HITACHI
Inspire the Next

Example of GC-rich sequence analysis using the DS3000 Compact CE Sequencer

Overview

Sequences rich in guanine (G) and cytosine (C) (GC-rich sequences) are often found in functionally important positions on the genome such as transcription control regions. GC-rich sequences inhibit polymerase extension reactions, resulting in failure to obtain expected products in the polymerase chain reaction (PCR) and the cycle sequencing reaction. A possible reason for this is that the double-stranded DNA does not dissociate well because guanine and cytosine form more stable base pairs than adenine and thymine. Therefore, additives (PCR additives) that promote the dissociation of DNA strands are sometimes added to the reaction mixture. Typical PCR additives include formamide (FA) and betaine (References 1 and 2).

Here, we used GC-rich sequences to examine the concentration of FA and betaine and the timing of including PCR additives. PCR reaction was successful in the presence of PCR additives even when the GC content exceeded 70 %. In contrast, the addition of PCR additives during cycle sequencing not only decreased the signal intensity, but also introduced noises which resulted in the misreading of the base (basecall errors).

DS3000
Compact CE Sequencer

Results

Effect of additives on Polymerase Chain Reaction

PCR additives were added to amplify *CEBPA* and *RB* according to the conditions shown in Table 3, and the purity of the resulting PCR products was confirmed (Figure 1). Formamide (FA) and betaine are additives known to improve the efficiency of PCR reaction on GC-rich sequence (Reference 1 and 2). In *CEBPA*, a signal of approximately 1,000 bp was confirmed in the sample without PCR additives. This signal was determined to be a PCR artifact because it was clearly different from the expected PCR product (587 bp). In the presence of PCR additives, expected PCR products were obtained in all of the systems. *RB* yielded the expected PCR product (466 bp) without the addition of any PCR additive. The amount of the product was higher by adding the FA(3%) and betaine, but the expected PCR product was not obtained when 10 % FA was added (Figure 1).

Effect of PCR additives on cycle sequencing reaction

A sample obtained by adding betaine was used as template DNA. PCR additives were added during the cycle sequencing reaction according to the conditions shown in Table 6. An example of the obtained data is shown in Figure 2. High-quality peak shapes can be seen without the addition of PCR additives. Please note that the base adjacent to the 3' end of the sequencing primer is counted as 1 in this document.

Next, we compared the effects of PCR additives in the cycle sequencing reaction. Figure 3 shows a comparison of QV20CRL (Contiguous Read Length) under each condition. QV20CRL is an index that indicates the size of the highly reliable region in the sequence obtained by electrophoresis. In Sanger sequencing, bases are estimated according to the peak colors and peak shapes obtained by electrophoresis. One index of the reliability of this process is the QV value. A QV value of 20 indicates an estimated reliability of 99 % for the base calling. QV20CRL calculates the average QV value every 20 bases, and that value becomes the maximum size of the sequence that maintains the value above 20. Figure 3 shows an example of the reduction in QV20CRL as a result of the PCR additives. Therefore, the quality of the reaction product deteriorated when the PCR additive was added to the cycle sequencing reaction in this experiment. Peak shapes were also compared for the 100-300 nt region of *CEBPA* (Figure 4). The addition of betaine reduced the signal intensity to approximately 1/10. As a result, the sample peak and noise overlapped, resulting in read errors (Figure 4, area outlined in black). A similar trend was observed in *RB* (data not shown).

Comparison of sequencing accuracy

As shown in Figure 5, sequencing accuracy decreases around 100 nt. This is because the peak shape was disturbed by the inclusion of the dye terminator (Reference 3). Dye terminators can be removed by purifying DNA after cycle sequencing with commercially available kits, such as the BigDye™ Xterminator Purification Kit (Thermo Fisher Scientific). Therefore, although there are differences depending on the gene, as shown here, there may be cases in which the reading accuracy decreases when PCR additives are added during cycle sequencing.

