Addition of RNA sequins to sample for RNA sequencing.

Version 1

Tim Mercer[1]

[1]Garvan Institute of Medical Research

dx.doi.org/10.17504/protocols.io.x8cfrsw

Tim Mercer

ABSTRACT

RNA sequencing can measure both gene or isoform expression, and reconstruct novel and complex spliced isoforms. However, the sheer size and complexity of the transcriptome, as well as technical bias, can confound analysis with RNA-seq. To assess the impact of these variables, we have developed a set of RNA sequins that represent synthetic genes that act as internal controls during RNA sequencing.

Each RNA sequin represents an individual isoform, with multiple isoforms forming artificial gene loci that are encoded within the *in silico* chromosome (chrIS). By modulating the relative abundance of individual sequin isoforms we can emulate alternative splicing, whilst modulating the abundance of multiple isoforms we can emulate gene expression. Accordingly, RNA sequins are mixed at different concentration to emulate differences in gene expression and alternative splicing.

By sequentially diluting sequins, we can establish a reference ladder across a range of gene expressions. We formulate multiple alternative mixtures that differ in the concentration of individual sequins. By comparing mixtures, we can emulate differential gene expression and alternative splicing between samples. By contrast, RNA sequins with invariant concentrations between mixtures provide static scaling factors that enable quantitative normalization between multiple RNAseq libraries.

The RNA sequin mixture is added to a user's RNA sample at a fractional concentration prior to library preparation. The combined sample and the sequins then together undergo sequencing. The sequins can then be distinguished in the output library by their synthetic sequence, and analyzed as internal controls.

For further detailed background on the design, validation and use of sequins, we refer users to 'Spliced synthetic genes as internal controls in RNA sequencing experiments' by Hardwick et al., (2016) NAture Methods DOI:10.1038/nmeth.3958\

EXTERNAL LINK

www.sequinstandards.com

THIS PROTOCOL ACCOMPANIES THE FOLLOWING PUBLICATION

Hardwick et. al., Spliced synthetic genes as internal controls in RNA sequencing experiments. (2016) Nature Methods.

PROTOCOL STATUS

Working

We use this protocol in our group and it is working

MATERIALS

NAME	CATALOG #	VENDOR
RNA sequins standards	View	Sequins

STEPS MATERIALS

NAME	CATALOG #	VENDOR
RNA sequins standards	View	Sequins

Re-suspension and storage of sequins

Re-suspension and storage of sequins,

1

RNA sequins standards
by Sequins
View

Upon receipt of RNA sequins, first check to ensure they have not thawed during shipment., and immediately transfer the RNA sequins to frozen storage at -80°C (sequins should not be stored in a -20°C frost-free freezer).

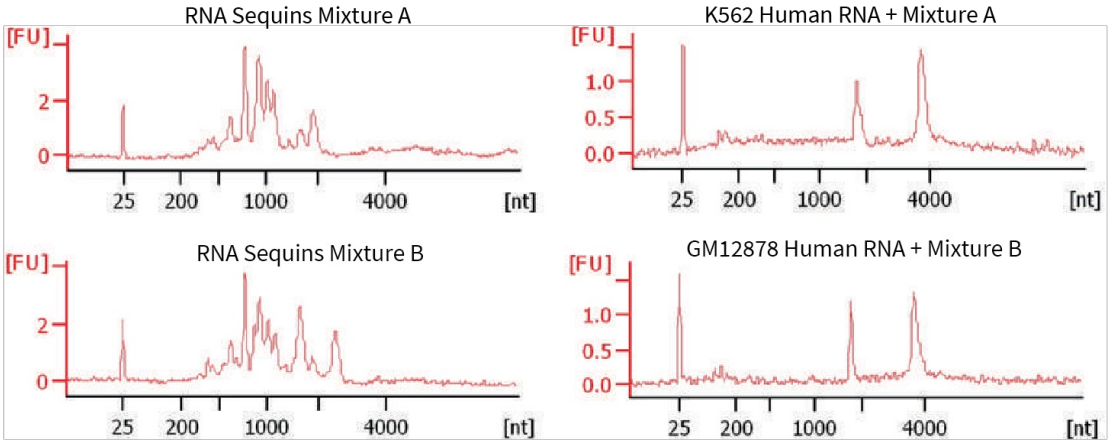

Figure 1. Example traces of RNA sequins using an 2100 BioAnalyzer with the RNA Nano Kit (Agilent Technologies) for (left upper) neat Sequin Mixture A and (left lower) neat Sequins Mixture B. Also shown are example traces for (right upper) K562 with Sequin Mixture A and (right lower) GM12878 with Sequins Mixture B.

