Guidance of the data required for genome assemble using ATOPlex SARS-CoV-2 sequencing on DNBSEQ platform

Introduction

ATOPlex SARS-CoV-2 sequencing is a 2-step multiplex-PCR based Massively Parallel Sequencing (mPCR-based MPS) performed on DNBSEQ platform. It is designed to detect and to assemble the SARS-CoV-2 genome from nasopharyngeal (NP) swabs, oropharyngeal (OP) swabs, plasma, and bronchoalveolar lavage (BAL) specimens of suspected or confirmed COVID-19 individuals. The multiplex PCR of ATOPlex is a short amplicon-based（106-199bp）PCR with high specificity and sensitivity. It allows whole SARS-CoV-2 virus genomes to be generated from samples with PCR ct values of up to 35. Partial genomes can be generated from samples with ct values of 40. ATOPlex SARS-CoV-2 sequencing can be used to obtain the full-length genome sequence of SARS-CoV-2 for population-scale virus detection and surveillance.

Accurate detection of a mutation requires at least 100X sequencing depth using amplicon-based method[1]. Adequate sequencing data is important for SARS-CoV-2 genome assembly. This guideline provides the data requirement for genome assembly using ATOPlex SARS-CoV-2 sequencing on DNBSEQ platform.

Performance of ATOPlex for clinical sample

Experimental methods

We validate ATOPlex SARS-CoV-2 sequencing in a dataset of 75 samples. Library preparation was performed with the ATOPlex RNA Library Prep Set(MGI, China). Briefly, 10ul total RNA of each sample was converted to the first-strand cDNA products by RNA reverse transcriptase with random hexamers (5'-NNNNNN-3'). 1st strand cDNA products were then amplified by ATOPlex SARS-CoV-2 full-length genome panel and universal barcode primer. After amplification, bead cleans up, and quantification, libraries were equimolar pooled and sequenced on DNBSEQ-G400* with PE100. Raw sequenced reads were analysed according to the SARS-CoV-2_Multi-PCR_v1.0 pipeline on github (https://github.com/MGI-tech-bioinformatics/SARS-CoV-2_Multi-PCR_v1.0).

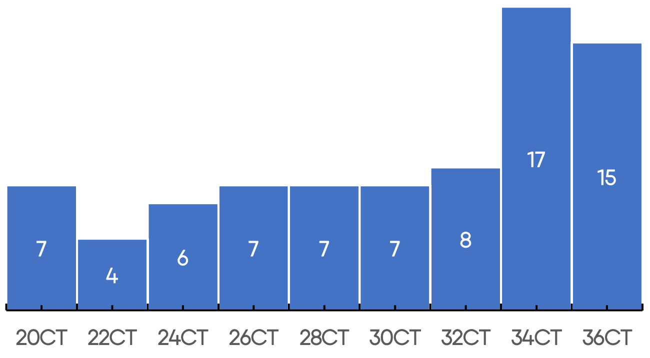

Figure1. Distribution of ct value

Results

The ct value of each sample were shown in figure1. 13-42 millions (mean=15M) PE reads was generated for each sample on DNBSEQ-G400*. After analyzed by ATOPlex multiplex pipeline, 60 samples get 99.5% genome coverage with 100X sequence depth, 65 samples get 95.0% genome coverage with 100X depth (Table1).

Data requirement for SASR-CoV-2 genome assembly

To evaluate the data requirement for genome at 100X sequencing depth, we set 95.0% and 99.5% genome coverage as the high-quality and super-quality genome and measured the minimum PE reads for samples with different ct value for genome assembly (Figure2).

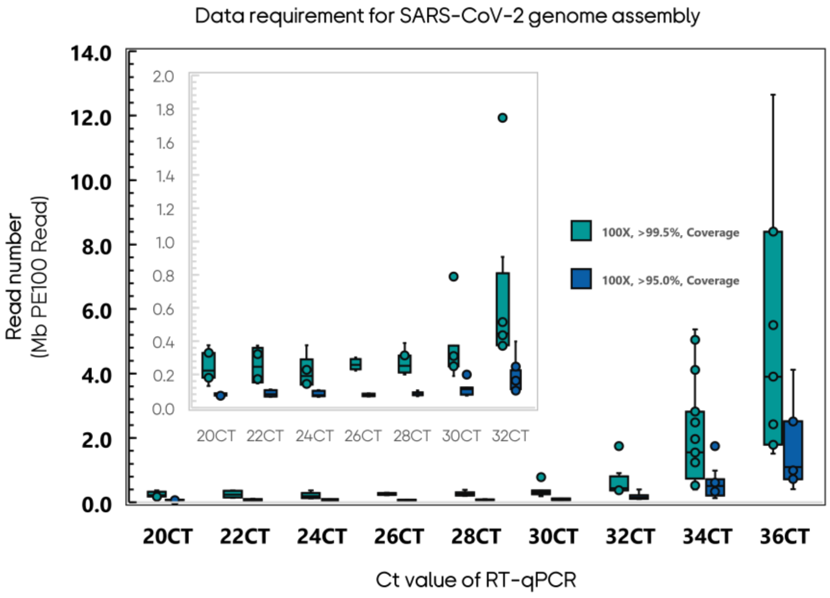

Figure2. Data requirement for SARS-CoV-2 genome assembly with different ct values

0.25M and 1.0M PE reads are sufficient for high-quality genome and super-quality genome in sample with ct values range from 20 to 30 , respectively. The mean sequencing depth for each sample range from 1200X to 6500X (Table2).

