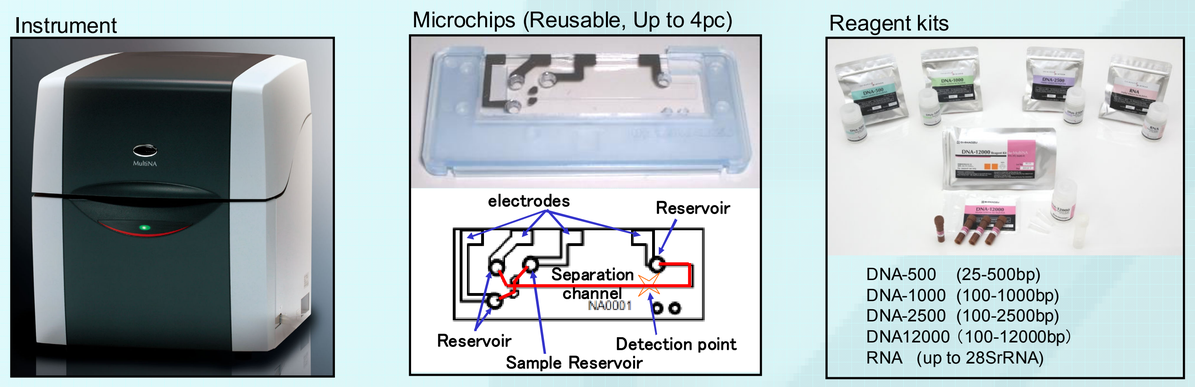
A method for quality control of NGS library using an automated electrophoresis platform, MCE-202 MultiNA.

Hidesato Kumagai[1],Akihiro Arai[1],Akira Harada[1],Taigo Nishida[1],Shin Nakamura[1],Takashi Watanabe[2], Yoko Kuroki[2],Osamu Ohara[2,3]
[1]Clinical & Biotechnology Business Unit, Life Science Business Department, SHIMADZU CORPORATION
[2]Laboratory for immunogenomics, RIKEN Research Center for Allergy and Immunology
[3]Department of Human Genome Res., Kazusa DNA Research Institute

