molecularloop

POSTER ABSTRACT

Capture and sequencing of SARS-CoV-2 RNA using molecular inversion probes

Authors: Eric D. Boyden, Patrick C. Saunders, Gregory J. Porreca – *Molecular Loop Biosciences, Inc.*

Background

One of the most effective tools for fighting the COVID-19 pandemic is widespread testing. Diagnostic tests must be both rapid and sufficiently sensitive to detect low viral titers. Subsequently, positive samples should be fully sequenced to identify and track new variants that could impact transmissibility, pathology, and vaccine efficacy.

Molecular Loop's molecular inversion probe (MIP)-based capture chemistry is a simple, highly scalable, and easily automatable system to enable targeted high throughput sequencing with a protocol that can be performed in less than 8 hours. Our MIPs contain unique molecular identifiers (UMIs) that are vital for deduplication of reads and estimating the number of captured molecules, and library amplification is performed using unique dual index (UDI) primers to minimize the impact of index hopping. We hypothesized that by integrating reverse transcription of cDNA and MIP hybridization into a single step, we could facilitate capture of SARS-CoV-2 RNA without increasing the workflow complexity (Figure 1), and that the workflow could be optimized for either rapid RNA detection or comprehensive sequencing by adjusting the hybridization time and sequencing parameters. Since tiled MIPs can capture multiple diagnostic loci and are highly resistant to variant allele dropout, we propose that they can be used for both detection of acute infection and ongoing strain monitoring within a population.

Methods and Results

We designed a set of ~1000 MIPs tiled across 99.6% of the SARS-CoV-2 genome, with most bases targeted by 7 – 8 MIPs. Capture was performed on variable amounts of synthetic RNA from either of two SARS-CoV-2 strains (Twist Bioscience) and 10 ng of human gDNA, using either a 16-hour or a 4-hour hybridization. Paired-end sequencing was performed on an Illumina HiSeq 2500, and data were analyzed using standard tools (Figure 2).

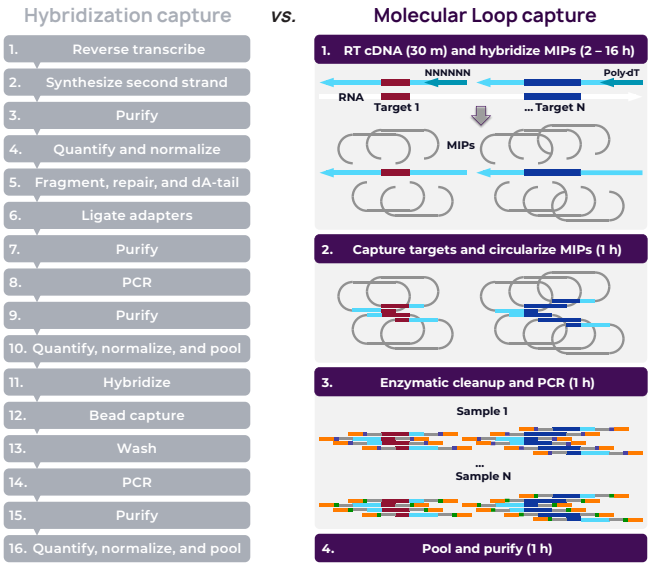

Figure 1. Molecular Loop's workflow is dramatically simpler

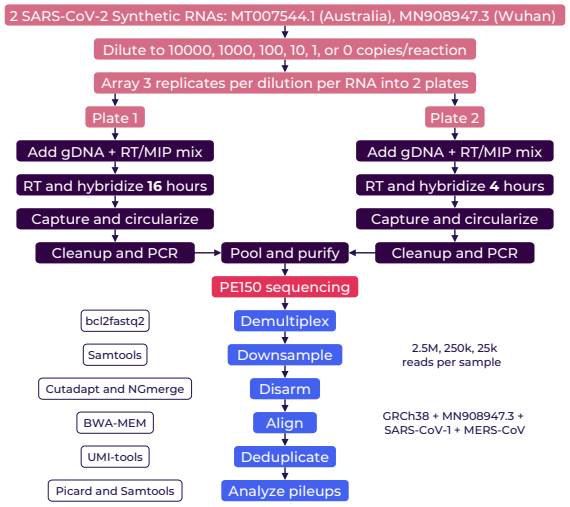

Figure 2. Experimental design and analysis

This study was funded by Molecular Loop Biosciences, Inc.