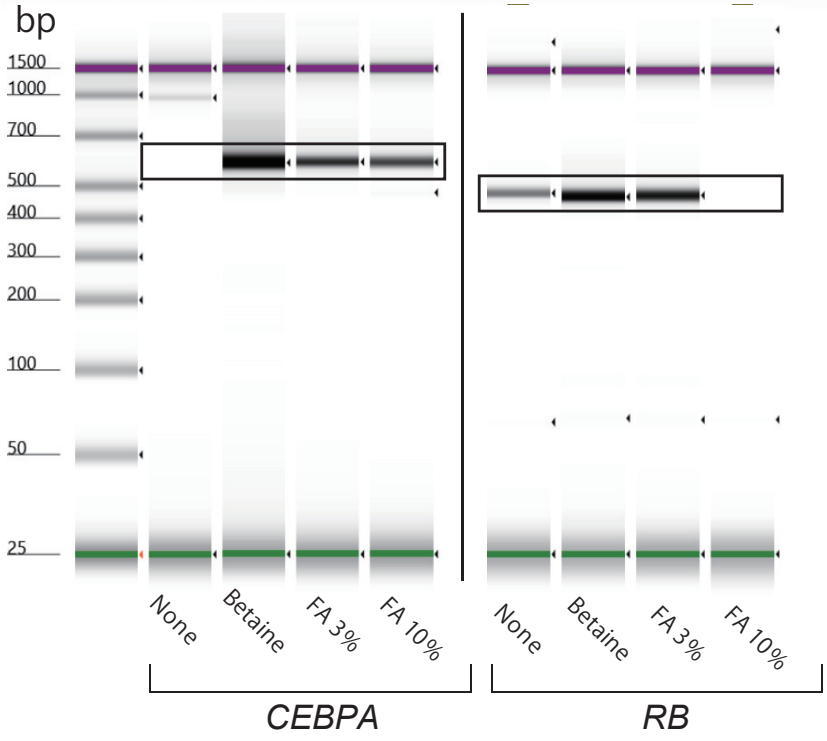

Figure 1: Confirmation of PCR products

The solutions following PCR were analyzed using an Agilent D1000 ScreenTape System (Agilent Technologies). Signals of the expected PCR products are represented in the black box. Expected PCR products could not be obtained for samples with no PCR additives in *CEBPA* (indicated by "None") and for samples with 10 % FA in *RB*. The base length of the PCR products was confirmed using the DNA ladder at the far left.

