Automated Size Selection using SPRIselect beads and the Biomek FX Liquid Handler

Introduction

The use of next-generation sequencing (NGS) has become the standard for research projects in genomics. There is a pressing need for automating all steps of sample manipulation for NGS from sample plating to library preparation to the final pooling before loading the sequencing instruments. Most of the steps are easily automatable; however, the size selection step which is crucial for many applications is more challenging. While most purification steps can be performed using Beckman Coulter's SPRI beads, many NGS applications require a tight distribution of fragments within a specific range. This cannot easily be achieved using Beckman Coulter's AMPure XP beads and size selection on agarose gel is not amenable to high throughput. Double size selection using SPRIselect beads can yield very tight fragment size distributions and is fully automatable in a 96-well format.

This application note describes the automation of the Size Selection using SPRIselect beads on the Biomek FX Liquid Handler. The Size Selection method is based on the protocol developed by Beckman Coulter, SPRIselect User Guide - SPRI Based Size Selection (PN B24965AA, October 2012). Using state-of-the-art gel sizing (for example Pippin Prep™ from Sage Sciences), approximately 10 hours are required to process 96 samples. The Biomek SPRIselect-based Size Selection method provides fully automated Size Selection for up to 96 samples per run in less than 2 hours. SPRIselect beads can selectively bind fragments based on the ratio of SPRIselect reagent to sample. The adjustment of ratios allows the efficient removal of fragments that are outside the desired size range. The workflow for the SPRIselect kit is easily adjustable and flexible for the removal of either small fragments, large fragments or both.

The method developed at the McGill University and Génome Québec Innovation Centre (MUGQIC) in collaboration with Beckman Coulter gives users the option of applying multiple reagent:sample ratios within a 96-well plate if libraries of multiple size ranges are desired. The individual Span-8 probes are used to transfer SPRIselect beads, ethanol and elution buffer whereas the multi-channel 96 pipetting head is used to remove the supernatant and to transfer the final eluted material.

Materials and Methods

The Biomek FX used in the experiments described below has a hybrid configuration, equipped with a 96-channel Pod (capacity of 200 µl) and a Span-8 Pod equipped with 1000 µl syringes. The complete method utilizes twelve positions on the deck (see Figure 1).

The automated size selection method offers 4 different protocols:

- Left – removal of genomic material on the left side of the size distribution (to keep higher molecular DNA weight)
- Right – removal of genomic material on the right side of the size distribution (to keep lower molecular DNA weight)
- Both – this size selection uses the Right Ratio to split the original "parent" sample input into 1 daughter plate containing smaller fragments, and 1 daughter plate containing larger fragments
- Double – removal of genomic material on the left and on the right to keep "size selected" material in the interval

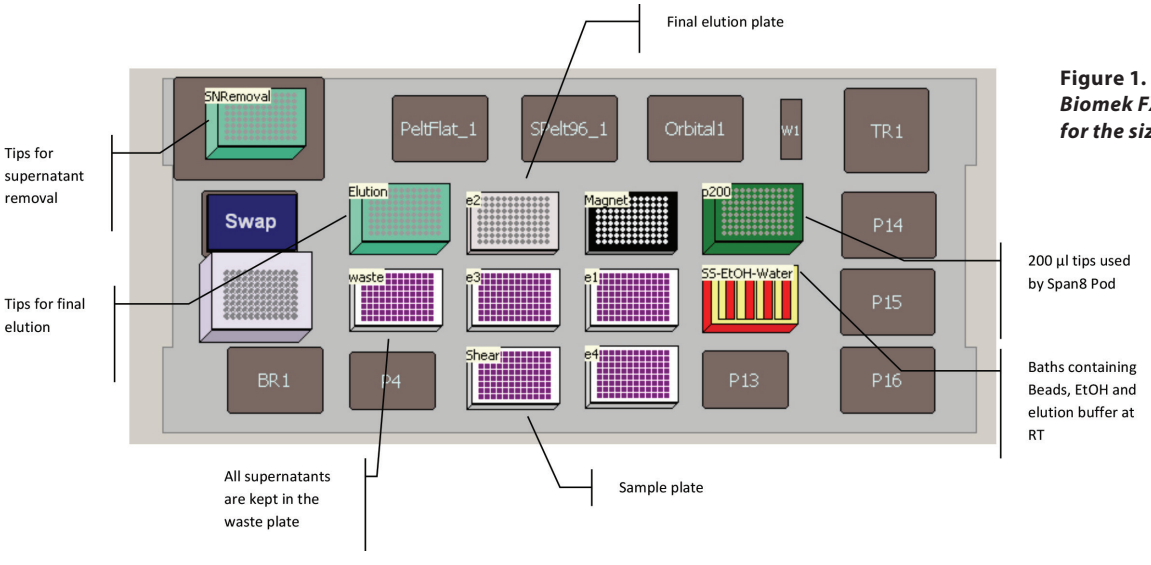

Figure 1. ***Diagram showing the Biomek FX deck configuration for the size selection method.***