Our assay was sensitive to 1 copy of SARS-CoV-2 RNA with only 25k reads, whereas 0 RNA copies produced no coverage even with 2.5M reads (Figures 3A and 4A). With 10000 RNA copies and 250k reads, genome coverage was 98% at 1X, 97% at 5X, and 96% at 10X depth using either workflow; with 1000 RNA copies, genome coverage was 97% (16 h) or 96% (4 h) at 1X, 89% or 84% at 5X, and 78% or 65% at 10X depth (Figures 3 and 4). Nearly all variant positions in the Brazilian, South African, and UK SARS-CoV-2 strains had ≥10X coverage with 10000 RNA copies and 250k reads (Table 1). All known variants in the MT007544.1 RNA were successfully captured (Table 2).

In conclusion, our MIP target capture technology is a robust strategy for simple, rapid, and sensitive detection or comprehensive sequencing of SARS-CoV-2 RNA.

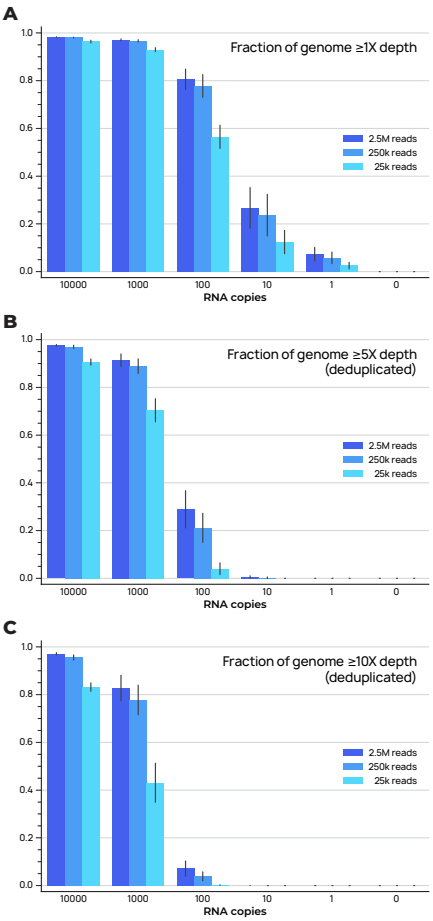

Figure 3. 16-hour hybridization

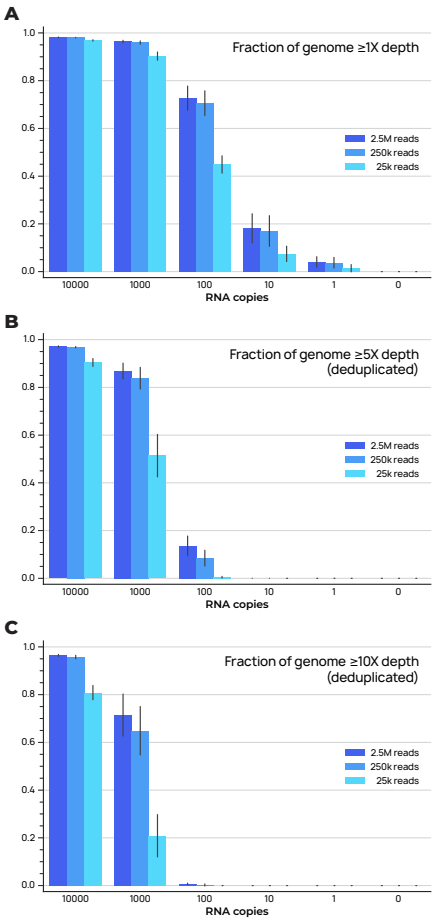

Figure 4. 4-hour hybridization

Table 2. MT007544.1 variant reads (deduplicated)