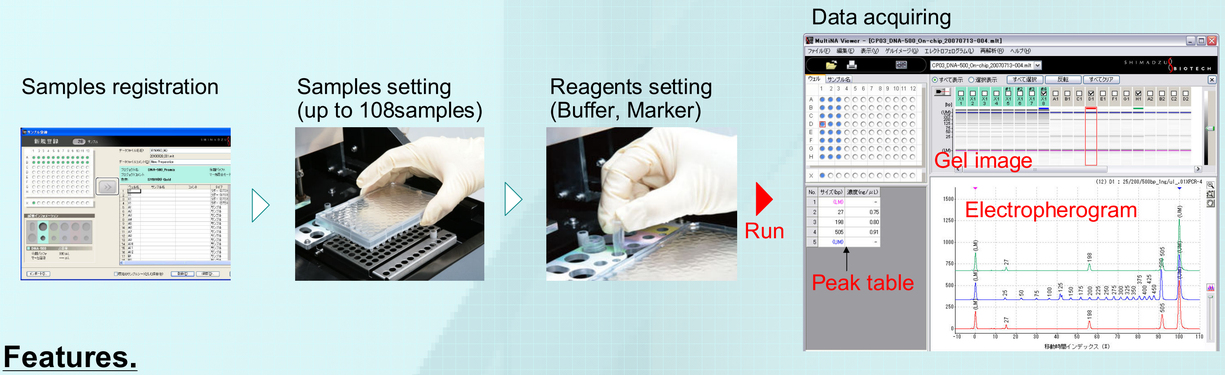
Abstract

Along with the rapid spread of next-generation sequencers (NGS) in recent years, the number of NGS libraries processed in the field continues to increase. Under these situations, the demand for quality control (QC) of NGS libraries is moving toward more rapid, fuss-free and cost-effective method. Here, we report a method of the QC using a fully automated instrument for electrophoresis, Shimadzu MCE-202 MultiNA (Fig.1), as an alternative for conventional QC methods using qPCR and/or Agilent Bioanalyzer.
Although MultiNA is originally designated for fully automated DNA fragment analysis up to 108 samples, it could not quantitatively analyze electropherograms with a smear pattern, which is inevitably required for QC of NGS library. Hence, we here developed a software for analyzing smear DNAs: By reanalyzing the data of the smear electropherogram acquired on MultiNA after defining the target size, the software estimates peak-top size, weight concentration, and molar concentration of DNAs of interest. In the preliminary experiment, we evaluated the accuracy of quantitative results of NGS libraries acquired by MultiNA by comparing with that of qPCR. The NGS libraries for RNA sequencing were prepared with a Clontech SMARTer Ultra Low RNA kit and an Illumina TruSeq DNA sample prep kit. The results indicated that the MultiNA method gave the quantification values which was proportional to those obtained by qPCR while absolute quantities estimated by the MultiNA method were about 75% of those by qPCR (CV: 11%), implying that both quantification and size-validation of NGS libraries can be assayed only by using MultiNA, even without qPCR. The accuracy of the QC data using MultiNA was further confirmed by NGS on an Illumina Hiseq1000 sequencer. Four mouse NGS libraries for RNA sequencing tagged with different index sequences were sequenced on a single lane of the flow cell. The sequencing quality was good enough and, more importantly, each mouse library was almost equally represented in the sequencing data. These results collectively showed that the MultiNA method can be used as a high throughput and cost-effective QC method for NGS library.

MCE-202 MultiNA Overview.
This microchip electrophoresis system quickly and easily performs DNA and RNA size confirmation and quantitation. This system uses microchip technology to conduct fully automated high-speed electrophoresis separation, and fluorescence detection to perform high-sensitivity analysis. The system is simply placed on a lab bench, instantly performs a run, and easily see the results.



MultiNA Analysis Flow.
All preparations are completed within 10 minutes. Operation is as follows, (i) create a analysis schedule, (ii) prepare samples, (iii) prepare MultiNA reagents, (iv) start analysis, and then (v) check the analysis results.



Features.
- High-speed automatic operation up to 108 samples
- Cost-effective analysis by reusable microchips and optimized reagent kits.

Fig.1 Introduction of MCE-202 MultiNA.

Methods and Results

Analysis Flow (Fig.2). MultiNA is applied to both quatification and smear DNAs check in the Illumina NGS flow. In this study, the analysis software was newly developed for smear patterns. The software estimates peak-top size (=Average size, bp), weight concentration (ng/μL), and molar concentration (nmol/L) of DNAs of interest.

Experiments for Evaluations of Sizing and Quantification (Table.1, 2). The MultiNA method with a DNA-1000 kit was compared to conventional methods using Agilent Bioanalyzer and qPCRs (Shimadzu and Roche). The library preparations are shown below.
- ➢Eight cDNA Libraries of mouse were prepared from total RNA with Clontech SMARTer Ultra Low kit.
- ➢After sharing of cDNA libraries according with illimina protocol, NGS libraries ware prepared with TruSeq DNA sample prep kit.