McGill
www.gqinnovationcenter.com

McGill University and Génome Québec Innovation Centre
740 Docteur Penfield Ave., Montréal, QC Canada H3A 0G1
T 514 398-7211 • infoservices@genomequebec.com

GenomeQuébec
GenomeCanada

Automated Size Selection using SPRIselect beads and the Biomek FX Liquid Handler

Materials and Methods (cont'd)

To evaluate the method, we used the "Double" process. Basically, a double size selection consists of a right into a left selection. The method begins with the removal of the larger fragments using the appropriate amount of SPRIselect bead to sample ratio in relation to the desired fragment size. The supernatant is retained to which a small volume of SPRIselect reagent is added, thereby increasing the ratio of reagent to sample. After removal of the supernatant and an ethanol wash, elution buffer is added and the eluted material is transferred into a clean plate (see Figure 2).

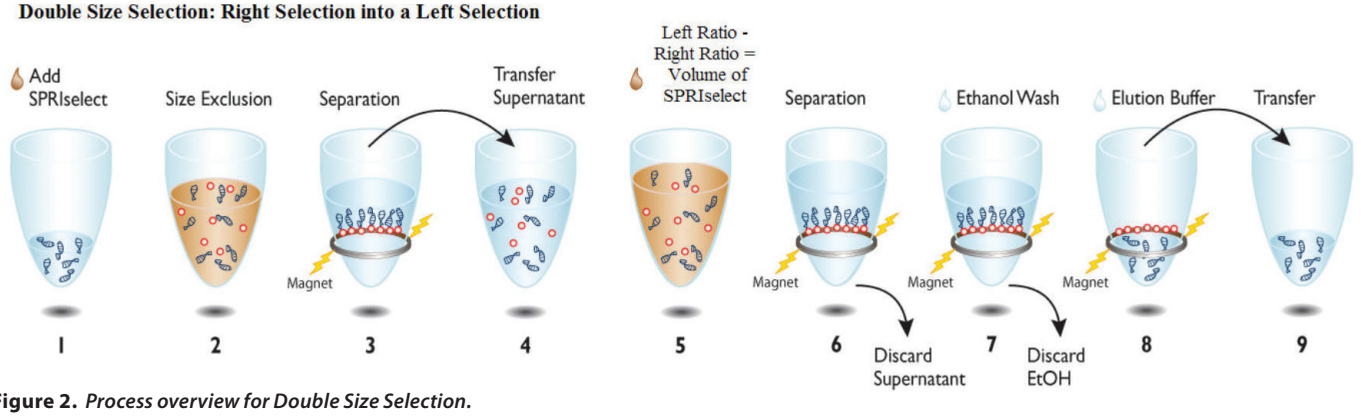

Figure 2. *Process overview for Double Size Selection.*

Using a .csv file, users can define different size selection ratios for each sample in a 96-well plate. The method's simple User Interface gives users control over selection type, sample input volume, binding and elution parameters (Figure 3).