Table 1 SARS-CoV-2 genome coverage in the samples with different ct value

	% of sample cover >99.5% genome with 100X depth					% of sample cover >95.0% genome with 100X depth				
Sample size	0.25M Reads	0.5M Reads	1.0M Reads	2.5M Reads	10M Reads	0.25M Reads	0.5M Reads	1.0M Reads	2.5M Reads	10M Reads
20CT(19-20), n=7;	57.1%	100.0%	100.0%	**100.0%**	100.0%	100.0%	100.0%	100.0%	**100.0%**	100.0%
22CT(21-22), n=4;	25.0%	100.0%	100.0%	**100.0%**	100.0%	100.0%	100.0%	100.0%	**100.0%**	100.0%
24CT(23-24), n=6;	33.3%	100.0%	100.0%	**100.0%**	100.0%	100.0%	100.0%	100.0%	**100.0%**	100.0%
26CT(25-26), n=4;	25.0%	100.0%	100.0%	**100.0%**	100.0%	100.0%	100.0%	100.0%	**100.0%**	100.0%
28CT(27-28), n=7;	0.0%	100.0%	100.0%	**100.0%**	100.0%	100.0%	100.0%	100.0%	**100.0%**	100.0%
30CT(29-30), n=7;	0.0%	85.7%	100.0%	**100.0%**	100.0%	100.0%	100.0%	100.0%	**100.0%**	100.0%
32CT(31-32), n=8;	0.0%	37.5%	75.0%	**100.0%**	100.0%	75.0%	100.0%	100.0%	**100.0%**	100.0%
34CT(33-34), n=17;	0.0%	0.0%	23.5%	**70.6%**	76.5%	17.6%	35.3%	64.7%	**88.2%**	88.2%
36CT(35-36), n=15;	0.0%	0.0%	0.0%	**20.0%**	26.7%	0.0%	6.7%	13.3%	**46.7%**	46.7%

Table 2 Sequencing depth of the samples with different ct value

	Mean depth of SARS-CoV-2 genome in different sequence data				
CT	0.25M Reads	0.5M Reads	1M Reads	2.5M Reads	10M Reads
20CT(19-20), n=7;	1640.2±24.2	3280.7±47.6	6560.8±95.8	**15302.8±2649.3**	47726.6±15056.9
22CT(21-22), n=4;	1647.3±2.9	3294.5±6.3	6588.8±11.9	**13986.6±2500.3**	36467.7±16108.5
24CT(23-24), n=6;	1628.4±21.4	3256.8±42.9	6514.2±85.9	**11295.3±5500.4**	29294.4±18705.8
26CT(25-26), n=4;	1592.5±34.4	3184.1±68.6	6369±136.3	**15697±391.1**	38046±8337.1
28CT(27-28), n=7;	1425.2±148.9	2851±297.7	5702.3±593.8	**13793.2±1674.5**	35670.8±10971.5
30CT(29-30), n=7;	1260.2±285.4	2519.7±571.2	5039.7±1142.3	**11087.5±2270.8**	29090.2±9475.6
32CT(31-32), n=8;	837.1±260	1674.5±519.7	3350.2±1039.9	**7789.6±2730.4**	21299.4±9066.3
34CT(33-34), n=17;	340.1±259.7	680.3±519.5	1361.3±1038	**3114.2±2656.4**	8302±7450.1
36CT(35-36), n=15;	124.2±63.6	249±127.6	498.5±253.9	**1061.8±564.6**	2333±1110.6

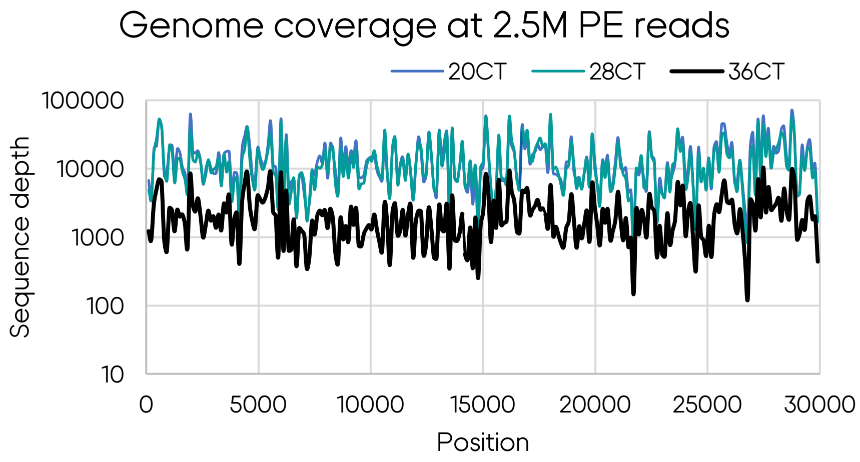

Figure3. SARS-CoV2-2 genome coverage

The ATOPlex SARS-CoV-2 sequencing is able to acquire the high-quality and super-quality genome in the sample with ct values of 32 and 30 with 1M PE reads. Even in samples with high ct values of 34 and 36, 88.2% and 46.7% of samples obtained the high-quality genome at 2.5M PE reads, respectively (Table1). ATOPlex also gets excellent genome uniformity in the sample with different ct values (Figure 3).

Summary

ATOPlex is a highly sensitive method for genome assembly in clinical sample with high ct value .

Guidance for sequencing on DNBSEQ platform

DNBSEQ-G400* and DNBSEQ-G50* are high-throughput sequencers and can generate more than 400M PE reads in one lane. Therefore, multiplex ATOPlex libraries can be pooled and parallel sequenced in one lane.

The number of samples parallelly sequenced on DNBSEQ sequencer depends on the amount of data required for each sample. Although ATOPlex able to obtain super-quality genome in the sample with the ct value of 30 or even 36 with 1.0 M PE reads, it is very difficult to accurately predict the amount of data required for a sample according to the ct value before sequencing.

Firstly, ct value is not a perfect predictor of amplification success as it can vary between different diagnostic methods and different sample types. Secondly, specimen collection, storage, transportation and RNA extraction affect the quality of the RNA and lead to a various quality of the ATOPlex library. Thirdly, even by pooling ATOPlex library with equimolar, library with different barcode has a various efficiency at ssCir and DNB making process that will lead to unequal amount of data .

For population-scale sequencing, quantification and normalization of ATOPlex libraries based on the ct value is time-consuming and costly . For super-quality genome, we recommend to pool 96 samples (~3M per sample) with equal volume regardless of ct value and sequence one lane on DNBSEQ-G400* or DNBSEQ-G50*. For high-quality genome, we recommend to pool 300 samples (~1M per sample) with equal volume regardless of ct value sequence one lane on DNBSEQ-G400* or DNBSEQ-G50*.

Pooling library with variable molar based on the ct value as the table below can be an alternative way.

Table 3. The amount of data required for super-quality genome

ct value of RT-qPCR	Data requirement	Max sample no. on DNBSEQ-G400*	Max samples no. on DNBSEQ-G50*
≤30	1.0M	300/lane	300/lane
30<ct≤34	2.5M	120/lane	120/lane
34<ct≤36	5.0M	60/lane	60/lane

MGI only provides 96 dual-barcode primers in the ATOPlex RNA Library Prep Set. More dual-barcode primers need to be synthesized by the customer from the third-party agent (eg. IDT, Twist, GENEWIZ, Sangon).

