

Application Note

High-performance germline variant analysis



The novel, end-to-end KAPA HyperPETE Target Enrichment Workflow enables high-performance germline variant analysis

Authors:

Nitya Furtado

Senior Scientist –
Reagent and Assay Development

Mingyang Cai*

Principal Bioinformatics Scientist II

Jingchuan Li

Principle Scientist I –
Reagent and Assay Development

Florence Crary-Dooley

Senior Scientist –
Reagent and Assay Development

Liu Xi

Principle Bioinformatics Scientist I

Brian Godwin

Director –
Reagent and Assay Development

Payal Shah*

Senior Product Support Specialist

Jimmy Tu

International Product Manager

Markos Mihalatos

International Business Leader –
Diagnostics

Roche Molecular Labs,
Pleasanton, CA, USA

Date of first publication: 2022

Date of this publication: 2024

*No longer employed by Roche.

Targeted NGS enables highly efficient, fast, and cost-effective analysis of germline variants. The KAPA HyperPETE Germline Workflow offers an end-to-end solution for flexible, high-performance analysis of inherited SNVs and indels with small panels. The single-day workflow offers specific advantages for focused, time-sensitive applications such as hereditary oncology and newborn screening.

Introduction

In addition to determining hundreds of human traits, germline (inherited) single-nucleotide variants (SNVs) and small insertions/deletions (indels) are associated with a plethora of genetic diseases. These include monogenic (Mendelian) disorders (e.g., cystic fibrosis, fragile X syndrome, and Huntington's disease), as well as common polygenic (complex) disorders, including heart disease, diabetes, and Alzheimer's disease.¹ Germline mutations also increase the predisposition for many types of cancers and impact cancer development and metastasis.²



Targeted next-generation sequencing (NGS) is an established and powerful tool for studying germline variants, offering the ability to survey tens to tens of thousands of targets (potential variants) in parallel—in a highly efficient, fast, and cost-effective manner.³

Roche has developed a novel Primer Extension Target Enrichment (PETE) technology (Figure 1), which enables robust, streamlined, and versatile analysis of both germline and somatic variants with performance comparable to that of gold-standard—but slower—hybridization capture methods. KAPA HyperPETE Workflows combine validated KAPA Library Preparation and KAPA Target Enrichment reagents for oncology, inherited disease, and other research applications. Optimized for small panels, the single-day, automation-friendly workflow supports up to 8-plex captures for germline samples with both catalog and custom panels. The KAPA HyperPETE Germline Workflow also offers specific advantages for focused applications such as hereditary oncology when fast turnaround times are critical.⁴

We have previously demonstrated the performance of the KAPA HyperPETE technology for somatic variant analysis from cell-free DNA, FFPE DNA and RNA, and reference cell line DNA and RNA samples.⁵ In this Application Note, we show sensitive and precise detection of germline SNVs and indels, achieving true positive rates ranging from >97.4 – 100% for SNVs and indels (after removal of low-complexity regions) from reference cell line samples.

