Gene, Transcription Start Site and Splice Site Identification Increased Through Use of Automated Preparative Gel Electrophoresis

Tim Cloutier[2], Chris Boles[1], Cristine Kinross[2], Anupama Khanna[2], Roy Sooknanan[2], John Hitchen[2], and Simran Singh[1]

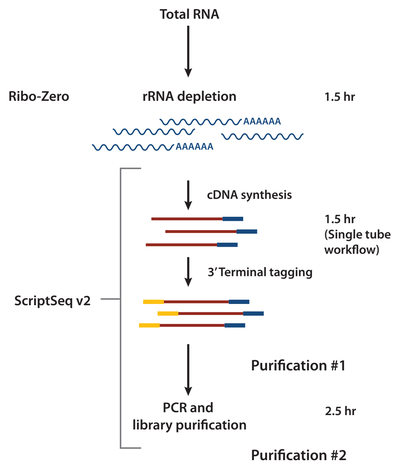
[1]Sage Science, Beverly, MA 01915, USA • [2]Epicentre® (an Illumina® company), 5602 Research Park Blvd., Suite 200, Madison, WI 53719, USA

Introduction

Gel purification of cDNA and RNA-Seq libraries is the gold standard for purification options but are time consuming and so most protocols have standardized on bead or column purification methods. Epicentre's ScriptSeq V2 kit provides high quality directional RNA-seq libraries with minimal hands-on time, and only two library purification steps. Studies were carried out to evaluate the Sage Science Pippin Prep system for the library purifications steps. ScriptSeq V2 library quality and sequencing results were compared between protocols using the Pippin Prep and protocols using standard column and bead cleanup methods. The number of genes, transcription start sites, isoforms, splice sites and promoters identified in sequencing results from libraries purified on the Pippin Prep increased 10% to 15% over bead based purification methods. The Pippin Prep is a preparative gel electrophoresis system that offers several advantages for this application including efficient removal of low molecular weight library contaminants (primer/adapter artifacts) and accurate size-selection of the library.

Methods Overview

Figure 1. ScriptSeq™ Complete Directional RNA-Seq workflow.

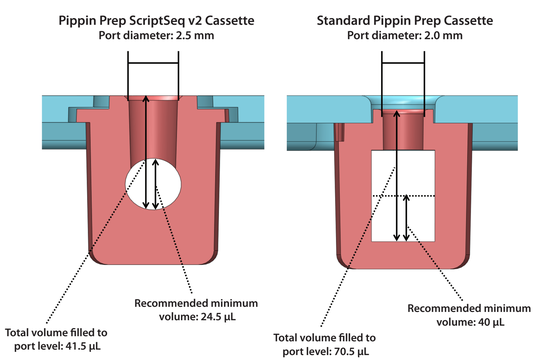


User-friendly ScriptSeq Complete workflow from intact or fragmented total RNA to cluster-ready RNA-Seq libraries in <1 day. Final library contains both polyA⁺ and non-polyadenylated RNA species.

Figure 2. The Pippin Prep™ system.

The Pippin Prep system is an automated, preparative electrophoresis system that includes a disposable five-channel precast agarose gel cassette and a computerized instrument that combines a power supply for electrophoresis with a fluorescence-based DNA detection unit. The cassette lanes are physically isolated from each other to prevent sample cross-contamination. DNA for library formation is collected in a membrane-delimited elution module. Fractionated DNA products are recovered in liquid buffer, and no gel extraction is required. Timing of DNA collection is determined by the onboard computer, which uses optical data from a DNA marker lane to determine the mobility of DNA through the cassette.

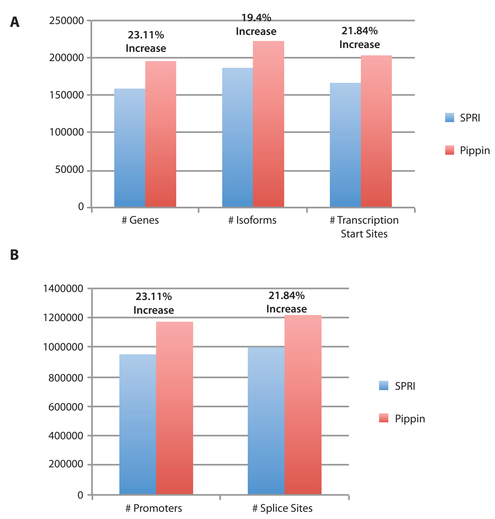
Figure 3. Comparison of elution modules for standard and ScriptSeq™ v2 Pippin Prep cassettes.



Sage has redesigned the elution modules of the ScriptSeq v2 cassette (shown at left, SSQ2010 from Sage) so that DNA fractions can be collected in as little as 25 µl of buffer. The condensed volume enables the entire collection of single stranded cDNA to be used in PCR, retaining all sample. For comparison, a standard Pippin Prep elution module (minimum recommended volume of 40 µl) is shown at right.