Dual-barcode primer used in ATOPlex

Dual-barcode primer mix in the 2nd PCR of ATOPlex includes barcode primer 1 and barcode primer 2. The barcode sequence in barcode primer 1 and barcode primer 2 is unique and paired (eg, 1-1,2-2,3-3…96-96). More barcode primers are listed in the appendix.

How to use the dual-barcode primer

Dual-barcode primer are diluted with NF-water or TE to 5uM. 4ul of the dual barcode primer 1 (5uM) and barcode primer 2 (5uM) are premixed and added into the 2nd PCR reaction. The schematic of the dual-barcode primer is shown in figure 4.

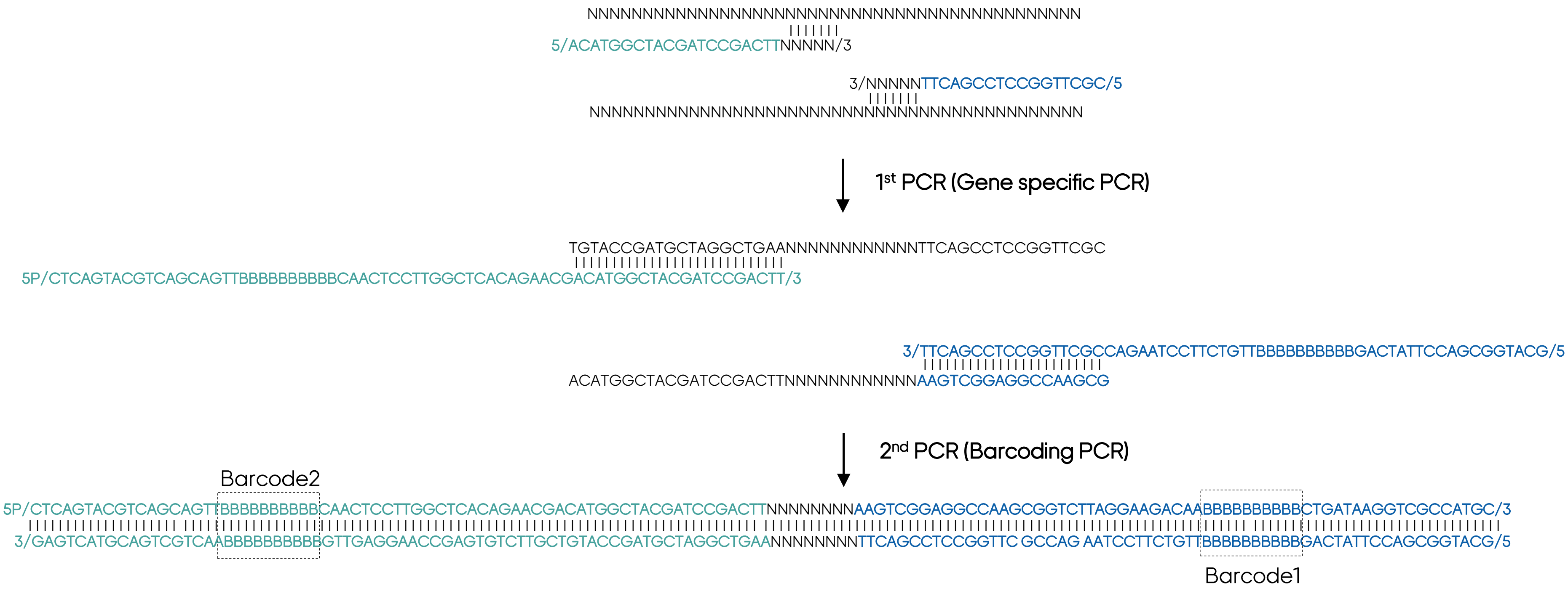

Figure4. Schematic of ATOPlex 2-step multiplex PCR

Reference

1. Xiao et al. Genome Medicine (2020) 12:57, https://doi.org/10.1186/s13073-020-00751-4

MGI Tech Co.,Ltd

Address: Building 11,Beishan Industrial Zone, Yantian District,

Shenzhen, CHINA 518083

Email: MGI-service@mgi-tech.com, MGI_ATOPlex@mgi-tech.com Website: en.mgi-tech.com

Tel: 4000-966-988

Version : March 2021

https://www.linkedin.com/company/mgi-bgi

https://twitter.com/MGI_BGI

Appendix：Dual-barcode primer sequence used in ATOPlex

Table 4. The sequence of the dual-barcode primer

PRIMER ID	Sequence
Barcode primer 1	5/GCATGGCGACCTTATCAGGAACTTGAATTTGTCTTCCTAAGACCGCTTGGCCTCCGACTT/3
Barcode primer 2	5P/CTCTCAGTACGTCAGCAGTTATTCAAGTTCCAACTCCTTGGCTCACAGAACGACATGGCTACGATCCGACTT/3#

#Phosphorylation of the first base at the 5 ends of barcode primer 2, 10bp barcode sequence marked blue. Substitute the barcode sequence (marked blue) in barcode primer1 and barcode primer2 with the barcode1 sequence and barcode2 sequence according to the following list.

Oligo purifies by PAGE or HPLC.

Table 5. Barcode sequences of the dual-barcode primer (ATOPlex 960-01#)