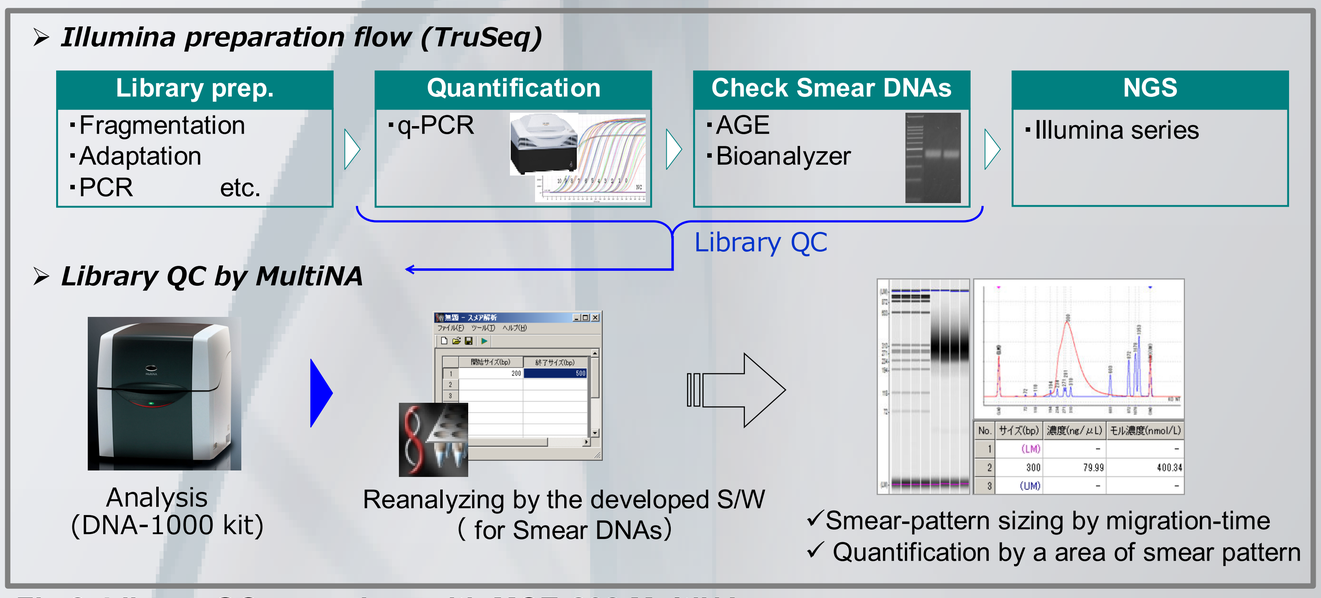


Fig.2 Library QC procedure with MCE-202 MultiNA.

Table.1 Comparison of average DNA size to a conventional Instrument.

SampleID	MultiNA [Average Size, bp]	Bioanalyzer [Average Size, bp]	Ratio [MultiNA/Bioanalyzer]
S668	291	314	92.7%
S669	290	313	92.7%
S670	288	317	90.9%
S671	290	316	91.8%
S672	285	313	91.1%
S674	292	317	92.1%
S676	287	307	93.5%
S679	293	313	93.6%
		Average	92.3% (CV 1.1%)

Table.2 Comparison of qualification to qPCR Instruments.

SampleID	MultiNA [nmol/L]	GVP-9600, Shimadzu [nmol/L]	LightCycler 480, Roche [nmol/L]	Ratio [MultiNA/GVP-9600]	Ratio [MultiNA/LightCycler]
S668	121.1	167.3	171.8	72.4%	70.5%
S669	126.9	145.1	160.6	87.5%	79.0%
S670	54.1	57.8	63.7	93.6%	84.9%
S671	214.0	292.3	279.3	73.2%	76.6%
S672	187.4	226.8	234.9	82.6%	79.8%
S674	166.3	229.3	217.7	72.5%	76.4%
S676	206.0	256.8	268.9	80.2%	76.6%
S679	158.7	181.9	286.1	87.3%	55.5%
			Average	81.2% (CV 9.9%)	74.9% (CV 11%)

Quality Control for NGS libraries only by using MultiNA (Table.3).
The MultiNA method was evaluated by Illumina Hiseq1000 for four NGS libraries tagged with different index sequences. Four cDNA Libraries of mouse were prepared with Illumina TruSeq RNA sample prep kit without size selection by gel-cutting. After library preparations, the library QCs were conducted only by MultiNA (DNA-1000 kit) without using conventional methods.

Table.3 Sequencing results of Hiseq1000 for four libraries via MultiNA method.

Cluster Density	548K / mm^2	Total Reads	151.5M reads
Total Reads trough Pass Filter	143.3M reads	≧Q30 Ratio	95.9 %
Reads Ratio sorted by Index seq.	21.2% - 26.2% (four libraries/1lane, Theoretical value:25%)		

Conclusions

In this study, the proportional accuracies of both sizing and quantification were demonstrated in comparing with conventional methods. Furthermore, the sequencing results of four libraries were almost equal with high-quality reads.

MultiNA method is applicable to use as a high throughput and cost-effective QC method for NGS library.