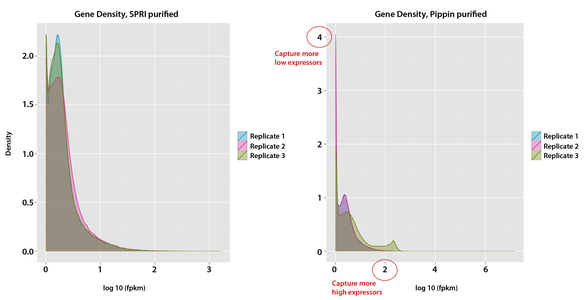
Results

Figure 4. RNA-Seq libraries 20% more informative when gel purified by Pippin Prep compared to bead purification.



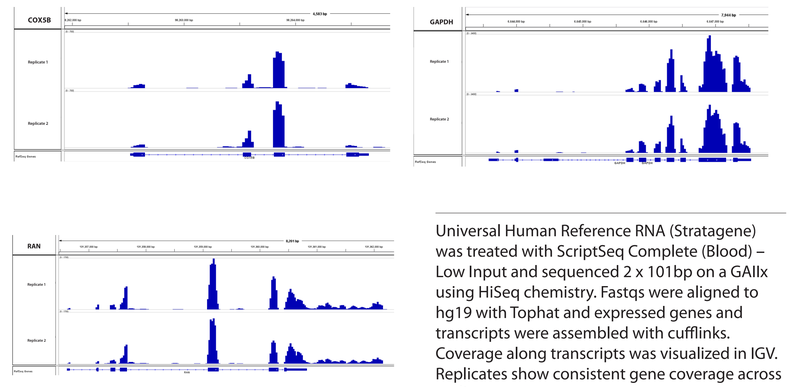
RNA-Seq libraries of Universal Human Reference RNA (Stratagene) contain more reads identified as (**A**) genes, isoforms, transcription start sites, (**B**) promoters or splice sites by Tophat and Cuffdiff when gel purified on the Pippin Prep compared to SPRI purification. Samples were sequenced on a 2 x 101bp paired-end GAIIx run using HiSeq chemistry. Reads were aligned by Tophat, analyzed for differential expression with cuffDiff and a cuffSet instance generated with cummeRbund.

Figure 5. Library diversity increased when purified by Pippin Prep.



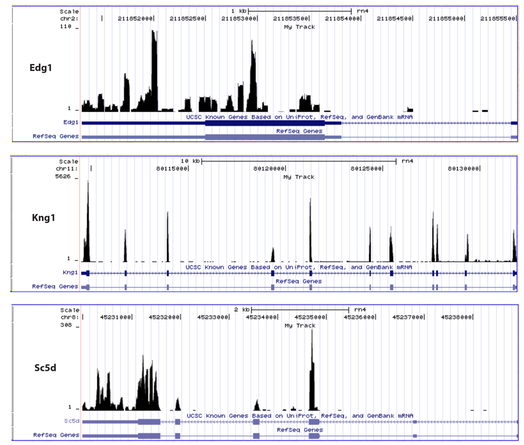
100 ng of Universal Human Reference RNA (Stratagene) was depleted of rRNA and mitochondrial rRNA and converted into a sequencing library with ScriptSeq Complete GOLD (Human/Mouse/Rat)–Low Input. Three replicates were purified by SPRI beads (Agilent) and three additional replicates were gel purified by Pippin Prep (Sage Science). The replicates purified by Pippin Prep contained greater diversity, particularly in low expressors and high expressors, to more accurately represent true sample content.

Figure 6. Coverage along transcript reproducible with Pippin purification.



Universal Human Reference RNA (Stratagene) was treated with ScriptSeq Complete (Blood) – Low Input and sequenced 2 x 101bp on a GAIIx using HiSeq chemistry. Fastqs were aligned to hg19 with Tophat and expressed genes and transcripts were assembled with cufflinks. Coverage along transcripts was visualized in IGV. Replicates show consistent gene coverage across replicates.

Figure 7. Coverage along entire transcript with ScriptSeq Complete and Pippin.



Universal Human Reference RNA (Stratagene) was treated with ScriptSeq Complete GOLD (Human/Mouse/Rat) – Low Input, purified by Pippin Prep, and sequenced 2 x 101bp on an Illumina GAIIx using HiSeq chemistry. Reads were aligned with Tophat and visualized in the UCSC Genome Browser.

Conclusions

- Retain sample information for discovery applications.
- Retain ~ 20% more genes, isoforms, transcription start sites, promoters and splice sites in ScriptSeq Complete libraries gel purified on a Pippin Prep.
- Retain lowly expressed and highly expressed genes with gel purification on a Pippin prep compared to column or bead purification.
- Consistent, reproducible coverage along the transcript.
- Simple workflow suitable for scientists of all levels of experience.

For additional information, please contact: Cris Kinross, Sr. Product Manager, Epicentre, an Illumina company

Phone: 608.442.6111 • Cristine.kinross@epicentre.com