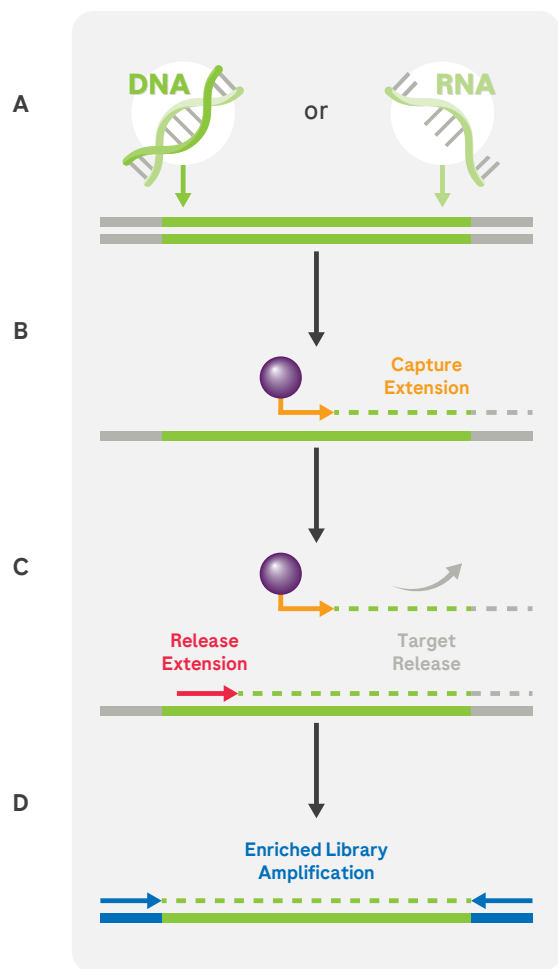


Figure 1. Overview of the KAPA HyperPETE (Primer Extension Target Enrichment) Technology. (A) Uniquely dual-indexed pre-capture libraries are generated from genomic DNA using the KAPA HyperPlus Library Preparation Kit, KAPA Universal Adapter, and KAPA UDI Primer Mixes. (B) Following a heated denaturation 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, specifically releasing target molecules from the beads into the supernatant (an intermediate wash step to remove non-hybridized primers is performed prior to the release extension). (D) Released products are amplified with universal primers (blue). Final libraries 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 germline variant analysis. To this end, unique dual-indexed libraries were prepared from human cell line genomic DNA with the KAPA HyperPlus Kit. Primer Extension Target Enrichment (PETE) was performed with two catalog KAPA HyperPETE Panels (with capture sizes ranging between 200 and 300 kb) according to the standard workflow for germline DNA. Sequencing was performed on the Illumina® platform using standard protocols. Data subsampling and analysis was performed using an internal Roche secondary analysis pipeline. Select sequencing and variant calling metrics are reported. A summary of the experimental design is given in Figure 2.

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

Germline DNA Workflow			
Samples	Reference cell line genomic DNA (n=7)		
LP	KAPA HyperPlus Kit (50 ng input) KAPA Universal Adapter, KAPA UDI Primer Mixes		
TE	<table border="1"> <tr> <td>KAPA HyperPETE Hereditary Oncology Panel 853 capture targets, 47 genes, 203 kb</td> <td>KAPA HyperPETE Newborn Screening Panel* 1,301 capture targets, 89 genes, 294 kb</td> </tr> </table>	KAPA HyperPETE Hereditary Oncology Panel 853 capture targets, 47 genes, 203 kb	KAPA HyperPETE Newborn Screening Panel* 1,301 capture targets, 89 genes, 294 kb
KAPA HyperPETE Hereditary Oncology Panel 853 capture targets, 47 genes, 203 kb	KAPA HyperPETE Newborn Screening Panel* 1,301 capture targets, 89 genes, 294 kb		
QC metrics	% reads on target • fold-80 base penalty % target bases $\geq 20X$, $\geq 30X$ • total duplication rate		
Variants	SNVs • Indels		

Figure 2. Summary of the experimental design for this study. Full details of cell line samples (inputs) are provided in Materials and methods. Uniquely dual-indexed libraries were prepared as described. Quality control (quantification and library fragment size assessment) of pre- and post-enrichment libraries is not shown in the diagram, but was performed as described in Materials and methods. Single (1-plex) and 8-plex captures were performed with the KAPA HyperPETE Hereditary Oncology and Newborn Screening* Panels, respectively. Data processing and variant calling were performed as described below. Full definitions of QC metrics are given in Results and Discussion. LP: Library Preparation, TE: Target Enrichment, UDI: Unique Dual Index.

*Not available for sale in the United States. Contact your local Roche representative for availability in your region.

Samples

Seven reference cell line DNA samples (purified genomic DNA) were obtained from the Coriell Institute for Medical Research. Sample details are given in Table 1.

Library Preparation and Target Enrichment

A total of 96 uniquely dual-indexed germline DNA libraries (Table 2) were prepared from 50 ng inputs for single (1-plex) or multiplexed (8-plex) target enrichment with two different KAPA HyperPETE Panels (Table 3).