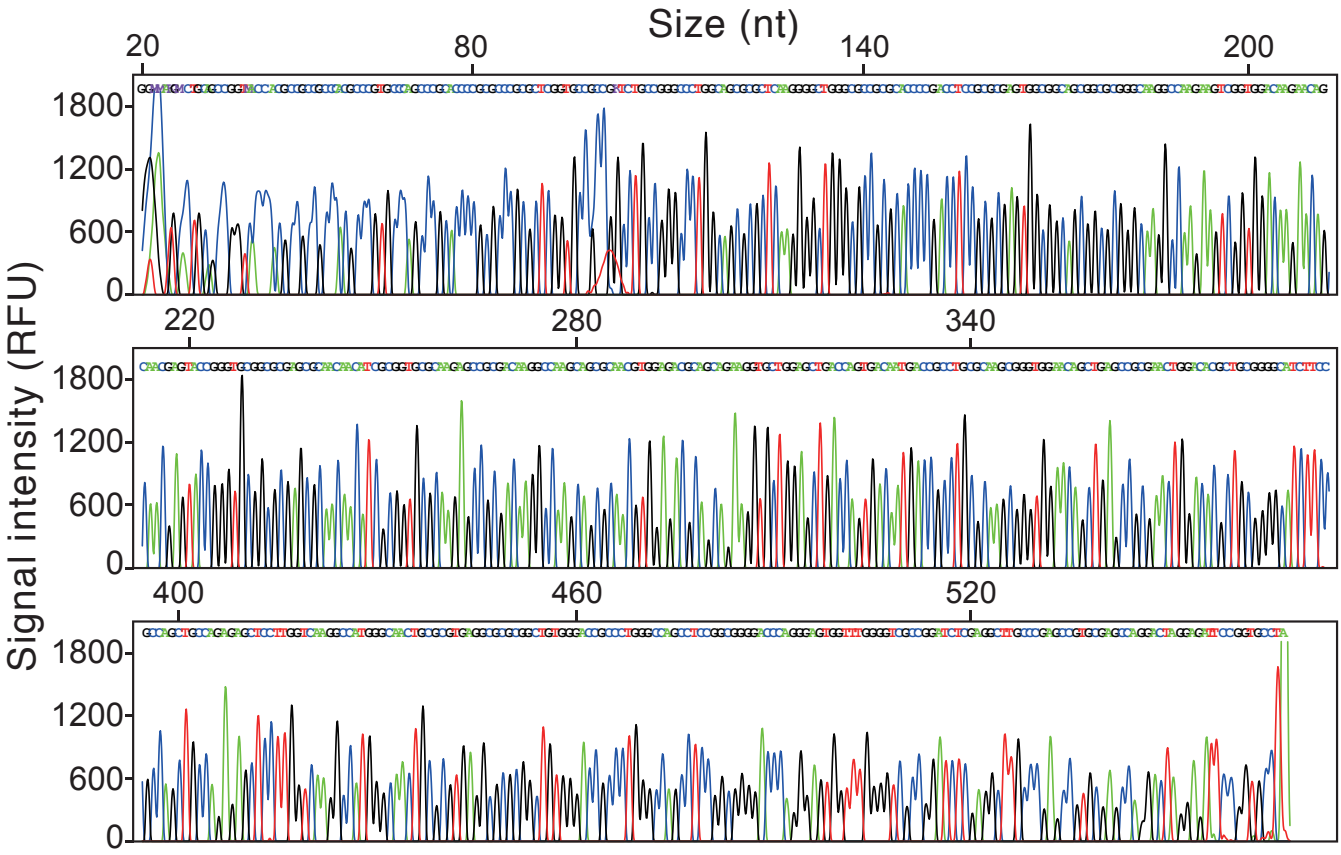

Figure 2: Peak shapes of GC-rich sequence

The figure shows a sample electropherogram using *CEBPA* as template DNA. The figure was drawn using sangerseqR (Reference 4) and information was added using Illustrator (Adobe Systems). For the size on the horizontal axis, the base adjacent to the 3' end of the sequencing primer is counted as 1.

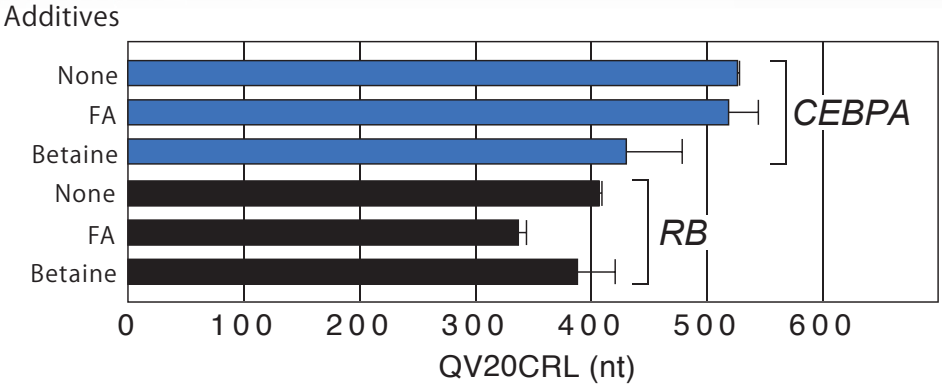

Figure 3: Effect of PCR additives in cycle sequencing reaction

QV20CRL is an index of the quality of sequencing data. See main text for more detail. In *RB*, QV20CRL is shorter than *CEBPA* under all conditions. This is because the PCR product of *RB* (466 bp) is shorter than that of *CEBPA* (587 bp). The bar graph shows the average value of eight replicates, and the error bar indicates the standard deviation.

