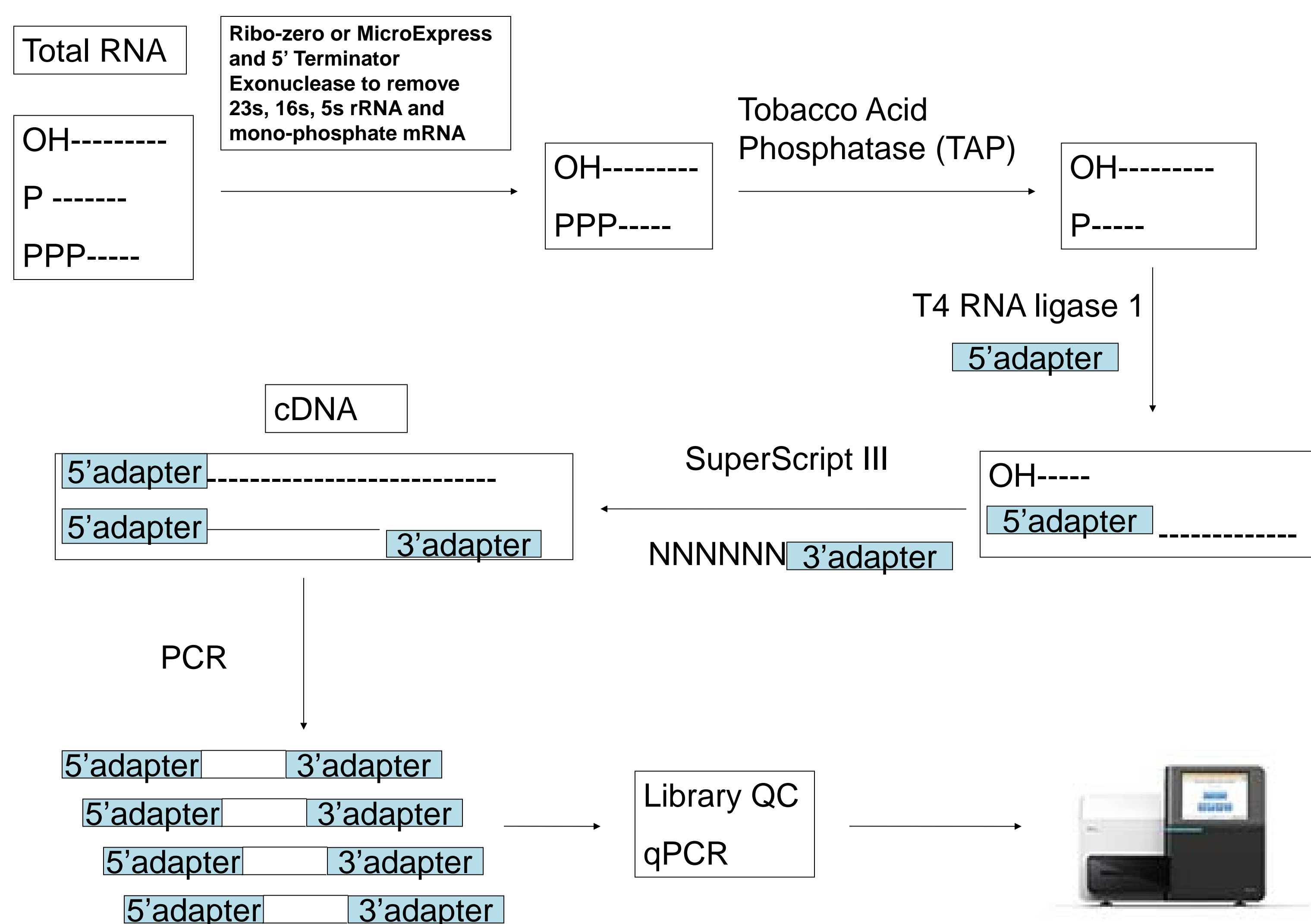


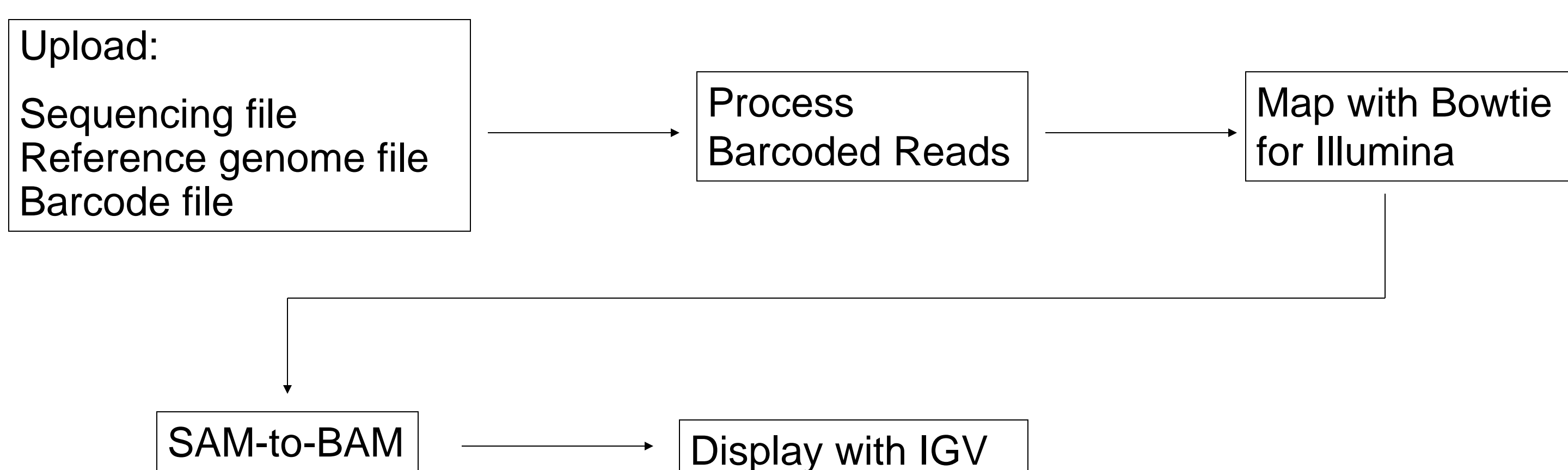
## Abstract

Next generation sequencing technology has dramatically changed the way gene expression profiles are carried out. Studying the 5'-transcript ends of prokaryotes genes has been difficult until recently due to technical difficulties in enriching for mRNAs that lack 3' poly(A) tails and 5' cap. Using E.coli as a model, we are trying to develop a simple process to sequence and analyze prokaryotic 5'-transcription start site (5'TSS) libraries. We have been testing construction parameters such as 1) the volume of RNA SPRI used for enzyme reaction clearance, 2) the concentrations of hexamer-3'adapter for reverse transcription (RT) as well as 3) different ways alone or combined to remove all kinds of rRNA. Our approach to library construction and the subsequent prediction of transcription start sites should contribute to genome annotation and cell biology research.