Libraries were prepared with the KAPA HyperPlus Kit (Roche PN: 07962401001 or 07962428001), KAPA Universal Adapter (Roche PN: 09063781001 or 09063790001), 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 Germline DNA Workflow v1.0 (or later) Instructions for use.⁶

Pre-capture Library QC: The concentration and fragment size distribution of amplified, pre-capture libraries were determined with an Agilent® 4200 TapeStation system and DNA High Sensitivity D1000 ScreenTape Assay (Agilent Technologies). All libraries were diluted 1:40 for analysis.

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), according to the standard protocol for 1- to 8-plex captures.

Table 1. Samples used in the KAPA HyperPETE Germline DNA Workflow

	Cell line	Origin/description	Source/reference
1	NA12878	CEPH/Utah pedigree 1463, International HapMap Project (female, age unknown)	https://www.corieell.org/0/Sections/Search/Sample_Detail.aspx?Ref=NA12878
2	NA24143	Personal genome project (female, 74 years)	https://www.corieell.org/0/Sections/Search/Sample_Detail.aspx?Ref=NA24143&Product=DNA
3	NA24149	Personal genome project (male, 90 years)	https://www.corieell.org/0/Sections/Search/Sample_Detail.aspx?Ref=NA24149&Product=DNA
4	NA24385	Personal genome project (male, 45 years)	https://www.corieell.org/0/Sections/Search/Sample_Detail.aspx?Ref=NA24385&Product=DNA
5	NA24631	Personal genome project (male, 33 years)	https://www.corieell.org/0/Sections/Search/Sample_Detail.aspx?Ref=NA24631&Product=DNA
6	NA24694	Personal genome project (male, 64 years)	https://www.corieell.org/0/Sections/Search/Sample_Detail.aspx?Ref=NA24694&Product=DNA
7	NA24695	Personal genome project (female, 63 years)	https://www.corieell.org/0/Sections/Search/Sample_Detail.aspx?Ref=NA24695&Product=DNA

Table 2. Breakdown of libraries prepared using KAPA HyperPETE Germline DNA Workflow

Panel	Cell line	Number of uniquely dual-indexed libraries and captures					
		Libraries in 1-plex capture	Libraries in 8-plex A	Libraries in 8-plex B	Libraries in 8-plex C	Total libraries	Total captures
KAPA HyperPETE Hereditary Oncology Panel	NA12878	12	0	4	8	24	24 x 1-plex + 3 x 8-plex
	NA24143	2	2	0	0	4	
	NA24149	2	2	0	0	4	
	NA24385	2	2	0	0	4	
	NA24631	2	2	0	0	4	
	NA24694	2	0	2	0	4	
	NA24695	2	0	2	0	4	
KAPA HyperPETE Newborn Screening Panel*	NA12878	12	0	4	8	24	24 x 1-plex + 3 x 8-plex
	NA24143	2	2	0	0	4	
	NA24149	2	2	0	0	4	
	NA24385	2	2	0	0	4	
	NA24631	2	2	0	0	4	
	NA24694	2	0	2	0	4	
	NA24695	2	0	2	0	4	
All panels	All cell lines	48	16	16	16	96	54

All libraries were prepared from 50 ng inputs of purified genomic DNA.

*Not available for sale in the United States. Contact your local Roche representative for availability in your region.

Table 3. KAPA HyperPETE Panels used in this study

Panel name	Panel type	Capture size	Details	Paired-end reads required per library
KAPA HyperPETE Hereditary Oncology Panel	Catalog	203 kb	853 capture targets, covering 47 genes related to hereditary breast and ovarian cancers (including <i>BRCA1</i> and <i>BRCA2</i>)	1 million
KAPA HyperPETE Newborn Screening Panel*	Catalog	294 kb	1,301 capture targets covering 89 related genes from the core conditions of the Hereditary Genetic Panel, plus 1,272 pathogenic SNPs in non-coding regions	1 million

Both panels are optimized to deliver high uniformity and specificity from germline DNA samples with pre-capture multiplexing capacity of up to 8 samples.