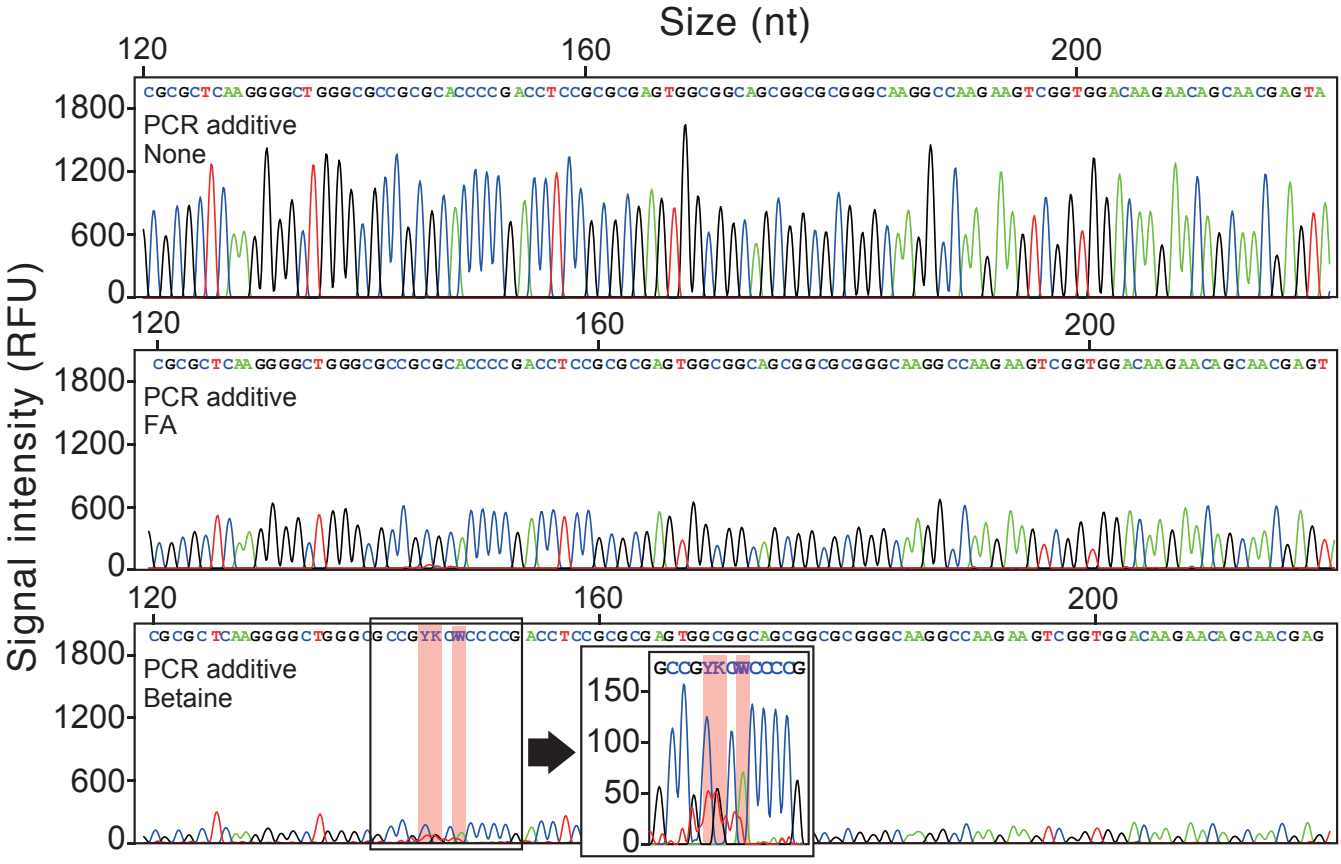

Figure 4: Comparison of peak shapes

The peak shape obtained using *CEBPA* as a template is shown. The PCR additives used are shown in the figure.
The red shaded areas indicate where read errors occurred. The arrow points to a magnified view of the region outlined in black.

