Online Method

Melanoma patient single cells WGA DNA

YUCLAT metastatic melanoma and Peripheral Blood Lymphocytes (PBL) were collected from a 64-yr-old male patient by the Tissue Resource Core of the Yale SPORE in Skin Cancer with the participant's signed informed consent according to Health Insurance Portability and Accountability Act (HIPAA) regulations with a Human Investigative Committee protocol as described.[1] Single cell WGA DNA was prepared in our previous work[2].

Single cell WGA DNA from DU145 *ex vivo* tree

DU145 *ex vivo* tree and single cell WGA DNA of this *ex vivo* tree were prepared with a modified RepliG Mini protocol in our previous work[2]. The profile of these single cells was retrieved from our cell lineage database.

Duplex MIPs Pipeline Reagents

Reagents	Catalog #
KOD Hot Start DNA Polymerase	Merck 71086
dNTP 25mM each	Bioline BIO-39053
Betaine solution 5M	Sigma B0306 1VL
Phusion High-Fidelity DNA Polymerase	NEB M0530L
Ampligase 10X Reaction Buffer 5ml	EPICENTRE A1905B
Ampligase DNA Ligase W/O Buffer	EPICENTRE A1905B
Exonuclease I (E.coli) - 15,000 units	NEB M0293L
Exonuclease III (E.coli) - 25000 units	NEB M0206L 1
RecJf - 1,000 units	NEB M0264S
Exonuclease T - 1,250 units	NEB M0265L
T7 Exonuclease - 5,000 units	NEB M0263L
Lambda Exonuclease	EPICENTRE LEO32K
NEBNext Ultra II Q5 Master Mix	NEB M0544S
MinElute PCR Purification Kit (250)	QIAGEN 28006
Qubit® dsDNA HS Assay Kit	Life Technologies Q32854
Agencourt Ampure XP Beads	Beckman Coulter A63881
2% Agarose, dye-free, w/ internal standards	Sage Science ES-BDF2003

KOD Hot Start Real Time PCR Mix 5X (KOD 5X MIX)

First, SYBR 1:100 was prepared by mixing 10 µl from stock SYBR green I (Lonza, 10,000X) and 990 µl Dimethyl Sulfoxide (DMSO) (Sigma).

The KOD 5X Mix was prepared according to the table below.

Reagents	Stock Conc.	Final Conc.	1X (µl)
DDW			0.27
KOD Buffer 10X (Merck)	10X	1X	2.5
$MgSO_4$ 25mM (Merck)	25mM	7.5mM	1.5
dNTP 25mM each (Bioline)	25mM	7.5mM	0.2
KOD Enzyme 1 U/µl (Merck)	1 U/µl	0.1 U/µl	0.5
SYBR 100X (Lonza)	100X	1X	0.025
Total Volume			5

Duplex MIPs Generation for OM6 and OM7

The same process were done for both OM6 and OM7

(A). PreAmp PCR on oligo pool:

Oligo pool of OM6 or OM7(order details in Supplemental File1&2) received from provider (Custom Array, Inc.) was diluted to 1ng/µl as PCR template, amplification primers designed to fit the universal adaptors were used for PreAmp PCR in LightCycler 480 (LC480, Roche) according to the setup shown below. SYBR green was added to track the amplification.

PreAmp PCR primers:

OM4_Mly_F: GTCTATGAGTGTGGAGTCGTTGC
OM4_Mly_R: CTAGCTTCCTGATGAGTCCGATG

PreAmp PCR condition:

Reagents	Stock Conc.	Final Conc.	1X (µl)
Template	1 ng/µl	0.2 ng/µl	1.8
OM4_Mly_F	10 pmol/µl	0.3 pmol/µl	1.35
OM4_Mly_R	10 pmol/µl	0.3 pmol/µl	1.35
KOD 5X MIX	5X	1X	9
DDW			31.5
Total Volume			45

PreAmp PCR program:

Temperature	Time	Cycles
95 ºC	2 min	
95 ºC	20 sec	X12 cycle
60 ºC	10 sec	
70 ºC	5 sec	
70 ºC	50 sec	
4 ºC	Hold	

PreAmp PCR product was purified by MinElute PCR purification kit (Qiagen), its concentration was measured by Qubit dsDNA HS Assay Kit (Life Technologies). The purified PreAmp PCR product was diluted to 1ng/µl as the template for next step, the production PCR.