*Not available for sale in the United States. Contact your local Roche representative for availability in your region.

Post-capture (PETE) Library QC: Sequencing-ready, post-capture (enriched) libraries were quantified and analyzed in the same manner as pre-capture libraries, but without any dilution.

Sequencing and Data Analysis

Sequencing: Libraries were pooled for multiplexed, paired-end Sequencing (2 x 151 bp) on an Illumina® NextSeq® 500, NextSeq 550Dx (RUO mode), or MiSeq® system. Pools were configured to meet the paired-end read requirement for each panel, as outlined in Table 3.

Data analysis: Raw sequencing data were demultiplexed with an internal Roche secondary analysis pipeline. The resulting sample-level fastq files were downloaded and randomly subsampled to 1 million reads per library using seqtk.⁷ Subsampled fastq files were subsequently processed (adapter trimming, alignment, position deduplication, SNV and indel calling). The performance of the KAPA HyperPETE technology with germline DNA, across two different levels of pre-capture multiplexing and both catalog panels, was evaluated by (i) assessing key sequencing QC metrics, and (ii) comparing VCF files to the truth data set for each cell line.⁸ Graphs were generated with ggplot2 (version 3.3.5) in R (version 3.4.4).⁹

Results and discussion

Library QC metrics

All pre-capture libraries met the criteria for target enrichment, namely:

- a concentration of ≥ 35 ng/ μ L in the region of 150 – 1000 bp.
- an average fragment size of ~ 300 bp.

For 8-plex captures, six individual Multiplex DNA Input Library Pools were created by pooling eight pre-capture libraries (as outlined in Table 2). The final concentration of every Multiplex DNA Input Library Pool exceeded the minimum of 67 ng/ μ L.

All post-capture (enriched) 1- and 8-plex libraries met the criteria for sequencing:

- molarity of ≥ 4 μ M in the region of 150 – 1000 bp.
- a mean fragment size in the range of 300 – 500 bp.

Table 4. Sequencing QC metrics reported in this study

Metric	Explanation/definition
% reads on target	<ul style="list-style-type: none"> • Also referred to as on-target rate • Defined as the percentage of unique reads overlapping the primary unpadded target by at least 1 base
Fold-80 base penalty	<ul style="list-style-type: none"> • A measure of coverage uniformity. Defined as the fold of additional sequencing required to ensure that the mean coverage is achieved for 80% of the target bases. • Calculated using the GATK CollectHsMetrics algorithm
% Panel exon region $\geq dX$ ($d = 20$ or 30)	<ul style="list-style-type: none"> • Percentage of all target bases (as defined by the capture BED file) achieving a coverage of at least d • Calculated using the GATK CollectHsMetrics algorithm
% total duplicates	<ul style="list-style-type: none"> • Total duplicate rate, or percentage of reads that exist as duplicates • In the Germline DNA Workflow, reflects position-based deduplication • Calculated using the GATK MarkDuplicates algorithm

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

Sequencing QC metrics

The internal Roche secondary analysis pipeline generates an extensive list of sample and analysis QC metrics for every run. A subset of these metrics (listed and defined in Table 4) is reported here. Results obtained with the two KAPA HyperPETE panels for libraries prepared from 50 ng genomic DNA inputs and processed as single or 8-plex captures are shown in Figure 3. Comparable results were obtained for 1-plex and 8-plex pre-capture Multiplex Pools, with the smaller Hereditary Oncology Panel performing slightly better than the Newborn Screening Panel.

On-target rates were high (66.8 – 77.5%, average of 72.6%) and highly reproducible across the entire data set.

Coverage uniformity was high and consistent across both levels of multiplexing, as evidenced by fold-80 base penalty scores ≤ 1.65 (1.45 – 1.65, average of 1.53). This confirmed that the KAPA HyperPETE design algorithm and novel enrichment strategy support high coverage uniformity from germline samples.

