

Naturalis Biodiversity Center
Elza Duijm
Darwinweg 2
2333CR Leiden
Netherlands

Project code/Order number

129554

Date of Report

2-Nov-2020

Version of Report

1

Authorization

Diana Van Rossen

*Manager Next Generation
Sequencing*



BaseClear B.V.

Galileiweg 4
2333 BD Leiden
The Netherlands
T +31 71 523 3917
E info@baseclear.com

Next Generation Sequencing project

Next Generation Sequencing project report



Project 129554
Client Elza Duijm
Date 2-Nov-2020 16:13 - V1

1. Introduction

This report contains a detailed overview of your next-generation sequencing experiment. The data and results apply to the following samples:

Sample Name	Illumina demultiplexing
rc_S18_20008_64053-111	X
rc_S34_20008_64053-127	X
rc_S62_20008_64053-155	X
rc_S69_20008_64053-162	X
rc_S79_20008_64053-172	X
rc_S92_20008_64053-185	X
S01_20008_64053-0	X
S02_20008_64053-1	X
S03_20008_64053-2	X
S04_20008_64053-3	X
S05_20008_64053-4	X
S06_20008_64053-5	X
S07_20008_64053-6	X
S08_20008_64053-7	X
S09_20008_64053-8	X
S10_20008_64053-9	X
S11_20008_64053-10	X
S12_20008_64053-11	X
S13_20008_64053-12	X
S14_20008_64053-13	X
S15_20008_64053-14	X
S16_20008_64053-15	X
S17_20008_64053-16	X
S19_20008_64053-18	X
S20_20008_64053-19	X
S21_20008_64053-20	X
S22_20008_64053-21	X
S23_20008_64053-22	X
S24_20008_64053-23	X
S25_20008_64053-24	X
S26_20008_64053-25	X
S27_20008_64053-26	X
S28_20008_64053-27	X
S29_20008_64053-28	X
S30_20008_64053-29	X
S31_20008_64053-30	X
S32_20008_64053-31	X
S33_20008_64053-32	X
S35_20008_64053-34	X
S36_20008_64053-35	X
S37_20008_64053-36	X
S38_20008_64053-37	X

S39_20008_64053-38	X
S40_20008_64053-39	X
S41_20008_64053-40	X
S42_20008_64053-41	X
S43_20008_64053-42	X
S44_20008_64053-43	X
S45_20008_64053-44	X
S47_20008_64053-46	X
S48_20008_64053-47	X
S49_20008_64053-48	X
S50_20008_64053-49	X
S51_20008_64053-50	X
S52_20008_64053-51	X
S53_20008_64053-52	X
S54_20008_64053-53	X
S55_20008_64053-54	X
S56_20008_64053-55	X
S57_20008_64053-56	X
S59_20008_64053-58	X
S60_20008_64053-59	X
S61_20008_64053-60	X
S63_20008_64053-62	X
S64_20008_64053-63	X
S66_20008_64053-65	X
S67_20008_64053-66	X
S68_20008_64053-67	X
S69_20008_64053-68	X
S70_20008_64053-69	X
S71_20008_64053-70	X
S73_20008_64053-72	X
S74_20008_64053-73	X
S75_20008_64053-74	X
S76_20008_64053-75	X
S77_20008_64053-76	X
S78_20008_64053-77	X
S80_20008_64053-79	X
S81_20008_64053-80	X
S82_20008_64053-81	X
S84_20008_64053-83	X
S85_20008_64053-84	X
S86_20008_64053-85	X
S87_20008_64053-86	X
S88_20008_64053-87	X
S89_20008_64053-88	X
S90_20008_64053-89	X
S91_20008_64053-90	X
S93_20008_64053-92	X
S94_20008_64053-93	X

2. Remarks and amendments

No data available for samples S46, S58, S65, S72, S83. Also, samples with reverse complement index have the following name adjustment: rc_<sample_name>.

3. Download of data and analysis files

Order placed via the order-portal

The data and analysis results can be downloaded via the [order-portal](#) after login with your credentials.

If you would like to download your data via a FTPs client (e.g. FileZilla), you will need to login to the [order-portal](#), navigate to 'Settings', check the 'Make your results also downloadable from our FTP server' box, and save your changes.

FTP connection details:

Login	order-portal username
Password	order-portal password
Host	ftps.baseclear.com
Port	990
Protocol	FTP – File Transfer Protocol
Encryption	Require implicit FTP over TLS

4. Illumina demultiplexing

Single-end or paired-end sequence reads were generated using the Illumina NovaSeq 6000 or MiSeq system. The sequences generated with the MiSeq system were performed under accreditation according to the scope of BaseClear B.V. (L457; NEN-EN-ISO/IEC 17025).

When paired-end sequencing is being performed, the "Number of reads" noted in this report is referring to read pairs. FASTQ read sequence files were generated using bcl2fastq2 version 2.18. Initial quality assessment was based on data passing the Illumina Chastity filtering. Subsequently, reads containing PhiX control signal were removed using an in-house filtering protocol. In addition, reads containing (partial) adapters were clipped (up to a minimum read length of 50 bp). The second quality assessment was based on the remaining reads using the FASTQC quality control tool version 0.11.5.

The final quality scores per sample are provided as enclosure.

