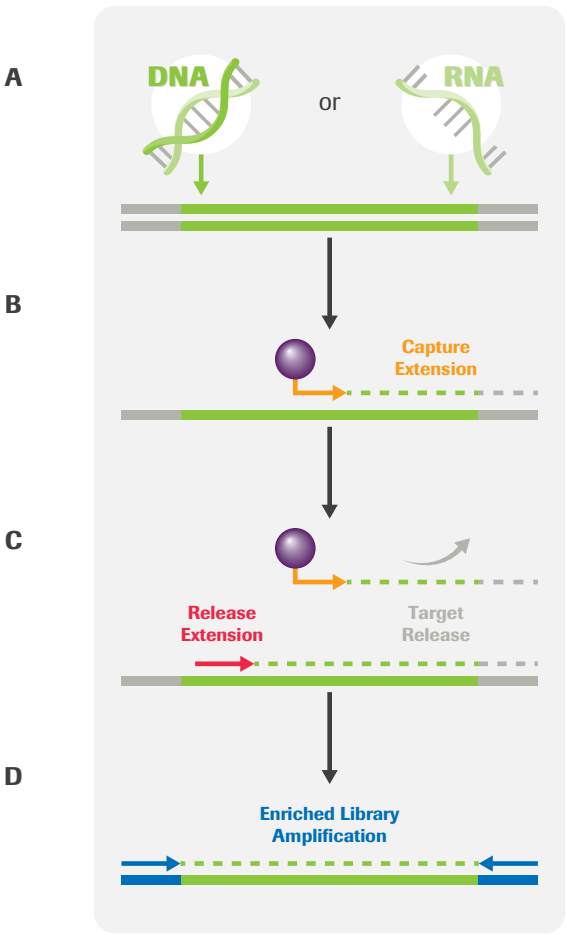

Figure 1. Overview of the KAPA HyperPETE (Primer Extension Target Enrichment) Technology. (A) Uniquely dual-indexed pre-capture libraries are generated from genomic DNA, FFPE DNA, cfDNA, or RNA using the appropriate (KAPA HyperPlus, KAPA HyperPrep, or KAPA RNA HyperPrep) library preparation kit, KAPA Universal Adapters (with or without UMI), and KAPA UDI Primer Mixes. **(B)** Following a heated denaturation step, biotinylated target-specific capture primers (orange) are annealed and extended using a DNA polymerase. Paramagnetic streptavidin beads (purple) are used to capture the hybrid molecules. Off-target, uncaptured library fragments are washed away (not shown). **(C)** Target-specific release primers (red) are hybridized and extended by a DNA polymerase, thereby releasing target molecules from the beads into the supernatant. Note that an intermediate wash step to remove non-hybridized primers is performed prior to the release extension. **(D)** Released targets are amplified with universal primers (blue) to generate final libraries that are purified, quantified, pooled, and prepared for sequencing. The workflow may be completed in 8 – 10 hours, depending on sample type and the number of samples that are processed.

Materials and methods

Experimental design

This study was designed to demonstrate the performance of the KAPA HyperPETE Workflow for somatic variant analysis with pre-capture multiplexing of four or eight libraries, as compared to the standard single-plex enrichment protocol. To this end, libraries were prepared from (i) cell line and plasma derived cfDNA, (ii) cell-line, tissue and xenograft-derived FFPE DNA, and (iii) cell-line and tissue derived FFPE RNA. Non-FFPE cell line controls were included in the Somatic Tissue DNA workflow.

Fragment libraries were prepared with the KAPA HyperPrep Kit (Plasma cfDNA Workflow), KAPA HyperPlus Kit (Tissue DNA Workflow), or KAPA RNA HyperPrep Kit (Tissue RNA Workflow). Primer Extension Target Enrichment (PETE) was performed with catalog KAPA HyperPETE Panels (capture sizes ranging from 18 – 301 kb) according to the standard workflow for each sample type and panel—with minor modifications for pre-capture multiplexing (see Appendix). Sequencing was performed on an Illumina® NextSeq® 500 instrument using standard protocols. Data analysis was performed with a Roche internal instance of the NAVIFY® Mutation Caller pipeline. Key sequencing and variant calling metrics are reported for each workflow. A summary of the experimental design is given in Figure 2.

Samples and DNA Extraction

Plasma cfDNA Workflow: Four reference materials (purified cell line ctDNA or ctDNA mixtures; Table 1, lines 1 – 4) and four plasma samples from healthy donors (Table 1, line 5) were obtained from a commercial supplier. For the KAPA HyperPETE Somatic Plasma cfDNA Workflow, cfDNA may be extracted with any appropriate method from blood collected in EDTA-containing collection tubes and handled according to standard procedures for plasma samples. For this study, cfDNA was extracted using the AVENIO® cfDNA Isolation Kit (Roche PN: 08036586190).

	Plasma cfDNA Workflow		Tissue DNA Workflow			Tissue RNA Workflow	
Inputs	Reference cell line ctDNA (n=4)	cfDNA from plasma (n=4)	Reference cell line DNA (FFPE, n=2) (non-FFPE, n=8)	FFPE DNA from xenografts (n=4)	FFPE DNA from surgical biopsies (n=2)	Reference cell line RNA (FFPE, n=3)	FFPE RNA from surgical biopsies (n=5)
Library Prep	**KAPA HyperPETE Somatic Plasma cfDNA Workflow (v1.0)[4]** KAPA HyperPrep Kit (10 or 50 ng input)		**KAPA HyperPETE Somatic Tissue DNA Workflow (v1.0)[5]** KAPA HyperPlus Kit (Q score-adjusted input for FFPE samples, 10 ng for non-FFPE samples)			**KAPA HyperPETE Tissue RNA Fusion Transcript Workflow (v1.0)[6]** KAPA RNA HyperPrep Kit (10 ng input)	
	1-plex, 4-plex or 8-plex pre-capture multiplex pooling		1-plex or 8-plex pre-capture multiplex pooling				
Target Enrichment & Sequencing	**KAPA HyperPETE Hot Spot Panel** 166 capture targets, 37 genes, 37 kb **13.5 M** paired-end reads		**KAPA HyperPETE Pan Cancer Panel** 1321 capture targets, 96 genes, 190 MSI loci, 301 kb **55 M** paired-end reads			**KAPA HyperPETE Lung Cancer Fusion Panel** 192 capture targets, 21 genes (17 lung cancer research, 4 housekeeping), 18 kb **1.8 M** paired-end reads	
QC metrics	% reads on target • % bases in 2-fold range % panel exon region ≥1000X, ≥3000X		% reads on target • % bases in 2-fold range % panel exon region ≥300X, ≥500X			% reads on target for: entire panel • fusion genes	
Variants	SNV • Indels		SNV • Indels • CNVs • MSI			Fusions	

