Fix the Fixation: Effect of Formalin Fixation on Targeted Sequencing and Variant Calling

#3616

Ravi Alla[1], Martina Lefterova[1], Shujun Luo[1], Elena Helman[1], Michael James Clark[1], Sean Michael Boyle[1], Scott Kirk[1], Parin Sripakdeevong[1], Mirian Karbelashvili[1], Michael Snyder[1, 2], John West[1], Richard Chen[1]

[1]Personalis, Inc. | 1330 O'Brien Dr., Menlo Park, CA 94025; [2]Department of Genetics, Stanford University, Palo Alto CA.

Contact:
ravi.alla@personalis.com

Introduction

Tumor biopsies are often Formalin-Fixed and Paraffin-Embedded (FFPE) for histological staining, genetic testing and archival purposes. Formalin treatment preserves tissue by crosslinking proteins, but also leads to mutation of the nucleic acid bases and poses a challenge to identification of true variants in the tumor using next-generation sequencing (NGS) methods. Studies have shown that it is possible to isolate high quality nucleic acids from good FFPE samples and to profile small variants using NGS. These studies used tissue fixed with 10% neutral buffered formalin for 24 hours, a standard protocol in the pathology field. In our initial handling of FFPE samples we found that the quality of the isolated DNA and subsequent sequencing results vary widely. We hypothesized that this may be due to deviations from the standard protocol, such as inaccurate logging of the fixation protocol, variation in the fixation time, and varied storage conditions of the samples. To understand the role of formalin fixation on the quality of variants called, we performed an augmented target enrichment and sequencing assay on fresh frozen (FF) and formalin treated reference cell line NA12878. We assessed raw DNA quality, library quality, sequencing metrics (alignment rate, duplication rate and on-target efficiency) and variant concordance profiles between FF and formalin treated cells. We also tested 6 matched FFPE and adjacent FF tumor samples, to assess DNA and library qualities, sequencing metrics and variant concordance profiles.

Methods

NA12878 cell line was subjected to formalin treatment using various protocols, as per the matrix in **Table 1**. Protocols varied fixation time, fixation temperature and the fixation buffer conditions (which mimic fresh buffered vs old stored formalin solutions). Following sample and library prep, Personalis ACE Cancer Plus Enrichment, sequencing and variant discovery/annotation (**Figure 1**) were performed on the orange shaded samples (**Table 1**). The matched FF and FFPE tumor samples, specimen were split in half and the mirror-image halves were either flash-frozen or formalin-fixed and paraffin-embedded using standard protocols, followed by variant profiling as described.

FIXATION TIME	1 DAY			3 DAYS		
FIXATION TEMP	RT	37 °C	45 °C	RT	37 °C	45 °C
Fixation Buffer %PBS	0	0	0	0	0	0
Fixation Buffer %PBS	20	20	20	20	20	20
Fixation Buffer %PBS	40	40	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

Table 1: Formalin Fixation Conditions for NA12878 Cells.

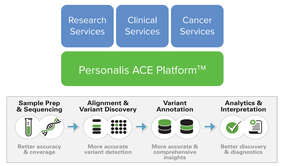

Figure 1: Personalis ACE Cancer Plus Pipeline.

Results

Formalin Damage Affects DNA and Library Qualities

Quality of DNA isolated from formalin treated NA12878 cells varies considerably between different protocols. Traditional formalin fixation protocol for 1 Day in 100% PBS buffered formalin at RT yields DNA that is high molecular weight and not fragmented (green asterisk, **Figure 2**). The library constructed from this preparation is also of good quality (green asterisk, **Figure 3**). Fixation protocols that use unbuffered formalin at higher temperatures and for prolonged periods of time have highly fragmented low molecular weight DNA (red asterisks, **Figure 2**). This poor DNA quality consequently led to poor library efficiency (red asterisks, **Figure 3**). Based on these results we believe that fixation using buffered formalin, and at the right temperature, is crucial for viable DNA for downstream NGS protocols.

