GERMLINE PHARMACOGENOMICS TESTING IN FORMALIN-FIXED PARAFFIN-EMBEDDED TUMOURS

by

Shyong Quin Yap

B.Sc. (Hons), Trent University, 2011

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Experimental Medicine Program)

The University of British Columbia
(Vancouver)

January 2017

Abstract

This document provides brief instructions for using the `ubcdiss` class to write a **UBC!**-conformant dissertation in LaTeX. This document is itself written using the `ubcdiss` class and is intended to serve as an example of writing a dissertation in LaTeX. This document has embedded Unique Resource Locators (URLS) and is intended to be viewed using a computer-based Portable Document Format (PDF) reader.

Note: Abstracts should generally try to avoid using acronyms.

Note: at **UBC!** (**UBC!**), both the Graduate and Postdoctoral Studies (GPS) Ph.D. defence programme and the Library's online submission system restricts abstracts to 350 words.

Preface

At **UBC!**, a preface may be required. Be sure to check the GPS guidelines as they may have specific content to be included.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

GPS Graduate and Postdoctoral Studies

PDF Portable Document Format

URL Unique Resource Locator, used to describe a means for obtaining some resource on the world wide web

Acknowledgments

Although this thesis only bears one name, its completion would be impossible without the contribution of many individuals. First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. Aly Karsan, for the opportunity to work with his team of diverse talents as well as his patience, guidance, and extensive knowledge in clinical informatics.

I would also like to thank my labmates from the bioinformatics team, Kieran and Rod, and members from the Centre of Clinical Genomics, Liz and Jill, for their insightful comments and help throughout my time in the lab. I would like to extend my gratitude to my supervisory committee members, Dr. Ryan Morin and Dr. Martin Hirst, for their knowledgeable feedback and continuous effort in asking me difficult questions which motivated me to widen my research perspective.

My sincere thanks also goes to my friends, who supported me and lifted my spirits through the tough times. Last but not least, I would like to thank my family: mom and dad, for always encouraging my interest in science and listening to my endless science talks and my sisters, for believing in my ability even when I doubt myself. This thesis is yours as much as it is mine.

Chapter 1

Introduction

1.1 Overcoming The Clinical Reality

Germline pharmacogenomic (PGx) variants can influence a patient's response to chemotherapy. Using next-generation sequencing (NGS) technologies, PGx variants can be screened to identify patients who are susceptible to toxicity risk, thereby preventing chemotherapy-associated morbidities. However, clinical NGS testing in oncology is challenging due to the lack of matched normal DNA while tumour biopsies are formalin-fixed paraffin-embedded (FFPE) for histologic examinations. Formalin induces DNA fragmentation and sequence artifacts, specifically C>T/A>G base transitions. Moreover, the ability of NGS approaches to interrogate genomic content at increased depth and breadth can result in detection of variant of unknown significance (VUS) and incidental findings of medical value. At present, there are limited guidelines available for management of these variant categories. Despite these challenges, would it be possible/feasible to harness a clinical targeted NGS panel that only sequence FFPE tumour DNA for germline PGx testing?

The work presented herein aims to address the aforementioned challenges as well as answer this question.

1.2 Genomics-Driven Oncology

1.2.1 Definitions

Genomics-driven oncology is defined as the use of genomic information to provide guidance for disease management and therapeutic intervention in oncologic care. The application of NGS to oncology, or genomics-driven cancer medicine, is conceptually logical and simple: First, the genome of a patients tumor is sequenced, and all genetic differences from the standard human reference genome are identified. Because all human beings have many normal genetic variants that differ from the reference genome, the tumor sequence is compared with the patients constitutional (germline) genome

to determine which alterations in the tumor are somatic (and therefore potentially pathogenic) and which are germline (and probably not cancer-related). Next, the somatic mutation list is filtered through a database of mutations that may render tumors sensitive to established and emerging anticancer drugs. Finally, an annotated list is provided to the treating physician to be used in clinical decision making and clinical research design. However, several technical and ethical challenges must be addressed before real-time application of NGS can become a reality in cancer medicine.

1.2.2 Genomic Alterations in Cancer Pathogenesis

1.2.3 Clinical Deployment of Targeted Cancer Therapies

1.2.4 Advances in DNA Sequencing Technologies

1.3 Next-Generation Sequencing Technologies

The Human Genome Project, which assembled the first human reference genome, was completed in 2003 at an expense of $2.7 billion within 13 years. This was a cost and turn-around time that would not be feasible for routine usage in research and clinical settings, thereby stimulating the advancement of DNA sequencing technologies. As a result, various MPS technologies have emerged and have been adapted to fit the needs of scientific research as well as clinical applications.

1.3.1 Sequencing by Ligation

1.3.2 Sequencing by Synthesis

1.4 Objectives

Current research in cancer genomics primarily focus on somatically acquired mutations that drive malignant transformation through conferring selective growth advantages to cells. These efforts are demonstrated by formation of large-scale collaborations such as the The Cancer Genome Atlas (TCGA) and the International Cancer Genome Consortium (ICGC), which aim to characterize and catalog the genomic landscapes of diverse tumour types. Understanding oncogenic mechanisms underlying driver somatic mutations have led to the development of targeted therapies, which resulted in improved clinical outcomes for various cancer subtypes. However, germline genetic variants can also influence cancer treatment by affecting drug targets and disposition, thereby causing interpatient differences in drug response. These germline variants, known as pharmacogenomic (PGx) variants, can assist with treatment selection, optimal drug dosing, and identifying toxicity risk to reduce cancer therapeutics-associated morbidities.

Advances in massively parallel sequencing (MPS) technologies have revolutionized genetic testing in clinical oncology through enabling surveillance of increased genomic depth and breadth with less DNA in a cost-effective and timely manner. Nevertheless, clinical application of MPS approaches to cancer medicine still encounter several challenges and financial barriers. One of these challenges is caused by formalin fixation of tumour biopsies. Tumour biopsies are routinely formalin-fixed paraffin-embedded (FFPE) to preserve morphology and cellular characteristics for histologic examination. Moreover, most clinical laboratories prefer storage of FFPE blocks at ambient temperature to avoid cost inflicted by maintaining fresh-frozen specimens. However, formalin fixation causes DNA fragmentation and base transition artifacts, which could result in false-negative or false-positive variant calls. These sequence artifacts are particularly concerning in a clinical setting because failure to detect or inaccurate detection of cancer biomarkers could have devastating consequences for patients and their families.

Another challenge in clinical MPS-based testing in oncology practice is the lack of matched normal DNA, which is not commonly collected in the clinic due to increased cost and logistical difficulties. Without matched normal DNA, determining the somatic or germline nature of the variant calls, which is essential for translating MPS data into clinically actionable information, rely heavily on filtering and interpretation using databases such as dbSNP, ExAC, and COSMIC. The bottleneck of MPS data generation to interpretation for clinical use is yet another limiting factor of clinical genomic sequencing. Despite the ability of MPS approaches to screen increased genomic content, these methods lead to higher rates of detecting variants of uncertain significance (VUS) that lack evidence of clinical utility. Conversely, incidental findings with medical value to patient care may arise while there are ethical controversies and very few guidelines on the management of this category of variants.

The main objective of this thesis is to investigate whether germline PGx variants can be accurately and sensitively detected in FFPE tumour DNA sequenced by a clinical targeted MPS panel. To achieve this objective, key challenges in clinical genomic sequencing that were briefly described above were addressed. This introductory chapter is organized into five sections to provide the necessary background knowledge: (1) Describes driving forces that led to emergence of genomics-driven oncology; (2) Introduces different applications of MPS to provide an overview of technologies behind sequence data generation; (3) Introduces bioinformatics pipeline for variant calling, which generated input data analyzed in this thesis; (4) Expands on key challenges in clinical genomic sequencing and potential solutions; (5) Emphasizes on the importance of germline PGx testing in oncology care.

1.5 Genomics-Driven Oncology

1.5.1 Definitions

Genomics-driven oncology is defined as the use of genomic information to provide guidance for disease management and therapeutic intervention in oncologic care. The application of NGS to oncology, or genomics-driven cancer medicine, is conceptually logical and simple: First, the genome of a patients tumor is sequenced, and all genetic differences from the standard human reference genome are identified. Because all human beings have many normal genetic variants that differ from the reference genome, the tumor sequence is compared with the patients constitutional (germline) genome to determine which alterations in the tumor are somatic (and therefore potentially pathogenic) and which are germline (and probably not cancer-related). Next, the somatic mutation list is filtered through a database of mutations that may render tumors sensitive to established and emerging anticancer drugs. Finally, an annotated list is provided to the treating physician to be used in clinical decision making and clinical research design. However, several technical and ethical challenges must be addressed before real-time application of NGS can become a reality in cancer medicine.

1.5.2 Genomic Alterations in Cancer Pathogenesis

1.5.3 Clinical Deployment of Targeted Cancer Therapies

1.5.4 Advances in DNA Sequencing Technologies

The first human reference genome was established in 2003 through completion of the Human Genome Project, which instigated major developments in DNA sequencing technologies and computational tools for large-scale genomic data analysis. As a result, MPS technologies have emerged with increased throughput, sensitivity, and cost efficiency, leading to the genomic characterization of a growing number of tumours. As the application of MPS technologies in cancer genomic studies continued to accelerate the progress of driver gene discoveries and drug development, the clinically feasible features of MPS have led to its rapid integration in oncology practice, giving rise to the genomics-driven oncology framework.

Genomics-driven oncology is defined as the use of genomic information to provide guidance for disease management and therapeutic intervention in oncologic care. One of the driving forces of this emerging approach is the expanding knowledge in tumour biology. A central focus of tumour biology research is elucidating oncogenic mechanisms driven by somatic mutations that confer selective growth advantages to cells. Translation of these findings into targeted therapies have demonstrated pronounced improvement in clinical outcomes, leading to the transition from morphology-based to genetic-based management of cancer. A well-known example is the treatment of BCR-ABL-

translocated chronic myeloid leukemia (CML) with the tyrosine kinase inhibitor imatinib, which targets the constitutively active ABL1 kinase as a result of the BCR-ABL fusion gene.