The average coverage of target regions was 98.5% (96.8 – 99.6%) and 97.9% (95.6 – 99.8%) at $\geq 20X$ and $\geq 30X$, respectively. No significant difference in coverage at either depth was observed between 1-plex and 8-plex captures.

Total duplication rates for 1-plex captures were ranged between 2.78% and 6.53% (average of 4.95%). Duplication rates for 8-plex captures were higher (between 7.08% and 12.2%, average of 10.8%). This was expected due to the lower per-library input for multiplexed captures.

Together, these results demonstrate that the KAPA HyperPETE Germline Workflow yields high and uniform sequencing coverage across different designs for both 1- and 8-plex captures, thereby supporting high-confidence variant calling (see below) from the recommended amount of paired-end reads for each panel (ref. Table 3)

Variant calling results

Characterized reference cell lines with validated mutations were used in this study in order to assess variant calling performance. True Positive

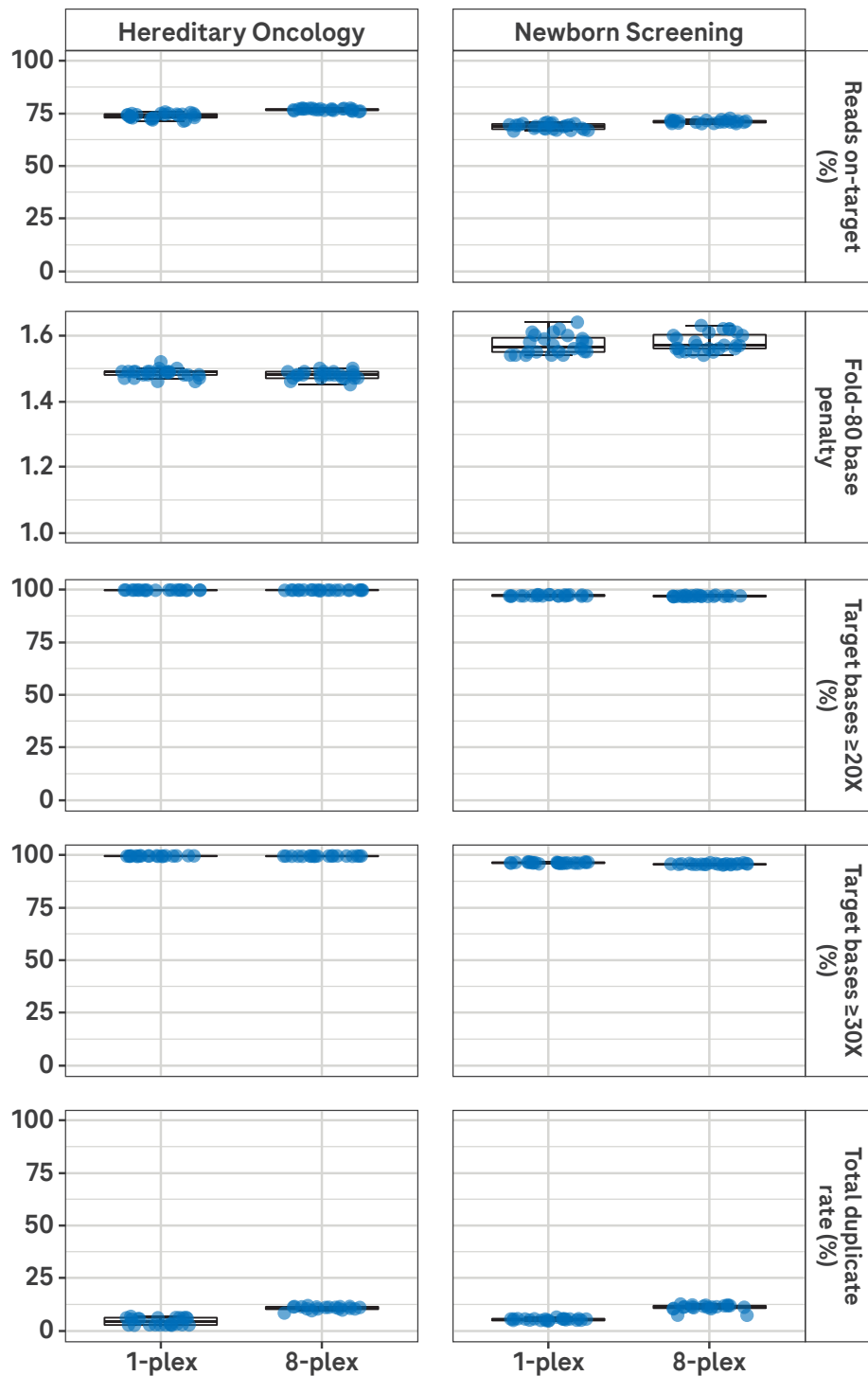


Figure 3. Key sequencing QC metrics achieved with the KAPA HyperPETE Germline DNA Workflow. Libraries were prepared from 50 ng inputs, enriched with the two different catalog panels, and sequenced as described in *Materials and methods*. Data were analyzed and QC metrics were generated with an internal Roche secondary analysis pipeline. A breakdown of the number of libraries and captures for each panel and level of multiplexing may be found in Table 3.

Rate (TPR, sensitivity) and Positive Predictive Value (PPV, precision) were calculated for the two most commonly assessed germline variant classes—SNVs and indels. Detailed results by cell line/capture panel are given in Tables 5 and 6.