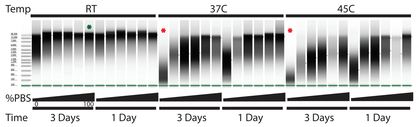

Figure 2: Raw DNA Size Distribution of Formalin Treated NA12878 Cells.

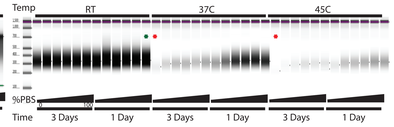

Figure 3: NGS Library Sizes of Formalin Treated NA12878 Cells.

Formalin Damaged Libraries have Low Alignment Metrics

Improperly fixed NA12878 samples show sub-par alignment metrics when compared to properly fixed and FF samples (**Figure 4**). In addition properly fixed tissue specimen have slightly lower, but still acceptable alignment metrics when compared to their matched FF counterparts (**Figure 5**).

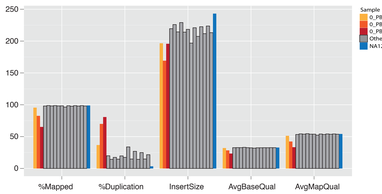

Figure 4: Alignment Metrics for FF vs Formalin treated NA12878 Cells.

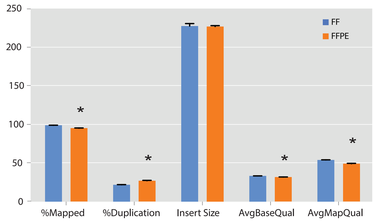

Figure 5: Alignment Metrics for matched FFPE and FF tumor specimen.

Formalin Damage Increases Low Allele Frequency Somatic Variants

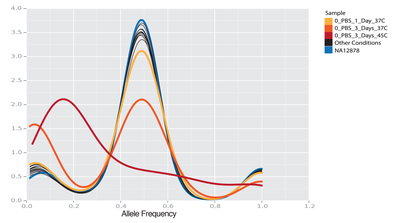

Figure 6: Allele Frequency Distribution of Somatic Variants in FF vs Formalin treated NA12878 Cells.

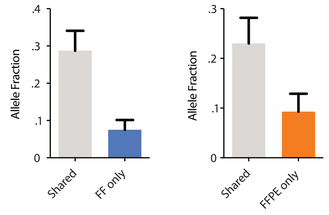

Figure 7: Average Allele Frequencies of FF vs FFPE Only Variants in matched Tumor Specimens.

Formalin Damage Decreases Somatic Variant Calling Accuracy

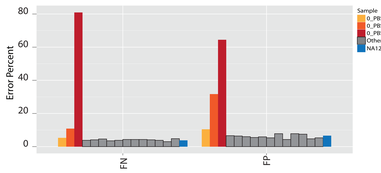

Figure 8: False Positive and False Negative Errors in Formalin Treated vs FF NA12878 Cells.

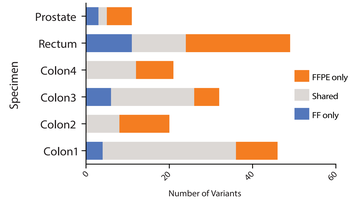

Figure 9: Somatic Variant Concordance between matched FF and FFPE Tumor Specimen.

Formalin Damage Increases Deamination (C->T) and Oxidation (G->A) Mutations

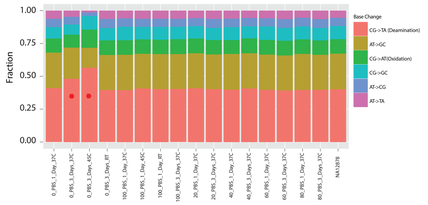

Figure 10: Somatic mutation profile in FF vs Formalin Treated NA12878 Cells.

Summary

Protocols that use long periods of fixation at high temperatures in unbuffered formalin adversely affect tissue DNA. In addition to lower DNA and library qualities, somatic variant discovery sensitivity and specificity (especially at low Allele Frequencies) are reduced too. Pathology labs which fix human tissue for histology and storage should adhere to proper fixation protocols, so that the tissue may be used in NGS diagnostic assays.

Personalis®
Pioneering Genome-Guided Medicine