Several successful applications of targeted therapy ensued the example of imatinib and BCR-ABL-translocated CML such as the use of anti-HER2 monoclonal antibody trastuzumab in treating HER2/neu-amplified breast cancer and BRAF inhibitor vemurafenib in treating advanced BRAF-mutated melanoma. The promising potential of targeted anti-cancer agents accelerated the progress of drug discovery and development as evident by the drastically decreasing timelines between driver mutation discovery and clinical proof-of-concept. For instance, it only took three years for the ALK inhibitor crizotinib to enter Phase II clinical trials since identification of ALK translocations in non-small cell lung carcinoma (NSCLC) whereas the Food and Drug Administration (FDA) approval of imatinib for treatment of BCR-ABL-translocated CML took 41 years since discovery of the Philadelphia chromosome. Consequently, there is an extensive compendium of targeted therapeutics with 19 listed as clinically approved by the National Cancer Institute in 2012 while approximately 150 compounds were listed as clinical candidates.

The enhanced understanding of oncogenic pathways and growing spectrum of targeted therapies have created the perfect opportunity for clinical screening of driver mutations to match patients with targeted treatments. Conversely, patients without specific mutations could be spared treatment-associated toxicities. For example, screening for KRAS mutations in codon 12 or 13 could prevent treating colorectal cancer (CRC) patients with anti-EGFR monoclonal antibody or EGFR inhibitors, which are associated with toxicity risk, as these patients are known to respond poorly. Despite the initial efficacy of targeted treatments, tumours could develop resistant mechanisms causing cancer relapse. One of the crucial realization from proceeding studies is that cancer is a heterogenous disease, in that a tumour can consist of multiple subclonal populations and resistant cancer cells may already exist at an early stage. Hence, treatment resistance occurs after the dominant clone has been wiped out allowing the resistant subclone to proliferate and metastasize. Complexities derived from the vast mutational profiles of tumours and intratumoural heterogeneity revealed that surveillance of multiple cancer genes with increased coverage depth during the course of a disease is essential for positive clinical outcomes. To achieve this, single gene assays using the Sanger method, also known as the dideoxynucleotide chain termination method, are not feasible due to time, labour, and cost constraints.

With the advent of MPS technologies, genome sequencing can be accomplished at a reduced cost of less than $5000 per genome within days. To put this into perspective, the first human reference genome sequenced using the Sanger method was completed at a cost of $2.7 billion over 13 years. Advances in MPS technologies have revolutionized cancer genomics by enabling international consortia such as TCGA and the ICGC to uncover the complex genomic architectures of various tumour types, thereby shedding insights into drug resistant mechanisms and potentiating therapeutic strategies against cancer relapse. The high-throughput nature of MPS and its ability

to generate robust genomic information in a timely and cost-efficient manner are also capable of overcoming the limitations of single gene assays in the clinic. Thus, various MPS approaches, most commonly targeted gene panels, have been rapidly adopted in clinical oncology to inform medical decision-making based on a patient's genomic make-up. Although the path to genomics-driven oncology was paved by a deeper mechanistic understanding of oncogenic pathways and the accelerated progress in targeted therapeutics development, the emergence of MPS technologies played a significant role in providing detailed insights into the cancer genome as well as a feasible method to generate genomic information for clinical use.

1.6 Applications of Massively Parallel Sequencing

1.6.1 Targeted Resequencing

Capture-based, amplicon-based etc.

1.6.2 Whole Exome Sequencing

1.6.3 Whole Genome Sequencing

1.7 Bioinformatics Tools for Variant Calling

1.7.1 Types of Genomic Alterations

There are different types of genomic alterations.

1.7.2 Variant Calling Pipeline

1.7.3 Variant Calling Algorithms

1.7.4 Variant Curation and Interpretation

1.8 Challenges in Clinical Genomics

1.8.1 DNA Damage by Formalin Fixation

- Fragmentation - Transition vs. transversion

1.8.2 Lack of Matched Normal DNA

1.8.3 Variant of Unknown Significance

1.8.4 Incidental Findings

1.9 Pharmacogenomics in Clinical Oncology

Cancer biomarkers, which are central to the genomics-driven oncology approach to medical decision-making, can be classified as diagnostic, prognostic, predictive, and pharmacogenomics (PGx). PGx markers are germline genetic variants that affect genes encoding drug targets as well as drug disposition proteins involved in absorption, distribution, metabolism, and excretion (ADME). Pharmacogenomics (PGx) applies genomic approaches to evaluate the interaction of genetic variants with drug response. These variations affect . The goals of PGx studies are to elucidate biological mechanisms underlying interpatient variability in drug efficacy and toxicity as well as identify PGx biomarkers with clinical utility, which would guide selection of treatment type, optimal dosage, and duration.

Cancer PGx takes into account tumour-associated somatic mutations and germline variants. Somatic mutations in driver genes promote malignant transformation through conferring selective growth advantage to the cells. Characterization of somatic driver mutations has provided an avenue for development of molecularly targeted drugs against specific tumour-defining somatic mutations. Hence, screening for these specific

somatic mutations serving as genomic predictors of tumour response and providing new leads for drug development germline variants optimize cancer drug dosing and predict the susceptibility of patients to the adverse side effects of these drugs - knowledge that can be used to improve benefit:risk ratio of cancer treatment for individual patients

1.9.1 Targeted Therapies

Tamoxifen etc.

1.9.2 Chemotherapy-Associated Morbidities

DPYD, MTHFR, GSTP1, TYMP, TYMS, UGT1A1

1.10 Summary

The advent of MPS technologies has refined analysis of the cancer genome at base-pair resolution,

Chapter 2

Materials and Methods

2.1 Patient Samples

Solid tumour biopsies were collected from 171 consented patients in The Oncopanel Pilot (TOP) study under a protocol approved by the British Columbia Cancer Agency (BCCA) Research Ethics Board (Protocol H12-00292). Details on tumour types are listed in Table 2.1. Excess tissue from tumuor biopsies after pathological evaluation were transported in serum-free medium at ambient temperature followed by washing in PBS with 6.5mM dithiothreitol (DTT). Fat, muscle, and necrotic areas were trimmed and approximately 50-100 mg of specimens were formalin-fixed at 4°C overnight before paraffin-embedding. All specimens were fixed within 2 hours and embedded within 24 hours of collection. For each patient in the TOP cohort, peripheral blood (PB) samples were also collected and processed to serve as germline DNA control for variant calling.

Table 2.1: Summary of FFPE Specimens from 171 TOP Patients

Tumour Type	Number of Specimens	Percentage
Colorectal	85	49.7
Lung	36	21.1
Melanoma	16	9.4
Other	30	17.5
Unknown	6	3.5

2.2 DNA Extraction, Library Preparation, and Illumina Sequencing

Tumour and germline DNA for 171 patients were extracted using the QIAGEN FFPE DNA extraction kit and xx respectively as per manufacturer's instructions. For specimens with sufficient DNA quantity, 250 ng of genomic DNA was used for library preparation. Genomic DNA was sheared

to generate fragment sizes of approximately 3 kbp, followed by PCR primer merging, amplicon generation, and adapter ligation using the RainDance Thunderstorm instrument (Figure 2.1). The complete list of primers used to generate the 429 amplicons for 26 genes screened by the OncoPanel is included in **??**. Six out of 26 genes are PGx genes namely *DPYD*, *GSTP1*, *MTHFR*, *TYMP*, *TYMS*, and *UGT1A1*. Libraries were pooled, ranging from 13-20 libraries per pool, and sequenced with the Illumina MiSeq system for paired end sequencing with a v2 250-bp kit. Pooling of libraries includes libraries from other studies, which is summarized in Figure 2.2.

2.3 Variant Calling Pipeline

Read alignment and variant calling were carried out by the BCCA Centre of Clinical Genomics (CCG) bioinformatics pipeline Figure 2.3. Raw reads from the MiSeq instrument were aligned to the GRCh37 human reference genome (hg19) using BWA (version 0.5.9, mem algorithm) and variant calling was performed using samtools mpileup (version 0.1.18) followed by VarScan2 (version 2.3.6). Variant calling in the six PGx genes were carried out using the following VarScan2 parameters: Variants were annotated as per the Human Genome Variation Society (HGVS) convention and interpreted with databases such as dbSNP, ExAC, COSMIC, and ClinVar using SnpEff (version 4.2). Gene reference models used for HGVS nomenclature are listed in Table A.2.

