Clinical Application of a Targeted Sequencing Panel for Germline Pharmacogenomic Testing in Formalin-fixed Paraffin-embedded Tumours

Because the tumour genome contains germline information, clinical laboratories can leverage tumour genomic testing to perform initial screening for clinically relevant germline variants such as variants in pharmacogenomic (PGx) genes. Subsequently, only patients with potential germline PGx variants would be subjected to downstream germline testing. While this procedure is more cost-effective than sequencing and analyzing normal DNA for every patient, the concordance of germline PGx variants between tumour and matched normal specimens must be evaluated to determine the feasibility of detecting germline PGx variants in tumour DNA.

# Background

Tumour profiling using next-generation sequencing (NGS) technologies has been increasingly integrated into oncologic care to detect targetable somatic mutations and personalize treatments for cancer patients. Although analysis of tumour-normal paired samples is required to accurately discriminate between somatic and germline variants, most clinical laboratories only sequence tumour samples to minimize cost and turnaround time . However, genomic analyses of tumours can also reveal secondary genomic findings, which are germline information that may have clinical implications for patients and their family members . In fact, several studies demonstrated that a germline cancer-predisposing variant is present in 3-10% of patients undergoing tumour-normal sequencing . Therefore, clinical laboratories providing tumour genomic testing must be equipped to perform germline confirmatory testing on potential germline variants or be prepared to refer such cases to external services.

Because the tumour genome contains germline information, clinical laboratories can leverage tumour genomic testing to perform initial screening for clinically relevant germline variants such as variants in pharmacogenomic (PGx) genes. Subsequently, a similar framework for validating secondary germline findings can be applied, in which only patients with potential germline PGx variants are subjected to downstream germline testing. This procedure for germline PGx testing is more cost-effective because it does not require processing, sequencing, and analysis of normal DNA for every patient. The ability to implement germline PGx testing at a reduced cost can significantly benefit patient care because these variants cause functional changes in drug targets and drug disposition proteins (proteins involved drug metabolism and transport), thereby contributing to inter-patient differences in chemotherapeutic response . Hence, such genomic information can be used to guide the selection of chemotherapeutic drugs and optimization of drug dosage for cancer patients, leading to improved safety and efficacy of treatment and reduced risk of toxicity .

Detection of genomic alterations in tumour DNA is also faced with technical challenges conferred by formalin-fixed paraffin-embedded (FFPE) tumour specimens . Tumour biopsies are often formalin-fixed to preserve tissue morphology for histological examination and to enable storage at room temperature; however, formalin fixation causes DNA fragmentation and base modifications, which pose difficulties in using DNA extracted from FFPE tumours for clinical genomic testing . Fragmentation damage caused by formalin fixation leads to reduced template DNA for PCR amplification, thereby affecting the efficiency of amplicon-based NGS testing . Furthermore, the degree of DNA fragmentation was shown to be higher in tissues from older FFPE blocks and tissues fixed with formalin of lower pH . Formalin fixation is also problematic because it gives rise to depurination, which generates abasic sites, and cytosine deamination resulting in CT/GA transitions . These forms of formalin-induced DNA damage contributes to the presence of sequence artifacts in FFPE specimens, which can be inaccurately identified as real genomic alterations.

In this study, we assessed the concordance of germline PGx variants between tumour and matched normal DNA by analyzing amplicon-based targeted NGS data from 213 patients with tumour-normal paired samples. While matched normal DNA was derived from peripheral blood, tumour DNA was extracted from FFPE tumour blocks; thus, we compared the quality metrics of sequencing data between tumour and blood specimens and evaluated the prevalence of formalin-induced DNA damages to address the impact of formalin fixation on amplicon-based NGS testing. We demonstrated that germline PGx variants can be identified with high sensitivity and precision in FFPE tumour DNA using a clinical targeted sequencing panel.

# Methods

## Patient Samples

Blood and FFPE tumour specimens were acquired from 213 patients who provided informed consent for The OncoPanel Pilot (TOP) study, a pilot study to optimize the OncoPanel, which is an amplicon-based targeted NGS panel for solid tumours, and assess its application for guiding disease management and therapeutic intervention. Patients in the TOP study are those with advanced cancers including colorectal cancer, lung cancer, melanoma, gastrointestinal stromal tumor (GIST), and other cancers (). The age of paraffin block for tumour specimens ranges from 18 to 5356 days with a median of 274 days.

## OncoPanel (Solid Tumour NGS Panel)

The OncoPanel is offered by the British Columbia Cancer Agency (BCCA) for clinical genomic testing of coding exons and clinically relevant hotpots of 20 cancer-related genes and six PGx genes: *DPYD*, *GSTP1*, *MTHFR*, *TYMP*, *TYMS*, and *UGT1A1*. Full list of genes and gene reference models for the OncoPanel is presented in . Primers were designed by RainDance Technologies (Lexington, MA) using the GRCh37/hg19 reference sequence to generate 416 amplicons between 100 bp and 250 bp in size, which interrogate 20 kb of target bases. Target regions of the six PGx genes in the OncoPanel were assayed by 49/416 amplicons. Complete lists of primers and amplicons are provided in the Supplemental Materials.

## Sample Preparation, Library Construction, and Illumina Sequencing

Genomic DNA was extracted from blood and FFPE tumour specimens using the Gentra Autopure LS DNA preparation platform and QIAamp DNA FFPE tissue kit (Qiagen, Hilden, Germany) respectively. The extracted DNA was sheared according to a previously described protocol to attain approximate sizes of 3 kb followed by PCR primer merging, amplification of target regions, and adapter ligation using the Thunderstorm NGS Targeted Enrichment System (RainDance Technologies, Lexington, MA) as per manufacturer’s protocol. Barcoded amplicons were sequenced with the Illumina MiSeq system for paired end sequencing with a v2 250-bp kit (Illumina, San Diego, CA).