Barcode primer 1 ID	Barcode1 sequence	Barcode primer 2 ID	Barcode2 sequence	Barcode primer 1 ID	Barcode1 sequence	Barcode primer 2 ID	Barcode2 sequence
BC1-131	GAACTTGAAT	BC2-131	ATTCAAGTTC	BC1-179	GTCGTACTCT	BC2-179	AGAGTACGAC
BC1-132	CGGAAGCCGT	BC2-132	ACGGCTTCCG	BC1-180	TACCTTCGAT	BC2-180	ATCGAAGGTA
BC1-133	TCTGCCATTG	BC2-133	CAATGGCAGA	BC1-181	CCAACGTCTG	BC2-181	CAGACGTTGG
BC1-134	ATCTGATGCC	BC2-134	GGCATCAGAT	BC1-182	AGTTACGAGC	BC2-182	GCTCGTAACT
BC1-135	GTTCAGGTTA	BC2-135	TAACCTGAAC	BC1-183	AAGTAGATTA	BC2-183	TAATCTACTT
BC1-136	TGATCTACGG	BC2-136	CCGTAGATCA	BC1-184	TCAGGTACGG	BC2-184	CCGTACCTGA
BC1-137	CCGATATACA	BC2-137	TGTATATCGG	BC1-185	CGTCCATGAC	BC2-185	GTCATGGACG
BC1-138	AACGGCCGAC	BC2-138	GTCGGCCGTT	BC1-186	GTGAGCGACA	BC2-186	TGTCGCTCAC
BC1-139	TCGTCATGCA	BC2-139	TGCATGACGA	BC1-187	CGTGCGAGTC	BC2-187	GACTCGCACG
BC1-140	TACGCTATGC	BC2-140	GCATAGCGTA	BC1-188	TACTCTCATA	BC2-188	TATGAGAGTA
BC1-141	CTACAGGCTT	BC2-141	AAGCCTGTAG	BC1-189	GTGAGCGTGT	BC2-189	ACACGCTCAC
BC1-142	AGTATCCAAG	BC2-142	CTTGGATACT	BC1-190	ACACTATCCG	BC2-190	CGGATAGTGT
BC1-143	GGACGTAGTT	BC2-143	AACTACGTCC	BC1-191	ACCAATGTAT	BC2-191	ATACATTGGT
BC1-144	GTTAGACCGG	BC2-144	CCGGTCTAAC	BC1-192	CATTGCTCGG	BC2-192	CCGAGCAATG
BC1-145	CCGGTCGTAA	BC2-145	TTACGACCGG	BC1-193	GTGCTACACA	BC2-193	TGTGTAGCAC
BC1-146	AACTAGTACC	BC2-146	GGTACTAGTT	BC1-194	TGAGAGAGAC	BC2-194	GTCTCTCTCA
BC1-147	GTGCGCAGCT	BC2-147	AGCTGCGCAC	BC1-195	ACTCGATCGA	BC2-195	TCGATCGAGT
BC1-148	TTCATTAATC	BC2-148	GATTAATGAA	BC1-196	CCGGCTTACT	BC2-196	AGTAAGCCGG
BC1-149	CAAGCATTGG	BC2-149	CCAATGCTTG	BC1-197	GGCAAGGTTG	BC2-197	CAACCTTGCC
BC1-150	ACATAGGCAA	BC2-150	TTGCCTATGT	BC1-198	TAATTCCGAC	BC2-198	GTCGGAATTA
BC1-151	CGTGAGCGTT	BC2-151	AACGCTCACG	BC1-199	CTCAGTAGTT	BC2-199	AACTACTGAG
BC1-152	GGCCGATCGG	BC2-152	CCGATCGGCC	BC1-200	ATTCCGATGG	BC2-200	CCATCGGAAT
BC1-153	TAGTTCCTAA	BC2-153	TTAGGAACTA	BC1-201	TAGGAACCAA	BC2-201	TTGGTTCCTA
BC1-154	ACTACTGACC	BC2-154	GGTCAGTAGT	BC1-202	GGATTCGACC	BC2-202	GGTCGAATCC
BC1-155	TGATCCGGCA	BC2-155	TGCCGGATCA	BC1-203	TTCGCGTATC	BC2-203	GATACGCGAA
BC1-156	GGTCATGAGT	BC2-156	ACTCATGACC	BC1-204	CTGGCCTGTG	BC2-204	CACAGGCCAG
BC1-157	ATGAGACTTG	BC2-157	CAAGTCTCAT	BC1-205	ACTCTAGTGT	BC2-205	ACACTAGAGT
BC1-158	CACATGTCAC	BC2-158	GTGACATGTG	BC1-206	GGAAGTACCA	BC2-206	TGGTACTTCC
BC1-159	CCAGGTACTT	BC2-159	AAGTACCTGG	BC1-207	AACTAGCTAT	BC2-207	ATAGCTAGTT
BC1-160	GTGCTATAGG	BC2-160	CCTATAGCAC	BC1-208	GAGCTTACGG	BC2-208	CCGTAAGCTC
BC1-161	ACTGACCGAA	BC2-161	TTCGGTCAGT	BC1-209	TCATGACGAA	BC2-209	TTCGTCATGA
BC1-162	TACTCGATCC	BC2-162	GGATCGAGTA	BC1-210	CGTAACGACC	BC2-210	GGTCGTTACG
BC1-163	CAGTGCCATA	BC2-163	TATGGCACTG	BC1-211	TCAACGGTAG	BC2-211	CTACCGTTGA
BC1-164	CCTACCTAAC	BC2-164	GTTAGGTAGG	BC1-212	ATGGTATCTC	BC2-212	GAGATACCAT
BC1-165	TGACATGTGT	BC2-165	ACACATGTCA	BC1-213	CACTATCGGT	BC2-213	ACCGATAGTG
BC1-166	ATCGTGAGCG	BC2-166	CGCTCACGAT	BC1-214	GGTCTCTACA	BC2-214	TGTAGAGACC
BC1-167	GACAGAGCTT	BC2-167	AAGCTCTGTC	BC1-215	GAGAGTAGTT	BC2-215	AACTACTCTC
BC1-168	AGATCTCTGG	BC2-168	CCAGAGATCT	BC1-216	TGTGAACTGG	BC2-216	CCAGTTCACA
BC1-169	GTGCAATCAC	BC2-169	GTGATTGCAC	BC1-217	ACCTGCGCAA	BC2-217	TTGCGCAGGT
BC1-170	TCTGTGAGCA	BC2-170	TGCTCACAGA	BC1-218	CTACCGAACC	BC2-218	GGTTCGGTAG
BC1-171	TCTAGATCAC	BC2-171	GTGATCTAGA	BC1-219	GGAATCCTAC	BC2-219	GTAGGATTCC
BC1-172	ACTATACAGA	BC2-172	TCTGTATAGT	BC1-220	TGCAACGACG	BC2-220	CGTCGTTGCA
BC1-173	GACGACGGTT	BC2-173	AACCGTCGTC	BC1-221	AAGCCTAGTT	BC2-221	AACTAGGCTT
BC1-174	CGACCTATCG	BC2-174	CGATAGGTCG	BC1-222	CCTGCGTCGA	BC2-222	TCGACGCAGG
BC1-175	TAGTTGCGTT	BC2-175	AACGCAACTA	BC1-223	GTCTGACGTT	BC2-223	AACGTCAGAC
BC1-176	ATATGTATGG	BC2-176	CCATACATAT	BC1-224	TATCAGATGG	BC2-224	CCATCTGATA
BC1-177	GTCGACTCAA	BC2-177	TTGAGTCGAC	BC1-225	ACGTGATACA	BC2-225	TGTATCACGT
BC1-178	CGGCCGGACC	BC2-178	GGTCCGGCCG	BC1-226	CTAGTTGCAC	BC2-226	GTGCAACTAG

#Barcode cross-talk analysis *in silico*. However, performance with these barcodes has not been validated in clinical samples.

