Automated TruSeq RNA Sample Preparation from FFPE tissue specimens utilizing the Biomek FX^P Liquid Handler

Bee Na Lee Ph.D., Senior Application Scientist, and Sarah Simons, Advanced Specialist- Field Systems Support, Beckman Coulter Life Sciences; Michaela Bowden Ph.D., Research Scientist, Ashley Rossi, Molecular Technician, and Ed Stack Ph.D., Associate Director, Center for Molecular Oncologic Pathology, Medical Oncology at DFCI; Svitlana Tyekucheva, Department of Biostatistics, DFCI Biostatistics & Computational Biology;and Massimo Loda MD., Director, Center for Molecular Oncologic Pathology, Medical Oncology at DFCI and Department of Pathology at Brigham and Women's Hospital

INTRODUCTION

Next Generation Sequencing (NGS) technology facilitates high throughput, high speed and cost-effective sequencing of DNA and/or RNA. One key application of NGS is to analyze whole transcriptome of cells. Analysis of transcriptome reveals the gene expression information, detects post transcriptional mutations, quantification of gene expression profiling in a certain disease development pathway. This application note describes automated RNA library construction using the Beckman Coulter Biomek FX^P Liquid Handler from archival FFPE tissue specimens. The RNA is extracted using the Beckman Coulter Agencourt FormaPure kit and the library preparation method is based on the Illumina TruSeq RNA sample preparation protocol (P/N 15026495 Rev C). The Biomek TruSeq RNA method comprises three parts; mRNA purification, fragmentation and cDNA synthesis, cDNA library construction (end repair, A-tailing and adaptor ligation), and PCR amplification and product purification. This automated method is suitable for up to ninety-six library constructions. The data presented here benchmarks the novel automated approach using a Biomek FX^P Liquid Handler against a standard manual preparation. The archival FFPE tissue specimens were pathologically reviewed and assigned as tumor or benign regions of interest RNA isolated from archival FFPE tissue is highly degraded due to a combination of fixation conditions, temperature, humidity and age. Therefore the RNA input material for a library construction from FFPE tissue is typically of lower yield and poorer quality than fresh frozen samples. The findings of our method development study were that libraries prepped on the Biomek automated platform were significantly superior to their manually prepped counterparts and that ultimately this provides end-users with a more reliable way to generate the libraries from precious clinical samples in a faster turnaround time.

MATERIALS AND METHODS

RNA was extracted according to the Agencourt FormaPure Kits (Beckman Coulter Life Sciences, PN#33342 - Agencourt® FormaPure™ Kit-96 Prep Kit for manual extraction and PN# A35556 - Software, FormaPure 96, Biomek NX Span 8 for Biomek FX^P automation method).

cDNA library size selection and cleanup was using Agencourt AMPure XP bead solution (Beckman Coulter Life Sciences, PN# A63881).

The cDNA library construction, for both the Biomek FX^P automation platform and the manual preparation, followed the guidelines that accompanied the TruSeq RNA Sample Preparationv2 (Illumina, part number rs-122-2001). All the incubation and PCR amplification cycling steps were performed on a bench top thermal cycler. However, users have the option to install a Biometra TRobot thermal cycler to perform a fully automated library construction process.

RESULTS AND DISCUSSION

RNA preparation

Three Tumor / Normal pairs of archival FFPE prostate tissue specimens were selected and pathological review was completed to determine Gleason score and TNM stage as show in Table 1. Tumor or benign regions on each FFPE block were identified for coring. Total RNA from these cored FFPE tissues were extracted according to the protocol of the Agencourt FormaPure kits (Beckman Coulter Life Sciences, PN# A33342 and PN# A35556).

Table 1: FFPE Tumor / Benign Pairs

Donor ID	Prostate Sample	Block Age [yrs]	Tumor Grade	TNM Staging
A	Tumor Benign	6	3+4	T2c, N0, Mx
B	Tumor Benign	6	3+4	T3a, N0, Mx
C	Tumor Benign	6	3+4	T2c, Nx, Mx

Each sample was eluted in 55 µL of nuclease free water, of which 1 µL was used to assess the quality of the total RNA with the Agilent RNA 6000 Nano Kit on the Bioanalyzer 2100 system (Agilent Technologies, PN# 5067-1511). Figure 1 shows the typical degraded profile of RNA isolated from archival FFPE tissue. The RIN# was ≤ 2.4, which indicates the high level of degradation associated with these tissues.