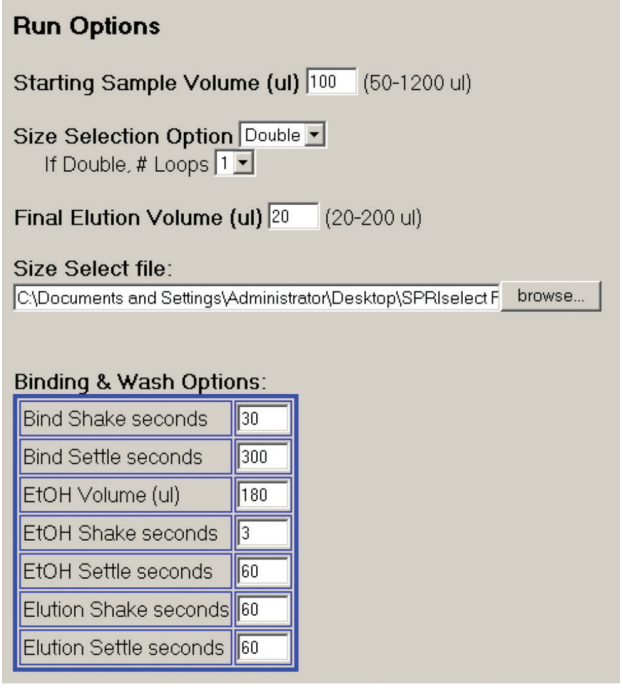

Figure 3. *User interface for the size selection method setup on the Biomek FX.*

DNA samples isolated from various species (including bacterial, yeast, plant and human) were sheared using a Covaris E210 instrument. Commercial human DNA samples purchased from Coriell were used as control in every experiment. All libraries were prepared using the KAPA HTP Library Preparation Kit for the Illumina® Platform (KAPA Biosystems, protocol ref# KR0426 – v3.13) using an automated method validated at the Innovation Centre, with the guidance of Beckman Coulter. Fragment sizes were analyzed using an Agilent BioAnalyzer 2100 instrument.

Results

The method was evaluated in a two-stage process. First, using sheared DNA samples, we tested thirty-six (36) different bead-to-sample ratios in triplicate to perform a clean double size selection. This step was used to determine the appropriate ratios between beads and sample volumes needed to isolate the desired fragment sizes (results showed in Table 1).

The second step was done using post-ligation PCR-amplified libraries to verify if the ratios determined in the first step were valid when applied on genomic DNA libraries. For every ratio tested, the same size distribution was observed between the sheared DNA and amplified libraries.

This size selection protocol yields a broad range of fragments. See Figure 4 for a comparison between a PCR amplification done on the sample library construction with and without size selection using SPRIselect beads on the Biomek FX.

Automated Size Selection using SPRIselect beads and the Biomek FX Liquid Handler

SPRIselect beads/sample Ratio	Start size (bp)	Size at Maximum (bp)	End size (bp)
0.850 - 0.875	183	283	483
0.840 - 0.865	187	291	491
0.830 - 0.855	190	296	500
0.820 - 0.845	193	300	503
0.810 - 0.835	201	319	518
0.800 - 0.825	204	331	533
0.790 - 0.815	209	341	541
0.780 - 0.805	222	351	545
0.770 - 0.795	228	356	550
0.760 - 0.785	233	362	554
0.750 - 0.775	240	371	555
0.740 - 0.765	243	376	560
0.730 - 0.755	252	381	569
0.720 - 0.745	259	391	585
0.710 - 0.735	268	398	595
0.700 - 0.725	270	405	605
0.690 - 0.715	283	424	624
0.680 - 0.705	285	439	633
0.670 - 0.695	295	448	649
0.660 - 0.685	315	462	664
0.650 - 0.675	320	480	697
0.640 - 0.665	337	496	728
0.630 - 0.655	340	509	774
0.620 - 0.645	348	519	812
0.610 - 0.635	384	547	844
0.600 - 0.625	391	564	870
0.590 - 0.615	404	589	900
0.580 - 0.605	422	608	923
0.570 - 0.595	441	636	947
0.560 - 0.585	456	651	973
0.550 - 0.575	472	676	990
0.540 - 0.565	486	692	1000

Table 1. *Reference bead-to-sample ratios to generate different library sizes.*

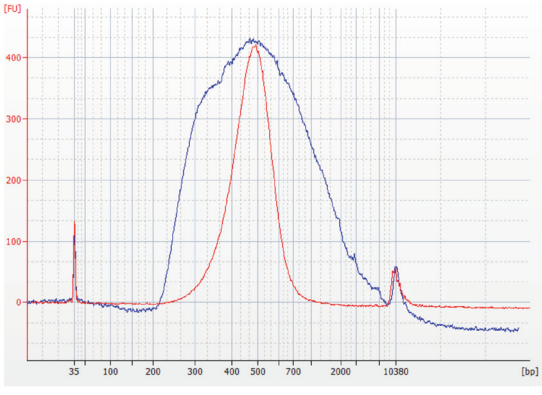

Figure 4. *Size selection targeted at 500 bp using right to left bead-to-sample ratios of 0.640 and 0.695 (size selected material is in red, original sample is in blue). Profiles were analyzed using a BioAnalyzer 2100 (Agilent).*

We then verified the reproducibility of this method. At the McGill University and Génome Québec Innovation Centre, we have generated over 1000 libraries for Illumina sequencing using the Biomek FX method in the past year. Libraries prepared with the described size selection method consistently showed similar profiles, samples of which are shown in Figure 5.

