

Preventive Control of Sequencing Through the Insert with the Agilent 5200 Fragment Analyzer System

Other

Kyle Luttgeharm (kyle.luttgeharm@agilent.com), Agilent Technologies

Library preparation in major next-generation sequencing (NGS) workflows is a procedure required to convert nucleic acid samples of interest into a platform-specific format. This generally includes fragmentation of input material, ligation of known sequences (adapters) to the resulting fragments (inserts), and amplification of the ligated constructs in the case of PCR-based protocols. Thus, a final library is multiple inserts of various length flanked by adapters meant to act as priming points for the sequencing chemistry reactions. When designing a sequencing experiment, it is important to consider both the library insert size and the desired sequencing read length. Ideally, the fragments composing the library should be optimized to an insert size greater than the anticipated sequencing read length. Exceeding the optimal insert length can result in sequencing through the insert, yielding in a partial or entire adapter sequence to be introduced into the read. Sequencing through the insert can have even more severe outcomes when the insert size is extremely short. In this case the sequencing reaction will continue beyond the adapter and into the surface of the flow cell. The imaging system will be detecting noise from the unincorporated bases that will cause low-quality basecalls at the end of the read. The Agilent Fragment Analyzer systems together with the Agilent NGS Fragment kits, and the companion Agilent ProSize data analysis software offer a well-established solution for sample quality control (QC) in the next-generation sequencing (NGS) workflows. Easy and comprehensive smear analysis enables the user to determine an optimal read length and thereby minimize the number of unwanted bases present in the sequencing reads, saving precious reagents and time for auxiliary data processing by estimating the percent of the library that will be sequenced through the insert based upon the planned sequencing run method.