Figure 2. Summary of the experimental design for this study. Full details of the inputs for each of the three workflows are provided in *Materials and methods*. Uniquely dual-indexed libraries were prepared as described in the Instructions for Use for each workflow. "Q score" refers to the normalized quality score for FFPE DNA samples, determined with a qPCR-based assay as described in the KAPA HyperPETE Somatic Tissue DNA Workflow. Quality control (QC) assays (quantification, library fragment size assessment) performed on pre- and post-enrichment libraries are not shown in the diagram, but were performed as described in *Materials and methods*. Single (1-plex), 4-plex, and/or 8-plex captures were performed as indicated. Sequencing was performed on an Illumina NextSeq 500 instrument. Data were randomly downsampled as indicated for each panel and analyzed with a Roche internal instance of the NAVIFY Mutation Caller pipeline. Full definitions of QC metrics are given in *Results and discussion*.

Data on file.

Table 1. Samples used in the Plasma cfDNA Workflow

	Sample name/ part number	Sample description	Geno-type	Sample type	Sample format	Source/reference
1	0710-0143	Seraseq® ctDNA Mutation Mix v2 AF0.125%*	Mut	Reference (cell line)	Purified ctDNA mix	https://www.seracare.com/Seraseq-ctDNA-Mutation-Mix-v2-AF0125-0710-0143/
2	0710-0141	Seraseq ctDNA Mutation Mix v2 AF0.5%	Mut	Reference (cell line)	Purified ctDNA mix	https://www.seracare.com/Seraseq-ctDNA-Mutation-Mix-v2-AF05-0710-0141/
3	0710-0140	Seraseq ctDNA Mutation Mix v2 AF1%	Mut	Reference (cell line)	Purified ctDNA mix	https://www.seracare.com/Seraseq-ctDNA-Mutation-Mix-v2-AF1-0710-0140/
4	0710-0144	Seraseq ctDNA Mutation Mix v2 WT	WT	Reference (cell line)	Purified ctDNA	https://www.seracare.com/Seraseq-ctDNA-Mutation-Mix-v2-WT-0710-0144/
5	Various (n=4)	Samples from healthy donors	WT	Plasma	Plasma samples; cfDNA previously extracted for this study	BioCollections Worldwide, Inc.

*Libraries prepared from the AF0.125% reference sample were only used to calculate the enrichment and sequencing metrics but not the variant detection metrics. The KAPA HyperPETE Plasma cfDNA Workflow has only been validated with up to 0.5% AF. Additional sequencing is needed for 0.125% AF.

Extracted cfDNA was quantified using a Qubit® Fluorometer and Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific). To confirm that cfDNA extracts were free of high-molecular weight genomic DNA, samples were subjected to electrophoretic analysis using a 4200 TapeStation system and Cell-free DNA ScreenTape Assay or 2100 Bioanalyzer system and High Sensitivity DNA Kit (Agilent® Technologies).

Tissue DNA Workflow: Ten reference samples from characterized cell lines were obtained from commercial suppliers or internal sources (Table 2, lines 1 – 10). These included purified genomic DNA (lines 7 – 10) or mixes of purified DNA (lines 3 – 6) from unprocessed cells (non-FFPE samples), as well as FFPE curls prepared from cell cultures (lines 1 – 2). Four FFPE xenograft samples (lines 11 – 14) were obtained from a commercial supplier, and two surgical tissue samples (line 15) from internal sources.

DNA was extracted from FFPE curls (Table 2, lines 1 – 2 and 11 – 15; one/two sections of up to 20 µm per extraction) with the KAPA NGS DNA Extraction Kit (Roche PN: 09189823001 or 09190023001) as described in the KAPA HyperPETE Somatic Tissue DNA Workflow; Instructions for use (Version 1.0).[5]

All DNA samples (FFPE and non-FFPE) were quantified using a Qubit® Fluorometer and Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific). The quality of all formalin-compromised DNA samples (Table 2, lines 1 – 2 and 11 – 15) was assessed with the qPCR-based KAPA NGS FFPE DNA QC Kit (Roche PN: 09217193001 or 09217207001) as described in the KAPA HyperPETE Somatic Tissue DNA Workflow; Instructions for use (Version 1.0).[5] DNA extracts were divided into high, mid or low quality based on the normalized Q (quality) score calculated for each sample.

Tissue RNA Workflow: Three reference samples (FFPE curls prepared from cultures of characterized cell lines) were obtained from commercial suppliers (Table 3, lines 1 – 3). Five surgical tissue samples (FFPE curls; Table 3, line 4) were obtained internally.

Total RNA was extracted from FFPE curls (one/two sections per extraction) with the High Pure FFPET RNA Isolation Kit (Roche PN: 06650775001).[7] All RNA samples were quantified using a Qubit Fluorometer and Qubit RNA HS Assay Kit (Thermo Fisher Scientific). RNA quality was assessed with a 2100 Bioanalyzer system and RNA 6000 Pico Kit (Agilent Technologies). Extracts were divided into high, mid, or low quality based on the calculated DV_{200} value (percentage of RNA fragments >200 nt) for each sample.[8]

Library Preparation and Target Enrichment

Plasma cfDNA Workflow: A total of 32 cfDNA libraries were prepared from eight unique samples (10 ng or 50 ng input, n=2 per input), as outlined in Table 4.

All cfDNA libraries were prepared with the KAPA HyperPrep Kit (Roche PN: 07962312001, 07962347001, or 07962363001), KAPA Universal UMI Adapter (Roche PN: 09329862001 or 09329889001), KAPA UDI Primer Mixes 1 – 96 (Roche PN: 09134336001), and KAPA HyperPure Beads (Roche PN: 08963835001, 08963843001, 08963851001, 08963878001, or 08963860001) as described in the KAPA HyperPETE Somatic Plasma cfDNA Workflow; Instructions for use (Version 1.0).[4]

Tissue DNA Workflow: A total of 32 genomic DNA libraries were prepared from sixteen unique samples (Q score-adjusted input for FFPE samples or 10 ng input for non-FFPE samples; n=2 per input), as outlined in Table 5. "Q score" refers to the normalized quality score (Q) determined with the qPCR-based KAPA NGS FFPE DNA QC Kit, as described in the KAPA HyperPETE Somatic Tissue DNA Workflow; Instructions for Use (Version 1.0).[5] The Q score-adjusted input ranged from 88.7 – 97.7 ng for low-quality (LQ) FFPE DNA extracts, from 39.8 – 46.2 ng for mid-quality (MQ) extracts, and was equal to 20 ng or 10 ng for high-quality (HQ) FFPE and non-FFPE DNA, respectively.