SNVs were detected with a high degree of sensitivity (TPR >97.4% for both panels) and precision (PPV of 99.8% and 99.2% for the Hereditary Oncology and Newborn Screening* Panels, respectively). With the

Hereditary Oncology Panel, indels were called with 100% precision and sensitivity, whereas indel calling sensitivity and precision with the Newborn Screening Panel* were 100% and 99%, respectively.

Variant calling performance for both SNVs and indels are compared for 1-plex vs. 8-plex captures in Tables 7 and 8, with no significant difference in results. Low-complexity regions (homopolymers of 10–13 bp) were excluded from the analysis due to high intrinsic error

Table 5. SNV calling performance for the KAPA HyperPETE Germline DNA Workflow

Panel	Cell line	Replicates	Expected number of SNVs per sample	True Positives (TP) per sample ¹	False Negatives (FN) per sample ¹	False Positives (FP) per sample ¹	True Positive Rate (TPR, %) ²	Positive Predictive Value (PPV, %) ³
KAPA HyperPETE Hereditary Oncology Panel	NA12878	24	89	87.4	1.63	0.00	98.2	99.8
	NA24143	4	117	115.5	1.50	0.25		
	NA24149	4	127	124.0	3.00	1.50		
	NA24385	4	123	120.0	3.00	0.50		
	NA24631	4	89	87.5	1.50	0.00		
	NA24694	4	94	92.5	1.50	0.25		
	NA24695	4	99	97.8	1.25	0.00		
KAPA HyperPETE Newborn Screening Panel*	NA12878	24	196	192.2	3.83	1.71	97.4	99.2
	NA24143	4	229	219.0	10.0	2.75		
	NA24149	4	227	217.8	9.25	2.00		
	NA24385	4	248	237.25	10.8	1.25		
	NA24631	4	194	190.5	3.50	0.75		
	NA24694	4	200	194.0	6.00	1.25		
	NA24695	4	204	200.5	3.50	1.50		

¹ Calculated across all replicates, therefore not always an integer.

² True Positive Rate (sensitivity, recall) = TP/(TP+FN)

³ Positive Predictive Value (precision) = TP/(TP+FP)

*Not available for sale in the United States. Contact your local Roche representative for availability in your region.