(B). Production PCR (48 reactions)
48 reactions of production PCR in a 96 well plate (Roche) were performed according to the setup below in LC480. The amplification was tracked with SYBER.

Reagents	Stock Conc.	Final Conc.	1X (µl)
Template	1 ng/µl	0.2 ng/µl	1.8
OM4_Mly_F	10 pmol/µl	0.3 pmol/µl	1.35
OM4_Mly_R	10 pmol/µl	0.3 pmol/µl	1.35
KOD 5X MIX	5X	1X	9
DDW			31.5
Total Volume			45

Production PCR program

Temperature	Time	Cycles
95 ºC	2 min	
95 ºC	20 sec	X12 cycle
60 ºC	10 sec	
70 ºC	5 sec	
70 ºC	50 sec	
4 ºC	Hold	

Every four wells of PCR product were merged together and purified by one MinElute column according to the manufacturer's protocol. Elution was in 45µl DDW, and all products were pooled; 1µl of the pool was used to check the concentration by NanoDrop spectrophotometers (Thermo Scientific) according to the manufacturer's protocol. The pool was diluted to ~30ng/µl based on measured concentration. 20µl of sample was kept for quality control; the rest was processed to next step.

(C). The diluted DNA from last step was mixed with reagents according to the table below and incubated at 37 ℃ overnight, then deactivated at 80 ℃ for 20min and finally kept at 4 ℃.

Reagents	Stock Conc.	Final Conc.	1X (µl)
DNA	30 ng/µl	25.2 ng/µl	84
10X NEB Smarter Buffer	10X	1X	10
MlyI	10 U/µl	0.6 U/µl	6
Total Volume			100

Digested DNA was cleaned by MinElute. If more than one PCR was performed, all elution samples were merged into one tube. Concentration was measured using by Qubit dsDNA HS (High Sensitivity) assay kit according to the manufacturer's protocol.

(D). Tape Station size check
The cleaned digested product was the final duplex MIPs. Its size (~105bp) was measured together with an undigested sample from step (C) (~150bp) by Tape Station (Agilent) (Supplemental Figure1).

(E). Duplex MIPs working solution preparation
Based on length of 105 bp and the measured concentration, the final duplex MIPs was diluted to 5.8ng/µl working solution, equivalent to 80nM (80fmol/µl). Adjusted to 8nM by 1:10 dilution where needed. Working solutions were stored in -20°C freezer.

Duplex MIPs-based targeted enrichment pipeline

(A). Hybridization
Single cell WGA product concentration is generally 100-200 ng/µl. 200~500 ng single cell WGA DNA (~2µl) was used as template for each reaction. Reaction mix was prepared according to the table below:

Reagents	Stock Conc.	Final Conc.	1X (µl)
Single Cell WGA DNA	100ng/µl	20 ng/µl	2
Duplex MIPs	8fmol/µl	0.8fmol/µl	1
Ampligase Buffer	10X	1X	1
Betaine	5M	0.9 M	1.8
DDW			4.2
Total Volume			10

For a big batch experiment, hybridization mix was prepared based on the above table without DNA according to the sample numbers. 8µl hybridization mix was distributed to a 96-well plate, 2µl DNA or DDW was added to each well and mixed by liquid handling system (Evoware, Tecan) or manually.

The reaction plate was put into a PCR machine with 100°C lid temperature, and then heated to 98°C for 3 minutes, followed by a gradual decrease in temperature of 0.01°C per second to 56°C and incubated at 56°C for 17 hours.

Optional: if your PCR machine could not decrease as slow as 0.01°C/second, alternative strategy could be applied: the reaction plates was heated to 98°C and kept for 3 minutes, decreased by 0.1°C every cycle as slow as possible and was kept 15 second at this temperature. Cycling until 56°C and incubated at 56°C for 17 hours.

An example in our PCR machine was shown below.

Step	Temperature	Time	Cycles
1	97.9 ºC	3 min	
2	97.9 ºC	15 sec	X420
	decrease as slow as 0.1ºC/sec		
	decrease by 0.1ºC/sec every cycle		
3	56 ºC	17 hour	
4	56 ºC	Pause for adding gap filing mix	

(B). Gap filling

The gap filling mix was prepared half an hour before hybridization finished as below and kept at 56°C on the heat block.