5. Results

A summary of the results is provided below. Sequence and project data are recorded digitally in our secure database and stored for backup purposes only. Data is stored for a period of one month. The result files which have been generated within this project for each analysis are as follows:

Illumina demultiplexing

The files are stored in the "raw_sequences" folder and include:

- The compressed sequence reads in FASTQ format (*.filt.fastq.gz)
- A text file containing the MD5 checksum (fingerprint) of each gzipped FASTQ file ("md5sum_*.txt")

Illumina demultiplexing report

Project 129554

Client Elza Duijm

Date 2-Nov-2020 16:13 - V1



Summary of the results - Illumina demultiplexing report

* More information about the average Q-score can be found at the Illumina website (<https://www.illumina.com/science/education/sequencing-quality-scores.html>)

Statistics after demultiplexing and filtering

Sample name	Number of read-pairs	Yield in mbp	Average quality
rc_S18_20008_64053-111	320045	57	35.38
rc_S34_20008_64053-127	430937	93	35.69
rc_S62_20008_64053-155	376661	90	35.41
rc_S69_20008_64053-162	335456	80	35.47
rc_S79_20008_64053-172	324381	77	35.35
rc_S92_20008_64053-185	803048	240	35.67
S01_20008_64053-0	754492	225	35.34
S02_20008_64053-1	512362	148	35.64
S03_20008_64053-2	1178842	345	35.41
S04_20008_64053-3	230326	42	34.94
S05_20008_64053-4	163007	35	35.45
S06_20008_64053-5	202448	40	35.4
S07_20008_64053-6	1180473	328	35.72
S08_20008_64053-7	233437	50	35.18
S09_20008_64053-8	379260	66	35.21
S10_20008_64053-9	326800	55	34.66
S11_20008_64053-10	209300	36	34.73
S12_20008_64053-11	439224	69	34.09
S13_20008_64053-12	333037	54	34.8
S14_20008_64053-13	395855	83	35.62
S15_20008_64053-14	341162	78	34.85
S16_20008_64053-15	710300	188	35.75
S17_20008_64053-16	438157	78	35.18
S19_20008_64053-18	1463064	424	35.52
S20_20008_64053-19	407714	78	34.66
S21_20008_64053-20	251820	54	35.16
S22_20008_64053-21	262448	55	35.01
S23_20008_64053-22	294697	55	34.98

Sample name	Number of read-pairs	Yield in mbp	Average quality
S24_20008_64053-23	266542	61	35.8
S25_20008_64053-24	172630	32	34.2
S26_20008_64053-25	515359	114	34.79
S27_20008_64053-26	338418	83	35.86
S28_20008_64053-27	1020257	268	35.39
S29_20008_64053-28	234916	40	34.29
S30_20008_64053-29	3312237	789	35.64
S31_20008_64053-30	169262	38	35.33
S32_20008_64053-31	908878	216	35.84
S33_20008_64053-32	132936	28	35.64
S35_20008_64053-34	574312	129	35.66
S36_20008_64053-35	480204	85	34.91
S37_20008_64053-36	221407	44	35.69
S38_20008_64053-37	599376	115	34.98
S39_20008_64053-38	55778	16	35.69
S40_20008_64053-39	968220	287	35.7
S41_20008_64053-40	1581221	449	35.69
S42_20008_64053-41	841741	236	35.83
S43_20008_64053-42	835720	237	35.84
S44_20008_64053-43	531955	151	35.77
S45_20008_64053-44	784216	222	35.66
S47_20008_64053-46	16499	3	35.38
S48_20008_64053-47	1151797	321	35.54
S49_20008_64053-48	120169	35	35.9
S50_20008_64053-49	116146	31	35.5
S51_20008_64053-50	57110	13	35.44
S52_20008_64053-51	1	0	36.5
S53_20008_64053-52	1	0	36.07
S54_20008_64053-53	8402	2	35.65
S55_20008_64053-54	769211	184	35.41
S56_20008_64053-55	263415	63	35.38
S57_20008_64053-56	1233165	296	35.4
S59_20008_64053-58	21	0	35.46
S60_20008_64053-59	218197	52	35.33
S61_20008_64053-60	393055	94	35.34
S63_20008_64053-62	1815613	503	35.62
S64_20008_64053-63	686954	165	35.68
S66_20008_64053-65	58	0	35.93
S67_20008_64053-66	6569966	1579	35.34
S68_20008_64053-67	482329	115	35.33
S69_20008_64053-68	2	0	28.33
S70_20008_64053-69	341644	84	35.98
S71_20008_64053-70	981823	236	35.38

Sample name	Number of read-pairs	Yield in mbp	Average quality
S73_20008_64053-72	450906	114	35.91
S74_20008_64053-73	415053	110	35.76
S75_20008_64053-74	17	0	35.11
S76_20008_64053-75	299760	72	35.45
S77_20008_64053-76	184655	44	35.92
S78_20008_64053-77	187996	45	35.33
S80_20008_64053-79	244953	59	35.97
S81_20008_64053-80	664383	159	35.7
S82_20008_64053-81	69972	16	35.91
S84_20008_64053-83	218094	52	35.75
S85_20008_64053-84	131520	31	35.4
S86_20008_64053-85	101867	24	35.59
S87_20008_64053-86	416664	100	35.39
S88_20008_64053-87	73303	17	35.66
S89_20008_64053-88	16923	4	35.78
S90_20008_64053-89	942254	282	35.33
S91_20008_64053-90	673141	201	35.54
S93_20008_64053-92	715669	214	35.46
S94_20008_64053-93	638890	190	35.98