2 Each tube contains RNA sequins provided in solution in 10 µL nuclease-free water at a concentration of 15 ng/µL. On first thaw, spin the tube down to collect the contents at the bottom of tube, and prepare smaller single-use aliquots to minimize subsequent freeze-thaw cycles. The exact amount of RNA sequins required for a single use aliquot depends on the sample input required for your preferred library preparation method.

Table 1 (below) provides guidance on the amount and dilution of sequins that should be used according to the sample RNA amount required.

Sample RNA.	Sequin Mass.	Sequin Volume (dilution from 15ng/ul stock)
20ng	0,2ng	1ul (1:75)
50ng	0.5ng	1ul (1:30)
100ng	1.0ng	1ul (1:15)'
500ng	5.0ng	1ul (1:3)
1000ng	10.0ng	1ul (2:3)

Table 1. Guidelines for diluting RNA sequins according to sample RNA amounts (recommended 1% spike-in).

Addition of sequins to sample, ,library preparation and sequencing

3 The diluted RNA sequins should then be added directly to the sample RNA prior to any subsequent processing steps (such as poly-A enrichment or rRNA depletion). Whilst this enables an assessment of these processing steps, the amount and dilution of RNA sequins added may need to be modified accordingly.

COMMENT | RNA sequins are provided in two alternative mixture formulations; Mix A and B. Each mixture contains the same sequin isoforms, however they have been formulated at molar ratios. This emulates fold-change differences in gene expression and alternative splicing between the two mixtures. If you are performing a gene-profiling RNAseq experiment to

identify differences in gene expression and splicing between two conditions, we suggest that mixture A and B are added alternatively added to separate samples from each condition being compared (ensure that you do not add both mixtures to a single sample). This enables the use of RNA sequins to assess the detection of fold-change differences between samples.

COMMENT | This protocol adds sequins at a relative fraction according to the amount of RNA in the sample. However, sequins can alternatively be added to your RNA sample at an absolute amount to enable absolute measurements of gene expression, fold-changes between samples and improved normalization. For further information on this approach, please refer to 'Measuring Absolute RNA Copy Numbers at High Temporal Resolution Reveals Transcriptome Kinetics in Development.' by Owens et. al., Cell Reports (2016) Jan 26;14(3):632-47.

4 Use the combined sample and sequins as input according to the protocol of your preferred library preparation kit

COMMENT | The downstream library preparation workflow may require user's to concentrate the sample RNA after the addition of the RNA sequins. RNA samples can be concentrated by either ethanol precipitation, SPRI® bead purification (e.g. RNAClean® XP, Beckman Coulter), column-based methods (e.g. RNA Clean & Concentrator™ Kit, Zymo Research), or using vacuum centrifugation.

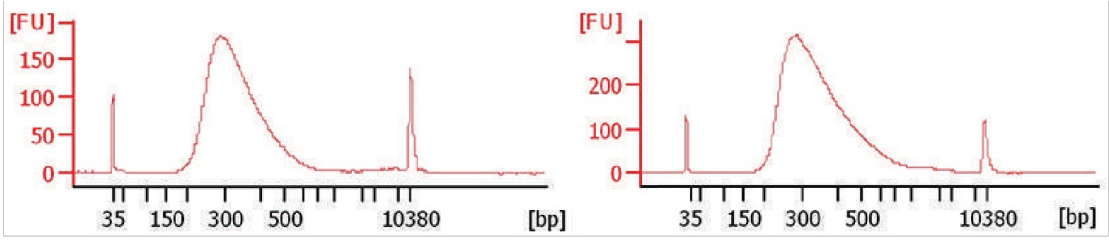

Figure 2. Successful sequin-containing (total) RNA Libraries. A, K562 with Sequins Mix A. B, GM12878 with Sequins Mix B. Samples analysed by Agilent 2100 BioAnalyzer trace (size distributions sequenced on an Illumina® HiSeq 2500 Instrument).

5 The library that is generated from the combined RNA sample and sequins is then sequenced per manufacturer's instructions. In this example, we have used the Illumina HiSeq 2500.