2.4 Data Analysis and Visualization

Coverage depth was measured using bedtools (version 2.25.0) and per-base metrics were obtained using bam-readcount (https://github.com/genome/). Statistical analyses and data visualization were performed using R (version 3.3.2) and associated open-source packages. Manual review of PGx variants were carried out using the Intergrative Genomics Viewer (IGV, version 2.3). *Note: be more specific on how the data is generated*

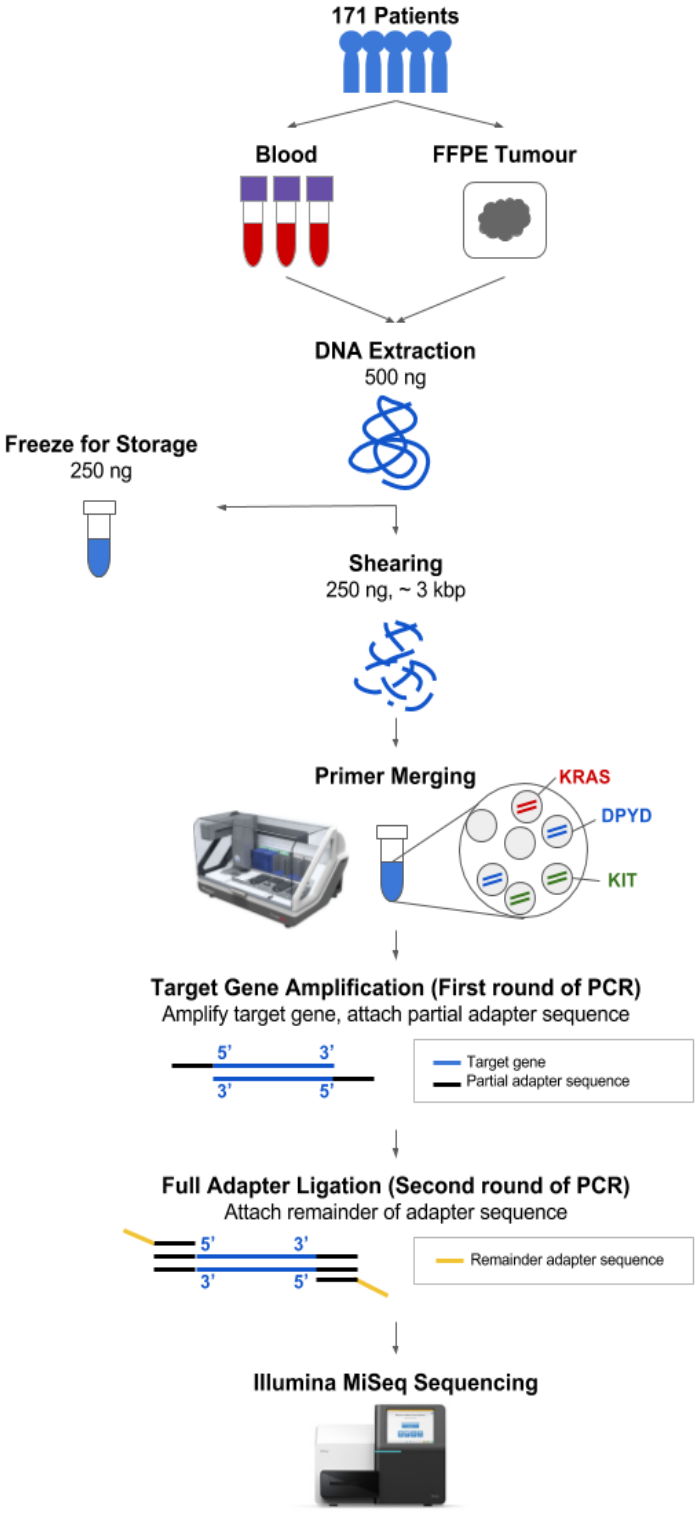

Figure 2.1: Workflow for Sample Processing, Library Preparation, and NGS Sequencing.

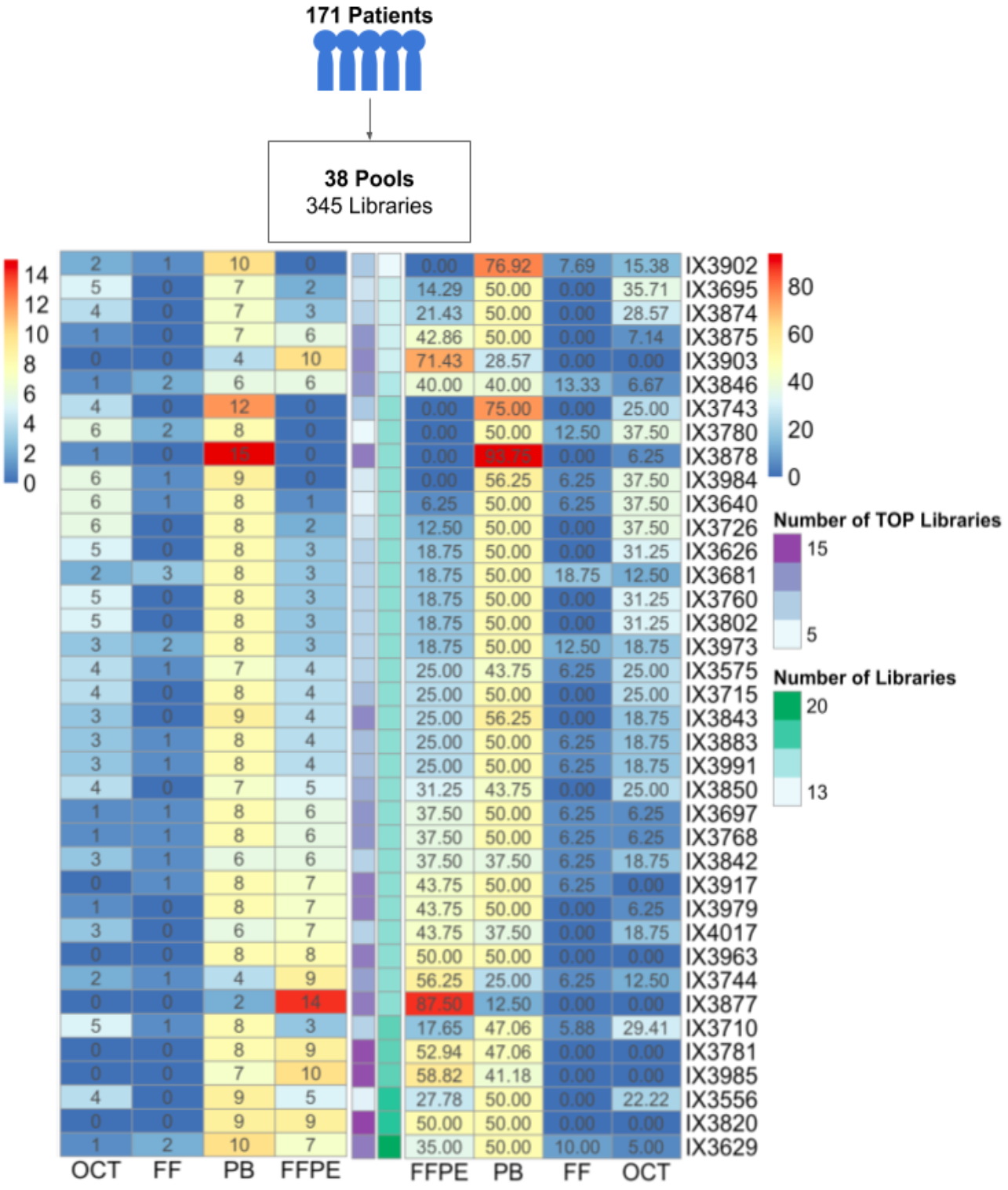

Figure 2.2: 345 TOP Libraries Distributed Across 38 Pools. Number of libraries is presented on the *left* whereas percentage of libraries per total libraries in a pool is presented on the *right*.

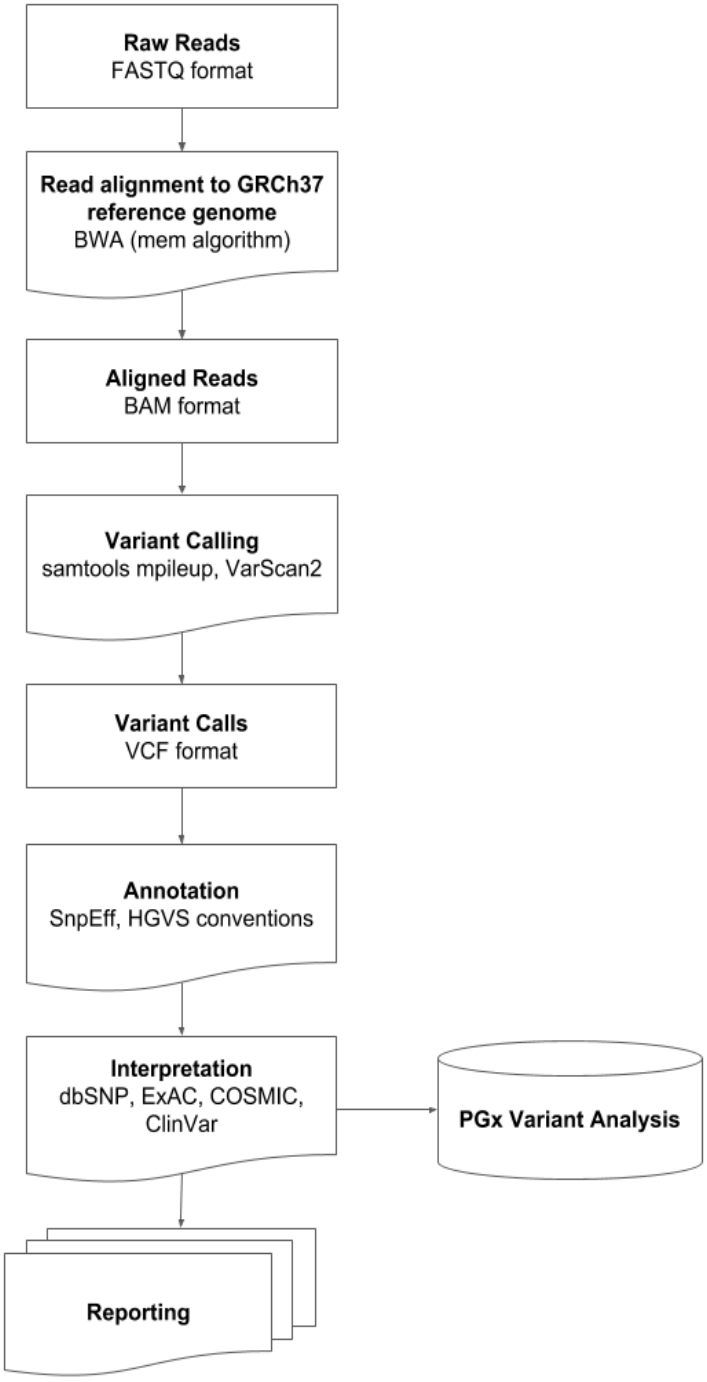

Figure 2.3: OncoPanel Pipeline for Variant Calling. Variants in PGx genes were filtered for downstream analysis.

