Sequencing Data

Laboratory: Molecular Characterization and Clinical Assay Development Laboratory
Revision Date: 2/3/2023 Page 1 of 4

MCCRD-SOP0064: Detection of Human Leukocyte Antigen (HLA) Typing from Whole Exome Sequencing Data

Effective Date: 2/3/2023

Please check for revision status of the SOP at

https://pdmr.cancer.gov/sops/

PDMR NCI Patient-Derived Models Repository An NCI Precision Oncology Initiative Resource

TABLE OF CONTENTS

1.0	PURPOSE/SCOPE	3
2.0	DESCRIPTION OF HLA TYPING DETECTION	3
3.0	CODE DESCRIPTION	3
4.0	EXAMPLE HLA PROFILE REPORTING IN THE NCI PDMR DATABASE:	

Sequencing Data

Laboratory: Molecular Characterization and Clinical Assay Development Laboratory
Revision Date: 2/3/2023 Page 2 of 4

APPROVALS

Laboratory Director:	Biswajit Das, Ph.D.
	Chris Karlovich. Ph.D.
	Mickey Williams, Ph.D.
Author/Owner:	Lily Chen, Ph.D.

VERSION INFORMATION

1. Change History

Revision	Description
	Internal SOP used by MOCHA Laboratory
2/3/2023	Standardize SOP for posting to PDMR internal site for use by designated NCI intramural laboratories

2. Related SOPs

MCCRD_SOP0011: Whole Exome Sequencing Data Analysis Pipeline and Specifications

3. Reference Documentation

Number	Title/Link
[1]	Szolek, A, Schubert, B, Mohr, C, Sturm, M, Feldhahn, M, and Kohlbacher, O (2014). OptiType: precision HLA typing from next-generation sequencing data Bioinformatics, 30(23):3310-6.
[2]	https://github.com/FRED-2/OptiType

Sequencing Data

Laboratory: Molecular Characterization and Clinical Assay Development Laboratory
Revision Date: 2/3/2023 Page 3 of 4

1.0 PURPOSE/SCOPE

This Standing Operating Procedure (SOP) describes procedures for detection of Human Leukocyte Antigen (HLA) typing using whole exome sequencing (WES) data for reporting in the NCI Patient-Derived Models database as performed by the Molecular Characterization Laboratory (MoCha) at the Frederick National Laboratory for Cancer Research. This SOP is for research purposes only and no clinical samples will be processed using this SOP.

2.0 DESCRIPTION OF HLA TYPING DETECTION

- 2.1 The mouse reads are first filtered out from whole exome sequence (WES) data following the WES data analysis pipeline in the SOP MCCRD_SOP0011.
- 2.2 HLA reads are extracted from paired-end human only sequencing data.
- **2.3** HLA type is estimated using OptiType $^{[1,2]}$.

3.0 CODE DESCRIPTION

3.1 HLA reads are extracted from paired-end sequencing data

```
razers3 -i 95 -m 1 -dr 0 -tc 20 -o ${sample}_1.bam

$OPTITYPE_HOME/data/hla_reference_dna.fasta ${sample}_hg19_R1.fastq.gz

razers3 -i 95 -m 1 -dr 0 -tc 20 -o ${sample}_2.bam

$OPTITYPE_HOME/data/hla_reference_dna.fasta ${sample}_hg19_R2.fastq.gz

samtools bam2fq ${sample}_1.bam >${sample}_1.fastq

samtools bam2fq ${sample} 2.bam >${sample} 2.fastq
```

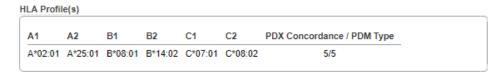
- 3.2 OptiType is used to detect HLA type for PDX sample
 OptiTypePipeline.py -i {sample}_1.fastq \${sample}_2.fastq --dna -v -o
 \${sample} -c \$OPTITYPE HOME/config.ini
- 3.3 HLA type is reported at model level with #concodance/#total PDX specimens in the model. The data for PDC and PDOrg is reported separately.

Sequencing Data

Laboratory: Molecular Characterization and Clinical Assay Development Laboratory
Revision Date: 2/3/2023 Page 4 of 4

4.0 EXAMPLE HLA PROFILE REPORTING IN THE NCI PDMR DATABASE:

4.1 All PDX tumors sequenced (n=5) are in concordance



4.2 In a small fraction of models, intra-model HLA variation in sequenced PDX tumors is observed.