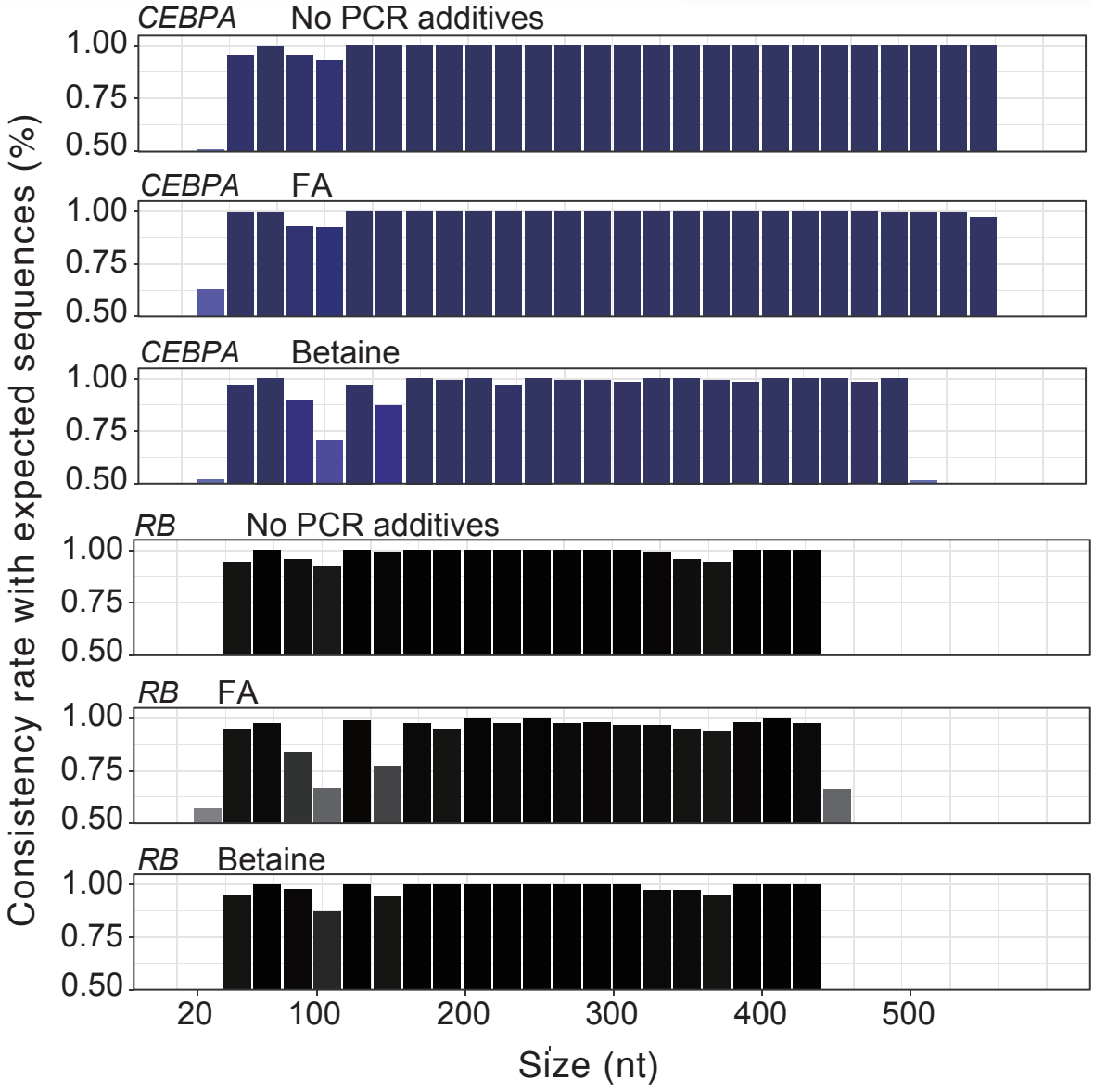

Figure 5: Sequencing accuracy for each position

Sequencing was performed eight times under each condition, and the ratio of the obtained bases matching the expected bases is displayed together for every 20 bases. Size is counted as 1 for bases adjacent to the 3' end of the sequencing primer.
One replicate with no significant signal in the sample to which betaine was added was excluded from the analysis.

Conclusions

The PCR additives were confirmed to improve the efficiency of PCR with both *CEBPA* and *RB*. In contrast, the cycle sequencing reaction exhibited degradation in data quality. PCR additives such as betaine and FA can inhibit polymerase extension reactions depending on the base sequence and concentration. In this experiment, PCR additives did not contribute to improving the quality of the reaction products in the cycle sequencing reaction.

If the GC-rich sequence does not provide the expected quality of sequencing data, consider using PCR additives. When using PCR additives, please refer to this document and adjust the type and concentration of the PCR additives and the timing of their addition.