Chapter 3

Results

3.1 Sequencing Depth is Comparable between FFPE and FF-PB Samples

Chapter 4

Discussion and Conclusion

Bibliography

Appendix A

Supporting Materials

Table A.1: A simple longtable example

First entry	Second entry	Third entry	Fourth entry
1	2	3	4
1	2	3	4
1	2	3	4
1	2	3	4
1	2	3	4
1	2	3	4
1	2	3	4
1	2	3	4
1	2	3	4
1	2	3	4
1	2	3	4
1	2	3	4
1	2	3	4

Gene Target ID	Chromosome	Target Start	Target End	Amplicon Start	Amplicon End	Forward Primer	Reverse
AKT1_b	14	105246450	105246458	105246450	105246458	GCCACAGAGAAGTTGTTGAG	GTACAT
AKT1_a	14	105246546	105246554	105246546	105246554	TCTTGAGGAGGAAGTAGCGT	AGGCA
ALK_a	2	29445212	29445271	29445212	29445271	GGAGATATCGATCTGTTAGAAACC	CCCAC
ALK_b	2	29443629	29443703	29443629	29443703	GGACTTGAGGTCTCCCC	CTTTGT
ALK_c	2	29446148	29448519	29446148	29446365	CCCATAGGGAGGGCTCTG	GCACC
ALK_c	2	29446148	29448519	29446336	29446587	GGAGCTTGCTCAGCTTGTA	CCATAT
ALK_c	2	29446148	29448519	29446519	29446774	CTGAGCTCTGAACCTTTCCA	TACTTC
ALK_c	2	29446148	29448519	29446696	29446959	ACTGGAGATGGGATTAGACC	CCCAA
ALK_c	2	29446148	29448519	29446940	29447186	CCATGACTCCCAGGAATTGG	GAAAC
ALK_c	2	29446148	29448519	29447129	29447369	AATTGGTGTCTGGGGATCTG	TTTGC
ALK_c	2	29446148	29448519	29447344	29447591	CCTAAAGAGCTCTACCAATGTG	GCTAA
ALK_c	2	29446148	29448519	29447522	29447779	CCTCCTCTATGCAATGGACC	CTCTTA

ALK_c	2	29446148	29448519	29447745	29447987	CTCAAGAGCCTTTCCCTCTG	TTCTAC
ALK_c	2	29446148	29448519	29447910	29448162	TCCTATCTCTCTGCCTGGAG	CAAGC
ALK_c	2	29446148	29448519	29448101	29448366	AGGAGGATACACACGGGG	GGTGA
ALK_c	2	29446148	29448519	29448288	29448519	ACACTATTCAGTCCTGCCTT	AAGTG
BRAF_a	7	140481395	140481418	140481395	140481418	ACTTACCATGCCACTTTCCC	TCGAG
BRAF_b	7	140453129	140453152	140453129	140453152	ATGGATCCAGACAACTGTTC	TTTTCC
BRAF_b	7	140453085	140453218	140453085	140453218	AAAATAGCCTCAATTCTTACCATC	TCATAA
DPYD_a	1	97915607	97915621	97915607	97915621	TCACCAACTTATGCCAATTCT	GAAAA
DPYD_a	1	97915608	97915621	97915608	97915621	ACATTCACCAACTTATGCCAA	CTGGA
DPYD_b	1	98348878	98348892	98348878	98348892	TCTTGTCTAATTTCTTGGCCG	CCTGG
DPYD_b	1	98348855	98349048	98348855	98349021	CAGTGGTACTTACAAAGCAGTT	GACAA
DPYD_c	1	97981336	97981350	97981336	97981350	AAAGTTTTGGTGAGGGCAAA	CTCCA
DPYD_c	1	97981418	97981425	97981418	97981425	GGTGAGGGCAAAACCCC	GCAGT
DPYD_d	1	97547940	97547954	97547940	97547954	AGCCAGAATCATTACAGGTCA	TGAGC
DPYD_d	1	97547941	97547954	97547941	97547954	TTCCAGCAGGATTCTTACCT	CAACG
DPYD_e	1	97981417	97981425	97981417	97981425	CGAATCATTGATGTGCTGGT	TTTCTC
DPYD_f	1	98039415	98039423	98039415	98039423	CCAGCACTGTACCTTTAGGA	TTGCTA
DPYD_f	1	98039416	98039423	98039416	98039423	GCTCCCAGCACTGTACC	AGGTG
DPYD_g	1	97770916	97770924	97770916	97770924	AATCCCATCAGACCTGAGAC	TTTTTC
DPYD_h	1	98165087	98165095	98165087	98165095	CAACTTATACTTGCAGGCCC	ACCATC
DPYD_h	1	98165088	98165095	98165088	98165095	AGGCCCAGCACCAAAAAG	ATTTTA
EGFR_18	7	55241614	55241736	55241614	55241733	TGACCCTTGTCTCTGTGTTC	TGCCA
EGFR_18	7	55241614	55241736	55241646	55241736	TTACACCCAGTGGAGAAGC	TATACA
EGFR_18	7	55241599	55241751	55241599	55241751	GAGGTGACCCTTGTCTCTG	AGACC
EGFR_19	7	55242415	55242513	55242415	55242513	GCATGTGGCACCATCTC	CCCCA
EGFR_19	7	55242400	55242528	55242400	55242528	CAGCATGTGGCACCATC	GCCATC
EGFR_20	7	55248986	55249171	55248986	55249101	CACTGACGTGCCTCTCC	CTTTGT
EGFR_20	7	55248986	55249171	55249054	55249171	GGCATCTGCCTCACCTC	GGATCC
EGFR_20	7	55248971	55249186	55248971	55249186	CCTCCTTCTGGCCACC	GATCCT
EGFR_21	7	55259412	55259567	55259412	55259545	CCCATGATGATCTGTCCCTC	GCCTCC
EGFR_21	7	55259412	55259567	55259508	55259567	CAGCATGTCAAGATCACAGAT	CCCAG
EGFR_21	7	55259397	55259582	55259397	55259582	CCCATGATGATCTGTCCCTC	CTGGTC
ERBB2_20	17	37880979	37881164	37880979	37881136	CTCAGCGTACCCTTGTCC	TCTGCA
ERBB2_20	17	37880979	37881164	37881030	37881164	CATATGTCTCCCGCCTTCTG	CAAAG
GSTP1_a	11	67352682	67352696	67352682	67352696	TGAATGACGGCGTGGAG	GCCCCT
HRAS_a	11	534281	534293	534281	534293	TCACCTCTATAGTGGGGTCG	ACGGA
HRAS_a	11	534282	534293	534282	534293	GGGTCGTATTCGTCCACA	GGAGC
HRAS_b	11	533870	533884	533870	533884	TACTGGTGGATGTCCTCAAA	GTGGTC
HRAS_b	11	533871	533884	533871	533884	CATGTACTGGTCCCGCAT	GATTCC
HRAS_c	11	533549	533557	533549	533557	AGGTCTCGATGTAGGGGATG	GATCAA
HRAS_c	11	533550	533557	533550	533557	TAGCTTCGGGCGAGGT	TCAAAC
HRAS_d	11	533462	533470	533462	533470	GTCAAGGGAGAGGGTCAGT	GCATCC
HRAS_d	11	533463	533470	533463	533470	GTCAGTGAGTGCTGCTCC	CTGTGC

