Impact of DNA Quality on PacBio® RSII Read Length

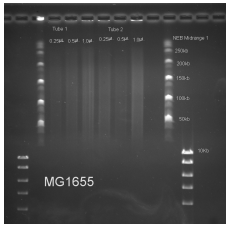
PACIFIC BIOSCIENCES®
Robert Bogden[1], Steve Kujawa[2], Amy Mraz[1], Travis Ruff[1], Karl Voss[2], Joan Wilson[2] & Jon Wittendorp[1]
[1]Amplicon Express, Inc., Pullman, WA
[2]Pacific Biosciences, Menlo Park, CA

Introduction

Maximizing next generation sequencing (NGS) read length facilitates de novo genome assembly. Currently, the Pacific Biosciences® RSII system leads the industry with respect to maximum possible NGS read lengths.

Amplicon Express specializes in preparation of high molecular weight, NGS-grade genomic DNA for a variety of applications, including next generation sequencing. This study was performed to evaluate the effects of gDNA quality on RSII read length.

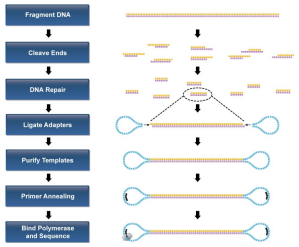
Purification of high quality gDNA



Dilutions of NGS grade gDNA preps on pulsed-field gel. Most of the gDNA is over 20 Kb in length (mass middle is around 100Kb).

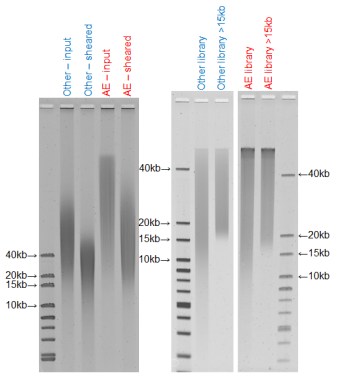
Key features of Amplicon's NGS grade gDNA prep services are: robust cell lysis (rough handling in strong detergents) followed by steps with gentle handling to prevent gDNA shearing. The gDNA is column purified with a traditional anion exchange resin and re-suspended in a proprietary solution maximizing DNA quantity and quality.

Preparation of long-insert libraries



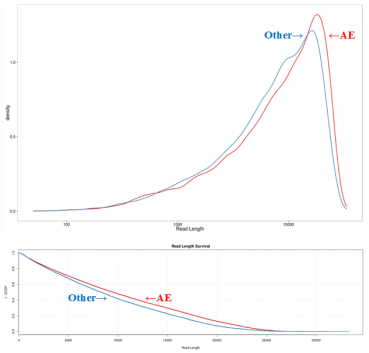
Genomic DNA was isolated from *E. coli* K12 strain MG1655 by Amplicon Express and another commercial supplier. Large insert libraries from both gDNA sources were prepared in parallel at Pacific Biosciences, according to the workflow shown here. Specifically, gDNA was sheared to >20kb using Covaris® g-TUBES®, and >15kb SMRTbell™ templates were purified by size-selection using Sage Science's BluePippin™.

Amplicon Express gDNA is less fragmented



Pulsed-field gel electrophoresis demonstrated that gDNA from E. coli K12 strain MG1655 prepared by Amplicon Express (AE) was less fragmented than that from another commercial supplier (Other). Sheared AE gDNA also exhibited larger fragment sizes, as did SMRTbell™ libraries prepared from AE gDNA. No difference in library yield, either before or after size-selection, was seen (not shown).

Amplicon Express gDNA is higher quality



RSII data demonstrated that the Amplicon Express library had longer inserts and less internal damage.

Top: Mean read length of the Amplicon Express library (AE) was 15% longer than that of a library from another commercial supplier (Other).

Bottom: Polymerase read length survival for the AE library was also significantly improved, demonstrating that this library had less DNA damage. Loading of both libraries was comparable (not shown). Chemistry was P5_C3, and movie length was 3 hours.

Maximize RSII performance with high quality gDNA from Amplicon Express

Average mapped read lengths of >12 kb (top) and N50s of >20kb (bottom) are routinely possible when libraries are prepared from high quality gDNA. SMRTbell™ libraries were prepared with Amplicon Express gDNA as described (see Figure 2). Greater than 20kb size-selection was performed using Sage Science's BluePippin™. After size-selection, a portion of the library was treated with an additional DNA damage repair reaction (Post-DDR). RSII sequencing chemistry was P6_C4, and movie length was four hours.

Conclusions

To produce optimal data from the Pacific Biosciences' RSII for de novo genome assembly, it is essential to use high quality genomic DNA, such as that prepared by Amplicon Express, during library production.

With high quality gDNA, P6_C4 chemistry and current longest library preparation methods, mean mapped read lengths exceeding 12 kb, and N50s of greater than 20 kb, are routinely possible.