## 1. 5'-TSS Library Construction Procedure



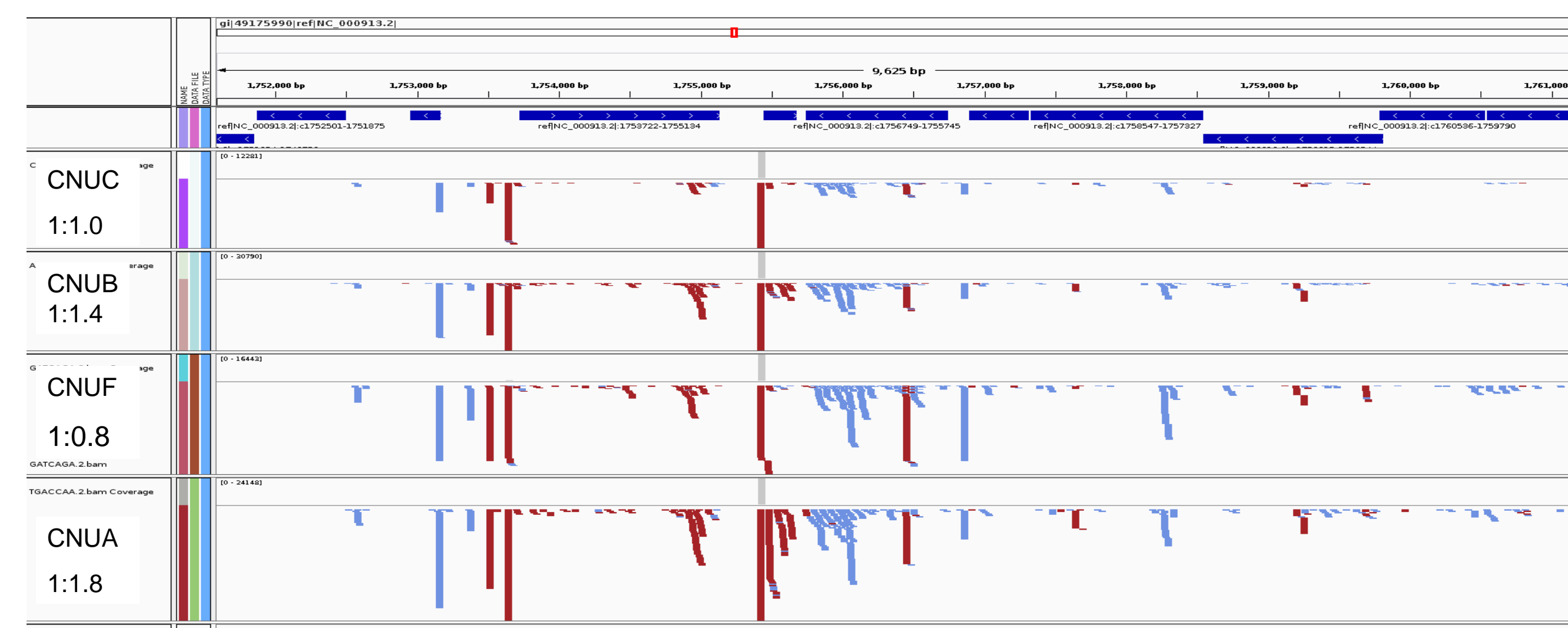
## 2. Procedure of Data Analysis using Galaxy Programs