IDH1_a	2	209113110	209113113	209113110	209113113	ACATGACTTACTTGATCCCCA	CCAAA
IDH1_a	2	209113111	209113113	209113111	209113113	ACATGACTTACTTGATCCCCA	TGGAA
IDH2_a	15	90631836	90631839	90631836	90631839	TGTACTGCAGAGACAAGAGG	AAAAA
IDH2_a	15	90631933	90631935	90631933	90631935	ACTAGGCGTGGGATGTTTTT	TGCAG
IDH2_b	15	90631932	90631935	90631932	90631935	ACTAGGCGTGGGATGTTTTT	C
IDH2_b	15	90631837	90631839	90631837	90631839	AAAGTCTGTGGCCTTGTACT	G
KIT_11	4	55593567	55593723	55593567	55593723	AGAGTGCTCTAATGACTGAGA	A
KIT_11	4	55593567	55593723	55593567	55593723	TTACAGGTAACCATTTATTTGTTCT	G
KIT_13	4	55594162	55594302	55594162	55594302	GCCAGTTGTGCTTTTTGCTA	A
KIT_14	4	55595486	55595666	55595486	55595636	ATGACCACCCTTGGGTATTT	T
KIT_14	4	55595486	55595666	55595612	55595666	GGAAGATCATGCAGAAGCTG	T
KIT_14	4	55595486	55595666	55595486	55595666	ATCTCACCTTCTTTCTAACCTTT	C
KIT_17	4	55599221	55599373	55599221	55599371	AAAATGAATTTAAATGGTTTTCTTTTCT	G
KIT_17	4	55599221	55599373	55599270	55599373	CTTGGCAGCCAGAAATATCC	G
KIT_18	4	55602649	55602790	55602649	55602790	GAGCTTCTGAATTAACATTATTGAC	A
KIT_9	4	55592008	55592231	55592008	55592160	TATGCCACATCCCAAGTGTT	A
KIT_9	4	55592008	55592231	55592096	55592231	CACCGTTTGGAAAGCTAGTG	G
KRAS_a	12	25398274	25398291	25398274	25398291	TGCATATTAAAACAAGATTTACCTCTAT	G
KRAS_a	12	25398275	25398291	25398275	25398291	TGCATATTAAAACAAGATTTACCTCTATT	A
KRAS_b	12	25380269	25380283	25380269	25380283	ATACACAAAGAAAGCCCTCC	G
KRAS_b	12	25380270	25380283	25380270	25380283	ATACACAAAGAAAGCCCTCC	G
KRAS_c	12	25378554	25378568	25378554	25378568	TCTGTATTTATTTCAGTGTTACTTACC	A
KRAS_c	12	25378555	25378568	25378555	25378568	TCTGTATTTATTTCAGTGTTACTTACC	G
KRAS_d	12	25378644	25378653	25378644	25378653	TGTCTACTGTTCTAGAAGGCAA	A
KRAS_d	12	25378645	25378653	25378645	25378653	GTGTCTACTGTTCTAGAAGGC	T
MAP2K1_a	15	66727449	66727486	66727449	66727486	AGCGAAAGCGCCTTGA	A
MAP2K1_a	15	66727450	66727486	66727450	66727486	GAGCTAGAGCTTGATGAGCA	G
MAPK1_1	22	22221602	22221740	22221602	22221638	GTGAAGTCCGGGTTCGAG	C
MAPK1_1	22	22221602	22221740	22221611	22221673	GGTCGCGGACACTCA	C
MAPK1_1	22	22221602	22221740	22221658	22221740	GATGTACGAGAGGTTGGTGT	C
MAPK1_1	22	22221602	22221740	22221602	22221645	GCGGTGAAGTCCGGG	G
MAPK1_1	22	22221602	22221740	22221620	22221687	GGACACTCACCACACCAT	G
MAPK1_1	22	22221602	22221740	22221653	22221740	GCCGATGTACGAGAGGTT	A
MAPK1_2	22	22161943	22162145	22161943	22162075	AAGGTTACCAAGCAGTGGAA	C
MAPK1_2	22	22161943	22162145	22162037	22162145	GTCTGAAGCGCAGTAAGATT	C
MAPK1_2	22	22161943	22162145	22161943	22162113	AAGGTTACCAAGCAGTGGAA	G
MAPK1_2	22	22161943	22162145	22161979	22162145	TTCATTTGCTCGATGGTTGG	G
MAPK1_3	22	22160129	22160338	22160129	22160157	ACATGTTTTTGGGTATTTCTGGT	T
MAPK1_3	22	22160129	22160338	22160143	22160266	AAATAACCTGGCTGACCTTG	C
MAPK1_3	22	22160129	22160338	22160241	22160338	AACCCTCTGAGGATCTGGTA	T
MAPK1_4	22	22153291	22153427	22153291	22153399	CAGAAAGTTCTCTTACTTACTGGA	A
MAPK1_4	22	22153291	22153427	22153365	22153427	GCCACATATTCTGTCAGGAAC	A
MAPK1_4	22	22153291	22153427	22153291	22153427	TTACTGGATTTCTTTTTATGCAAAC	T

MAPK1_5	22	22142973	22143107	22142973	22143079	GATGTAAGCTGTACAAACTTGAG	G
MAPK1_5	22	22142973	22143107	22143071	22143107	CCAGAATGCAGCCTACAGA	G
MAPK1_6	22	22142536	22142687	22142536	22142645	TCAGGCTAGTGACCTAACAAT	C
MAPK1_6	22	22142536	22142687	22142567	22142687	TTTGGAGTCAGCATTTGGGA	C
MAPK1_6	22	22142536	22142687	22142536	22142687	GCGAAGCTAAGCCTAAGAAA	G
MAPK1_7	22	22127152	22127281	22127152	22127281	CTGGCTGATCTATGTCCCTG	T
MAPK1_7	22	22127152	22127281	22127152	22127281	CTGGCTGATCTATGTCCCTG	C
MAPK1_8	22	22123483	22123619	22123483	22123602	CCTTGCTAGAGCTCACTGTAT	T
MAPK1_8	22	22123483	22123619	22123520	22123619	CTGTATCCTGGCTGGAATCT	A
MAPK1_8	22	22123483	22123619	22123483	22123619	CTCACTGTATTAAACTCCAGGT	A
MTHFR_a	1	11856371	11856385	11856371	11856385	AAAGAAAAGCTGCGTGATGA	A
MTHFR_a	1	11856372	11856385	11856372	11856385	TCAAAGAAAAGCTGCGTGAT	A
MTHFR_b	1	11854469	11854483	11854469	11854483	GGTTCTCCCGAGAGGTAAAG	G
MTHFR_b	1	11854470	11854483	11854470	11854483	GGTTCTCCCGAGAGGTAAAG	G
mTOR_0	1	11319295	11319476	11319295	11319413	CACACATCCCACAATGACTG	C
mTOR_0	1	11319295	11319476	11319381	11319476	TTTCCTCATTCCGGCTCTTT	T
mTOR_0	1	11319295	11319476	11319424	11319476	GACGCTCACATTGCTAGATG	G
mTOR_1	1	11318532	11318660	11318532	11318632	CACAAAGGTGAGTGTGTTGT	C
mTOR_1	1	11318532	11318660	11318570	11318660	AAGATGCCACCTTTCCTCTC	A
mTOR_10	1	11298449	11298684	11298449	11298553	CCCAATTGTCCTAAGCTCCT	A
mTOR_10	1	11298449	11298684	11298494	11298652	CAACTACGAGCAGTTTGCTA	C
mTOR_10	1	11298449	11298684	11298629	11298684	TCTCCTTGTGCTCACTGTTC	A
mTOR_11	1	11297890	11298115	11297890	11298007	TCCTCAAGCAGTTCTCACAG	G
mTOR_11	1	11297890	11298115	11297952	11298108	AAAGGCAGGGTTCATGCTA	C
mTOR_11	1	11297890	11298115	11298035	11298115	CAAACAAGGCCTGCAAGTT	T
mTOR_11	1	11297890	11298115	11297890	11298044	TTCCTCAAGCAGTTCTCACA	C
mTOR_11	1	11297890	11298115	11297959	11298115	AGGGTTCATGCTACTGAGTC	G
mTOR_12	1	11294190	11294332	11294190	11294300	TCCCCATATGAGCTGATGAC	T
mTOR_12	1	11294190	11294332	11294200	11294332	AGCCTCTCGGTTTGTGTTAC	T
mTOR_13	1	11293445	11293554	11293445	11293550	CCCTATCTCCGCTATGGAAA	G
mTOR_13	1	11293445	11293554	11293523	11293554	GTTTGGATCAGGGTCTGGAT	A
mTOR_13	1	11293445	11293554	11293445	11293554	CCCTATCTCCGCTATGGAA	A
mTOR_14	1	11292483	11292595	11292483	11292594	GGGGTCTGTCTTGCTCAATC	T
mTOR_14	1	11292483	11292595	11292499	11292595	GCAGTAATACTCACCTGCCT	T
mTOR_15	1	11291347	11291501	11291347	11291465	AATGCTGCAACCATCTCTCT	T
mTOR_15	1	11291347	11291501	11291416	11291501	ATTCAGTAGCACCTCAAGCA	T
mTOR_15	1	11291347	11291501	11291347	11291501	CTCTTGCCATCGTCCCA	G
mTOR_16	1	11290972	11291121	11290972	11291087	GGCTCTGTGAGTGAGAACTT	T
mTOR_16	1	11290972	11291121	11291016	11291121	GACTTGGATTCTGACAGGCT	C
mTOR_17	1	11288715	11288985	11288715	11288817	AAGAGAAGGATTGGGGTTTG	A
mTOR_17	1	11288715	11288985	11288803	11288918	GCAGGAACTGCACACATTT	A
mTOR_17	1	11288715	11288985	11288886	11288985	GACTGGTCTCGGAAGATCC	T
mTOR_18	1	11276195	11276301	11276195	11276296	TTCAAACTTGCTTCTGAGCC	C

