SPOTLIGHT feature

Proteomics, Genomics & Microarrays

# Quality Control During Library Construction for de novo Next-Gen Sequencing Platforms

*Brian Matlock and Andrew F. Page, PhD, Thermo Fisher Scientific*

**The arrival of Next Generation (Next-Gen) sequencing technologies has revolutionised the field of genomics by scaling up the sequencing process and allowing rapid acquisition of large amounts of data. The most common Next-Gen sequencing platforms currently available are (1) the Genome Analyzer systems by Illumina®, (2) the SOLiD™ systems by Applied Biosystems, and (3) the 454 Life Sciences™ systems by Roche. All three platforms require the construction of a randomly fragmented nucleic acid library with a highly specific fragment size range for de novo sequencing. There are a number of steps involved in library construction, and it is useful to monitor and track the process carefully to ensure that the majority of the fragments are the correct structure and the optimal size for the sequencing platform used. In most cases library preparation is an expensive process. Thus, it is prudent to include quality control steps throughout the library construction process to ensure the quality of downstream data.**

*Figure 1* outlines the library construction steps for the most common Next-Gen platforms. In this note we describe the workflow for each platform, and highlight suitable points (denoted by red diamonds in *Figure 1*) where the inclusion of QC steps using a NanoDrop 2000c is valuable.

The Thermo Scientific NanoDrop 2000c spectrophotometer is well suited for QC validation of Next-Gen sequencing library construction. Three features make the NanoDrop 2000c™ micro-volume spectrophotometer uniquely suitable for QC testing during library construction:

(1) Conservation of precious sample by requiring as little as 0.5µl for measurement.

(2) Highly accurate concentration measurements.

(3) Purity assessment by determining the 260/280 and 260/230 ratios and displaying full spectral data for contaminant detection.

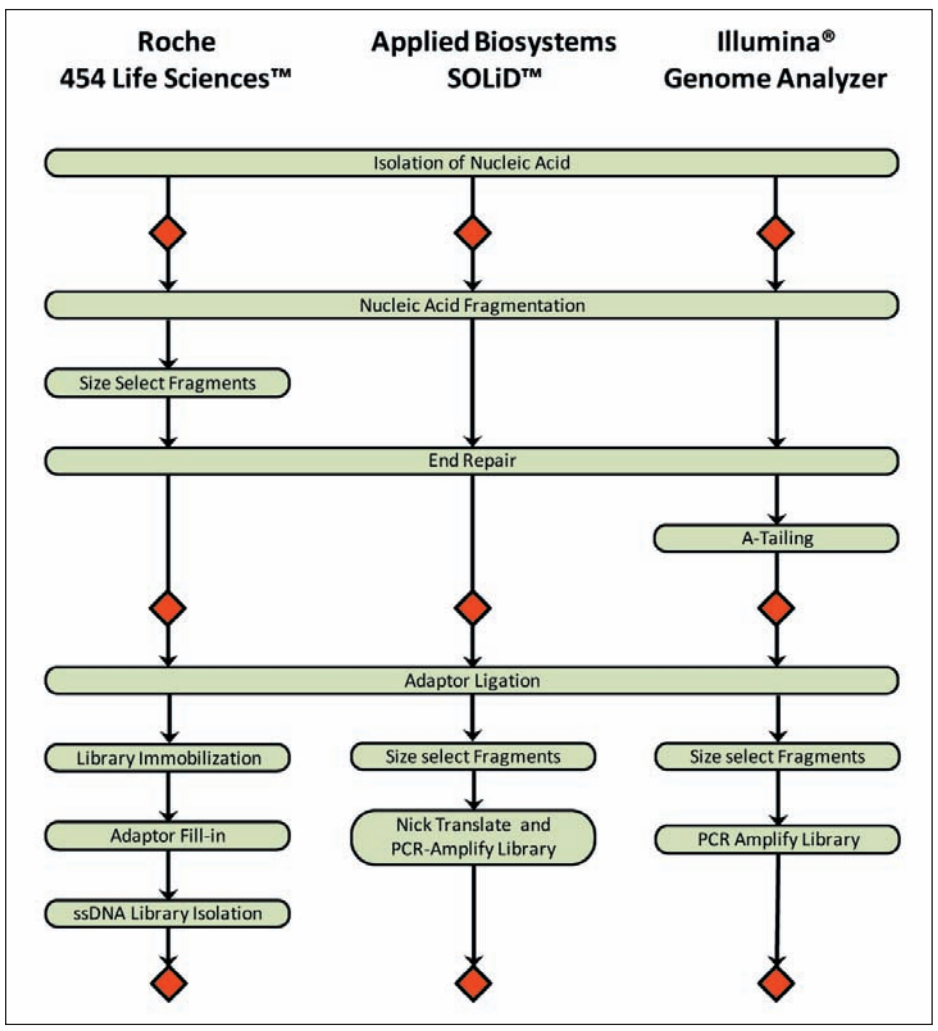


*Figure 1. Next-Gen de novo Library Construction Workflows. The red diamonds denote points where QC steps using the NanoDrop 2000c may prove useful.*