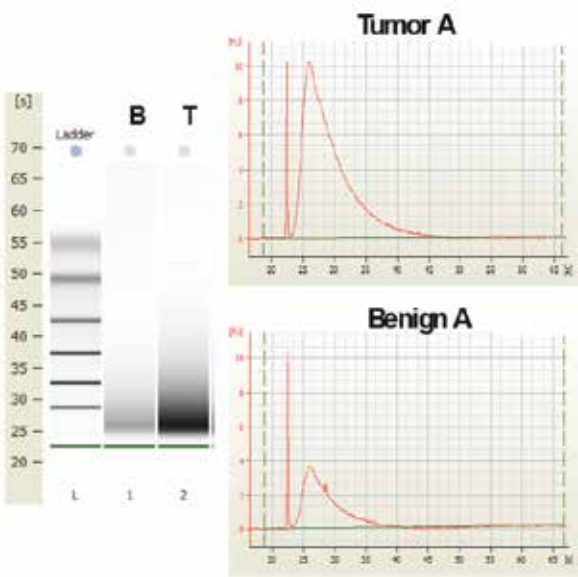

Figure 1: Representative RNA QC plots with associated elution bands on the left for Tumor / Benign pair, A

RNA concentration was determined by Quant-iTTM RiboGreen RNA Assay Kit (Invitrogen TM, R11490) and analyzed on a 96-well plate reader spectrophotometer (Perkin-Elmer Victor TM X3). Table 2 summarizes the concentration and yield ranges determined for the total RNA extracted from the FFPE blocks. The standard RNA input amount for both the manual and Biomek automated library construction procedure was 500ng of total RNA including the ribosomal RNA.

Table 2:

# Samples	RNA Concentration Range [ng/µL]	Total RNA Range (ng)	Elution Volume (µL)
8	24.4 - 61.1	1344 - 3359	55

TruSeq cDNA library preparation

Fragmented RNA was reverse-transcribed to make the first strand cDNA. The 1st strand cDNA was then used as a template to generate double stranded DNA. The dsDNA template was blunt-ended and adenlyated at the 3' ends to enforce strict ligation of the adapters. The adapter modified templates were subjected to 15 rounds of amplification to enrich and selectively amplify DNA containing adapter molecules. After the final clean up step, the purified DNA samples were eluted in 30µL of Resuspension Buffer. 1µL of the library was used to assess the quality of the library generated using Agilent High Sensitivity DNA kit on the Bioanalyzer 2100 system (Agilent Technologies, PN# 5067-4626).

cDNA library quality data clean up

Initial library quality showed an average peak size for the amplified cDNA library at approximately 260bp, however, there was a 126bp non-specific adaptor amplified peak also present in all 12 samples. These represent adaptor-dimer contamination, which can lead to junk reads when sequencing. In some cases, an additional fragment greater than 400bp was also detected, which could represent single-stranded library products that have self-annealed. In order to improve the library quality, a two-step size selection cleanup protocol was implemented. In the first cleanup step, a 0.7X ratio of AMPure XP bead solution was added to each sample. Under this condition, only larger fragment greater than 500bp DNA were bound to the beads. The bead bound DNA was discarded and the supernatant treated with a 1.1X volume of AMPure beads to rebind the DNA fragments between 150bp-400bp. Figure 2 shows the overlaid QC plots of the pre- and post-cleanup libraries for tumor / benign pair A with a side-by-side comparison of the Biomek and manual libraries. The overall coverage was greater when the library was constructed on the Biomek system when the same amount of total RNA was input. The post-cleanup plots show an absence of the 126bp peak in the Biomek libraries and only a reduction in the manual preparations, the latter could lead to a higher percentage of mis-aligned reads.