Reagents	Stock Con.	Final Conc.	1X (µl)
dNTP	2mM	0.3 mM	1.5
NAD	10mM	2 mM	2
betaine	5M	1.1 M	2.2
Ampligase buffer	10X	1X	1
Ampligase	5U/µl	0.5 U/µl	1
Phusion	2U/µl	0.8 U/µl	0.4
DDW			1.9
Total Volume			10

The reaction plate was transferred from the PCR machine to a 56°C heat block when the hybridization step finished. 10µl of gap filling mix was added to each well, carefully mixed by pipette, sealed tightly and quickly and put it back to the PCR machine for 56°C incubation for 4 hour, then 68°C for 20 minutes and 4°C until next step.

Optional: After the gap filling, the reaction plate can be stored at 4°C fridge for up to two days.

(C). Digestion of linear DNA:

The digestion mix was prepared as below 15 min before gap filling step finished.

Reagents	Stock Con. (U/µl)	Final Conc. (U/µl)	1X (µl)
exo I	20	3.5	0.175
exo III	100	18	0.18
exo T7	10	4	0.4
exo T	5	0.4	0.08
RecJf	30	3	0.1
lambda exo	10	0.2	0.02
DDW			1.045
Total Volume			2

The reaction plate was taken off from PCR machine, the cover was carefully removed. 2µl of the digestion mix was added to each well and mixed. The reaction plate was spin down and sealed, then incubated at 37°C for 60 minutes, 80°C for 10 minutes and 95°C for 5 minutes.

Pause Point: the reactions can be stored at -20°C for months after the digestion step.

Sequencing library preparation

(A).Sample specific barcoding PCR

The barcoding PCR was using the same dual-index barcoding primers designed and published by Biezuner *et al.* [2] with a modification of the PCR enzyme to NEBNext Ultra II Q5. Here just showed the structure of the dual-index Illumina barcoding primers used in the experiments, details of their sequences could be found in the supplemental information by Biezuner *et al.* [2].

i5-index-primer: AATGATACGGCGACCACCGAGATCTACAC[i5-8bp-index]ACACTCTTTCCCTACACGACGCTCTTCCG;
i7-index-primer: CAAGCAGAAGACGGCATACGAGAT[i7-8bp-index]GTGACTGGAGTTCAGACGTGTGCTCTTCCG;

2µl product from last step and unique barcoding pair of primers for each sample were used in the barcoding PCR according to the table below.

Reagents	Stock Conc.	Final Conc.	1X (µl)
Template	NA	NA	2
dual-index Illumina primers	5 pmol/µl each	0.5 pmol/µl each	2
NEBNext Ultra II Q5 Master Mix	2X	1X	10
SYBER 10X	10X	0.5X	1*
DDW			5
Total Volume			20

* Final Conc. 0.25X SYBER used in early calibration experiments.

Barcoding PCR program

Temperature	Time	Cycles
98 ºC	30 sec	
98 ºC	10 sec	X5 cycle
56 ºC	30 sec	
65 ºC	45 sec	
98 ºC	10 sec	X15 cycle
65 ºC	75 sec	
65 ºC	5 min	
4 ºC	Hold	

(B). Diagnosis Sequencing Sample pooling and Purification

Barcoding PCR product was cleaned in 96-well plate by 0.8x AMPure XP SPRI magnetic beads (Beckman Coulter) according to manufactory's manual by Tecan liquid handling system, eluted in

40µl DDW. Equal volume of purified samples was pooled by Echo (Echo550, Labcyte). The pool was concentrated by MinElute according to manufactory into 35µl DDW.

(C). Size Selection for Diagnosis Sequencing
3ul of the concentrated pool was kept for later quality control. 30µl of the concentrated pool was used on a 2% V1 cassette BluePippin (Sage Science) with setting range 240-340bp according to manufactory's protocol. The size selected elution was collected and cleaned by MinElute into 15µl DDW. The concentration was measured by Qubit dsDNA HS (High Sensitivity) assay kit. The size distribution of the concentrated pool before and after BluePippin was measured by Tape Station dsDNA chip (Supplemental Figure5). The size selected pool with a single peak around 300bp was used to prepare 12µl of 4nM (4fmol/µl) library for Illumina NGS calculated based on the measured concentration and the average size on Tape Station.

