Detection of rare variants in degraded FFPE samples using HaloPlex PCR

Fredrik Roos, Eva Åström, Patrik Eriksson, Linus Forsmark, Filip Karlsson, Marcus Danielsson and Henrik Johansson
Agilent Technologies, Genomics Division

Introduction

We demonstrate the complete workflow of using Agilent HaloPlex PCR on formalin-fixed paraffin-embedded (FFPE) tissue derived DNA samples. The protocol consists of optimized workflow for DNA extraction and sample quality control. Furthermore it includes design and experimental setup optimizations aimed to improve data quality obtained from sequencing of degraded samples. We demonstrate this workflow on a set of matched fresh-frozen and FFPE samples showing 99.5% coverage at 1x, 95% coverage at 20x and high SNV concordance to fresh frozen controls. In addition we show detection of known variants down to 3% allele frequency.

Workflow overview

Tissue sample

Snap frozen | **Formalin-fixed paraffin-embedded**

---2 years storage---

QiaAmp DNA Micro Kit | **QiaAmp DNA FFPE Tissue Kit**

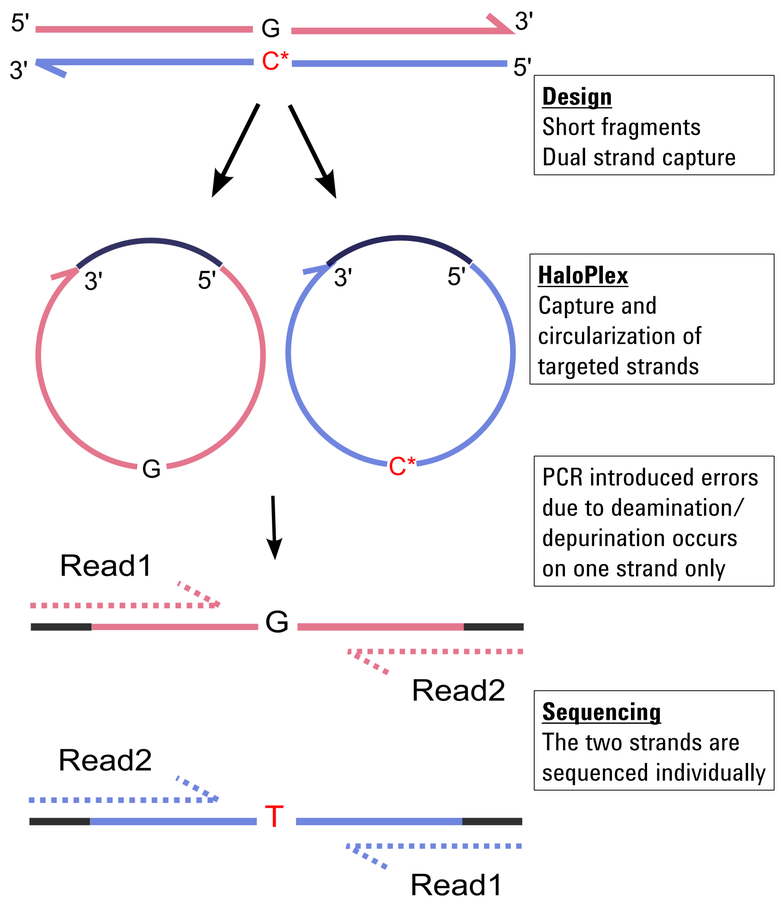
Quality control
- TapeStation gDNA ScreenTape
- Mutliplex PCR QC

HaloPlex enrichment
- 50 kb Cancer Panel, triplicates
- FFPE optimized design

Methods

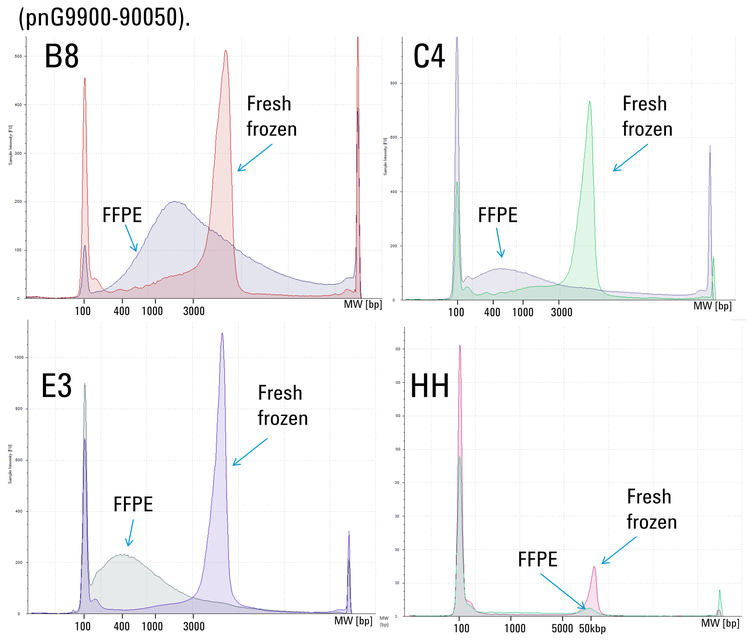
For the HaloPlex fast protocol, 200 ng genomic DNA is split up in eight different restriction digestion reactions. Following digestion the reactions are pooled and mixed with HaloPlex probes that hybridize to restriction fragments containing sequence present in the ROI. Following hybridization the bait/target hybrids are captured on solid phase and thereby separated from off-target DNA. In the last step prior to amplification the captured sequences are released from the beads using NaOH denaturation. The eluted fragments are then amplified by PCR for 18-25 cycles depending on the size of the captured region.

