Automated Walk-Away NGS Sample Preparation Using a Flexible Library Prep System with Enhanced Error Correction

Saharnaz Bigdeli[1], Bernd Buehler[1], Jan Godoski[1], Denise Rhodes[1], Brian Vaughan[1], Ryan Aguilar[1], Chai-Siong Siow[2], Ji Zhu[3], Bahram Arezi[1], Ronda Allen[1]

1 Agilent Tech. La Jolla, CA; 2 Agilent Tech. Penang, Malaysia; 3 Agilent Tech. Santa Clara, CA

ASHG2022
POSTER #
PB2913

Agilent
Trusted Answers

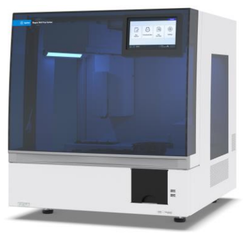
Introduction

Next-generation sequencing (NGS) target enrichment protocols are in high demand for preparation of Illumina sequencing-ready libraries from samples of varying quality and input. Automation of these complex protocols eliminates the need for highly trained specialists while considerably improving the NGS library prep turnaround time, quality, and reproducibility. A fully automated walk-away protocol on the Agilent Magnis NGS Prep system was developed and optimized for a wide range of DNA inputs (10 to 200 ng) and sample types (for example, high quality, formalin-fixed paraffin-embedded [FFPE], circulating tumor DNA samples). It includes different shearing methods within the automated protocol (mechanical or enzymatic fragmentation) using 192 unique dual indexes (UDI) to minimize the effects of possible index hopping and contamination. Further, incorporation of inline duplex molecular barcodes filters out PCR or sequencing errors by making consensus calls using barcode information. The automated workflow on Magnis, referred to as the SSEL-DNA-XTHS2-ILM protocol, is easily set up from pre-aliquoted, single-use reagents to minimize contamination risk. It delivers up to eight target-enriched NGS libraries per run, in nine to 10 hours, without the need of user intervention. To optimize library yields, catalog and custom small, medium, and large probes can be used and PCR cycles can easily be adjusted from the touch screen. The expanded Magnis SureSelect XT HS2 DNA protocol now includes dual indexing capabilities and a new method to process initial and poor-quality FFPE RNA. This secondary automated procedure, the SSEL-RNA-XTHS2-ILM protocol, can be set up and run separately using RNA that is first converted to cDNA and then further processed using similar steps as the SSEL-DNA-XTHS2-ILM protocol to create a final NGS library.

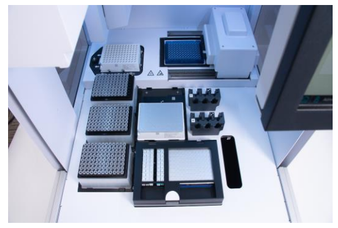
Methods

SureSelect XT HS2 DNA libraries were generated from over 80 runs (640 libraries) on six Magnis instruments using high-quality HapMap and FFPE DNA at 10 to 200ng inputs with small, medium, and large bait panels. Ultrasonication methods were used to manually shear the DNA prior to the runs, and short and long fragments were created by changing the shearing time. In addition, the enzymatic fragmentation option was used and is programmed in the protocol to shear the genomic DNA on-deck. All sequencing results shown in Figure 2 are from libraries generated from DNA using the large capture V8 probe set and were down sampled to 100X raw sequencing depth. Single nucleotide polymorphism (SNP) calling was completed using the Agilent SureCall Software (Figure 3). Inter- and intrarun contamination assays were performed using 50 ng of input DNA for all eight samples (15,000 copies). Every other sample was spiked with one million copies of Alien 6 DNA fragment (239 bp) which contains a specific DNA sequence absent from the human genome and can be explicitly detected using a sensitive qPCR assay. A Taqman qPCR assay specific to the Alien 6 DNA was used for the pre- and postcapture samples. It was used postrun to rule out intrarun contamination and in the following run to rule out inter-run contamination. Quantification of copies per µL sample was completed using a standard curve (spiked samples were pre-diluted at 100X; data not shown). The sequencing data from each pool was aligned to all 192 indexes. When fewer than eight samples per run were sequenced, alternate samples were sequenced. Any sample-to-sample or index contamination would likely be higher over a shorter distance and thus more likely to contaminate neighboring wells. Therefore, sequencing alternate samples allowed for capture of contamination throughout the process, from mixing the index pairs and plating into the strip tubes, through the Magnis library prep and target enrichment steps, sample collection, and quantification (Figure 4).