Table 6. Barcode sequences of the dual-barcode primer (ATOPlex 960-02[#])

Barcode primer 1 ID	Barcode1 sequence	Barcode primer 2 ID	Barcode2 sequence	Barcode primer 1 ID	Barcode1 sequence	Barcode primer 2 ID	Barcode2 sequence
BC1-227	ACGATCCATA	BC2-227	TATGGATCGT	BC1-275	TGGTTAATCT	BC2-275	AGATTAACCA
BC1-228	CCATTCAGCT	BC2-228	AGCTGAATGG	BC1-276	CGGCATAGCT	BC2-276	AGCTATGCCG
BC1-229	GGCCAGTTGG	BC2-229	CCAACTGGCC	BC1-277	GACACGTCTG	BC2-277	CAGACGTGTC
BC1-230	TATGCTGCAC	BC2-230	GTGCAGCATA	BC1-278	ATTGCCGAGC	BC2-278	GCTCGGCAAT
BC1-231	ATTGGTCGTT	BC2-231	AACGACCAAT	BC1-279	CCAAGGCTAA	BC2-279	TTAGCCTTGG
BC1-232	TGCCGAATGG	BC2-232	CCATTCGGCA	BC1-280	GACGTTCGTG	BC2-280	CACGAACGTC
BC1-233	GAATCATACA	BC2-233	TGTATGATTC	BC1-281	TCATACGAGA	BC2-281	TCTCGTATGA
BC1-234	CTGAAGGCAC	BC2-234	GTGCCTTCAG	BC1-282	ATTCGATCAC	BC2-282	GTGATCGAAT
BC1-235	CAAGCCAGCC	BC2-235	GGCTGGCTTG	BC1-283	AGATGCAGAG	BC2-283	CTCTGCATCT
BC1-236	CGTCTCTATG	BC2-236	CATAGAGACG	BC1-284	ACATTCACAG	BC2-284	CTGTGAATGT
BC1-237	GTCAATGTGT	BC2-237	ACACATTGAC	BC1-285	CTCAAGGATT	BC2-285	AATCCTTGAG
BC1-238	ACGTGACCAA	BC2-238	TTGGTCACGT	BC1-286	TATCCATTGC	BC2-286	GCAATGGATA
BC1-239	TCCGAAGGTT	BC2-239	AACCTTCGGA	BC1-287	GACGATCTCA	BC2-287	TGAGATCGTC
BC1-240	GGATGGTCGG	BC2-240	CCGACCATCC	BC1-288	TCGACAGCTT	BC2-288	AAGCTGTCGA
BC1-241	TTGATGCTAA	BC2-241	TTAGCATCAA	BC1-289	CGGCTGCACA	BC2-289	TGTGCAGCCG
BC1-242	AATCCTAACC	BC2-242	GGTTAGGATT	BC1-290	GTTGGTTGGC	BC2-290	GCCAACCAAC
BC1-243	GAGACGACGC	BC2-243	GCGTCGTCTC	BC1-291	TGCTGTACTA	BC2-291	TAGTACAGCA
BC1-244	GGCTAGGATT	BC2-244	AATCCTAGCC	BC1-292	GTACACATCC	BC2-292	GGATGTGTAC
BC1-245	ACAGGCTTCG	BC2-245	CGAAGCCTGT	BC1-293	CATACATGGT	BC2-293	ACCATGTATG
BC1-246	CTACTTCGAA	BC2-246	TTCGAAGTAG	BC1-294	ACGCTGGAAG	BC2-294	CTTCCAGCGT
BC1-247	TGTACCGGTT	BC2-247	AACCGGTACA	BC1-295	CCAGCGCGTT	BC2-295	AACGCGCTGG
BC1-248	CATTGAATGG	BC2-248	CCATTCAATG	BC1-296	ATGAGTTCGG	BC2-296	CCGAACTCAT
BC1-249	TTGCTACCAA	BC2-249	TTGGTAGCAA	BC1-297	GACTACGACA	BC2-297	TGTCGTAGTC
BC1-250	ACCGATTACC	BC2-250	GGTAATCGGT	BC1-298	TGTGTACTAC	BC2-298	GTAGTACACA
BC1-251	CGTTCGTCGT	BC2-251	ACGACGAACG	BC1-299	ACACACAGAG	BC2-299	CTCTGTGTGT
BC1-252	GTTGAGACGT	BC2-252	ACGTCTCAAC	BC1-300	CATGCTCGAT	BC2-300	ATCGAGCATG
BC1-253	TACAGCGTTG	BC2-253	CAACGCTGTA	BC1-301	GGCTTGGTTC	BC2-301	GAACCAAGCC
BC1-254	ACGACTTGCC	BC2-254	GGCAAGTCGT	BC1-302	TTGAGATCGA	BC2-302	TCGATCTCAA
BC1-255	CAACTACTAA	BC2-255	TTAGTAGTTG	BC1-303	CTTGATTACT	BC2-303	AGTAATCAAG
BC1-256	ACCGACGATG	BC2-256	CATCGTCGGT	BC1-304	TCGTCGCTTG	BC2-304	CAAGCGACGA
BC1-257	GTGTGACGAA	BC2-257	TTCGTCACAC	BC1-305	AGCAGAGACA	BC2-305	TGTCTCTGCT
BC1-258	TGACTTAACC	BC2-258	GGTTAAGTCA	BC1-306	GAACTCACGC	BC2-306	GCGTGAGTTC
BC1-259	TCCTTAGTCT	BC2-259	AGACTAAGGA	BC1-307	ACGTCATCAC	BC2-307	GTGATGACGT
BC1-260	ACGATCACGC	BC2-260	GCGTGATCGT	BC1-308	GCGTCTCGGC	BC2-308	GCCGAGACGC
BC1-261	CATCAGTGTG	BC2-261	CACACTGATG	BC1-309	TTCGACGATT	BC2-309	AATCGTCGAA
BC1-262	GTAGCTCTAA	BC2-262	TTAGAGCTAC	BC1-310	CATCGGATCG	BC2-310	CGATCCGATG
BC1-263	TGAAGTGATT	BC2-263	AATCACTTCA	BC1-311	CGAAGCGTTA	BC2-311	TAACGCTTCG
BC1-264	CTGTGATCGG	BC2-264	CCGATCACAG	BC1-312	TACGTACGGT	BC2-312	ACCGTACGTA
BC1-265	GACGCGCACA	BC2-265	TGTGCGCGTC	BC1-313	GGTATGTAAG	BC2-313	CTTACATACC
BC1-266	AGTCACAGAC	BC2-266	GTCTGTGACT	BC1-314	ATACATACCA	BC2-314	TGGTATGTAT
BC1-267	GTCGCTGTCT	BC2-267	AGACAGCGAC	BC1-315	CGCTAATGGC	BC2-315	GCCATTAGCG
BC1-268	CAGGTCTACT	BC2-268	AGTAGACCTG	BC1-316	AGTCTACCGG	BC2-316	CCGGTAGACT
BC1-269	TCTAAGCGTG	BC2-269	CACGCTTAGA	BC1-317	GTGAGTGATT	BC2-317	AATCACTCAC
BC1-270	AGACGAACGC	BC2-270	GCGTTCGTCT	BC1-318	TAAGTGATCA	BC2-318	TGATCACTTA
BC1-271	CCTTAGTTAA	BC2-271	TTAACTAAGG	BC1-319	ACGACCGTAT	BC2-319	ATACGGTCGT
BC1-272	AACTGACGTG	BC2-272	CACGTCAGTT	BC1-320	TCATCTAGTG	BC2-320	CACTAGATGA
BC1-273	GTAATCACAC	BC2-273	GTGTGATTAC	BC1-321	CTTCAGCCAA	BC2-321	TTGGCTGAAG
BC1-274	TGGCCTGAGA	BC2-274	TCTCAGGCCA	BC1-322	GACGGCTACC	BC2-322	GGTAGCCGTC