Libraries were prepared with the KAPA HyperPlus Kit (Roche PN: 07962380001, 07962401001, or 07962428001), KAPA Universal UMI Adapter (Roche PN: 09329862001 or 09329889001), KAPA UDI Primer Mixes 1 – 96 (Roche PN: 09134336001), and KAPA HyperPure Beads (Roche PN: 08963835001, 08963843001, 08963851001, 08963878001, or 08963860001) as described in the KAPA HyperPETE Somatic Tissue DNA Preparation Workflow; Instructions for use (Version 1.0).[5] Formalin-compromised samples (Table 2, lines 1 – 2 and lines 11 – 15) were subjected

Table 2. Samples used in the Tissue DNA Workflow

	Sample name/ part number	Sample description	Geno-type	Sample type	Sample format	Source/reference
1	HD200	Quantitative Multiplex Reference Standard (FFPE)	Mut	Reference (cell line)	FFPE curls (10 – 15 µm), DNA extracted previously for this study	https://horizondiscovery.com/en/reference-standards/products/quantitative-multiplex-reference-standard-ffpe?catalognumber=HD200
2	HD789	Structural Multiplex Reference Standard (FFPE)	Mut	Reference (cell line)		https://horizondiscovery.com/en/reference-standards/products/structural-multiplex-reference-standard-ffpe
3	Custom part no.	Custom gDNA Mutation Mix (Custom Sample 1)	Mut	Reference (cell line)	Custom purified DNA mix (non-FFPE)	LGC SeraCare
4	Custom part no.	Custom gDNA Mutation Mix, (Custom Sample 3)	Mut	Reference (cell line)		LGC SeraCare
5	Custom part no.	Custom gDNA Mutation Mix, (Custom Sample 4)	Mut	Reference (cell line)		LGC SeraCare
6	0710-0412	Seraseq® Breast CNV Mix, +6 copies	Mut	Reference (cell line)	Purified DNA mix (non-FFPE)	https://www.seracare.com/Seraseq-Breast-CNV-Mix--6-copies-0710-0412/
7	MOLT-4	MSI cell line DNA	Mut	Reference (cell line)	Sequence-verified, purified DNA (non-FFPE)	Internal
8	DLD-1	MSI cell line DNA	Mut	Reference (cell line)		Internal
9	A-172	MSS cell line DNA	Mut	Reference (cell line)		Internal
10	A549	MSS cell line DNA	Mut	Reference (cell line)		Internal
11	CT116	Xenograft, MSI	Mut	Xenograft	FFPE curls (10 – 15 µm) DNA extracted previously for this study	CrownBio
12	SW48	Xenograft, MSI	Mut	Xenograft		CrownBio
13	BT474	Xenograft, MSS	Mut	Xenograft		CrownBio
14	MDA-MB-453	Xenograft, MSS	Mut	Xenograft		CrownBio
15	Various (*n*=2)	Surgical biopsy	WT	Tissue		Internal

Table 3. Samples used in the Tissue RNA Workflow

	Sample name	Sample description	Geno-type	Sample type	Sample format	Source/reference
1	HD784	ALK RET ROS RNA fusion	Mut	Reference (cell line)	FFPE curls (10 – 15 µm), RNA extracted in this study	https://horizondiscovery.com/en/reference-standards/products/alk-ret-ros-rna-fusion
2	0710-0496	Seraseq® FFPE Tumor Fusion RNA v4 Reference Material	Mut	Reference (cell line)	FFPE curls (10 – 15 µm), RNA extracted in this study	https://www.seracare.com/Seraseq-FFPE-Fusion-RNA-RM-v4-0710-0496/
3	0710-1031	Seraseq FFPE NTRK Fusion RNA Reference Material	Mut	Reference (cell line)	FFPE curls (10 – 15 µm), RNA extracted in this study	https://www.seracare.com/Seraseq-FFPE-NTRK-Fusion-RNA-Reference-Material-0710-1031/
4	Various (*n*=5)	Surgical biopsy, normal	WT	Tissue	FFPE curls (10 µm), RNA extracted in this study	Internal

to DNA Polishing with the KAPA NGS FFPE DNA Polishing Kit (Roche PN: 09217215001 or 09217223001) prior to enzymatic fragmentation.

Tissue RNA Workflow: A total of 16 RNA libraries were prepared from eight unique samples (10 ng input, *n*=2), as outlined in Table 5. RNA libraries were prepared with the KAPA RNA HyperPrep Kit (Roche PN: 08098093702 or 08098107702), KAPA Universal UMI Adapter (Roche PN: 09329862001 or 09329889001), KAPA UDI Primer Mixes 1 – 96 (Roche PN: 09134336001), and KAPA Pure Beads (included in the KAPA RNA HyperPrep Kit) as described in the KAPA HyperPETE Tissue RNA Fusion Transcript Workflow; Instructions for use (Version 1.0).[6]

Pre-capture Library QC: Amplified pre-capture libraries were diluted 1:40 for analysis of fragment size distribution, performed with a 4200 TapeStation system and DNA High Sensitivity D1000 ScreenTape Assay (Agilent® Technologies). The same diluted material was used for library quantification using a Qubit® Fluorometer and Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific).

Data on file.
For Research Use Only. Not for use in diagnostic procedures.