Figures



A: Magnis Exterior



B: Magnis Interior and Consumables

C: Magnis Touchscreen

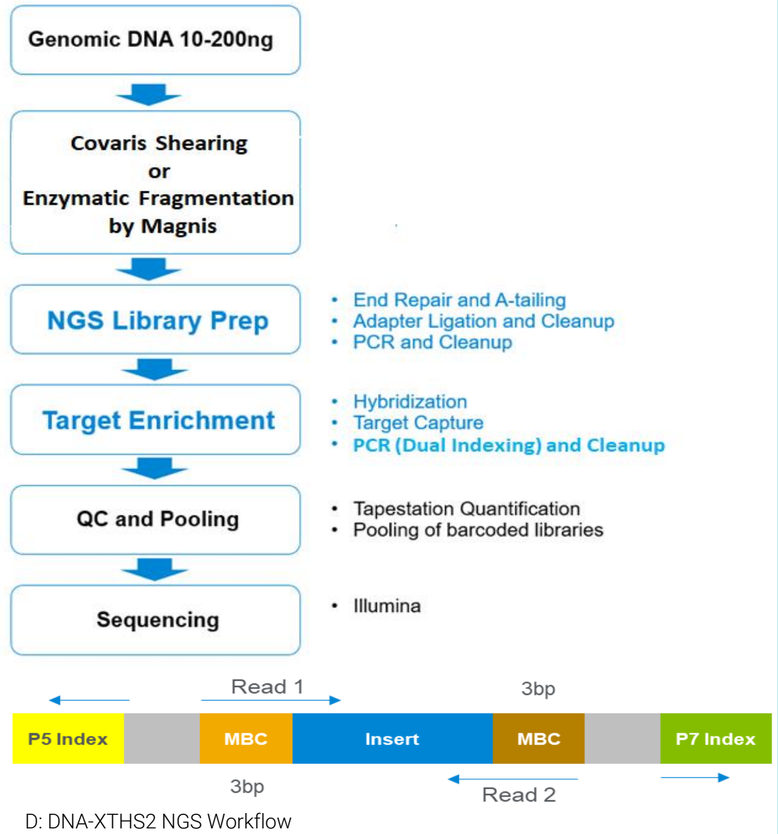
Modules:

1. Gantry with an eight-channel pipette
2. Barcode reader
3. Heater Shaker Magnet (HSM)
4. Thermal Cycler
5. Chiller block
6. Tube holder for reagents requiring special handling and storage
7. Four tip box holding positions (96/box)
8. Beads and buffer position with a built-in plate for camera calibration
9. Touchscreen UI
10. UV decontamination system

Special Features:

- Auto-teach system using the barcode reader's camera for all module positions, built-in calibration plate with special marks, and an algorithm to calculate the accurate teaching position for each module
- Barcoded labware for traceability and confirmation of user setup
- Barcoded, ready-to-use, pre-aliquoted reagents
- Disposable labware, PCR lid, and liners to prevent cross-contamination
- Pre-programmed and validated protocol for SureSelect DNA-XTHS2 Reagent kit

Figure 1. Magnis NGS Library Prep system



D: DNA-XTHS2 NGS Workflow

Assay 5, Exome V8 2x150 bp 6Gb	HiSeq4000 2x150 bp			
	10ng 15 min EF		200ng 2x60s Covaris	
Average Per run	Run A	Run B	Run C	Run D
Samples	4	4	4	4
Percent duplicate reads:	9.1%	9.3%	4.1%	4.0%
Percentage reads in targeted regions:	65.9%	66.1%	65.6%	66.4%
Uniformity (1/2 mean with upper tail):	91.8%	91.8%	92.4%	92.3%
at least 30 reads:	94.8%	94.8%	95.7%	95.6%
HS_LIBRARY_SIZE (optical duplicates included)	62,718,072	62,164,280	180,459,014	192,264,197
HS_LIBRARY_SIZE (w/o optical duplicates)	74,337,502	73,503,064	350,263,746	384,185,123
FOLD_80_BASE_PENALTY	1.43	1.44	1.42	1.43
AT_DROPOUT	4.10	4.02	1.32	1.11
GC_DROPOUT	0.40	0.44	1.98	2.21

Figure 2. DNA-XTHS2 Magnis Yields, CVs, & Sequencing Metrics

Run	Variant	Number of replicates with at least 3 variant reads	Average actual allele frequency	Actual / Expected
A	T790M	8	0.98	1.01
	L858R	8	4.14	1.18
	G12D	8	6	1.09
B	T790M	8	1.01	1.04
	L858R	8	4.4	1.26
	G12D	8	5.79	1.06