[#] Barcode cross-talk analysis *in silico*. However, performance with these barcodes has not been validated in clinical samples.

MGI Tech Co.,Ltd

Address: Building 11,Beishan Industrial Zone, Yantian District,

Shenzhen, CHINA 518083

Email: MGI-service@mgi-tech.com, MGI_ATOPlex@mgi-tech.com Website: en.mgi-tech.com

Tel: 4000-966-988

Version：March 2021

https://www.linkedin.com/company/mgi-bgi

https://twitter.com/MGI_BGI

Table 7. Barcode sequences of the dual-barcode primer (ATOPlex 960-03#)

Barcode primer 1 ID	Barcode1 sequence	Barcode primer 2 ID	Barcode2 sequence	Barcode primer 1 ID	Barcode1 sequence	Barcode primer 2 ID	Barcode2 sequence
BC1-323	TCTACGTCCT	BC2-323	AGGACGTAGA	BC1-371	TATCGTCTCG	BC2-371	CGAGACGATA
BC1-324	TTCGGTTCAG	BC2-324	CTGAACCGAA	BC1-372	TCACAGGACC	BC2-372	GGTCCTGTGA
BC1-325	CGACACGTTC	BC2-325	GAACGTGTCG	BC1-373	AGCTCAAGTT	BC2-373	AACTTGAGCT
BC1-326	GAGTCACGGA	BC2-326	TCCGTGACTC	BC1-374	CTGATCTCGA	BC2-374	TCGAGATCAG
BC1-327	ACAGTGAATT	BC2-327	AATTCACTGT	BC1-375	GTAGATGTAT	BC2-375	ATACATCTAC
BC1-328	ATCCTTCAGG	BC2-328	CCTGAAGGAT	BC1-376	GACAGATGTG	BC2-376	CACATCTGTC
BC1-329	CAGTAAGGAA	BC2-329	TTCCTTACTG	BC1-377	CCGTTCCAGA	BC2-377	TCTGGAACGG
BC1-330	GGTAGCATCC	BC2-330	GGATGCTACC	BC1-378	AGTGCGACAC	BC2-378	GTGTCGCACT
BC1-331	CATATCGTCG	BC2-331	CGACGATATG	BC1-379	GTGCATAGCA	BC2-379	TGCTATGCAC
BC1-332	CTCGACCGTC	BC2-332	GACGGTCGAG	BC1-380	CTCTATGCAC	BC2-380	GTGCATAGAG
BC1-333	GGACCTTCGT	BC2-333	ACGAAGGTCC	BC1-381	AGTACGCATT	BC2-381	AATGCGTACT
BC1-334	TCGTCGTTAA	BC2-334	TTAACGACGA	BC1-382	TAGGTCCTGG	BC2-382	CCAGGACCTA
BC1-335	ACTCGGAATT	BC2-335	AATTCCGAGT	BC1-383	TCATGAGCTT	BC2-383	AAGCTCATGA
BC1-336	TGGTTACAGG	BC2-336	CCTGTAACCA	BC1-384	CAACTCTAGG	BC2-384	CCTAGAGTTG
BC1-337	AACGAAGCAA	BC2-337	TTGCTTCGTT	BC1-385	ACCGGATGCA	BC2-385	TGCATCCGGT
BC1-338	GTAAGTAGCC	BC2-338	GGCTACTTAC	BC1-386	GGTACGATAC	BC2-386	GTATCGTACC
BC1-339	AAGGCCATTA	BC2-339	TAATGGCCTT	BC1-387	AATTCGATAC	BC2-387	GTATCGAATT
BC1-340	TACCGGTCAC	BC2-340	GTGACCGGTA	BC1-388	ACGTAGTACG	BC2-388	CGTACTACGT
BC1-341	CCTAATCGGT	BC2-341	ACCGATTAGG	BC1-389	CGAGTACGTT	BC2-389	AACGTACTCG
BC1-342	ATATCAGACG	BC2-342	CGTCTGATAT	BC1-390	GTCAGTGCGA	BC2-390	TCGCACTGAC
BC1-343	GGAATGCGTT	BC2-343	AACGCATTCC	BC1-391	TCACACAGTT	BC2-391	AACTGTGTGA
BC1-344	GGTCTCTCGG	BC2-344	CCGAGAGACC	BC1-392	TGTGCATCGG	BC2-392	CCGATGCACA
BC1-345	CTCGGAATCA	BC2-345	TGATTCCGAG	BC1-393	CTGCGTCACA	BC2-393	TGTGACGCAG
BC1-346	TCGTATGAAC	BC2-346	GTTCATACGA	BC1-394	GACATCGTAC	BC2-394	GTACGATGTC
BC1-347	TCGGTACCGG	BC2-347	CCGGTACCGA	BC1-395	CCGGAACGAG	BC2-395	CTCGTTCCGG
BC1-348	TAATGAAGTT	BC2-348	AACTTCATTA	BC1-396	TCCTGGCACG	BC2-396	CGTGCCAGGA
BC1-349	ATCAAGTTCC	BC2-349	GGAACTTGAT	BC1-397	GATATCGCTT	BC2-397	AAGCGATATC
BC1-350	CGTCCTGTAA	BC2-350	TTACAGGACG	BC1-398	AGACAGTTGC	BC2-398	GCAACTGTCT
BC1-351	GGTAACGATT	BC2-351	AATCGTTACC	BC1-399	CTGACCATTA	BC2-399	TAATGGTCAG
BC1-352	GAACTGCCGG	BC2-352	CCGGCAGTTC	BC1-400	GGTCCTAGGT	BC2-400	ACCTAGGACC
BC1-353	CTCGGTTACA	BC2-353	TGTAACCGAG	BC1-401	TAATGTGCAA	BC2-401	TTGCACATTA
BC1-354	ACGTCCAGAC	BC2-354	GTCTGGACGT	BC1-402	ATCGTATACC	BC2-402	GGTATACGAT
BC1-355	GATCAGTCTA	BC2-355	TAGACTGATC	BC1-403	CCTGTGTGTG	BC2-403	CACACACAGG
BC1-356	TCATGGAATC	BC2-356	GATTCCATGA	BC1-404	GACAGCACTA	BC2-404	TAGTGCTGTC
BC1-357	AGCACTGTGT	BC2-357	ACACAGTGCT	BC1-405	ATATCAGTGT	BC2-405	ACACTGATAT
BC1-358	CTGGTACGCG	BC2-358	CGCGTACCAG	BC1-406	AGGCATTACC	BC2-406	GGTAATGCCT
BC1-359	TCTAGCGTAT	BC2-359	ATACGCTAGA	BC1-407	TCTGTGCTAT	BC2-407	ATAGCACAGA
BC1-360	GAGTATCCGG	BC2-360	CCGGATACTC	BC1-408	CGATGTCAGG	BC2-408	CCTGACATCG
BC1-361	CGCGCAAGCA	BC2-361	TGCTTGCGCG	BC1-409	GTCACAGCAA	BC2-409	TTGCTGTGAC
BC1-362	ATACTCTAAC	BC2-362	GTTAGAGTAT	BC1-410	TAGCACAGCC	BC2-410	GGCTGTGCTA
BC1-363	ACGGTTAGGC	BC2-363	GCCTAACCGT	BC1-411	GGCAGTGCCT	BC2-411	AGGCACTGCC
BC1-364	GGTAATTCAT	BC2-364	ATGAATTACC	BC1-412	GTTAACGACG	BC2-412	CGTCGTTAAC
BC1-365	TAACCGGTTG	BC2-365	CAACCGGTTA	BC1-413	ACATTACTTC	BC2-413	GAAGTAATGT
BC1-366	CTAAGGCACA	BC2-366	TGTGCCTTAG	BC1-414	TAGCACTGGA	BC2-414	TCCAGTGCTA
BC1-367	TCCTTAAGTT	BC2-367	AACTTAAGGA	BC1-415	CTTGCGATAT	BC2-415	ATATCGCAAG
BC1-368	CACGACTCGG	BC2-368	CCGAGTCGTG	BC1-416	CAACTGCGTG	BC2-416	CACGCAGTTG
BC1-369	GGTTCACTAA	BC2-369	TTAGTGAACC	BC1-417	TCCGGATCAA	BC2-417	TTGATCCGGA
BC1-370	ATGCGCGACC	BC2-370	GGTCGCGCAT	BC1-418	AGGTCTAAGC	BC2-418	GCTTAGACCT