Table 4. Breakdown of libraries prepared with the KAPA HyperPETE Somatic Plasma cfDNA Workflow

Panel	Capture strategy	cfDNA samples (see Table 1)					
		Reference cfDNA (10 ng input)	Plasma cfDNA (10 ng input)	Reference cfDNA (50 ng input)	Plasma cfDNA (50 ng input)	Total libraries	Total captures
KAPA HyperPETE Hot Spot Panel	1-plex	4 (rep 1)	4 (rep 1)	4 (rep 1)	4 (rep 1)	16	16 x 1-plex
	4-plex	4 (rep 2)	4 (rep 2)	4 (rep 2)	4 (rep 2)	16	4 x 4-plex
	8-plex	4 (rep 1)	4 (rep 1)	4 (rep 1)	4 (rep 1)	Same 16 as for 1-plexes	2 x 8-plex
Total (unique libraries)		**8**	**8**	**8**	**8**	**32**	**22**

Table 5. Breakdown of libraries prepared with the KAPA HyperPETE Somatic Tissue DNA Workflow

Panel	Capture strategy	DNA samples (see Table 2)					
		LQ FFPE (Q-adjusted input)	MQ FFPE (Q-adjusted input)	HQ FFPE (20 ng input)	Non-FFPE (10 ng input)	Total libraries	Total captures
KAPA HyperPETE Pan Cancer Panel	1-plex	3 (rep 1)	3 (rep 1)	2 (rep 1)	8 (rep 1)	16	16 x 1-plex
	8-plex	3 (rep 2)	3 (rep 2)	2 (rep 2)	8 (rep 2)	16	2 x 8-plex
Total (unique libraries)		**6**	**6**	**4**	**16**	**32**	**18**

Table 6. Breakdown of libraries prepared with the KAPA HyperPETE Tissue RNA Fusion Transcript Workflow

Panel	Capture strategy	RNA samples, 10 ng input (see Table 3)				
		Seraseq FFPE	Horizon FFPE	Normal FFPE	Total libraries	Total captures
KAPA HyperPETE Lung Cancer Fusion Panel	1-plex	2 (rep 1)	1 (rep 1)	5 (rep 1)	8	8 x 1-plex
	8-plex	2 (rep 2)	1 (rep 2)	5 (rep 2)	8	1 x 8-plex
Total (unique libraries)		**4**	**2**	**10**	**16**	**18**

Target enrichment: Primer Extension Target Enrichment (PETE) was performed with the KAPA HyperPETE Reagent Kit (Roche PN: 09211624001 or 09211683001) and KAPA HyperCapture Bead Kit (Roche PN: 09075780001 or 09075798001). Single-plex captures were performed as described in the standard protocol for the Plasma cfDNA, Tissue DNA, and Tissue RNA Fusion Transcript Workflows, respectively.[4-6] For 4- or 8-plex captures, the amount of library DNA per capture was adjusted to 1 µg each (i.e., a total of 4 µg for 4-plex pools, and 8 µg for 8-plex pools) and a bead-based concentration step was included to reduce the volume of the pool to 15 µL for addition to the Capture Extension Master Mix. All modifications needed to adapt the standard protocol for multiplexed capture are described in detail in the Appendix.

The number of captures performed with each panel is outlined in Tables 4 – 6. Details of the three the KAPA HyperPETE Panels used in this study are summarized in Table 7.

Post-capture (PETE) Library QC: The concentration and fragment size distribution of sequencing-ready, post-capture (enriched) libraries (from single-plex captures) or library pools (from multiplexed captures) were determined with an Agilent 4200 TapeStation system and DNA High Sensitivity D1000 ScreenTape Assay (Agilent Technologies).

Sequencing and Data Analysis

Sequencing: Libraries were pooled for multiplexed, paired-end Sequencing (2 x151 bp) on an Illumina® NextSeq® 500 system. Pools were configured to meet the paired-end read requirement for each panel, as outlined in Table 7.

Data analysis: Secondary data analysis was performed with a Roche internal instance of the NAVIFY® Mutation Caller pipeline, verified to routinely generate equivalent results to the NAVIFY Mutation Caller (data on file). Analysis included demultiplexing,

Table 7. KAPA HyperPETE Catalog Panels used in this study

Workflow	Panel name	Capture target	Details	Paired-end (PE) reads required per library
Plasma cfDNA	KAPA HyperPETE Hot Spot Panel	37 kb	▪ 166 capture targets across 37 genes from the genomic 'hot spot' regions that are frequently mutated in human cancer genes ▪ Panel is optimized to deliver high uniformity and specificity from low-quality, low-input cfDNA and FFPE DNA samples	15 M*
Tissue DNA	KAPA HyperPETE Pan Cancer Panel	301 kb	▪ 1321 capture targets, covering 86 cancer-related genes and 190 MSI loci which have been found relevant to somatic oncology research applications ▪ Panel is optimized to deliver high uniformity and specificity from low-quality, low-input cfDNA and FFPE DNA samples	55 M
Tissue RNA	KAPA HyperPETE Lung Cancer Fusion Panel	18 kb	▪ Designed for use with the novel PETE technology to detect known/unknown fusions and fusion partners ▪ Covers both 5'- and 3'- ends of 144 exons across 17 oncogenes related to lung cancer research (*ALK*, *AXL*, *BRAF*, *EGFR*, *FGFR1*, *FGFR2*, *FGFR3*, *MET*, *NRG1*, *NTRK1*, *NTRK2*, *NTRK3*, *PDGFRA*, *PDGFRB*, *PPARG*, *RET*, *ROS1*) ▪ Includes 4 housekeeping genes (*RAB7A*, *VCP*, *CHMP2A*, *HPRT1*) as positive controls	1.8 M

*The recommended amount of sequencing for the KAPA HyperPETE Hot Spot Panel in the Plasma cfDNA Workflow is 15 M paired-end reads. However, only 13.5 M reads were obtained for one of the samples in this study. To ensure consistency and accurate comparisons, data for all samples captured with the Hot Spot Panel were therefore downsampled to 13.5 M reads.

adapter trimming, and alignment; as well as barcode and position deduplication, and the detection of single nucleotide variants (SNV), short insertions and deletions (indels), RNA fusions, microsatellite instability (MSI) and copy number variation (CNV). The performance of the KAPA HyperPETE technology across different sample types, inputs, and panels was evaluated by assessing (i) key sequencing QC metrics for each workflow, and (ii) variant calling results for characterized reference samples. Plots (Figures 3 – 5) were generated with JMP®, Version 16, SAS Institute Inc., Cary, NC, 1989–2021.

Results and discussion

Library QC metrics

All of the pre-capture libraries prepared with each of the three workflows met the criteria for target enrichment, and all post-capture (enriched) libraries/library pools met the criteria for sequencing (data not shown; refer to Instructions for Use for acceptable metrics).

Sequencing performance metrics

The NAVIFY® Mutation Caller pipeline and the internal instance used in this study generate an extensive list of sample and analysis QC metrics for every run. A subset of metrics are reported for each of the three workflows. These are listed and defined in Table 8.