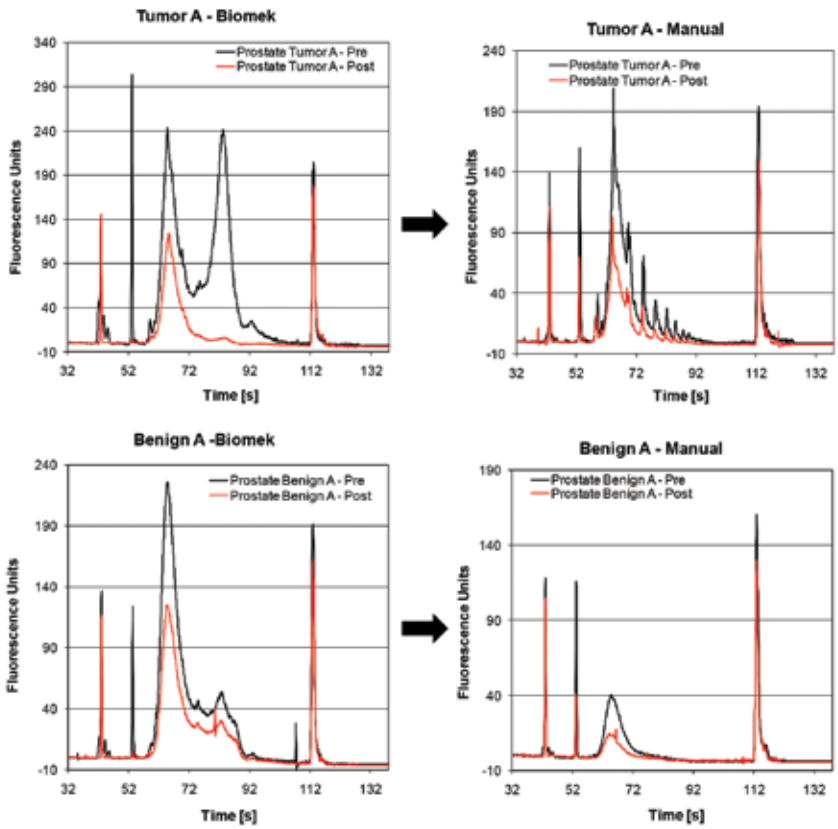

Figure 2: QC plots of Pre- and Post-Cleanup Library Preparations for Tumor / Benign pair, A

Post-cleanup libraries for the remaining samples are shown in Figure 3. It is evident that tumor / benign pair C is the most poorly defined library preparation, irrespective of whether the protocol was carried out manually or on the Biomek platform. All manual, post-cleanup library constructions have a prominent 126bp peak that is eliminated in the Biomek preparations. The overall coverage is also superior when the Biomek platform was utilized.

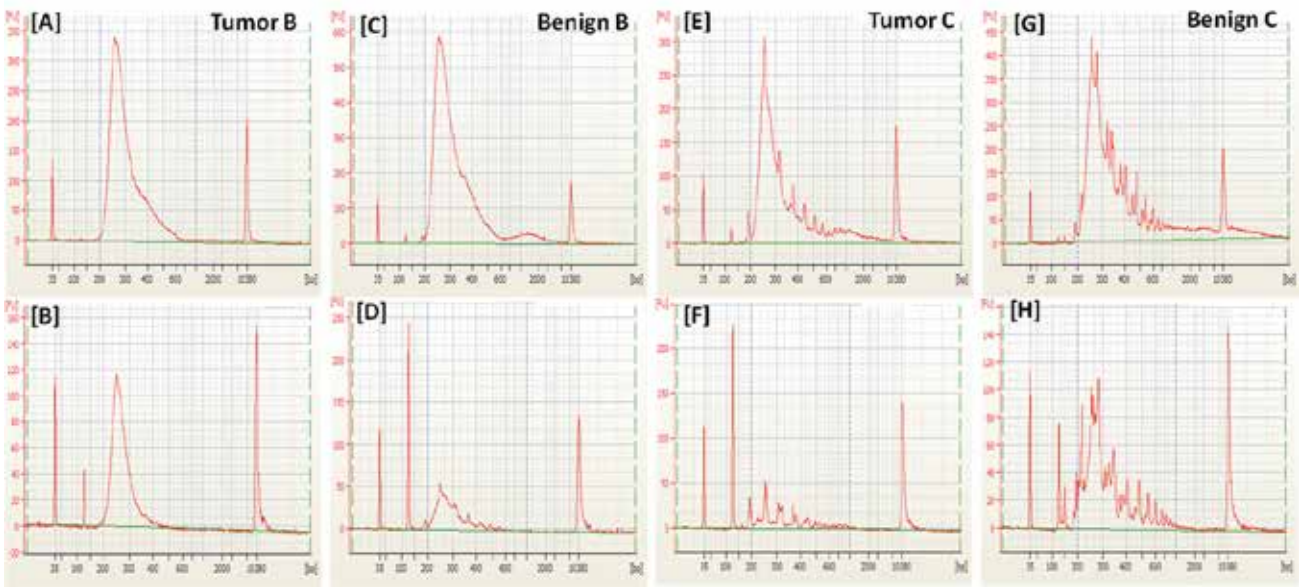

Figure 3: Post-Cleanup Electrophoretic QC plots for Tumor / Benign pairs, B and C, where top panels [A], [C], [E] and [G] represent Biomek and bottom panels [B], [D], [F] and [H], represent manual library preparations.

