Universal mouse reference RNA as a standard metric for evaluating bulk RNA-seq data

Marie Dinh (marie.t.dinh@nasa.gov), NASA Ames Research Center, **Marie Dinh**, NASA Ames Research Center, **Amanda Saravia-Butler**, NASA Ames Research Center, **Valery Boyko**, GeneLab, **Yi Chun Chen**, NASA GeneLab, **Charles Houseman**, SeqMatic LLC, **Lovorka Degoricija**, NASA Ames Research Center, **Maha Ulhaq**, GeneLab, **Samrawit Gebre**, NASA Ames Research Center, **Lauren Sanders**, NASA Ames Research Center, **Sylvain Costes**, NASA Ames Research Center, **Jonathan Galazka**, NASA Ames Research Center

Bulk RNA sequencing (RNA-seq) is a powerful approach to identify gene expression changes in organisms exposed to unique environments such as spaceflight. One of the challenges of evaluating RNA-seq data across different space-relevant studies is the ability to identify and control for technical differences, including the use of different library preparation kits, sequencing parameters, person-to-person variation in sample preparation, and instrument drift. To help address these issues, the NASA GeneLab Sample Processing Laboratory utilizes various controls, including universal mouse reference RNA (UMRR), as part of our standard operating procedures for every in-house RNA-seq study. Sequencing data from UMRR control samples are evaluated using a variety of metrics such as gene body coverage, strandedness, fragment size, read distribution, and gene expression to assess sample quality and reproducibility across experiments. Here will discuss the use, benefits, and limitations of using UMRR samples as a standard measurement to generate quality control information for RNA-seq studies conducted at GeneLab.