Figure 3. DNA-XTHS2 Magnis Variant Calling (AIO-Lung)

Well	Run A		Run B		Run C		Run D	
	Magnis 1		Magnis 2		Magnis 3		Magnis 4	
	Alien 6 copies detected							
Well A	ND	ND	ND	ND	ND	ND	ND	ND
Well B	ND	ND	ND	ND	ND	ND	ND	ND
Well C	ND	ND	ND	ND	ND	ND	ND	ND
Well D	ND	ND	ND	ND	ND	ND	ND	ND
Well E	ND	ND	ND	ND	ND	ND	ND	ND
Well F	ND	ND	ND	ND	ND	ND	ND	ND
Well G	ND	ND	ND	ND	ND	ND	ND	ND
Well H	ND	ND	ND	ND	ND	ND	ND	ND

Figure 5. Magnis Inter-Run Contamination Assay - PrePCR

Pool A	A	B	C	D	E	F	G	H	I	J	K	L
A	5,122,813	2	1	1	4,373,408	0	4,838,350	0	4,878,309	0	5,063,221	4,792,921
B	0	4,783,242	4,484,731	5,253,840	0	4,736,733	0	4,317,655	0	4,097,481	9	45
C	4,400,472	0	5,409,572	0	4,466,626	0	5,018,033	2	4,145,455	0	3,864,767	4,578,836
D	0	4,530,294	28	4,220,667	5	4,640,999	0	4,983,004	0	4,747,340	26	28
E	5,720,111	0	5,236,734	1	4,460,446	0	4,258,533	0	4,624,413	0	5,515,970	4,466,523
F	0	5,111,245	0	4,628,447	0	4,810,115	0	5,230,129	0	4,770,761	9	8
G	4,961,188	0	4,644,188	0	4,694,319	0	4,237,746	0	5,378,198	0	4,723,957	5,081,356
H	0	4,784,374	4	5,068,320	0	5,562,029	0	5,086,110	0	3,749,053	23	9

Sample not added to the pool

NTC sample added to the sequencing pool

Cluster reads mapped back to all 192 indices

Pool A	1	2	3	4	5	6	7	8	9	10	11	12
A	0	0	4,940,471	5,493,072	4,665,441	4,473,171	5	5	5,498,037	5,832,957	0	0
B	1	0	4,127	301	62	102	0	0	21	14	0	0
C	0	0	5,602,519	4,745,253	3,970,070	4,307,939	5	7	6,320,015	4,505,297	0	0
D	0	0	152	835	122	146	0	0	33	17	0	0
E	1	0	5,694,112	5,073,087	5,038,200	4,427,275	4	7	5,338,560	4,670,039	0	0
F	0	0	386	780	211	117	0	0	26	16	0	0
G	1	0	5,351,334	5,149,142	5,305,803	5,135,137	4	5	5,645,723	5,237,682	0	0
H	0	0	279	297	127	371	0	0	9	8	0	0

Figure 4. Magnis Contamination Assay

Results

For all sequencing metrics, the Magnis NGS Prep system is equivalent in generating output libraries for all input DNA quality grades, regardless of shearing method (Figure 2).

Similarly, in all Magnis-generated libraries using AIO-Lung bait, mutations for all Input DNA types are detected (Figure 3).

No reads above noise level were detected for no template control (NTC) samples, or samples that were not added to the sequencing pool (Figure 4).

No interrun contamination above noise level was detected (Figure 5).

No intrarun contamination above noise level was detected. (Data not shown.)

Conclusions

The Agilent Magnis NGS Prep system consistently generates sequence-ready libraries from small, medium, and large probe panels using a simple setup and without the need for extensive training.

No sample-to-sample or run-to-run contamination was detectable for the dual indexed libraries.

SNPs down to one percent allele frequency were detected with Magnis prepared libraries using the AIO-Lung Probe set.

Sample processing using the fully automated Magnis instrument gave high library complexity even for challenging samples. (Data not shown.)

Acknowledgments

Agilent Technologies Diagnostics & Genomics Group (DGG), Mohd-Aizuddin Kamaruddin, Oi-Chu Lim, Jia-Wei Lee, Sim-Siang Teh, Yvonne Lee, Schryl Alonzo Castaneda, Rajesh Bagga, John Halladay, Kelly Flesher, Sarah Jandle, Dawn Lesperance, Carlos Pabon, Gilbert Amparo, Neelima Mehendale