Single End Sequencing Performance

Sequencing was performed on the Illumina HiSeq platform. 50bp single reads were mapped using [1]Tophat and transcript abundance in FPKM units (Fragments per Kilobase of mRNA per 10^6 reads) calculated using [1]Cufflinks. Multiplexed sequencing was performed in such that each lane contained 3 samples. RNA-Seq library construction / sequencing were successful in 12 /12 samples with aligned reads ranging from 21% - 72% relative to the total # reads, as shown in Table 3.

In Table 3 some of the basic sequencing metrics of performance, including % alignment and a compositional breakdown of sequencing regions are tabulated. Experimental details are also included, which allow comparisons to be drawn, not only between the Biomek platform and manual preparation but between inter-lane variability and adaptor ligand selection.

1 Tophat (Institute of Genetic Medicine, Johns Hopkins University Center for Bioinformatics and Computational Biology (CBCB), Baltimore, Md.)

Table 3: Summary of basic sequencing performance metrics

Donor ID	A				B				C			
Tissue	Tumor	Benign	Tumor	Benign	Tumor	Benign	Tumor	Benign	Tumor	Benign	Tumor	Benign
Library Prep	Biomek		Manual		Biomek		Manual		Biomek		Manual	
Flowcell #	1	1	3	3	2	2	4	4	1	2	3	4
Adaptor ID	AR012	AR006	AR012	AR006	AR006	AR012	AR006	AR012	AR005	AR019	AR005	AR019
Sequence	CTTGTA (A)	GCCAAT (A)	CTTGTA (A)	GCCAAT (A)	GCCAAT (A)	CTTGTA (A)	GCCAAT (A)	CTTGTA (A)	ACAGTG (A)	GTGAAA (C)	ACAGTG (A)	GTGAAA (C)
# Total Reads	17687022	29046556	48645830	41657865	59938621	49836308	50991836	46441369	19964567	41801701	57399460	38203892
% Aligned	58.55	59.61	50.94	65.81	47.60	57.18	62.94	70.35	21.97	71.88	42.79	66.03
% Intergenic	15.22	15.15	9.13	13.57	14.94	14.67	11.12	12.76	15.01	13.58	9.06	10.22
% Intronic	10.11	12.81	14.45	13.74	10.25	11.20	14.31	12.60	10.45	9.61	13.86	12.66
% Exon	68.46	66.09	24.31	51.02	68.98	71.35	37.92	58.72	62.69	73.32	23.07	44.70
% rRNA	0.01	0.02	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.02	0.03	0.02

Reproducibility was measured by calculating the Pearson Correlation from transcript abundance values, presented as Log_2[FPKM] values. TruSeq libraries prepped by the Biomek automated platform and manually on the bench, exhibited good concordance, as evidenced in the scatterplots depicted in Figure 4.