Materials and Methods

• Sample preparation

Human CCAAT enhancer binding protein alpha (*CEBPA*) and human retinoblastoma susceptibility protein (*RB*) were selected as examples of GC-rich sequences, which contain 71 % and 72 % of G/C, respectively. PCR was performed using 2800M Control DNA (Promega®) as the template according to the information provided in Tables 1 to 4. The obtained PCR products were purified using DNA Clean & Concentrator -5 (Zymo Research). To confirm purity, 1/20 of the amount of the each PCR product was analyzed using an Agilent D1000 ScreenTape System (Agilent Technologies).

• Cycle sequencing reaction

The cycle sequencing reaction was performed according to the information provided in Tables 5 to 7. After the reaction, excess dye terminator and salt were removed by ethanol precipitation.

• Electrophoresis

Electrophoresis was performed using the conditions shown in Table 8.

Table 1. Target genes and primer sequences

Target gene	Primer	Primer sequence	Base length of PCR product (bp)	GC content	Annealing temperature (℃)
CEBPA	CEBPA-F	ACCTGCAGTTCCAGATCGC	587	71%	61
	CEBPA-R	AGGCACCGGAATCTCCTAGT			
RB	RB-F	CGGTTTTTCTCAGGGGACGT	466	72%	61
	RB-R	CTCGCCCAAGAACCCAGAAT			

Table 2. PCR mix

Reagent	Volume (μL)
10 μM Primer F	0.5
10 μM Primer R	0.5
2800 M DNA (10 ng/μL)	0.5
Phusion™ Taq High-Fidelity Polymerase (2 U/μL)	0.1
5X Phusion™ HF Buffer	2.0
2.5 mM dNTPs	0.8
PCR additive	see Table 3
PCR-grade water	see Table 3

Table 3. PCR additives

Condition	Reagent	Volume (μL)
No PCR additives	PCR-grade water	5.6
3% FA	PCR-grade water	5.3
	100% FA	0.3
10% FA	PCR-grade water	4.6
	100% FA	1.0
1.5 M Betaine	PCR-grade water	2.6
	5 M Betaine	3.0

Table 4. PCR conditions

Temperature (°C)	Time (s)	No. of cycles
98	30	×1
98	20	×30
Table 1	30	
72	30	
72	300	×1
4	hold	

Table 5. Cycle sequencing mix

Reagent	Volume (μL)
10 μM Primer F	0.4
PCR product (20 ng/μL)	1.0
BigDye™ Terminator 3.1(x2.5) Ready Reaction Mix	4.0
BigDye™ Terminator v1.1 & v3.1 5X Sequencing Buffer	2.0
PCR additive	see Table 6
PCR-grade water	see Table 6

Table 6. PCR additives used for cycle sequencing

Condition	Reagent	Volume (μL)
No PCR additives	PCR-grade water	12.6
3% FA	PCR-grade water	12.0
	100% FA	0.6
1.5 M Betaine	PCR-grade water	6.6
	5 M Betaine	6.0

Table 7. Cycle sequencing reaction

Temperature (°C)	Time (s)	No. of cycles
96	60	×1
96	10	×25
50	5	
60	240	
4	hold	

Table 8. Electrophoresis conditions

Run module	Run time	Run voltage	Sample injection time	Sample injection voltage
AB_Seq_36_Std	Approx. 60 min	7.5 kV	4 s	1.2 kV

References
(1) Grassi M., Volpe E., Colizzi V., & Miriani F. An improved, real-time PCR assay for the detection of GC-rich and low abundance templates of Mycobacterium tuberculosis testing. J. Microbiol. Methods. 64, 406-410 (2006)
(2) Henke W., Herdel K., Jung K., Schnorr D., & Loening S.A. Betaine improves the PCR amplification of GC-rich DNA sequences. Nucleic Acids Res. 25, 3957-3958 (1997)
(3) "Noise occurs around HTB-DS-TS-J001 100 bases"
https://biz.hitachi-hightech.com/sinavi/asi_articledetail?kijiId=SU210863&viewLanguage=EN
Please contact your agent for details. Browsing requires registration (free). P lease register at the following URL:
https://biz.hitachi-hightech.com/sinavi/ASI_TopPage?=EN
(4) Hill, J.T., et al. Poly peak parser: Method and software for identification of unknown indels using sanger sequencing of polymerase chain reaction products. Developmental Dynamics. 243, 1632-1636 (2014)

