Effects of formalin fixation variables on DNA integrity for genomic applications in cancer

#297 | 2052

M. I. Lefterova[1], R.K. Alla[1], S. Luo[1], E. Helman[1], M.J. Clark[1], S.M. Boyle[1], M. Morra[1], S. Kirk[1], P. Sripakdeevong[1], M. Karbelashvili[1], D.M. Church[1], M. Snyder[2], J. West[1], R. Chen[1]
[1]Personalis, Menlo Park, CA, United States; [2]Stanford University, Stanford, Ca, United States

FINANCIAL DISCLOSURE: R.K. Alla, M.J. Clark, S.M. Boyle, S. Kirk, M. Snyder, J. West, and R. Chen are stockholders of Personalis, Inc.
M. I. Lefterova, R.K. Alla, S. Luo, E. Helman, M.J. Clark, S.M. Boyle, M. Morra, S. Kirk, P. Sripakdeevong, M. Karbelashvili, J. West, and R. Chen are employees of Personalis, Inc.

Contact:
martina.lefterova@personalis.com

Introduction

Formalin fixation and paraffin embedding (FFPE) is routinely used for tissue preservation. However, it leads to DNA fragmentation and mutation artifacts, which are exacerbated by variability of FFPE processing protocols. This impacts nucleic acid recovery for molecular diagnostics, particularly applications with high DNA requirements such as next-generation sequencing (NGS), and specimen types that are genetically heterogeneous such as tumors.

To understand the role of formalin fixation on NGS-based somatic variant calling, we compared matched fresh frozen (FF) and FFPE samples from different sources on a clinical NGS cancer panel. First, we simulated different formalin fixation protocols on the NA12878 reference cell line, whose genomic profile is well characterized. We then tested seven tumor FFPE samples with matched adjacent FF tissue, and assessed raw DNA quality, library quality, sequencing metrics and variant concordance profiles.

Experimental Design

The NA12878 cell line was subjected to formalin treatment, varying fixation time and time, temperature and buffering conditions (to mimic freshly prepared vs. old stored formalin buffered solutions), according to the matrix in Table 1.

Tumor specimens were split in half and the mirror-image halves were either fresh frozen (FF) or fixed in formalin and paraffin embedded (FFPE). Fixation was performed in 10% formalin (PBS-buffered) at room temperature (RT) for 16 hours or at 4C for 72 hours.

Following nucleic acid extraction and library preparation, DNA fragmentation was assessed using Agilent Tapestation. Sequencing (on Illumina Hiseq) and variant discovery/annotation were performed using the Personalis ACE CancerPlus assay, which tests 181 cancer-related genes (with solid tumor focus). For the NA12878 experiments, only the green shaded samples in Table 1 were sequenced and analyzed. Sequencing reads were aligned to human reference assembly hs37d5.

Table 1: Formalin Fixation Conditions for NA12878 Cell Line

Fixation Time	1 Day			3 Days		
Fixation Temp	RT	37C	45C	RT	37C	45C
Fixation Buffer %PBS	0	0	0	0	0	0
Fixation Buffer %PBS	20	20	20	20	20	20
Fixation Buffer %PBS	40	20	40	40	40	40
Fixation Buffer %PBS	60	60	60	60	60	60
Fixation Buffer %PBS	80	80	80	80	80	80
Fixation Buffer %PBS	100	100	100	100	100	100

Results

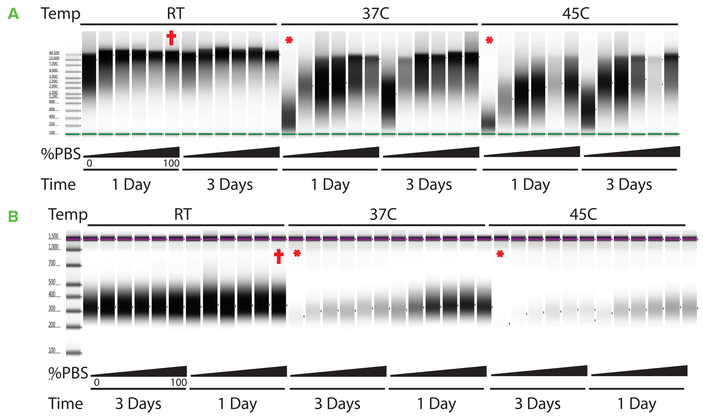


Figure 1: Formalin damage affects the quality of extracted DNA and NGS library preparations. (A) Size distribution of DNA extracted from NA12878 cells fixed with formalin according to the conditions in Table 1. Traditional fixation for 1 Day in 10% formalin buffered with 100% PBS at RT yields DNA that is high molecular weight and not fragmented (dagger). Fixation protocols that use unbuffered formalin at higher temperatures and for prolonged periods of time have highly fragmented low molecular weight DNA (asterisk). (B) Size distribution of next generation sequencing libraries prepared from the DNA extracts shown in (A). The dagger indicates a library from properly fixed NA12878 cells (daggers in (A)), asterisks indicate libraries from poor-quality DNA extracts (asterisks in (A)).