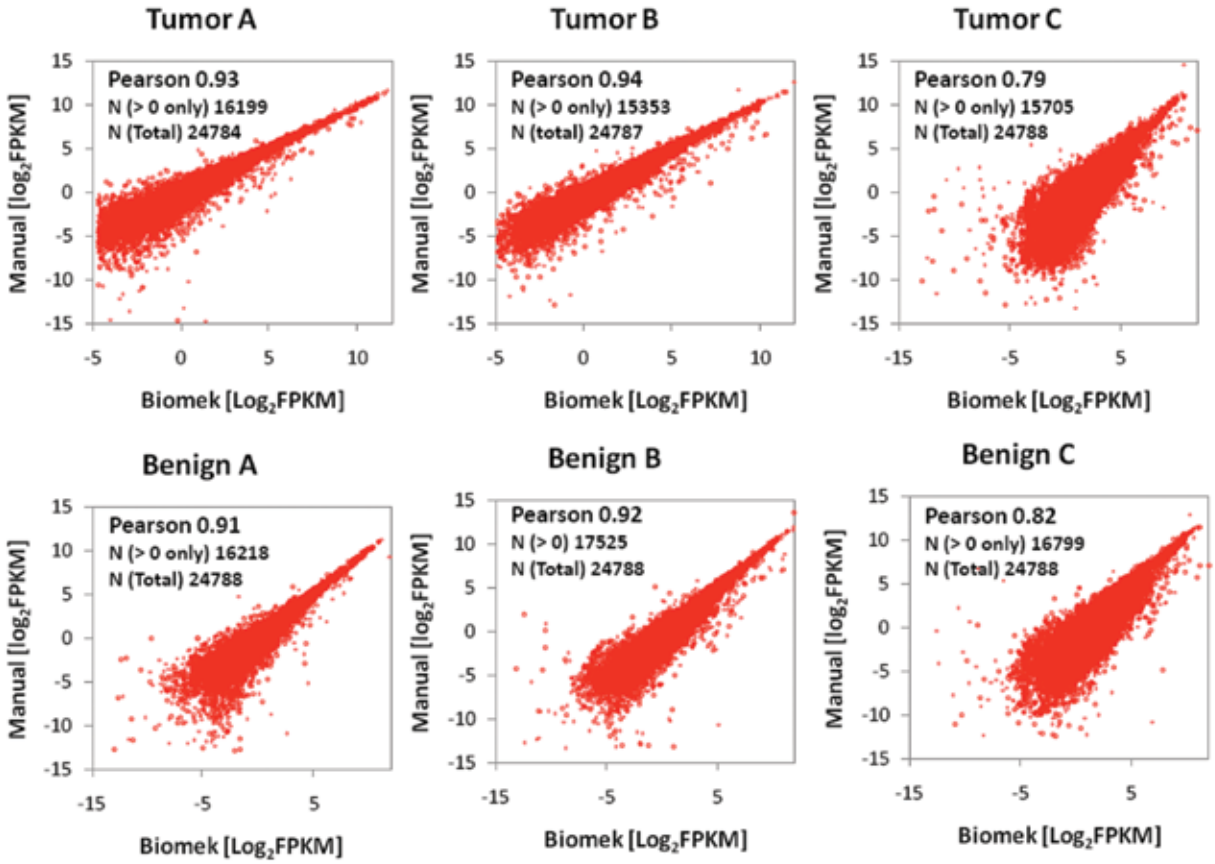

Figure 4 Scatterplots comparing method preparation for each individual sample. Log_2[FPKM] values were utilized to make the correlation.

Transcripts with zero reads are excluded from analysis. The total # of transcripts is also displayed in Figure 4. Pearson correlations across samples (A and B relative to C) are lower than across methods (Biomek v manual). In particular the poor library quality of tumor / normal pair, C, is verified in the corresponding Pearson correlation coefficients, which could suggest a mechanism of determining the level of degradation of the starting material.