#Barcode cross-talk analysis *in silico*. However, performance with these barcodes has not been validated in clinical samples.

MGI Tech Co.,Ltd

Address: Building 11,Beishan Industrial Zone, Yantian District,

Shenzhen, CHINA 518083

Email: MGI-service@mgi-tech.com, MGI_ATOPlex@mgi-tech.com Website: en.mgi-tech.com

Tel: 4000-966-988

Version：March 2021

https://www.linkedin.com/company/mgi-bgi

https://twitter.com/MGI_BGI

Table 8. Barcode sequences of the dual-barcode primer (ATOPlex 960-04#)

Barcode primer 1 ID	Barcode1 sequence	Barcode primer 2 ID	Barcode2 sequence	Barcode primer 1 ID	Barcode1 sequence	Barcode primer 2 ID	Barcode2 sequence
BC1-419	CCAATCTGTC	BC2-419	GACAGATTGG	BC1-467	TCTGCAGGTA	BC2-467	TACCTGCAGA
BC1-420	CGGCCAATTC	BC2-420	GAATTGGCCG	BC1-468	AGCGAACGCA	BC2-468	TGCGTTCGCT
BC1-421	ATTGATGCGT	BC2-421	ACGCATCAAT	BC1-469	CAATGGATGT	BC2-469	ACATCCATTG
BC1-422	GACTGTCACG	BC2-422	CGTGACAGTC	BC1-470	GTGATCTCAG	BC2-470	CTGAGATCAC
BC1-423	TTCAAGTTAA	BC2-423	TTAACTTGAA	BC1-471	ACGCTCTATC	BC2-471	GATAGAGCGT
BC1-424	ACGGTCGCGT	BC2-424	ACGCGACCGT	BC1-472	CTCTGTGCGT	BC2-472	ACGCACAGAG
BC1-425	TGATGACACA	BC2-425	TGTGTCATCA	BC1-473	GGACAGCAAC	BC2-473	GTTGCTGTCC
BC1-426	GATCCGAGAG	BC2-426	CTCTCGGATC	BC1-474	TATACTATCG	BC2-474	CGATAGTATA
BC1-427	TAAGTCTCGC	BC2-427	GCGAGACTTA	BC1-475	CGGTATGATG	BC2-475	CATCATACCG
BC1-428	TCGACGACTC	BC2-428	GAGTCGTCGA	BC1-476	ACAGCATCTC	BC2-476	GAGATGCTGT
BC1-429	CTTCGAGTCT	BC2-429	AGACTCGAAG	BC1-477	GACCTGCTGT	BC2-477	ACAGCAGGTC
BC1-430	GGCTGTCGAG	BC2-430	CTCGACAGCC	BC1-478	TTAAGCAGCA	BC2-478	TGCTGCTTAA
BC1-431	AGCAAGTATA	BC2-431	TATACTTGCT	BC1-479	CATACGGTAT	BC2-479	ATACCGTATG
BC1-432	GAGTCTCTGT	BC2-432	ACAGAGACTC	BC1-480	ACCTACTAGG	BC2-480	CCTAGTAGGT
BC1-433	ACTCAAGGAA	BC2-433	TTCCTTGAGT	BC1-481	GTTGGTCCAA	BC2-481	TTGGACCAAC
BC1-434	CTAGTCAACG	BC2-434	CGTTGACTAG	BC1-482	TGGCTAAGCC	BC2-482	GGCTTAGCCA
BC1-435	AGCACTTACT	BC2-435	AGTAAGTGCT	BC1-483	GTGTATCACG	BC2-483	CGTGATACAC
BC1-436	ACGCAGTATA	BC2-436	TATACTGCGT	BC1-484	TCCAGGAATG	BC2-484	CATTCCTGGA
BC1-437	GAAGGACTGG	BC2-437	CCAGTCCTTC	BC1-485	CAAGTAGTGT	BC2-485	ACACTACTTG
BC1-438	CCTTACGGAC	BC2-438	GTCCGTAAGG	BC1-486	AGTAGCTGAC	BC2-486	GTCAGCTACT
BC1-439	TTAATGACTT	BC2-439	AAGTCATTAA	BC1-487	AGACCGTCTA	BC2-487	TAGACGGTCT
BC1-440	GATCCAACGG	BC2-440	CCGTTGGATC	BC1-488	TCGTCACTGT	BC2-488	ACAGTGACGA
BC1-441	TGCGTCCGAA	BC2-441	TTCGGACGCA	BC1-489	GTTCACGCAA	BC2-489	TTGCGTGAAC