mTOR_18	1	11276195	11276301	11276275	11276301	CAAAGGACACCAACATTCCC	A
mTOR_19	1	11273446	11273633	11273446	11273548	TTTAGCTGATCACCCAGGGA	A
mTOR_19	1	11273446	11273633	11273543	11273633	GGGGCAGGTAGAGCTTAAA	T
mTOR_19	1	11273446	11273633	11273446	11273633	AGAAAATCTCTCTGGAGGATGA	T
mTOR_2	1	11316980	11317232	11316980	11317112	ATGACACCATCGTTCCCC	T
mTOR_2	1	11316980	11317232	11317060	11317200	CTCGCTTCACCTCAAATTCC	G
mTOR_2	1	11316980	11317232	11317143	11317232	TCCATGACAACTGGGTCATT	G
mTOR_2	1	11316980	11317232	11316980	11317147	AGGAAGCAAAAGACCCCC	C
mTOR_2	1	11316980	11317232	11317080	11317232	ACGTACTCAGCGGTAAAAGT	T
mTOR_20	1	11272843	11272975	11272843	11272955	GCTTAAAGATTGCTAGTCCCA	C
mTOR_20	1	11272843	11272975	11272892	11272975	CTTCAGGGGCATCAAACAAC	G
mTOR_21	1	11272359	11272541	11272359	11272507	CAGAAGAGCACCTGTCTGT	G
mTOR_21	1	11272359	11272541	11272430	11272541	CTGTGGAGCGCAGTTCT	C
mTOR_22	1	11270861	11270973	11270861	11270963	ACACAAACTCCCATAGCCAA	T
mTOR_22	1	11270861	11270973	11270922	11270973	CGCTGATGATTGATTCGGTG	T
mTOR_22	1	11270861	11270973	11270861	11270973	GACAACAGGGACTTCAGAAC	A
mTOR_23	1	11269359	11269525	11269359	11269490	GAGCTCTCCTTTCCCAGT	C
mTOR_23	1	11269359	11269525	11269435	11269525	TTTCCACTGGTCCACTAGC	T
mTOR_23	1	11269359	11269525	11269359	11269520	GAGCTCTCCTTTCCCAGT	A
mTOR_24	1	11264608	11264770	11264608	11264708	CCAGCCCCTTGATTATTACTTC	C
mTOR_24	1	11264608	11264770	11264689	11264770	GGCGATGATGAGTCCTTCAG	T
mTOR_24	1	11264608	11264770	11264608	11264770	GCTGATCTTCTCCACCCG	T
mTOR_25	1	11259588	11259770	11259588	11259723	CCAGCTGGTTCCCTGTC	A
mTOR_25	1	11259588	11259770	11259642	11259770	AGCCAAGTTTAAGAGGGTCT	A
mTOR_26	1	11259305	11259470	11259305	11259461	AGGCAGCAATTAAAAAGGGT	T
mTOR_26	1	11259305	11259470	11259320	11259470	GGGTTTATGGCCTACCTGAT	C
mTOR_26	1	11259305	11259470	11259305	11259470	GTCATTTTGCATGAAGGCAG	C
mTOR_27	1	11227489	11227584	11227489	11227584	CAAGTCTCTACCTCCTGCTT	T
mTOR_28	1	11217199	11217358	11217199	11217356	GCCACACATGCCATCATTC	A
mTOR_28	1	11217199	11217358	11217284	11217358	TCCTTGTTGGTGTCCATTTTC	G
mTOR_28	1	11217199	11217358	11217199	11217358	GCCACACATGCCATCATTC	T
mTOR_29	1	11210173	11210293	11210173	11210293	ACAGGGTGCCTGTGAG	G
mTOR_29	1	11210173	11210293	11210173	11210293	ACAGGGTGCCTGTGAG	T
mTOR_3	1	11316039	11316259	11316039	11316161	TTTAGGCCAGGTGATTCTCT	T
mTOR_3	1	11316039	11316259	11316119	11316259	AATCAGACAGGCACGAAGG	A
mTOR_3	1	11316039	11316259	11316039	11316196	CCTTTAGGCCAGGTGATTCT	C
mTOR_3	1	11316039	11316259	11316182	11316259	CGGCCACAAAAATGTTGTCA	C
mTOR_30	1	11206723	11206858	11206723	11206839	TTCCTCACTGAGAGATCTGG	T
mTOR_30	1	11206723	11206858	11206763	11206858	AAGGAGAAGAGGTCCTGATG	A
mTOR_31	1	11205015	11205112	11205015	11205112	CCTCAAAAATGACAATGTGCAG	C
mTOR_32	1	11204695	11204822	11204695	11204822	GTAGGGGTAGGTGGGTGAA	T
mTOR_32	1	11204695	11204822	11204695	11204822	GGTGAACTGGGGCTTTCT	T
mTOR_33	1	11199580	11199725	11199580	11199712	CCTGCCCATGTGGGTG	A

mTOR_33	1	11199580	11199725	11199636	11199725	TTGCATACTTGAGCCAGGTT	G
mTOR_34	1	11199351	11199502	11199351	11199473	ATTCTGGAGAAGGTGGTCTG	A
mTOR_34	1	11199351	11199502	11199432	11199502	GTGAACTGTTGGCAGAGGAT	A
mTOR_34	1	11199351	11199502	11199351	11199502	GGAGAAGGTGGTCTGTTCTG	T
mTOR_35	1	11194398	11194533	11194398	11194513	TGACAGGGCTGGAATATGAC	C
mTOR_35	1	11194398	11194533	11194456	11194533	TATGCTGCTGGTCCTCAGTA	A
mTOR_36	1	11193127	11193264	11193127	11193233	TGTAACCACGAGCACACAG	T
mTOR_36	1	11193127	11193264	11193145	11193264	GCTTCTGATCACCTTGTACC	T
mTOR_37	1	11190576	11190844	11190576	11190648	CACTTCAGATACAGCCTCAG	A
mTOR_37	1	11190576	11190844	11190635	11190792	GGGGCTGTTCTCGGTG	T
mTOR_37	1	11190576	11190844	11190680	11190844	TCGGTGCTGGCAGTG	G
mTOR_37	1	11190576	11190844	11190576	11190748	TTCCATTTCTCAGAGAGCCT	C
mTOR_37	1	11190576	11190844	11190662	11190844	TCGCTCTCACTGTTGCTG	C
mTOR_38	1	11189785	11189905	11189785	11189886	CCTACCAGAGTTGCATCCTT	T
mTOR_38	1	11189785	11189905	11189849	11189905	AGGAGATGGAACGGAAGAAG	A
mTOR_39	1	11188902	11189018	11188902	11189018	ACACAGAAGAGAGACTTGGA	A
mTOR_39	1	11188902	11189018	11188902	11189018	CCACCTTCACCTGTAACCA	C
mTOR_4	1	11313886	11314040	11313886	11313987	TCGCTCACAGAATGGTACAC	G
mTOR_4	1	11313886	11314040	11313941	11314040	ACCAGCTCGTTAAGGATCAA	C
mTOR_40	1	11188501	11188619	11188501	11188619	GGTTAGATGAGAAACTGCCC	A
mTOR_40	1	11188501	11188619	11188501	11188619	GGTTAGATGAGAAACTGCCC	A
mTOR_41	1	11188051	11188193	11188051	11188177	AAGGGACCAGGGTCTATGAA	G
mTOR_41	1	11188051	11188193	11188111	11188193	GTGCTCACACATGTTCTTCA	A
mTOR_42	1	11187671	11187873	11187671	11187806	GGCAGAATATTTCTACAGGGTT	G
mTOR_42	1	11187671	11187873	11187765	11187873	GCACCTCAAACATGCCTTT	T
mTOR_42	1	11187671	11187873	11187671	11187860	ACAGGGTTATGTCCTTTCGT	T
mTOR_42	1	11187671	11187873	11187776	11187873	CATGCCTTTCACGTTCCTTT	T
mTOR_43	1	11187057	11187211	11187057	11187173	GGACTATAATGACAGTTAACCCTG	T
mTOR_43	1	11187057	11187211	11187104	11187211	AGATTCGTCGGAACACATGA	C
mTOR_44	1	11186669	11186863	11186669	11186769	GAGAAGTGGGTGACAGAAGT	C
mTOR_44	1	11186669	11186863	11186754	11186863	TGCTATGGACTGAATGCGAA	C
mTOR_45	1	11184545	11184700	11184545	11184691	CAAATTGTTGCCATTTCAGGG	G
mTOR_45	1	11184545	11184700	11184548	11184700	TGTTGCCATTTCAGGGTTTC	C
mTOR_46	1	11182026	11182193	11182026	11182145	GGAAGGGGCACTAGCTCT	T
mTOR_46	1	11182026	11182193	11182047	11182193	ATCACATACCCGCAACATGA	C
mTOR_47	1	11181293	11181435	11181293	11181428	TATTGCGAGTGGGGGTTC	T
mTOR_47	1	11181293	11181435	11181362	11181435	GTCGTCCCCAGCTGTATTAT	T
mTOR_48	1	11177051	11177153	11177051	11177133	ACCATTCAGGAAAACTACAATGG	C
mTOR_48	1	11177051	11177153	11177133	11177148	CGGGTATAATTGGTTCTTCGG	T
mTOR_48	1	11177051	11177153	11177144	11177153	GTTCTTCGGTCAAACCACAC	A
mTOR_48	1	11177051	11177153	11177051	11177153	AAAACTACAATGGAGAAAGAAGAC	T
mTOR_49	1	11175443	11175535	11175443	11175535	GAAGATGAGGTTGGGGTTCTA	A
mTOR_5	1	11307866	11308161	11307866	11307970	AAGTGAGGTGTGGAGCTTAG	G

mTOR_5	1	11307866	11308161	11307966	11308116	GGAGGTCCCAAATCCCAT	A
mTOR_5	1	11307866	11308161	11308056	11308161	AAACTGGTGAAGGGGGTAAT	T
mTOR_5	1	11307866	11308161	11308015	11308161	CCCACCAAGGCATTTGAC	T
mTOR_50	1	11174860	11174954	11174860	11174954	GATCCCATTTGGAAGCAGC	A
mTOR_51	1	11174365	11174520	11174365	11174511	AGCTCCCAGGCACTTGA	G
mTOR_51	1	11174365	11174520	11174387	11174520	ATACTCACTGTCCATCAGCC	T
mTOR_52	1	11172899	11172984	11172899	11172984	TTGCGACCTCCCGTG	A
mTOR_52	1	11172899	11172984	11172899	11172984	CTTTGCGACCTCCCGTG	A
mTOR_53	1	11169696	11169796	11169696	11169796	GCTGCTATTTTCTTAATGAGCTA	G
mTOR_53	1	11169696	11169796	11169696	11169796	AGTCACTGGTGCGGTT	G
mTOR_54	1	11169337	11169437	11169337	11169437	TGCTCAGATTTTATGTCCCTTTT	C
mTOR_54	1	11169337	11169437	11169337	11169437	TGCTCAGATTTTATGTCCCTTTT	G
mTOR_55	1	11168228	11168353	11168228	11168342	GCTTGGGACCTGATTGCTTA	T
mTOR_55	1	11168228	11168353	11168251	11168353	CACTCACCAGCCAATATAGC	T
mTOR_56	1	11167532	11167567	11167532	11167567	TTTCTCACCATGGTTTCAGT	G
mTOR_56	1	11167532	11167567	11167532	11167567	TCACCATGGTTTCAGTTTAGTG	T
mTOR_6	1	11307672	11307800	11307672	11307799	CTGTTCCCTGTTTACCCTGA	C
mTOR_6	1	11307866	11308161	11307866	11307891	CGGGGCAACAAATTAAGGAT	T
mTOR_7	1	11303161	11303367	11303161	11303298	TTCCTCTCCAACCAAATGGA	G
mTOR_7	1	11303161	11303367	11303234	11303367	GCACGCGAGGCAAATAG	T
mTOR_7	1	11303161	11303367	11303161	11303349	CCAACCAAATGGAGTGGAAG	T
mTOR_8	1	11301600	11301748	11301600	11301728	CAAGCCTCACGCTGATACA	T
mTOR_8	1	11301600	11301748	11301656	11301748	CAGCAGCTCCTTGATATCCT	C
mTOR_9	1	11300350	11300614	11300350	11300487	AAGTTTCCAGCATCTCTCAC	C
mTOR_9	1	11300350	11300614	11300412	11300571	AAGAGTGATGCTGCCCAC	C
mTOR_9	1	11300350	11300614	11300515	11300614	GTTTGTGCATAAGGACCAGG	T
mTOR_9	1	11300350	11300614	11300350	11300513	CTTTCCCAAAGTTTCCAGCA	G
mTOR_9	1	11300350	11300614	11300407	11300614	AGGGCAAGAGTGATGCTG	C
NRAS_a	1	115258737	115258754	115258737	115258754	TGGATTAGCTGGATTGTCAGT	A
NRAS_a	1	115258738	115258754	115258738	115258754	ATCCGACAAGTGAGAGACAG	A
NRAS_b	1	115256522	115256536	115256522	115256536	ATTGGTCTCTCATGGCACTG	A
NRAS_b	1	115256523	115256536	115256523	115256536	TCCGCAAATGACTTGCTATT	A
NRAS_c	1	115252286	115252294	115252286	115252294	TGGAATCCCGTAACTCTTGG	A
NRAS_d	1	115252199	115252207	115252199	115252207	GCAAACTCTTGCACAAATGC	T
PDGFRA_12	4	55140993	55141155	55140993	55141155	GCACTGGGACTTTGGTAATTC	A
PDGFRA_14	4	55144048	55144188	55144048	55144188	AGCTCAGCTGGACTGATATG	A
PDGFRA_18	4	55151993	55152145	55151993	55152145	AGATGGCTTGATCCTGAGTC	G
PIK3CA_a	3	178936076	178936102	178936076	178936102	AGCTCAAAGCAATTTCTACAC	G
PIK3CA_a	3	178936047	178936198	178936047	178936198	ACAGACTAGCTAGAGACAATGA	G
PIK3CA_b	3	178952078	178952092	178952078	178952092	AAAACTGAGCAAGAGGCTTT	T
PIK3CA_b	3	178952079	178952092	178952079	178952092	AGCAAGAGGCTTTGGAGTAT	G
PTEN_a	10	89692985	89692999	89692985	89692999	GGGCAAATTTTTAAAGGCACA	A
PTEN_b	10	89717666	89717680	89717666	89717680	TGGTATGTATTTAACCATGCAGA	G