Plasma cfDNA Workflow: Results obtained with KAPA HyperPETE Hot Spot Panels for libraries prepared from different inputs (10 ng and 50 ng) and sample types (plasma and cell line cfDNA), are shown in Figure 3. On-target rates were high and consistent across the 1-plex (66.8% – 79.6%), 4-plex (68.4% – 79.3%), and 8-plex (66.4% – 80.6%) captures, irrespective of the input into library construction.

Similarly, coverage uniformity (% bases in 2-fold range) was high and comparable across multiplexing level (87.1% – 98.8% for 1-plexes, 86.8% – 98.9% for 4-plexes, and 86.2% – 98.6% for 8-plexes), with slightly lower uniformity for 50 ng inputs. This was expected from the available number of reads, which is insufficient for the number of genome equivalents in 50 ng inputs. Deeper sequencing will result in higher uniformity at high inputs.

Table 8. Sequencing QC metrics reported in this study

Metric	Explanation/definition	Workflow details
% reads on target	▪ Also referred to as on-target rate ▪ Defined as the percentage reads aligning to the capture panel's BED coordinate marked areas	▪ Reported for Plasma cfDNA and Tissue DNA Workflows ▪ For Tissue RNA Workflow, on-target rates for entire panel, and the fusion genes in the design are reported separately
% bases in 2-fold range	▪ Measure of coverage uniformity ▪ % bases covered at a depth ranging between 0.5x and 2x of the median coverage	▪ Reported for Plasma cfDNA and Tissue DNA Workflows
% Panel exon region ≥*d*X	▪ % of panel positions with a coverage depth of at least *d*	▪ Plasma cfDNA Workflow: *d*=1000X or 3000X ▪ Tissue DNA Workflow: *d*=300X or 500X

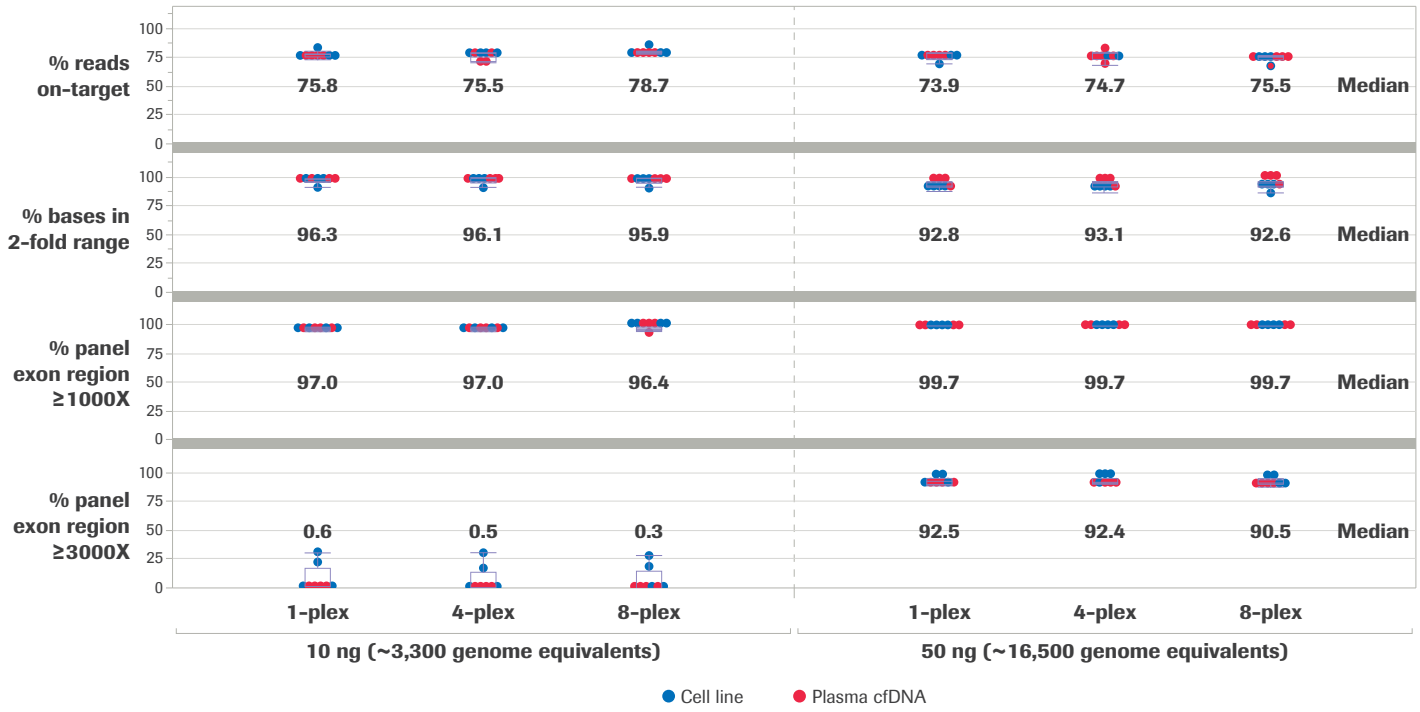

Figure 3. Key sequencing performance metrics for the Plasma cfDNA Workflow (37 kb KAPA HyperPETE Hot Spot Panel, 13.5 M paired-end reads per library). Libraries were prepared from plasma (red) or cell line (blue) samples, enriched, and sequenced as outlined in *Materials and methods*. Data were analyzed as described. The number of libraries prepared and analyzed from each input/sample type is summarized in Table 4.

Coverage of exonic regions at a depth ≥1000X was also high and reproducible across sample types, input, and multiplexing level (95.0% – 99.8% for 1-plexes, 95.3% – 99.8% for 4-plexes, and 93.9% – 99.7% for 8-plexes), with slightly better performance for 50 ng input libraries. At a depth ≥3000X, coverage of exonic regions dropped significantly (0% – 30.7% for 10 ng libraries and 87.7% – 95.1% for 50 ng input) across all multiplexing levels. This was expected due to the limited number of genome equivalents (~3300 for 10 ng inputs). Higher read depth is expected to improve performance, especially at the 50 ng input level.

Tissue DNA Workflow: DNA extracted from FFPE samples were divided into three quality groups, based on the quality score (Q) determined with the KAPA NGS FFPE DNA QC Kit: high-quality (HQ) samples had a Q score ≥0.5, whereas 0.22 ≤ Q <0.5 applied to mid-quality (MQ) samples, and Q ranged between 0.042 and 0.200 for low-quality (LQ) samples.