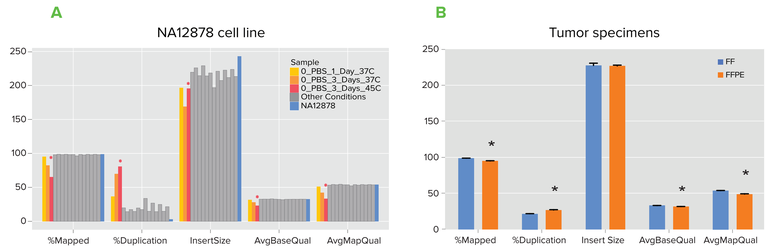
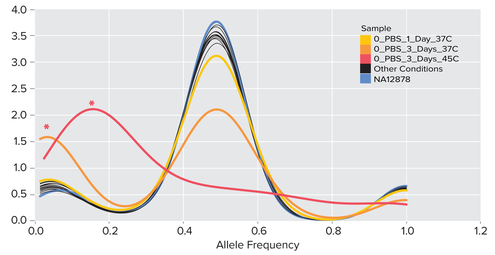


Figure 2: Formalin damage negatively impacts sequencing metrics. Samples were tested on the Personalis ACE CancerPlus platform and the following average sequencing metrics were compared among samples: mapping rates, duplication rate, insert sizes, recalibrated (GATK) base quality scores. (A) NA12878 cell line (samples shaded green in Table 1). Red asterisks indicate results from the poor quality DNA indicated with asterisks in Figure 1, contrasted with the good quality DNA (shown in blue). (B) seven tumor matched FFPE-FF pairs. * $p<0.05$.

Figure 3: Formalin damage increases low allele frequency somatic variants. Distribution of allele frequencies across NA12878 cells fixed according to the conditions shaded in green in Table 1. Asterisks indicate results from the poor quality DNA indicated with asterisks in Figure 1.



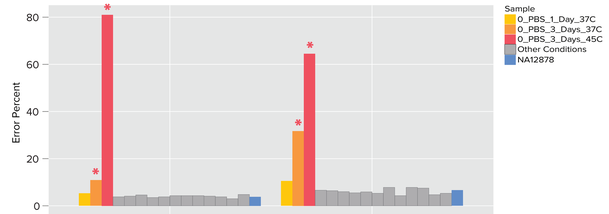


Figure 4: Formalin damage decreases somatic variant calling accuracy. False positive (FP) and false negative (FN) errors in NA12878 samples fixed according to the conditions shaded in green in Table 1. Red and asterisks indicate results from the poor quality DNA indicated with asterisks in Figure 1, contrasted with the good quality DNA (shown in blue).

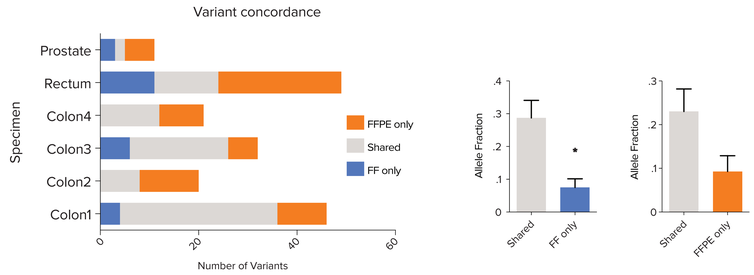


Figure 5: Formalin treatment affects somatic variant calling accuracy at low allele frequencies. (A) Somatic variant concordance between fresh frozen (FF) and formalin fixed paraffin embedded (FFPE) samples. Shown are the number of variants that are shared between matched specimens or unique to FF or FFPE conditions. (B) Average allele frequencies in FF specimens for variants that are shared with FFPE or unique to FF. * $p < 0.05$ (C) Average allele frequencies in FFPE specimens for variants that are shared with FF or unique to FFPE.

Conclusion

- Protocols that use long periods of fixation at high temperatures in unbuffered formalin adversely affect NGS-based tests
- Even specimens with well-controlled fixation conditions show mildly reduced sequencing quality metrics
- Formalin damage leads to reduced accuracy of variant calling, primarily at low allele fractions

Acknowledgements

- Amer S. Abouhamze and Erin L. Kelly from CTSI Biorepository at the University of Florida
- Members of the Personalis team

Personalis®
Pioneering Genome-Guided Medicine