(D). Diagnosis Sequencing
10pM library was sequenced on MiSeq 1M nano or 4M micro flow cell with 151x2 pair-end run parameters according to manufactory manual, the default sequencing primers were used.

(E). Based on the diagnosis sequencing result, the volume for each sample to equalize the reads was calculated to create production sequencing Echo pooling table. According to this table, the purified samples from step (B) were pooled by Echo550 and concentrated by into 35µl DDW. The pool was used to prepare the production-sequencing library similar process in step (C).

(F). Production sequencing
1.8~2.2pM library was sequenced on NextSeq500 flow cell with 151x2 pair-end run parameters according to manufactory manual, the default sequencing primers were used.

(Optional) If the production sequencing did not generate enough reads for some samples, another round of NextSeq could be conducted using the same library to get more reads.

Target Specific duplex MIPs design and preparation

Hyper-mutable STRs selection criteria: Several types of STRs were chosen based on the hg19 reverence human genome annotation. AC-type STRs longer than 10 repeats, AG-type STRs longer than 10 repeats, A-type STRs longer than 6 repeats and G-type STR longer than 6 repeats were selected. SNVs targets were chosen based on cancer related highly mutable regions or known cancer associated regions.
Amplicon criteria: Amplicons contains TTAA sequence were ruled out to fit Ampli1 WGA kit.
Amplicon size was designed around ~150bp.

Primer-3 based python script was used to design the targeting primers. Only top-scored primers were chosen as candidates for duplex MIPs precursor design. From these candidates, several more filters were used in the designing: both forward primer and reverse primer shall be unique across the human genome; precursors with MlyI digestion recognition site shall be ruled out; the total

length of the precursors shall not be longer than 150bp, which is the maximum limit of the oligo synthesis provider.

Duplex MIPs structure:

Mly1_F='TATGAGTGTG**GAGTC**GTTGC'

Mly1_R='GCTTCCTGAT**GAGTC**CGATG'

5`[Mly1_F]+[FW_PRIMER_SEQUENCE]+[NNN]+[AGATCGGAAGAGCACACGTCTGAACTCTTTCCCTACAC GACGCTCTTCCGATCT]+[NNN]+[Reverse Complement(RV_PRIMER_SEQUENCE)]+[Reverse Complement(Mly1_R)]-3`

Duplex MIPS precursors sequence example:

5`TATGAGTGTGGAGTCGTTGCTACTTGGTGGCTAATTCAGCAGGNNNAGATCGGAAGAGCACACGTCTGAACTC TTTCCCTACACGACGCTCTTCCGATCTNNNTTGCAAGCTCCCTCTGAAAAGTTCATCGGACTCATCAGGAAGC3`

Database Management System (DBMS)

Our computational workflow can be divided into two parts, (1) Sample documentation from Individual to NGS barcoded sample that later annotates the reconstructed cell lineage trees (2) Analysis workflow that starts with raw genomic sequencing data and ends with cell lineage trees, including mutation analysis and lineage tree reconstruction (Supplemental Figure7).
All computational tools mentioned in the text can be found in https://github.com/ofirr/clineage
Although this git folder is, still being improving by adding detailed introductions of the whole pipeline systematically.

Ex-vivo reconstruction parameters for Figure 3.

AA data[2] was genotyped with a minimal coverage of 30X reads, a confidence threshold of 0.05 (correlation above 0.95) between the measured histogram and the reported model. Duplex MIPs data was genotyped with a minimal coverage of 5X and the same confidence threshold as AA.
In both attempts, low coverage samples and loci were filtered out, leaving the top 75% of the cells and loci as input. In both attempts, the reconstruction was performed using the Neighbor Joining algorithm with the absolute distance function.

1. Krauthammer, M. et al. Exome sequencing identifies recurrent mutations in NF1 and RASopathy genes in sun-exposed melanomas. *Nat Genet* **47**, 996-1002 (2015).
2. Biezuner, T. et al. A generic, cost-effective, and scalable cell lineage analysis platform. *Genome Res* **26**, 1588-1599 (2016).