## Variant Calling Pipeline

Reads that passed the Illumina Chastity filter were aligned to the GRCh37/hg19 human reference genome using the BWA mem algorithm (version 0.5.9) with default parameters, and the alignments were processed and converted to the BAM format using SAMtools (version 0.1.18). Variant calling was performed with the SAMtools mpileup function (samtools mpileup -BA -d 500000 -L 500000 -q 1) to generate pileup files for all target bases followed by the VarScan2 mpileup2cns (version 2.3.6) function with parameter thresholds of variant allele frequency 10% and Phred base quality 20 (–min-var-freq 0.1 –p-value 0.01 –strand-filter 0 –output-vcf –variants –min-avg-qual 20). Variant calls were filtered using the VarScan2 fpfilter function with fraction of variant reads from each strand 0.1 and default thresholds for other parameters. SnpEff (version 4.2) was used for variant annotation and effect prediction whereas the SnpSift package in SnpEff was used to annotate variants with databases such as dbSNP (b138), COSMIC (version 70), 1000 Genomes Project, ClinVar, and ExAC (release 0.3) for interpretation.

## Data Analysis

Specifics on how each analysis was performed.

# Results

## Formalin fixation affects amplicon generation and sequencing

## Coverage depth and quality are highly variable in FFPE tumours

## Formalin-induced sequence artifacts are prevalent at low variant allele frequency

## Germline PGx variants can be detected with high sensitivity and precision in FFPE tumours

## Discordant germline PGx variants are caused by low coverage sites

# Discussion

# Conclusions

# Competing interests

The authors declare that they have no competing interests.

# Author’s contributions

Text for this section …

# Acknowledgements

Text for this section …

# Figures

# Tables

Distribution of cancer types in the TOP cohort.

|  |  |  |  |
| --- | --- | --- | --- |
| Cancer Type | Number of Cases | Percentage (%) |  |
| Colorectal | 97 | 46 |  |
| Lung | 59 | 28 |  |
| Melanoma | 18 | 8 |  |
| Other\* | 17 | 8 |  |
| GIST | 7 | 3 |  |
| Sarcoma | 4 | 2 |  |
| Neuroendocrine | 4 | 2 |  |
| Cervical | 2 | 0.9 |  |
| Ovarian | 2 | 0.9 |  |
| Breast | 2 | 0.9 |  |
| Unknown | 1 | 0.5 |  |

This category includes thyroid, peritoneum, sarcomatoid carcinoma of lung, Fallopian tube, gastric, endometrial, anal, salivary gland, pancreas, and small bowel cancers.

Gene Reference Models for Genes in the OncoPanel.

|  |  |  |
| --- | --- | --- |
| Gene | Protein | Reference Model |
| AKT1 | Protein kinase B | NM\_001014431.1 |
| ALK | Anaplastic lymphoma receptor tyrosine kinase | NM\_004304.3 |
| BRAF | Serine/threonine-protein kinase B-Raf | NM\_004333.4 |
| DPYD | Dihydropyrimidine dehydrogenase | NM\_000110.3 |
| EGFR | Epidermal growth factor receptor | NM\_005228.3 |
| ERBB2 | Receptor tyrosine-protein kinase erbB-2 | NM\_001005862.1 |
| GSTP1 | Glutathione S-rransferase pi 1 | NM\_000852.3 |
| HRAS | GTPase HRas | NM\_005343.2 |
| IDH1 | Isocitrate dehydrogenase 1 | NM\_005896.2 |
| IDH2 | Isocitrate dehydrogenase 2 | NM\_002168.2 |
| KIT | Tyrosine-protein kinase Kit | NM\_000222.2 |
| KRAS | KRas proto-oncogene GTPase | NM\_033360.2 |
| MAPK1 | Mitogen-activated protein kinase 1 | NM\_002745.4 |
| MAP2K1 | Mitogen-activated protein kinase kinase 1 | NM\_002755.3 |
| MTHFR | Methylenetetrahydrofolate reductase | NM\_005957.4 |
| MTOR | Serine/threonine-protein kinase mTOR | NM\_004958.3 |
| NRAS | Neuroblastoma RAS viral oncogene homolog | NM\_002524.3 |
| PDGFRA | Platelet-derived growth factor receptor alpha | NM\_006206.4 |
| PIK3CA | Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha | NM\_006218.2 |
| PTEN | Phosphatase and tensin homolog | NM\_000314.4 |
| STAT1 | Signal transducer and activator of transcription 1 | NM\_007315.3 |
| STAT3 | Signal transducer and activator of transcription 3 | NM\_139276.2 |
| TP53 | Tumor protein P53 | NM\_000546.5 |
| TYMP | Thymidine phosphorylase | NM\_001113755.2 |
| TYMS | Thymidylate synthetase | NM\_001071.2 |
| UGT1A1 | Uridine diphosphate (UDP)-glucuronosyl transferase 1A1 | NM\_000463.2 |

# Additional Files

## Additional file 1 — Sample additional file title

Additional file descriptions text (including details of how to view the file, if it is in a non-standard format or the file extension). This might refer to a multi-page table or a figure.

## Additional file 2 — Sample additional file title

Additional file descriptions text.