## Pre-Fragmentation of Nucleic Acid

Prior to the fragmentation step, most library construction workflows require a purity assessment of the starting material in order to ensure there are no chemical or protein contaminants in the sample. This is achieved by determining the absorbance ratios (A260/A280 and A260/A230). Criteria for acceptable ratios are often specifically listed in the library construction protocols. Two common methods used for nucleic acid fragmentation are nebulisation and adaptive focused acoustics. These specialised techniques fragment DNA randomly and need to be carried out under tightly controlled conditions. Variables such as the type of nucleic acid, the viscosity, and the concentration of sample can affect these fragmentation processes. Changes in these variables must be kept to a minimum in order to generate a reproducible distribution of genomic fragments. Thus it is important to determine an accurate concentration and the purity ratios of the sample prior to the fragmentation step. The NanoDrop 2000c is ideal for this type of QC because it requires very little sample to get full spectral data.

## Pre-Linker/Adaptor Ligation

It is valuable to know the concentrations of both the linkers and the genomic DNA inserts before the ligation step. Knowing the concentrations of both insert and linker ensures that the proper molar ratios of these components are used in the ligation reaction. A micro-volume quantification step using the NanoDrop 2000c can be used to ensure that the correct molar ratio of linker to genomic DNA is used in the ligation reaction.

## Pre-Library Sequencing

Careful quality and quantity assessment of the library is important at this point, because (i) the sequencing steps following library construction are costly, and (ii) the quality of the sequencing data relies solely on the uniformity and the quality of the DNA library. Most researchers use a combination of the following techniques to assess the quantity and quality of the DNA library: absorbance at 260 nm (A260), fluorescence-based or microfluidics-based DNA quantification assays, densitometry (gel electrophoresis using a mass ladder) and/or quantitative assays (qPCR). Each of these techniques has its own advantages for characterising a sample. An absorbance measurement at 260nm quantifies total nucleic acid content, provides spectral data that can be used to assess chemical or protein contaminants, is easy to use and does not require standards. Fluorescence-based nucleic acid quantification assays provide a sensitive and highly specific method to quantify nucleic acids. Microfluidics-based nucleic acid quantification assays and densitometry can provide information about the structure of the nucleic acid, while qPCR can determine the number of correct library molecules (i.e. insert with ligated adaptors). An absorbance measurement on the NanoDrop 2000c, in conjugation with some of the other DNA quantification methods described above, will provide valuable information about the constructed library.

## Important Considerations

The nucleic acid concentrations obtained from A260 and fluorescence-based methods may not be the same. This is primarily due to the inherent differences between these techniques. The fluorophores used in fluorescence-based assays bind very specifically to either single stranded or double stranded nucleic acid while A260 absorbance provides a measure of total nucleic acid and does not discriminate between single or double stranded nucleic acid. Each fluorophore also has different affinities for nucleic acid molecules depending on the size of the molecule. Thus, when comparing quantitation results from an absorbance measurement which is based on total nucleic acid content and the results from fluorescence based assays differences are seen because of the specificity of these assays. For instance, a sample that contains ssDNA as well as dsDNA may show a higher reading based on A260 measurement than it would with a fluorescence-based measurement.

## Conclusion

Next-Generation Sequencing technologies have made large-scale genomic studies widely accessible to researchers. All current Next-Gen sequencing technologies require the construction of specific DNA libraries involving diverse molecular manipulations. The integrity of the sequencing data can be directly affected by the efficiency and quality of the library construction process. Although these new sequencing technologies have reduced the sequence's cost per base, the cost per run is still substantial. Thus, careful QC evaluation during the library construction process is warranted. The NanoDrop 2000c is uniquely suitable for QC assessment because it requires very little sample to obtain a highly accurate, full spectrum analysis of nucleic acid samples and it has utility at various stages of the workflow.

Thermo SCIENTIFIC