Main specifications for the DS3000

■ Main unit specifications

Item	Details
Number of capillaries	4
Capillary length	36 cm
Sample format	8-tube strip × 4
Device control	Touch panel PC
Number of Dyes	6
Application	Sequencing analysis / Fragment analysis
Size	400 (W) × 600 (D) × 600 (H) mm
Weight	45 kg
Performance guarantee temperature	15 – 30˚C
Performance guarantee humidity	20 – 80% RH (no condensation)
Power input	100 – 240 ±10% VAC, 50/60 Hz
Rated power	260 VA
Supported secondary analysis software	• Mutation Surveyor (SoftGenetics, LLC, sold separately) • GeneMarker (SoftGenetics, LLC, sold separately) • GeneMarker HID (SoftGenetics, LLC, sold separately)

■ Run module specifications

Run Module	Application	Polymer type	Contiguous Read Length*1 (bp, QV20 CRL)	Average run time (minutes)
Fast_Sequence36_Polymer7	Sequencing analysis	Polymer7	≥600	≤32
Standard_Sequence36_Polymer7	Sequencing analysis	Polymer7	≥700	≤60
BDx_Fast_Sequence36_Polymer7	BDx sequencing analysis	Polymer7	≥600	≤32
BDx_Standard_Sequence36_Polymer7	BDx sequencing analysis	Polymer7	≥700	≤60

Run Module	Application	Polymer type	Average run time (minutes)	Sizing precision*2 (bp, 50 – 400 bp)
Fragment_Analysis36_Polymer7	Fragment analysis	Polymer7	≤35	NA
Fragment_Analysis36_Polymer4	Fragment analysis	Polymer4	≤44	<0.16

*1 Contiguous Read Length (bp, QV20 CRL) is measured with BigDye™ Terminator v3.1. Sequencing Standard Kit (Thermo Fisher Scientific, sold separately)

*2 Sizing precision (bp, 50 - 400 bp) is measured with PowerPlex™ ESI17 Fast Allelic Ladder and WEN ILS 500 ESS. (Promega®, sold separately)

· BigDye™ is a registered trademark of Thermo Fisher Scientific Inc.

· Promega®, PowerPlex™ is a registered trademark of Promega Corporation.

■ Consumables specifications

Product name	Part number	Details	Maximum number of uses	Remarks
Capillary Cartridge 36 cm	613-0330	1 pcs	300 injections/unit	Storage temperature: 15 – 30˚C
Buffer	613-0252	Anode Buffer × 2 cartridges Cathode Buffer × 2 cartridges	80 injections/unit	Storage temperature: 2 – 10˚C
Polymer7	613-0251	4 cartridges	16 injections/unit	Storage temperature: 2 – 10˚C
	613-0291	4 cartridges	24 injections/unit	Storage temperature: 2 – 10˚C
Polymer4	613-0250	4 cartridges	16 injections/unit	
	613-0290	4 cartridges	24 injections/unit	
Septa for Cathode Buffer Cartridge	613-7231	10 pcs	Single-use	
Retainer for Cathode Buffer Cartridge	613-7233	4 pcs	–	
Septa for 8 well tubes	613-7230	24 pcs	Single-use	
Base and Retainer for 8 well tubes	613-7232	4 pcs	–	
Anode Electrode Assembly	613-7263	1 pc	–	

Specifications in this catalog are subject to change with or without notice,
as Hitachi High-Tech Corporation continues to develop the latest technologies and product for its customers.

CAUTION:For correct operation, follow the instruction manual when using the instrument.

NOTICE:The system is For Research Use Only, and is not intended for any animal or human therapeutic or diagnostic use.
Hitachi High-Tech does not guarantee the performance on the document with every possible sample under every possible condition.

Hitachi High-Tech Corporation

Tokyo, Japan
www.hitachi-hightech.com/global/science/
Toranomon Hills Business Tower, 1-17-1 Toranomon, Minato-ku,
Tokyo 105-6409, Japan
customercenter.ev@hitachi-hightech.com

HTB-DS-E014 2023.10