Gene	Variant		16 h 2.5M	16 h 250k	4 h 2.5M	4 h 250k
orf1ab	19065T>C (L6267P)	%	100	99.3	100	100
		#	61	39	28	20
spike	22303T>G (S247R)	%	99.9	100	99.9	99.8
		#	798	273	596	399
orf3a	26144G>T (G251V)	%	99.8	99.9	99.8	99.9
		#	711	474	353	255
	29750-29759del	%	100	97.5	97.5	100
		#	42	21	24	14

Table 1. Variant position mean depth (deduplicated)

Gene	P.1 (Brazil) Variant	16 h 2.5M	16 h 250k	4 h 2.5M	4 h 250k
orf1ab	733T>C (silent)	513	372	398	296
	2749C>T (silent)	152	105	125	89
	3828C>T (S1188L)	205	131	92	66
	5648A>C (K1795Q)	336	216	189	136
	11288-11296del	299	182	170	126
	12778C>T (silent)	107	76	96	67
	13860C>T (silent)	144	89	88	62
	17259G>T (E5665D)	413	270	247	178
spike	21614C>T (L18F)	35	20	57	41
	21621C>A (T20N)	33	19	56	38
	21638C>T (P26S)	31	19	53	36
	21974G>T (D138Y)	196	108	148	105
	22132G>T (R190S)	642	247	483	343
	22812A>C (K417T)	23	9	23	15
	23012G>A (E484K)	239	141	155	104
	23063A>T (N501Y)	353	200	253	178
	23525C>T (H655Y)	186	127	144	103
	24642C>T (T1027I)	307	200	167	121
orf8	28167G>A (E92K)	418	255	301	210
	28269-28273ins	639	405	367	257
N	28512C>G (P80R)	367	260	361	263

Gene	B.1.351 (SA) Variant	16 h 2.5M	16 h 250k	4 h 2.5M	4 h 250k
orf1ab	1059C>T (T265I)	132	87	74	55
	5230G>T (K1655N)	488	318	215	156
	8660C>T (H2799Y)	200	115	139	94
	8964C>T (S2900L)	153	97	117	82
	10323A>G (K3353R)	202	120	197	136
	13843G>T (D4527Y)	136	86	81	60
	17999C>T (T5912I)	136	76	98	66
spike	21614C>T (L18F)	35	20	57	41
	21801A>C (D80A)	548	364	447	346
	22206A>G (D215G)	321	132	270	176
	22286-22294del	808	292	650	443
	22287T>A (L242H)	813	294	653	445
	22299G>T (R246I)	753	278	600	410
	22813G>T (K417N)	21	8	21	14
	23012G>A (E484K)	239	141	155	104
	23063A>T (N501Y)	353	200	253	178
	23664C>T (A701V)	161	111	115	86
orf3a	25563G>T (Q57H)	1472	914	809	589
	25904C>T (S171L)	43	21	28	19
E	26456C>T (P71L)	63	11	99	66
N	28887C>T (T205I)	379	262	216	154

Gene	B.1.1.7 (UK) Variant	16 h 2.5M	16 h 250k	4 h 2.5M	4 h 250k
orf1ab	3267C>T (T1001I)	122	78	93	64
	5388C>A (A1708D)	657	437	314	229
	6954T>C (I2230T)	17	4	27	18
	11288-11296del	299	182	170	126
spike	21765-21770del	251	162	263	182
	21991-21993del	172	94	115	82
	23063A>T (N501Y)	353	200	253	178
	23271C>A (A570D)	114	72	76	53
	23604C>A (P681H)	146	103	127	94
	23709C>T (T716I)	135	89	92	67
	24506T>G (S982A)	119	70	133	94
	24914G>C (D1118H)	76	44	40	26
orf8	27972C>T (Q27X)	307	179	247	167
	28048G>T (R52I)	364	202	275	192
	28111A>G (Y73C)	344	214	196	143
N	28280GAT>CTA (D3L)	668	422	398	278
	28977C>T (S235F)	729	522	391	293

molecularloop
Molecular Loop Biosciences, Inc.
90 Bridge Street, Newton, MA 02458
www.molecularloop.com | sales@molecularloop.com