Table 6. Indel calling performance for the KAPA HyperPETE Germline DNA Workflow

Panel	Cell line	Replicates	Expected number of indels per sample	True Positives (TP) per sample ¹	False Negatives (FN) per sample ¹	False Positives (FP) per sample ¹	True Positive Rate (TPR, %) ²	Positive Predictive Value (PPV, %) ³
KAPA HyperPETE Hereditary Oncology Panel	NA12878	24	1	1.00	0.00	0.00	100	100
	NA24143	4	1	1.00	0.00	0.00		
	NA24149	4	2	2.00	0.00	0.00		
	NA24385	4	0	0.00	0.00	0.00		
	NA24631	4	2	2.00	0.00	0.00		
	NA24694	4	4	4.00	0.00	0.00		
	NA24695	4	2	2.00	0.00	0.00		
KAPA HyperPETE Newborn Screening Panel*	NA12878	24	9	9.00	0.00	0.21 ⁴	100	99.0
	NA24143	4	9	9.00	0.00	0.00		
	NA24149	4	9	9.00	0.00	0.00		
	NA24385	4	7	7.00	0.00	0.00		
	NA24631	4	10	10.0	0.00	0.00		
	NA24694	4	10	10.0	0.00	0.00		
	NA24695	4	9	9.00	0.00	0.00		

¹ Calculated across all replicates, therefore not always an integer.

² True Positive Rate (sensitivity, recall) = TP/(TP+FN)

³ Positive Predictive Value (precision) = TP/(TP+FP)

⁴ The higher FP rate for the Newborn Screening Panel is partially attributed to the ~45% larger panel size (as compared to the Hereditary Oncology Panel), but same number of reads (1 M)

*Not available for sale in the United States. Contact your local Roche representative for availability in your region.

Data on file with Roche.

For Research Use Only. Not for use in diagnostic procedures.

rates on the Illumina[®] sequencing platform.¹⁰⁻¹² This demonstrates that the KAPA HyperPETE Germline Workflow and KAPA HyperPETE Panels support an increase in sample throughput and lower per-sample cost without sacrificing performance.

Conclusions

The novel KAPA HyperPETE technology from Roche enables a new class of end-to-end targeted sequencing solutions for high-performance germline variant analysis using small (≤ 300 kb) panels. KAPA HyperPETE Workflows combine deep content expertise and innovative panel design with industry-leading library preparation and target enrichment reagents.

Key features of KAPA HyperPETE Germline Workflow include:

- Comparable performance to gold-standard hybrid capture workflows with a convenient, single-day protocol.
- Flexibility to process germline samples with high efficiency, using optimized catalog or easy-to-order custom panels.

Like the KAPA HyperPETE Somatic Workflow, the KAPA HyperPETE Germline Workflow:

- Employs the existing user-friendly [HyperDesign Tool](#) for custom panel design. The tool (which was originally developed to support custom probe design for the hybridization-based KAPA HyperCap Workflows) has been updated to include a new design/selection algorithm for KAPA HyperPETE primer panels.
- Is automation-friendly, and compatible with all liquid handlers commonly used in NGS sample preparation pipelines.
- Is supported from end-to-end with fully integrated service and technical support from a single vendor.

This Application Note demonstrates that the KAPA HyperPETE Germline Workflow delivers the high on-target rates, coverage depth, and uniformity needed to support economical, sensitive, accurate, and reproducible analysis of germline SNVs and indels.

Table 7. Comparison of SNV calling performance for 1-plex vs. 8-plex captures

Panel	Capture plex level	Number of captures	Expected number of SNVs per sample ¹	True Positives (TP) per sample ¹	False Negatives (FN) per sample ¹	False Positives (FP) per sample ¹	True Positive Rate (TPR, %) ²	Positive Predictive Value (PPV, %) ³
KAPA HyperPETE Hereditary Oncology Panel	1-plex	24	98.58	96.9	1.67	0.08	98.3	99.9
	8-plex	3		96.7	1.92	0.33	98.1	99.7
KAPA HyperPETE Newborn Screening Panel*	1-plex	24	206.50	201.0	5.50	1.79	97.4	99.1
	8-plex	3		201.0	5.50	1.50	97.4	99.3

¹ Calculated across all cell lines/replicates, therefore not always an integer.

² True Positive Rate (sensitivity, recall) = $FP/(FP+TN)$

³ Positive Predictive Value (precision) = $TP/(TP+FP)$

*Not available for sale in the United States. Contact your local Roche representative for availability in your region.