PTEN_b	10	89717667	89717680	89717667	89717680	GTATGTATTTAACCATGCAGATCC	A
PTEN_c	10	89717765	89717785	89717765	89717785	GCCGTTACCTGTGTGTGG	C
PTEN_c	10	89717766	89717785	89717766	89717785	TCTTCCACAAACAGAACAAGA	C
STAT1_0	2	191874592	191874739	191874592	191874698	GGCCCCAAGTCACTTAATCA	C
STAT1_0	2	191874592	191874739	191874634	191874739	TCTAACCACTGTGCCAGGTA	C
STAT1_1	2	191873679	191873843	191873679	191873783	GCCTTCCATAAACATGAGAACA	C
STAT1_1	2	191873679	191873843	191873719	191873843	ACGCTTGCTTTTCCTTATGTT	C
STAT1_1	2	191873679	191873843	191873679	191873843	TGAAGAAAACTGCCTTCCAT	A
STAT1_10	2	191851755	191851804	191851755	191851781	CCCAAAATGTTGAACTTCCTAAA	T
STAT1_10	2	191851755	191851804	191851799	191851804	TTCTCTCATTCACATCTCTGC	
STAT1_10	2	191851755	191851804	191851755	191851804	GAAATGCTGAAAAGTCTTCCAA	
STAT1_11	2	191851570	191851683	191851570	191851602	CAATGTGCCAAAAAGGGCT	
STAT1_11	2	191851570	191851683	191851596	191851649	TACCAGGTGCCGAAATTCA	
STAT1_11	2	191851755	191851804	191851755	191851779	GTGGACTCCTCCATGTTCA	
STAT1_12	2	191850335	191850396	191850335	191850396	TGGGCCCATTCACAACATAA	
STAT1_13	2	191849026	191849129	191849026	191849127	CAATTAAAAGTAAAAATAATGAAGTTTTCCA	
STAT1_13	2	191849026	191849129	191849091	191849129	TGGGTTTCAAAACTAAGGGAG	
STAT1_13	2	191849026	191849129	191849026	191849129	TTTCCAACTCGGGACCAT	
STAT1_14	2	191848358	191848476	191848358	191848459	CTCTGCTTAACCCTGGGAC	
STAT1_14	2	191848358	191848476	191848377	191848476	GTTTTCCATACCCTGGGTTC	
STAT1_15	2	191847099	191847254	191847099	191847201	GCAGAGGGGAAAAGAGCAA	
STAT1_15	2	191847099	191847254	191847199	191847254	CCAACTCAGCACTTCTGAAA	
STAT1_15	2	191847099	191847254	191847099	191847252	AGAGGGGAAAAGAGCAATTAGA	
STAT1_16	2	191845336	191845405	191845336	191845405	ACTTAGAGAGCATAAAACCCAG	
STAT1_17	2	191844488	191844602	191844488	191844591	ATACTGAAGCTGGACTCAGG	
STAT1_17	2	191844488	191844602	191844574	191844602	CCAAAGCCAGAAGGGAAAAT	
STAT1_17	2	191844488	191844602	191844488	191844602	ACCTCGCAGCACTAAAAATA	
STAT1_18	2	191843572	191843737	191843572	191843706	AAGAGGGACTTCACACACAT	
STAT1_18	2	191843572	191843737	191843622	191843737	TCCACCCATGTGAATGTGAT	
STAT1_19	2	191841556	191841761	191841556	191841670	CCCTCATCAGGAAAGACTGT	
STAT1_19	2	191841556	191841761	191841652	191841761	CTTCAGGGGATTCTCAGGAATA	
STAT1_19	2	191841556	191841761	191841556	191841734	CCCCTCATCAGGAAAGACTG	
STAT1_19	2	191841556	191841761	191841698	191841761	GCGAATGATGTCAGGGAAAG	
STAT1_2	2	191872279	191872397	191872279	191872357	ATCATTGCTTTGACATGGGC	
STAT1_2	2	191872279	191872397	191872352	191872367	CCTTCAGACAGCTGTAAATGAT	
STAT1_2	2	191872279	191872397	191872368	191872397	GTAAATGATCATAGACATCTGGATT	
STAT1_2	2	191872279	191872397	191872279	191872397	TGACATGGGCCCTAATAGT	
STAT1_20	2	191840528	191840623	191840528	191840623	GAGGTTTGTAAACATGTCACTCT	
STAT1_21	2	191839546	191839668	191839546	191839668	CTGAGCACACACACTTATTGA	
STAT1_21	2	191839546	191839668	191839546	191839668	CTTATTGAGAGCTACACACAGG	
STAT1_22	2	191835419	191835453	191835419	191835453	CTGTCGCCAGAGAAGATGAA	
STAT1_3	2	191865790	191865899	191865790	191865899	CCCCTACAGAAAGTTTCAGAATAA	
STAT1_3	2	191865790	191865899	191865790	191865899	CCCAAGCAATTGAAACCTTTTT	

STAT1_4	2	191864342	191864440	191864342	191864440	AACGGGCACCACTTCA
STAT1_4	2	191864342	191864440	191864342	191864440	AACGGGCACCACTTCA
STAT1_5	2	191862933	191863044	191862933	191863044	GTGGCATGCTATTCTGGAAA
STAT1_5	2	191862933	191863044	191862933	191863044	TGGCTATAATTTTTCCTCTCTTCTA
STAT1_6	2	191862572	191862743	191862572	191862689	CAGTCAGCTGCCAGTTTTC
STAT1_6	2	191862572	191862743	191862671	191862743	TCCACTCCACTAGTTCATCATT
STAT1_6	2	191862572	191862743	191862689	191862743	TTAATCAGGGCATTCTGGGT
STAT1_7	2	191859777	191859955	191859777	191859908	AGACCGATTACAGAAGGTACAA
STAT1_7	2	191859777	191859955	191859820	191859955	GCTGGAAAAGACTGAAGGTG
STAT1_7	2	191859777	191859955	191859777	191859942	AGACCGATTACAGAAGGTACAA
STAT1_7	2	191859777	191859955	191859821	191859955	CTGGAAAAGACTGAAGGTGC
STAT1_8	2	191855944	191856056	191855944	191856056	TGACAGGTGATGTATGGGAT
STAT1_8	2	191855944	191856056	191855944	191856056	GACAGGTGATGTATGGGATG
STAT1_9	2	191854331	191854410	191854331	191854393	CAGCTAGAAATCTGCTTATTTAGT
STAT1_9	2	191854331	191854410	191854395	191854410	TCAGCTCTTGCAATTTCACC
STAT1_9	2	191854331	191854410	191854331	191854410	TCTGTGCTTGAGTAACAAAATC
STAT3_0	17	40500397	40500544	40500397	40500497	ATGGAACAGCAAGGCATGA
STAT3_0	17	40500397	40500544	40500432	40500544	TTGACTCTCAATCCAAGGGG
STAT3_1	17	40498577	40498741	40498577	40498688	AGAACACTAACACCCGACTC
STAT3_1	17	40498577	40498741	40498670	40498741	CGGCTATACTGCTGGTCAA
STAT3_1	17	40498577	40498741	40498577	40498728	AGAACACTAACACCCGACTC
STAT3_1	17	40498577	40498741	40498713	40498741	ACCAAAGTGGCATGTGATTC
STAT3_10	17	40481755	40481804	40481755	40481804	CGGAACAAAAGGAAGCCT
STAT3_10	17	40481755	40481804	40481755	40481804	GTAGCCGGAGGATGAAGTTA
STAT3_11	17	40481562	40481675	40481562	40481672	CAGGTGTCCTGTGAGGC
STAT3_11	17	40481562	40481675	40481653	40481675	GTGTTTGTGCCCAGAATGTT
STAT3_11	17	40481562	40481675	40481562	40481675	CCAGAGGCCCTTTGTGAA
STAT3_12	17	40481418	40481485	40481418	40481485	CTCTCTCCCTCAAGGAAAAC
STAT3_12	17	40481418	40481485	40481418	40481485	TTTACCCCTCTCTCCCTCAA
STAT3_13	17	40478124	40478227	40478124	40478227	GTTTTTGTCCTGAGTCACCC
STAT3_14	17	40476971	40477089	40476971	40477089	GGGCACCAACTAAAAGGAGG
STAT3_14	17	40476971	40477089	40476971	40477089	GGGGCACCAACTAAAAGGA
STAT3_15	17	40476719	40476874	40476719	40476762	CAGTAGACATGGCCCAAATG
STAT3_15	17	40476719	40476874	40476735	40476874	TGAAATGCGGACCCAAGA
STAT3_15	17	40476719	40476874	40476719	40476874	CCGGATCCCTTTTCTGGG
STAT3_16	17	40475581	40475653	40475581	40475653	GGTGAGCATTCCCATTCC
STAT3_17	17	40475268	40475382	40475268	40475382	CAGGGGACTTGGTTACATCT
STAT3_18	17	40475012	40475171	40475012	40475170	GTGGGGTGGGTGGGA
STAT3_18	17	40475012	40475171	40475051	40475171	CCTTCTCCACCCAAGTGAAA
STAT3_19	17	40474290	40474522	40474290	40474390	TCCTCCAAGGATCCCAAAAT
STAT3_19	17	40474290	40474522	40474364	40474495	TACTTTCCGAATGCCTCCTC
STAT3_19	17	40474290	40474522	40474489	40474522	CTGCTGCTTTGTGTATGGTT
STAT3_19	17	40474290	40474522	40474290	40474460	TCCTCCAAGGATCCCAAAAT

