



**University
of Victoria**

Graduate Studies

Notice of the Final Oral Examination
for the Degree of Master of Science

of

BONNIE-JEAN ROBERT

BSc (Vancouver Island University, 2013)

**“A Pipeline for Differential Expression Analysis of RNA-seq Data and
The Effect of Filter Cutoff on Performance”**

Department of Mathematics and Statistics

Tuesday, August 15th, 2017

1:00 P.M.

Human and Social Development Building
Room A250

Supervisory Committee:

Dr. Mary Lesperance, Department of Mathematics and Statistics, University of Victoria (Supervisor)

Dr. Julie Zhou, Department of Mathematics and Statistics, UVic (Member)

External Examiner:

Dr. Caren Helbing, Department of Biochemistry and Microbiology, UVic

Chair of Oral Examination:

Dr. Martin Farnham, Department of Economics, UVic

Dr. David Capson, Dean, Faculty of Graduate Studies

Abstract

RNA sequencing is a powerful new approach to analyzing differential expression of transcripts between treatments. Many statistical methods are now available to test for differential expression, each one reports results differently. This paper presents a workflow of five popular methods and discusses the results. A pipeline was built in the R language to analyze four of these packages using a real RNA-seq dataset.

At present, researchers must prepare RNA-seq data prior to analysis to achieve reliable results. Filtering is a necessary preparatory step in which transcripts exhibiting low levels of genetic expression are removed from further analysis. Yet, little research is available to guide researchers on how best to choose this threshold. This paper introduces a study designed to determine if the choice of filter threshold has a significant effect on individual package performance. Increasing the filtering threshold was shown to decrease the sensitivity and increase the specificity of the four statistical methods studied.