FFPE design strategy



DNA quality

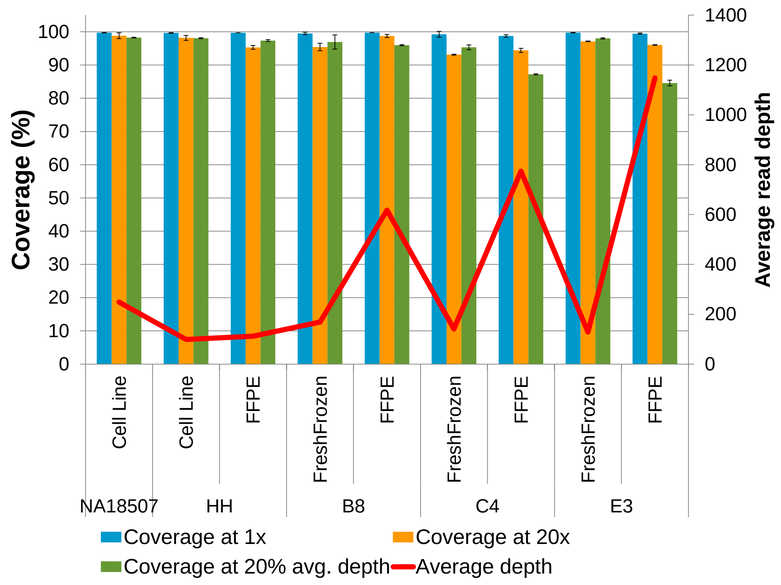
The extracted DNA was run on Agilent TapeStation gDNA ScreenTape to visualize the level of degradation. TapeStation show significant degradation of the FFPE samples compared to its matched fresh-frozen control. Average sizes correlate to QC scores obtained from Agilent's FFPE DNA Quality Assessment for HaloPlex. (pnG9900-90050).



Sample	Source / Treatment	Average size (bp)	PCR QC Score (1 = no degradation)	Input (ng)	Targeted average seq depth
NA18507	Cell line	>60.000	1.0	200	200x
B8	FF	26619	0.98	200	200x
	FFPE	1854	0.52	200	400x
C4	FF	28877	0.94	200	200x
	FFPE	639	0.14	500	800x
E3	FF	28055	0.4	200	200x
	FFPE	429	0.03	500	1200x
HH	Cell line	>60.000	0.95	200	200x
	FFPE	52795	0.93	200	200x

Sequencing results

All samples were pooled and sequenced on Illumina MiSeq. Uniformity (coverage at 0.2x of average depth) is lower in FFPE samples but the increased average depth compensated for this and ensured >95% coverage at 20x.



The sequencing data was analyzed using Agilent SureCall software and variant calls made in the fresh-frozen samples were compared to the matched FFPE samples.

	B8	C4	E3	HH
TOTAL_CALLS	41.7	35.7	38.0	62.3
DBSNP_OVERLAP	41.7	34.3	37.3	51.3
DBSNP_OVERLAP_RATE	100.0	96.3	98.2	82.4
DBSNP_OVERLAP_CONCORDANCE	100.0	100.0	97.3	100.0
FRESH_FROZEN_OVERLAP	40.0	33.0	30.0	64.0
HET_CONCORDANCE	100.0	100.0	58.3	98.7
HET_SENSITIVITY	100.0	100.0	70.6	93.5
HOM_VAR_CONCORDANCE	97.0	94.1	83.3	100.0
HOM_VAR_SENSITIVITY	100.0	100.0	92.3	100.0

Sample set "HH" contains 11 known variants with allele frequencies verified by digital PCR (HorizonDX, p/n HD500). The samples were sequenced to an average of 1500x and observed frequencies for these positions were recorded.

Gene	Mutation	Chr	Position	Digital PCR	HaloPlex Cell Line	HaloPlex FFPE
EGFR	G719S	chr7	55241707	24.5%	13.9%	19.4%
PIK3CA	H1047R	chr3	178952085	17.5%	17.5%	13.4%
KRAS	G13D	chr12	25398281	15.0%	14.4%	16.3%
NRAS	Q61K	chr1	115256530	12.5%	11.7%	11.6%
BRAF	V600E	chr7	140453136	10.5%	10.9%	9.5%
KIT	D816V	chr4	55599321	10.0%	11.4%	21.1%
PIK3CA	E545K	chr3	178936091	9.0%	6.4%	10.6%
KRAS	G12D	chr12	25398284	6.0%	4.7%	5.8%
EGFR	L858R	chr7	55259515	3.0%	3.3%	2.3%
EGFR	T790M	chr7	55249071	1.0%	0.2%	0.1%

Conclusions

Fragmented FFPE samples can successfully be used in HaloPlex target enrichment assays using the presented workflow. The decrease in sequencing coverage uniformity caused by unavailability of long targets can be predicted by the sample QC results. By compensating with increased targeted sequencing depth, >95% coverage at 20x is achieved without need of resequencing. Known variants down to 3% were successfully observed in 1500x depth sequencing data with intact allele frequencies. The 1% mutation had dropped out but could be observed in a few reads.