STAT3_19	17	40474290	40474522	40474350	40474522	TCTCTGGCCGACAATACTTT
STAT3_2	17	40497567	40497685	40497567	40497685	CCAGACCAGGGATTTGTTTT
STAT3_2	17	40497567	40497685	40497567	40497685	CCCAGACCAGGGATTTGTTT
STAT3_20	17	40469190	40469252	40469190	40469252	AGGGATAACTGAGGATATTAGAAAT
STAT3_20	17	40469190	40469252	40469190	40469252	AGAAATGAAGGCAAAACGGG
STAT3_21	17	40468797	40468929	40468797	40468897	GTGAGAGCATCACACAAAGG
STAT3_21	17	40468797	40468929	40468883	40468929	AATGAATCTAAAGTGCGGGG
STAT3_22	17	40467753	40467828	40467753	40467828	GTCGTATCTTTCTGCAGCTT
STAT3_3	17	40491322	40491437	40491322	40491430	TTAATGAAAGCTCCCTGCC
STAT3_3	17	40491322	40491437	40491334	40491437	GCTTGTAACTTGCATCACCT
STAT3_4	17	40490739	40490840	40490739	40490840	ACTCAACAACACAAACTCACTT
STAT3_4	17	40490739	40490840	40490739	40490840	CCGCCCGCCTTAAGAT
STAT3_5	17	40489771	40489885	40489771	40489885	GCCACAAGACGCTGAAATC
STAT3_5	17	40489771	40489885	40489771	40489885	GCAAGTGAGCGAGACAC
STAT3_6	17	40489443	40489614	40489443	40489544	GTACCAATTCTGTGGGCCT
STAT3_6	17	40489443	40489614	40489511	40489614	CCAATGCAGGCAATCTGTT
STAT3_6	17	40489443	40489614	40489443	40489614	AGCTTTCGAGAAAGAAAGGAA
STAT3_7	17	40485899	40486077	40485899	40486023	ACTCTACCACGTGAGTCTTT
STAT3_7	17	40485899	40486077	40485981	40486077	TGCTGTACAATGGGGTCC
STAT3_7	17	40485899	40486077	40485899	40486051	ACTCTACCACGTGAGTCTTT
STAT3_7	17	40485899	40486077	40485946	40486077	AAACAGCTCCACGATTCTCT
STAT3_8	17	40485681	40485793	40485681	40485786	CCTAACAGTGTCCCTCAGTAA
STAT3_8	17	40485681	40485793	40485717	40485793	TTAGTAGTGAACTGGACGCC
STAT3_9	17	40483480	40483559	40483480	40483559	TGTCTGTCAAAGTTCTCATTTTTC
TP53_0	17	7579829	7579922	7579829	7579922	AGGGGGCTGGGGTTG
TP53_0	17	7579829	7579922	7579829	7579922	GGCCTGCCCTTCCAAT
TP53_1	17	7579690	7579731	7579690	7579731	GCATCAAATCATCCATTGCTT
TP53_1	17	7579690	7579731	7579690	7579731	GGTGAAAAGAGCAGTCAGAG
TP53_2	17	7579302	7579600	7579302	7579440	GGCATTGAAGTCTCATGGAA
TP53_2	17	7579302	7579600	7579399	7579514	TAGGTTTTCTGGGAAGGGAC
TP53_2	17	7579302	7579600	7579461	7579598	GTGTAGGAGCTGCTGGTG
TP53_3	17	7578361	7578564	7578361	7578483	CCTGGGCAACCAGCC
TP53_3	17	7578361	7578564	7578437	7578564	CCGTCATGTGCTGTGACT
TP53_4	17	7578167	7578299	7578167	7578259	CAGCAGGAGAAAGCCCC
TP53_4	17	7578167	7578299	7578259	7578299	CCACACGCAAATTTCCTTCC
TP53_4	17	7578162	7578304	7578162	7578304	CAAATAAGCAGCAGGAGAAAG
TP53_5	17	7577489	7577618	7577489	7577618	CAGGCCAGTGTGCAGG
TP53_6	17	7577009	7577165	7577009	7577042	CTCCTCCACCGCTTCTTG
TP53_6	17	7577009	7577165	7577014	7577045	CGCTTCTTGTCCTGCTTG
TP53_6	17	7577009	7577165	7577057	7577165	AGCTCGTGGTGAGGC
TP53_6	17	7577009	7577165	7577009	7577165	CCTCCACCGCTTCTTGT
TP53_7	17	7576843	7576936	7576843	7576936	GCATTTTGAGTGTTAGACTGG
TP53_8	17	7573917	7574043	7573917	7574043	CCAACCTAGGAAGGCAGG

TP53_8	17	7573917	7574043	7573917	7574043	AGTAGGGCCAGGAAGGG
TP53_9	17	7572917	7573018	7572917	7573018	CCCAAAACCCAAAATGGCA
TP53_9	17	7572917	7573018	7572917	7573018	GCTGTCAGTGGGGAACAA
TYMP_0	22	50967915	50968148	50967915	50968030	GGGACCCAAAGTCTCTCG
TYMP_0	22	50967915	50968148	50967989	50968121	CCCCTGATGTCCGCTTC
TYMP_0	22	50967915	50968148	50968098	50968148	CGGAGAAGTCACCAGGC
TYMP_1	22	50967555	50967777	50967555	50967682	GACTGGTTGCTGCATGTG
TYMP_1	22	50967555	50967777	50967623	50967765	CTGACCTTGTCACCCACA
TYMP_1	22	50967555	50967777	50967738	50967777	GTCTCCTCCAGATCCATGC
TYMP_2	22	50966931	50967049	50966931	50967042	CCACCAGTGATCTTTTAGTGA
TYMP_2	22	50966931	50967049	50966997	50967049	AGACTCCAGCTTATCCAAGG
TYMP_3	22	50966007	50966156	50966007	50966156	GTGAACATGCAGAAGCAGG
TYMP_4	22	50965584	50965722	50965584	50965722	GCATCAAGACGCTTGCC
TYMP_5	22	50964995	50965177	50964995	50965108	CCTCCGCTCCCCTACA
TYMP_5	22	50964995	50965177	50965033	50965177	GTGACCAGGTCCCTTAAGTC
TYMP_5	22	50964995	50965177	50965126	50965177	TTGTCCATGGCGGTCAG
TYMP_6	22	50964665	50964915	50964665	50964793	CGGGAAGGGAAGGGGAT
TYMP_6	22	50964665	50964915	50964757	50964915	GGACTTCCCGAGCACAG
TYMP_6	22	50964665	50964915	50964665	50964876	CGGAAGGACGGGGACT
TYMP_6	22	50964995	50965177	50964995	50965016	CCAGCATCCGCTCGAA
TYMP_7	22	50964420	50964580	50964420	50964526	GCGAGGGGCTGTTAGAG
TYMP_7	22	50964420	50964580	50964465	50964580	CACGTCGACCAGCAGC
TYMP_7	22	50964420	50964580	50964420	50964570	GCGAGGGGCTGTTAGAG
TYMP_8	22	50964189	50964357	50964189	50964333	CCGCCCAAGCACTGAC
TYMP_8	22	50964189	50964357	50964204	50964357	GCGGCAAAGGAGCTTTATT
TYMS_a	18	673437	673451	673437	673451	TTTCACAAGCTATTCCCTCAA
TYMS_a	18	673438	673451	673438	673451	TTTCACAAGCTATTCCCTCAA
UGT1A1_a	2	234668849	234668909	234668849	234668909	TTGTGGACTGACAGCTTTT
UGT1A1_a	2	234668807	234668958	234668807	234668958	CTGAAAGTGAACTCCCTGCT
UGT1A1_a	2	234668807	234668958	234668807	234668958	CTGAAAGTGAACTCCCTGCT

Table A.2: My caption

Table A.3: Gene Reference Models for HGVS Nomenclature. PGx genes are highlighted in blue.

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