Results obtained with the Tissue DNA Workflow (summarized in Figure 4) confirmed that the KAPA HyperPETE technology yields high and uniform sequencing coverage across designs, sample types and quality, and pre-capture multiplexing levels—thereby supporting high-confidence variant calling (see below) from the recommended amount of paired-end reads for each panel:

- On-target rates ranged from 73.9% – 84.6% for 1-plexes and 77.3% – 85.6% for 8-plexes.
- Coverage uniformity (% bases in 2-fold range) ranged from 73.2% – 95.2% for 1-plexes and 72.8% – 95.0% for 8-plexes.
- Coverage of exonic regions at a depth ≥300X ranged from 95.9% – 99.4% for 1-plexes and 95.6% – 99.4% for 8-plexes. At a depth of ≥500X, coverage dropped only slightly, ranging from 87.8% – 99.1% for 1-plexes and 87.3% – 98.9% for 8-plexes.

Tissue FFPE RNA Workflow: Like FFPE DNA, RNA extracts from formalin-compromised samples were divided into three quality groups, based on the degree of RNA fragmentation (DV_{200} score): high-quality (HQ) samples had a DV_{200} ≥70%; whereas 50% ≤DV_{200} <70% applied to mid-quality (MQ) samples, and 30% ≤DV_{200} <50% to low-quality (LQ) samples.

Results obtained with small KAPA HyperPETE Lung Cancer Fusion Panel for libraries prepared from 10 ng RNA inputs are shown in Figure 5. On-target rates for the entire panel were high and consistent across 1-plex (92.5% – 97.2%) and 8-plex (92.0% – 97.3%) captures, irrespective of RNA quality.

On-target rates for the fusion genes in the design (on-target rate for entire panel *minus* on-target rate for housekeeping genes) ranged from 49.9% – 67.5% and 48.7% – 66.4% for 1-plex and 8-plex captures, respectively. On-target rates for fusion genes are expected to be more variable than for housekeeping genes due to intrinsic differences between the expression levels of fusion genes. All on-target rates were calculated after removal of reads mapping to rRNA (2.0% – 10.2% across the 16 libraries).

Variant calling results

Characterized reference materials with validated mutations and allele frequencies were included in all workflows in order to assess variant calling performance for different somatic variant classes.

Plasma cfDNA Workflow: Variant calling results for cfDNA samples are given in Table 9. Short variants (SNV and indels) were detected in reference cell line samples with high true positive rates, across all levels of pre-capture multiplexing. With 50 ng library inputs, 100% true positive rates were achieved for an expected AF of 1% or 0.5%, across 1-plex, 4-plex and 8-plex captures.

Figure 4. Key sequencing performance metrics for the Tissue DNA Workflow (301 kb KAPA HyperPETE Pan Cancer Panel, 55 M paired-end reads per library). Libraries were prepared from low (blue), medium (red), or high (green) quality samples, enriched, and sequenced as outlined in *Materials and methods*. Data were analyzed as described. The number of libraries prepared and analyzed from each input/sample type is summarized in Table 5.

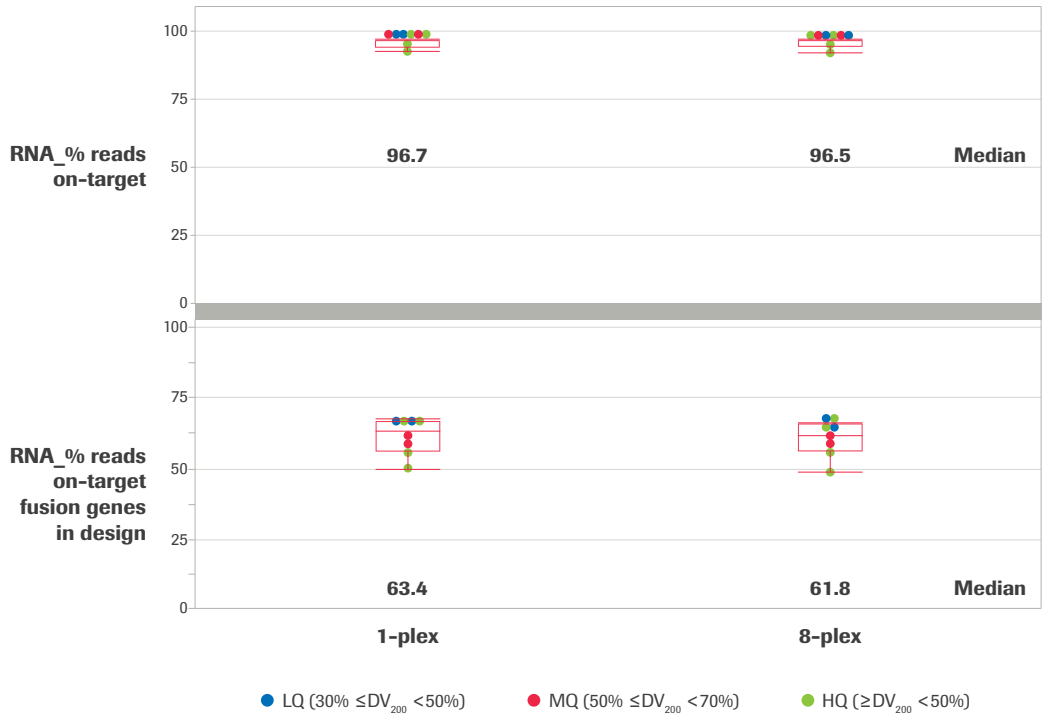

Figure 5. Key sequencing performance metrics for the Tissue RNA Workflow (18 kb KAPA HyperPETE Lung Cancer Fusion Panel, 1.8 M paired-end reads per library). Libraries were prepared from low (blue), medium (red), or high (green) quality samples, enriched, and sequenced as outlined in *Materials and methods*. Data were analyzed as described. The number of libraries prepared and analyzed from each input/sample type is summarized in Table 6.

Data on file.
For Research Use Only. Not for use in diagnostic procedures.

With 10 ng library inputs:

- for an expected AF of 1%, one indel in one of the two 4-plex captures and one SNV in the 8-plex captures were not detected
- for an expected AF of 0.5%, a slightly lower true positive rate was obtained for 4-plex captures, as a result of missing one more variant as compared to single- and 8-plexes.