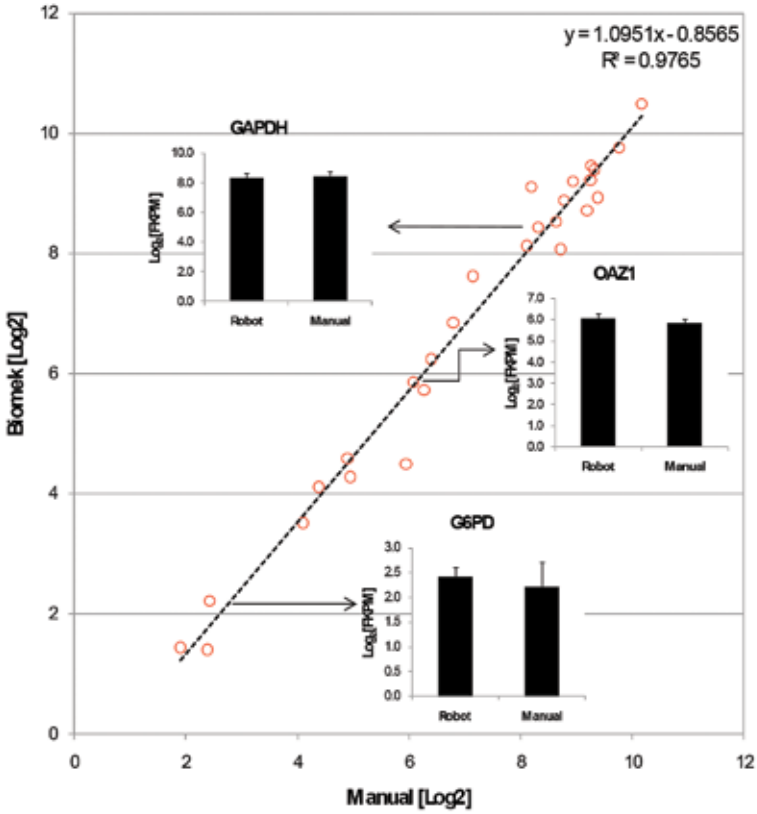

Figure 5 Scatterplots of panel of endogenous control as measured in Biomek and manual sequenced libraries (Insets: Log_2[FPKM] values plotted for low, medium and higher abundance endogenous controls, G6PD, OAZ1 and GAPDH, respectively

The panel is comprised of 15 commonly used and well-known housekeeping genes and 12 candidate genes as endogenous controls, the latter of which have been shown to exhibit superior stability from a large gene array meta-analysis study[1]. Good concordance was found between the manual and Biomek sequenced libraries, with a Pearson coefficient of 0.99, calculated, see Figure 5. The insets are plots of the average Log_2[FPKM] values for all samples for each library preparation method, which highlights the stability of these endogenous controls across all sample types, irrespective of preparatory method or tissue morphology.

CONCLUSION

This application note describes RNA-Seq library preparation carried out on the automated high throughput Biomek platform, resulting in sensitive and reproducible sequencing data, which facilitates biomarker discovery in archival FFPE tissue. Initially concordance with known prostate cancer biomarkers of disease progression will be undertaken. However it is envisaged that new coding and non-coding transcripts, as well as gene signaling networks that strongly associate with prostate cancer progression will be identified. Novel transcripts will be subject to independent validation.

The Biomek automation platform, capable of constructing 96 libraries in seven hrs, offers a viable high throughput alternative to traditional manual bench preparation of RNA libraries for sequencing. The Biomek FX^P automation platform provides a high throughput workflow to end-users with a more reliable way to generate libraries from precious clinical samples in a faster turnaround time. In addition, same method can be used for fresh tissue or other sample type of library construction.

Appendix: The labware and devices required for the Biomek method are:

Part number	Manufacturer	Description
717252	Beckman Coulter	Biomek AP96 P250 Tips, Pre-sterile
A21582	Beckman Coulter	Biomek P50 Tips, Pre-sterile with Barrier
379503	Beckman Coulter	Biomek Span-8 P250 Tips, Pre-sterile with Barrier
379502	Beckman Coulter	Biomek Span-8 P250 Tips, Pre-sterile
987936	Beckman Coulter	Biomek Span-8 P1000 Conductive Tips, Pre-sterile
372786	Beckman Coulter	Half Modular Reservoir
372790	Beckman Coulter	Quarter Modular Reservoir
A32782	Beckman Coulter	SPRIPlate 96R - Ring Super Magnet Plate
A83054	Beckman Coulter	Blue Chilled Tube Block
372795	Beckman Coulter	Reservoir Frame
A93938	Beckman Coulter	Static Peltier ALP
79448	Beckman Coulter	Orbital Shaker
210244-100	Seahorse Bioscience	Pyramid Bottom Reservoir
HSP-9601	BioRad	MJ Res HS 96-Well PCR plates
16466-042	VWR	VWR Screw-Cap Tubes (2mL Conical Bottom)
AB-1127	Thermo Fisher Scientific	ABgene 1.2mL Square Well Storage Plate
951020401	Eppendorf	Skirted 96-well PCR plates

Reference

1. De Jonge, HJM et al., Evidence Based Selection of Housekeeping Genes, PLoS One, **2007**, 2(8), e898.

Warranty Disclaimer

Beckman Coulter makes no warranties of any kind whatsoever express or implied, with respect to the method, including but not limited to warranties of fitness for a particular purpose or merchantability or that the method is non-infringing. All other warranties are expressly disclaimed. Our use of the method is solely at your own risk, without recourse to Beckman Coulter

BECKMAN COULTER

For Beckman Coulter's worldwide office locations and phone numbers, please visit **www.beckmancoulter.com/contact**

B2013-14020 IB-17929A www.beckmancoulter.com

PRINTED IN U.S.A.