Table 8. Comparison of indel calling performance for 1-plex vs. 8-plex captures

Panel	Capture plex level	Number of captures	Expected number of indels per sample ¹	True Positives (TP) per sample ¹	False Negatives (FN) per sample ¹	False Positives (FP) per sample ¹	True Positive Rate (TPR, %) ²	Positive Predictive Value (PPV, %) ³
KAPA HyperPETE Hereditary Oncology Panel	1-plex	24	1.55	1.55	0.00	0.00	100	100
	8-plex	3			0.00	0.00	100	100
KAPA HyperPETE Newborn Screening Panel*	1-plex	24	9.00	9.00	0.00	0.08	100	99.2
	8-plex	3			0.00	0.13	100	98.8

¹ Calculated across all cell lines/replicates, therefore not always an integer.

² True Positive Rate (sensitivity, recall) = $FP/(FP+TN)$

³ Positive Predictive Value (precision) = $TP/(TP+FP)$

*Not available for sale in the United States. Contact your local Roche representative for availability in your region.

References

1. Jackson M, Marks L, May G, Wilson, JB. The genetic basis of disease. *Essays Biochem.* 2018;62(643–723). doi: [10.1042/EBC20170053](https://doi.org/10.1042/EBC20170053).
2. Milanese J-S, Wang, E. Germline mutations and their clinical applications in cancer. *Breast Cancer Manag.* 2019;8. doi: [10.2217/bmt-2019-0005](https://doi.org/10.2217/bmt-2019-0005).
3. Di Resta C, Galbiati S, Carrera P, et al. Next-generation sequencing approach for the diagnosis of human diseases: open challenges and new opportunities. *EJIFCC.* 2018;29(4-14). PMID: [29765282](https://pubmed.ncbi.nlm.nih.gov/29765282/).
4. Roche Sequencing Solutions. KAPA HyperPETE Brochure. Accessed July 2024.
5. Furtado N, Cai M, Li J, et al. KAPA HyperPETE: a novel, end-to-end target enrichment workflow for high-performance somatic variant analysis. Roche Application Note, MC-US-15954. Accessed July 2024.
6. Roche Sequencing Solutions. KAPA HyperPETE Germline DNA Workflow v1.0 (or later) Instructions for use.
7. <https://github.com/lh3/seqtk>. Accessed 25 March 2022.
8. <https://ftp-trace.ncbi.nlm.nih.gov/giab/ftp/release>. Accessed July 2024.
9. <https://github.com/tidyverse/ggplot2>. Accessed July 2024.
10. Laehnemann D, Borkhardt A, McHardy AC. Denoising DNA deep sequencing data-high-throughput sequencing errors and their correction. *Brief Bioinform.* 2016;17(154-179). doi: [10.1093/bib/bbv029](https://doi.org/10.1093/bib/bbv029).
11. Stoler N, Nekrutenko A. Sequencing error profiles of Illumina sequencing instruments. *NAR Genomics and Bioinformatics* 2021;3(lqab019). doi: [10.1093/nargab/lqab019](https://doi.org/10.1093/nargab/lqab019).
12. Data on file with Roche.

Acknowledgments

The authors wish to thank Liesl Joubert, Marius Snyman, Etienne Slabbert, and Jo-Anne Penkler (Roche Molecular Labs, Cape Town) and Seng Saelee (Roche Molecular Labs, Pleasanton), for their contributions to important precursor studies which enabled the work presented in this project.

Published by:

Roche Sequencing and Life Science
9115 Hague Road
Indianapolis, IN 46256

sequencing.roche.com

Data on file with Roche.

For Research Use Only. Not for use in diagnostic procedures.

HYPERCAP, HYPERCAPTURE, HYPERDESIGN, KAPA, KAPA HYPERPETE, KAPA HYPERPLUS, and KAPA HYPERPURE are trademarks of Roche.

All other product names and trademarks are the property of their respective owners.

© 2024 Roche Sequencing Solutions, Inc. All rights reserved.