All missed variants were reported in the variant call format (VCF) file, but were labeled as "fail" since the number of supporting reads for the variant allele was lower than the fixed default threshold set in the variant calling pipeline. Slightly worse performance with 4-plex pools was not regarded as significant; metrics for 4-plexes are anticipated to be comparable to those for 1- and 8-plexes when analyzing additional data from more samples.

Tissue DNA Workflow: Variant calling results for four variant classes for tissue DNA samples (FFPE and non-FFPE) are given in Table 10. All variants were detected with a true positive rate of 100% in both 1- and 8-plex captures, and microsatellite instability status was correctly classified.

Tissue RNA Workflow: Variant calling results (fusion transcripts) for tissue RNA samples are given in Table 11. As for the tissue DNA Workflow, all variants were detected with a true positive rate of 100%, irrespective of the level of pre-capture multiplexing.

Conclusions

Roche's innovative KAPA HyperPETE technology enables fast and easy target enrichment for somatic[2] and germline[3] variant analysis using small (<300 kb capture target) panels, with performance comparable to overnight hybridization-based capture. KAPA HyperPETE Workflows support high coverage uniformity, as well as detection of all major somatic variant types (SNV, indels, CNV, MSI status, and fusions—including novel transcript partners).

KAPA HyperPETE Workflows for somatic applications (plasma cfDNA, tissue DNA, and tissue RNA samples) are robust and allow for simultaneous enrichment of four to eight libraries in a single PETE reaction, with sequencing and variant detection performance comparable to that of single-plex capture. Pre-capture multiplexing is enabled through minor modifications to the standard protocol (see Appendix), and offers significant benefits in terms of sample processing, throughput, turnaround time, and per-sample reagent cost in targeted sequencing workflows.

In addition, KAPA HyperPETE Workflows:

- are automation-friendly and compatible with all liquid handlers commonly used in NGS sample preparation pipelines.
- are supported from end-to-end with fully integrated service and technical support from a single vendor.

Table 9. Variant detection performance (true positive rate) for the Plasma cfDNA Workflow

Variant type	cfDNA input	Expected allele frequency (AF)	Expected variants/sample	True positive rate (%)		
				1-plex	4-plex	8-plex
Short variants (SNV, indels)	50 ng	~1%	28	100	100	100
		~0.5%	25	100	100	100
	10 ng	~1%	28	100	96.4	96.4
		~0.5%	25	96.0	92.0	96.0

Table 10. Variant detection performance (true positive rate) for the Tissue DNA Workflow

Variant type	Number of samples	Expected allele frequency (AF)	Expected number of variants	True positive rate (%)	
				1-plex	8-plex
Short variants (SNV, indels)	11	~5%	61	100	100
Copy number variants (CNV)	3	~6 copies	4	100*	100*
MSS/MSI (Microsatellite status)	7	N/A	4 MSS, 3 MSI	100	100

*The *MET* CNV in FFPE sample HD789 was a low-confidence call by the CNV caller.

Table 11. Variant detection performance (true positive rate) for the Tissue RNA Workflow

Variant type	RNA input amount	Number of samples	Expected number of variants	True positive rate (%)	
				1-plex	8-plex
Fusion transcripts	10 ng	3 reference samples	30	100*	100*

*The *EGFR-SEPT14* variant was manually curated as the fusion caller identified a different *EGFR* partner that has a highly homologous sequence to *SEPT14*.

Appendix

Detailed protocol for pre-capture multiplexing in KAPA HyperPETE workflows for somatic applications

Overview

The Instructions for Use for KAPA HyperPETE Workflows for somatic applications,[4–6] describe standard, validated protocols for the processing of plasma or cell line cfDNA, and tissue or cell line (fresh frozen or FFPE) DNA or RNA by means of single-plex target enrichment (one library per PETE reaction). Results presented in this Application Note demonstrate comparable performance between pre-capture multiplexing and single-plexing in KAPA HyperPETE Workflows for somatic applications. Nevertheless, the protocol modifications outlined here have not been fully validated and are not fully supported.

Pre-capture multiplexing is enabled through two simple modifications:

- more accurate quantification of pre-capture libraries, and
- addition of a Multiplex Pool Preparation and Concentration step.

Accurate Quantification of Pre-capture Libraries

In standard protocols, **Library Preparation** concludes with the assessment of library concentration and fragment size distribution using a 4200 TapeStation system and DNA High Sensitivity D1000 ScreenTape Assay (Agilent® Technologies), as described in:

- Chapter 3, Step 6 for the Plasma cfDNA Workflow;
- Chapter 4, Step 8 for the Tissue DNA Workflow; and
- Chapter 3, Step 8 for the Tissue RNA Workflow.

For pre-capture multiplexing, standard protocols are followed without modification up to this point. However, since precise pooling is critical for pre-capture multiplexing, a **fluorometric** assay is required for accurate quantification of pre-capture libraries. Library concentrations determined with the TapeStation assay are **not** used to calculate the volumes of libraries needed to create pre-capture pools. Instead, diluted libraries (1:40 dilutions prepared for electrophoretic analysis) are quantified with a Qubit® Fluorometer and Qubit dsDNA HS Assay Kit; Thermo Fisher Scientific), and measured concentrations are used to calculate the concentration of undiluted libraries.

- Undiluted pre-capture libraries must have a concentration of at least 45 ng/µL to support multiplexed capture in modified workflows for somatic applications. Since the additional quantification step does not consume any library material, approximately 23 µL of the 25 µL pre-capture library recovered after the 1X Post-amplification Purification should be available for pre-capture pooling to meet the minimum input requirement (≥45 ng/µL x 23 µL ≥1 µg).
- Once quantification has been completed, store pre-capture libraries at -20°C for up to 1 month, or proceed to the next step.

Multiplex Pool Preparation and Concentration

For pre-capture multiplexing, this additional step is performed before any of the **Primer Extension Target Enrichment (PETE)** steps of the standard protocol, i.e.,:

- between Chapters 3 and 4 for the Plasma cfDNA and Tissue RNA Workflows.
- between Chapters 4 and 5 for the Tissue DNA Workflow.

Detailed Protocol

1. Remove the KAPA HyperPure Beads from cold storage at least 30 minutes prior to use (in Step 4) to bring them to room temperature. Vortex beads until thoroughly resuspended right before use.
2. If necessary, thaw the indexed pre-capture libraries that will be pooled for capture on ice.

   NOTE: Make sure that all of the libraries selected for a capture pool were prepared with different sequencing indexes (KAPA UDI Primer Mixes).