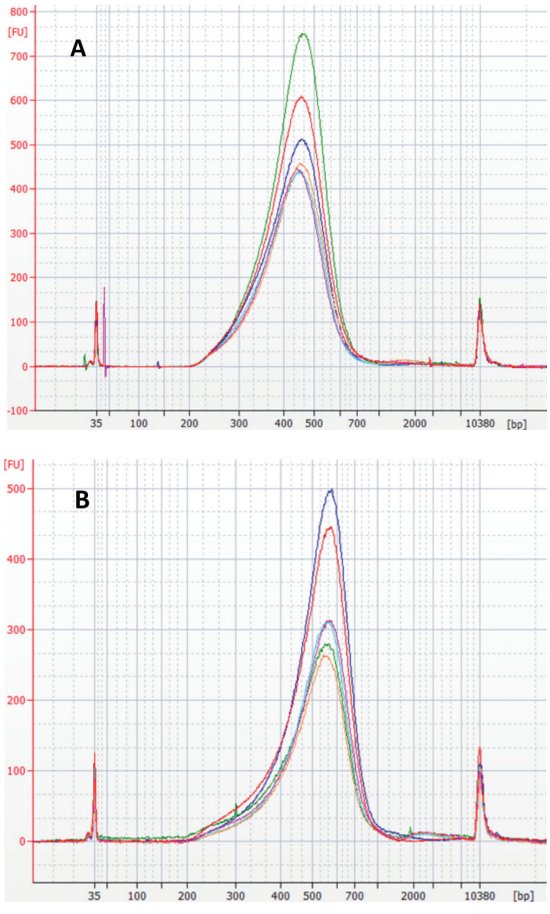

Figure 5. *Profile of 6 different library preparations of the same sample after size selection on beads at 450 bp using bead-to-sample ratios of 0.670 and 0.705 (A) and at 550 bp using 0.600 and 0.645 ratios (B).*

After comparing 50 whole-genome libraries, we obtained on average 225 ng using SPRIselect size selection and 270 ng using SAGE Pippin prep size selection. For both methods, over 95% of the libraries generated 150-500 ng of material, which is usually optimal for the sequencing on any Next Generation Sequencing (NGS) instrument. All libraries were sequenced using an Illumina HiSeq 2000 instrument for 200 cycles in paired-end mode. For every library tested, more than 200 000 reads could be obtained per lane with over 90% of reads aligning to the genome of reference and less than 5% PCR duplicates (by including reads equivalent to 10x genomic coverage) indicating high diversity within the library. Both methods of size selection gave similar results. Over 90% of the reads obtained using the SAGE Pippin method were within 100 bp of the targeted size and, using the automated SPRIselect method, on average 75% of DNA fragments were also within that range. Although the resolution of detection of large insertions and deletions will be impacted by a wider range of fragments, both methods yielded equivalent results for alignment or genome assembly.

Automated Size Selection using SPRIselect beads and the Biomek FX Liquid Handler

Discussion

A high throughput automated size selection method for genomic DNA libraries was developed using the Biomek FX liquid handler. By varying the bead-to-sample ratios, it was possible to select very specific size ranges for NGS libraries and a different size range can be selected for each well of a 96-well plate. Highly reproducible results were obtained after size selection and high-quality data could be obtained after sequencing on a HiSeq instrument. All the steps for the preparation of whole-genome libraries can be performed using the Biomek FX including library construction, clean-up (using AMPure XP), size selection, library QC (qPCR) and library pooling for loading onto a NGS flow cell, with the exception of DNA shearing.

Here, we compared two methods for NGS library size selection. Both yielded equivalent results, but the main advantages of the automated SPRIselect method over an electrophoresis-based approach is time and cost. To perform size selection on 96 libraries, it takes about two hours using the automated SPRIselect method on one Biomek FX compared to 10 hours using four SAGE Pippin instruments. In addition, the cost of consumables is approximately 5-7 times lower with the Biomek SPRIselect method.

Appropriate size selection can contribute to more efficient sequencing yields, as smaller and larger fragments do not contribute equally to the clustering or emulsion PCR step; and more accurate genome assemblies as a large fragment range can lead to incorrect alignment/assembly if the paired read interval is not well defined.

This approach can also be expanded for PCR clean-up or for any other type of DNA selection/purification steps.

Acknowledgements

We would like to thank Mrs. Mary Blair for Method Development, Dr. Louie Lamorte for Application support and Mr. Frédéric Massé for consultancy from Beckman Coulter.

Client Management Office: 514 398-7211
infoservices@genomequebec.com

Assistant Scientific Director:
Alexandre Montpetit, PhD 514 398-3311
alexandre.montpetit@mail.mcgill.ca ext. 00913

Alexandre Bélisle, Vicky Arsenault, Belisle Tir, Alfredo Staffa, Sylvie LaBoissière and Alexandre Montpetit

McGill University and Génome Québec Innovation Centre; Montréal, Québec, Canada.

5070_GQ15 (13-11)