BC1-442	CTGTGTGTCC	BC2-442	GGACACACAG	BC1-490	CACGTTAGCC	BC2-490	GGCTAACGTG
BC1-443	ACTCGGTTCG	BC2-443	CGAACCGAGT	BC1-491	GATGTGTGTC	BC2-491	GACACACATC
BC1-444	TATCGGTATA	BC2-444	TATACCGATA	BC1-492	AGCCGTAAGC	BC2-492	GCTTACGGCT
BC1-445	GTGTAACGGT	BC2-445	ACCGTTACAC	BC1-493	CCGAACGTCT	BC2-493	AGACGTTCGG
BC1-446	CGAATCGCAC	BC2-446	GTGCGATTCG	BC1-494	TTAGGACCAG	BC2-494	CTGGTCCTAA
BC1-447	TTCGCTACTT	BC2-447	AAGTAGCGAA	BC1-495	CAATCGGTTA	BC2-495	TAACCGATTG
BC1-448	AACGTCAAGG	BC2-448	CCTTGACGTT	BC1-496	ATGAACTGGT	BC2-496	ACCAGTTCAT
BC1-449	GCATCACGCA	BC2-449	TGCGTGATGC	BC1-497	GCTCTACCAA	BC2-497	TTGGTAGAGC
BC1-450	CGGAATGTAC	BC2-450	GTACATTCCG	BC1-498	TGCTCTAACG	BC2-498	CGTTAGAGCA
BC1-451	ACACAGAGCT	BC2-451	AGCTCTGTGT	BC1-499	CATATCACGT	BC2-499	ACGTGATATG
BC1-452	GTAACGTCAG	BC2-452	CTGACGTTAC	BC1-500	GCTCAGATCA	BC2-500	TGATCTGAGC
BC1-453	CACTGCGTTC	BC2-453	GAACGCAGTG	BC1-501	AGAGCATGTG	BC2-501	CACATGCTCT
BC1-454	TCGGTTCAGA	BC2-454	TCTGAACCGA	BC1-502	TGCCGTGAAC	BC2-502	GTTCACGGCA
BC1-455	GGTTCAGCTT	BC2-455	AAGCTGAACC	BC1-503	CTGTATCCTT	BC2-503	AAGGATACAG
BC1-456	TGTATATAGG	BC2-456	CCTATATACA	BC1-504	GCCAGCTAGG	BC2-504	CCTAGCTGGC
BC1-457	CACGACCGAA	BC2-457	TTCGGTCGTG	BC1-505	TAATCGCGAA	BC2-505	TTCGCGATTA
BC1-458	ATGCGTATCC	BC2-458	GGATACGCAT	BC1-506	ATGGTAGTCC	BC2-506	GGACTACCAT
BC1-459	CACGGTCTCT	BC2-459	AGAGACCGTG	BC1-507	TTACCTATCA	BC2-507	TGATAGGTAA
BC1-460	AATGATGACT	BC2-460	AGTCATCATT	BC1-508	TACATGATAT	BC2-508	ATATCATGTA
BC1-461	GGAACCTGTG	BC2-461	CACAGGTTCC	BC1-509	ACGTACTGTG	BC2-509	CACAGTACGT
BC1-462	TCGTTAACGC	BC2-462	GCGTTAACGA	BC1-510	CGTGTAGCGC	BC2-510	GCGCTACACG
BC1-463	ATACGGTTAA	BC2-463	TTAACCGTAT	BC1-511	GACAGTCATT	BC2-511	AATGACTGTC
BC1-464	GTTCCAAGTG	BC2-464	CACTTGGAAC	BC1-512	GTAGGCTCGG	BC2-512	CCGAGCCTAC
BC1-465	CGCTTCGCAA	BC2-465	TTGCGAAGCG	BC1-513	AGGCCAGGCA	BC2-513	TGCCTGGCCT
BC1-466	TCGAAGCAGC	BC2-466	GCTGCTTCGA	BC1-514	CCTTAGCAAC	BC2-514	GTTGCTAAGG

#Barcode cross-talk analysis *in silico*. However, performance with these barcodes has not been validated in clinical samples.

MGI Tech Co.,Ltd

Address: Building 11,Beishan Industrial Zone, Yantian District,

Shenzhen, CHINA 518083

Email: MGI-service@mgi-tech.com, MGI_ATOPlex@mgi-tech.com Website: en.mgi-tech.com

Tel: 4000-966-988

Version：March 2021

https://www.linkedin.com/company/mgi-bgi

https://twitter.com/MGI_BGI