3. Using concentrations determined with the Qubit assay, calculate the volume of each pre-capture library equivalent to 1 µg of library material. Taking care to pipette accurately, transfer the appropriate volume of each library to a PCR tube to construct a 4- or 8-plex pool.

   NOTE: If the total volume of the pre-capture multiplex pool adds up to ≤110 µL, pool the libraries in a 0.2 mL PCR tube, as this will facilitate recovery in a small elution volume (Step 15).

   If the total volume of the pre-capture multiplex pool adds up to >110 µL, pool the libraries in a 0.5 mL PCR tube to provide sufficient space for the KAPA HyperPure Beads to be added in the next step.

4. Concentrate the pooled libraries by performing a 1.5X "purification" with KAPA HyperPure Beads, as described in Steps 5 – 18.

   For example, if the volume of the pooled libraries is 80 µL, add 120 µL bead reagent.

5. Mix thoroughly by vortexing tubes for 1 minute using an IKA MS 3 Vortexer set to 2400 rpm. Quickly spin down the liquid without pelleting the beads.
6. Incubate at room temperature for **10 minutes**.
7. Place the tube(s) on a magnet to capture the beads. Incubate until the liquid is clear.
8. Carefully remove and discard the supernatant.

9. Keeping the tube(s) on the magnet, wash the pellet by adding 200 µL of 80% ethanol. Incubate at room temperature for ≥30 seconds.

   NOTE: 80% Ethanol must be prepared fresh daily.

10. Carefully remove and discard the ethanol.
11. Repeat the ethanol wash for a second wash (Steps 9 – 10).
12. Spin the tubes down quickly to bring residual ethanol to the bottom.
13. Place tubes on the magnet and remove residual ethanol using a P20 pipette without disturbing the beads.
14. Leave the tubes open on the magnet to dry the beads for 3 – 5 minutes or until all of the ethanol has evaporated.

    NOTE: Do not over dry. Over drying the beads may lead to reduction in yield. Beads are dry when they are not shiny in appearance. Avoid over drying the bead pellet by resuspending before the pellet begins to crack.

15. Remove the tubes from the magnet and resuspend the beads in 15 µL 10 mM Tris-HCl, pH 8.0, and mix thoroughly by pipetting until the pelleted beads are completely resuspended.

    NOTE: Avoid touching the pellet with the pipette tip.

16. Incubate for **5 minutes** at room temperature.
17. Briefly spin down the samples, place the tube(s) on a magnet to capture the beads, and incubate until the liquid is clear.

    NOTE: Visually confirm that the beads are pelleted.

18. Transfer the eluate into a new 0.2 mL PCR tube. The eluate contains the concentrated pre-capture multiplex library pool.
19. Proceed to the **Primer Extension Target Enrichment (PETE)** Chapter without delay.
    - No further modifications of standard protocols are needed for multiplexed capture.
    - In the Capture Extension Reaction step, the Capture Extension Reaction Master Mix will be prepared without any water (i.e., in a total of 35 µL) and added to the 15 µL concentrated pre-capture multiplex library pool recovered in Step 18 above.
    - The 15 µL concentrated pre-capture multiplex library pool will contain a total of 4 µg of library for 4-plex pools, and 8 µg of library material for 8-plex pools.
    - In the Pooling and Sequencing step (Chapter 4, Step 13 for Plasma cfDNA and Tissue RNA Workflows; Chapter 5, Step 13 for Tissue DNA Workflow), the sequencing pool will be constructed from post-capture library pools (instead of individual post-capture libraries when single-plex capture is performed). The total paired-end read requirements given in this section still applies on a per-library basis; i.e., must be calculated based on the total number of individual libraries in the sequencing pool, irrespective of whether enrichment was performed in single- or multiplexed format.

References

1. Bewicke-Copley F, Arjun Kumar E, Palladino G, et al. Applications and analysis of targeted genomic sequencing in cancer studies. *Comput. Struct. Biotechnol. J.* 2019;17(1348–1359). doi: 10.1016/j.csbj.2019.10.004.
2. Furtado N, Cai M, Li J, et al. https://sequencing.roche.com/content/dam/rochesequence/worldwide/resources/SPR-App-Note_KAPA_HyperPETE_Somatic_Variant_Analysis-MC--08463.pdf. KAPA HyperPETE: A novel, end-to-end target enrichment workflow for high-performance somatic variant analysis. Roche Application Note, MC--08463. Accessed 10 Aug 2022.
3. Furtado N, Cai M, Li J, et al. https://sequencing.roche.com/en/contactus.html. The novel, end-to-end KAPA HyperPETE Target Enrichment Workflow enables high-performance germline variant analysis. Roche Application Note, MC--09247. Accessed 10 Aug 2022.
4. Roche Sequencing Solutions. KAPA HyperPETE Somatic Plasma cfDNA Workflow; Instructions for use, Version 1.0. https://pim-eservices.roche.com/eLD/api/downloads/b616daab-a841-ec11-0d91-005056a71a5d?countryIsoCode=pi. Accessed 10 Aug 2022.
5. Roche Sequencing Solutions. KAPA HyperPETE Somatic Tissue DNA Workflow; Instructions for use, Version 1.0. https://pim-eservices.roche.com/eLD/api/downloads/124d7ceb-5033-ec11-0d91-005056a71a5d?countryIsoCode=pi. Accessed 10 Aug 2022.
6. Roche Sequencing Solutions. KAPA HyperPETE Tissue RNA Fusion Transcript Workflow; Instructions for use, Version 1.0. https://pim-eservices.roche.com/eLD/api/downloads/737ff2ff-b941-ec11-0b91-005056a772fd?countryIsoCode=pi. Accessed 10 Aug 2022.
7. Roche Diagnostics GmbH. High Pure FFPET RNA Isolation Kit; Instructions for Use, Version 06 (Content version: October 2020). https://pim-eservices.roche.com/eLD/web/pi/en/documents/download/4278ada1-ad12-eb11-fe90-005056a772fd. Accessed 10 Aug 2022.
8. Agilent Technologies. Simplified DV_{200} Evaluation with the Agilent 2100 Bioanalyzer System. https://www.agilent.com/cs/library/technicaloverviews/public/5991-8287EN.pdf. Accessed 10 Aug 2022.

Published by:

Roche Sequencing Solutions, Inc.
4300 Hacienda Drive
Pleasanton, CA 94588

sequencing.roche.com

MC--09862 08/2022