Galaxy tools is available at <https://galaxy.jgi-psf.org>

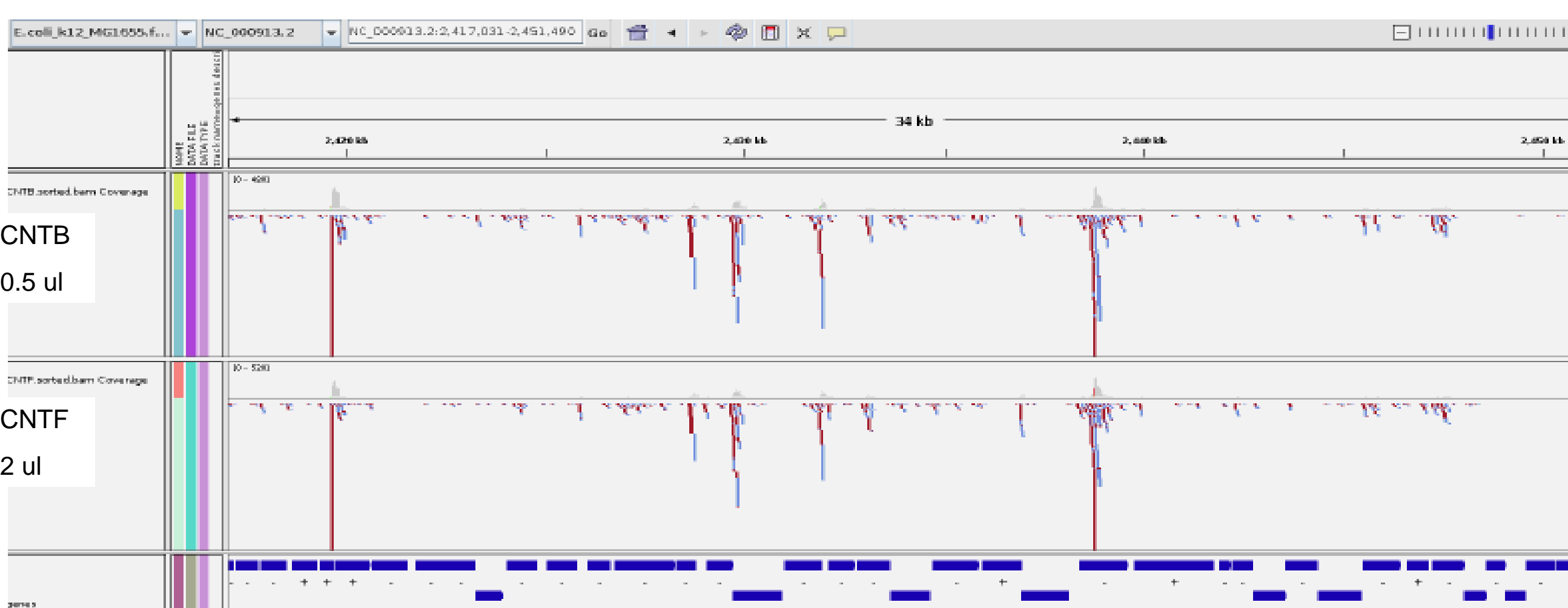
\* E-mail: zpeng@lbl.gov

## 3. Comparison of Different Ratios of RNA SPRI Beads for reaction Clearance



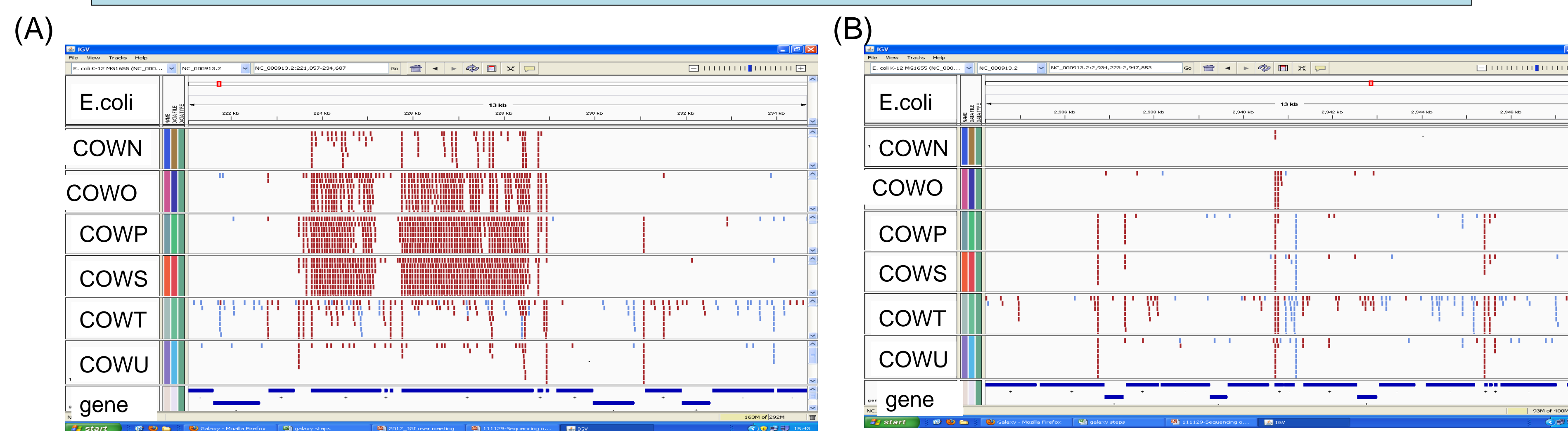
Different volume of RNA SPRI bead were used after enzyme reactions: (1)CNUC, 1:1.0; (2)CNUB,1:1.4; (3) CNUF, 1:0.8; and (4) CNUA, 1:1.8). mRNA recovery, small RNA such as 5S RNA and tRNA removal were compared. The results show that 1:1.8 volume of RNA beads get the best RNA recover.

## 4. Comparison of Different Concentrations of Hexmer-3'adapter Used in RT



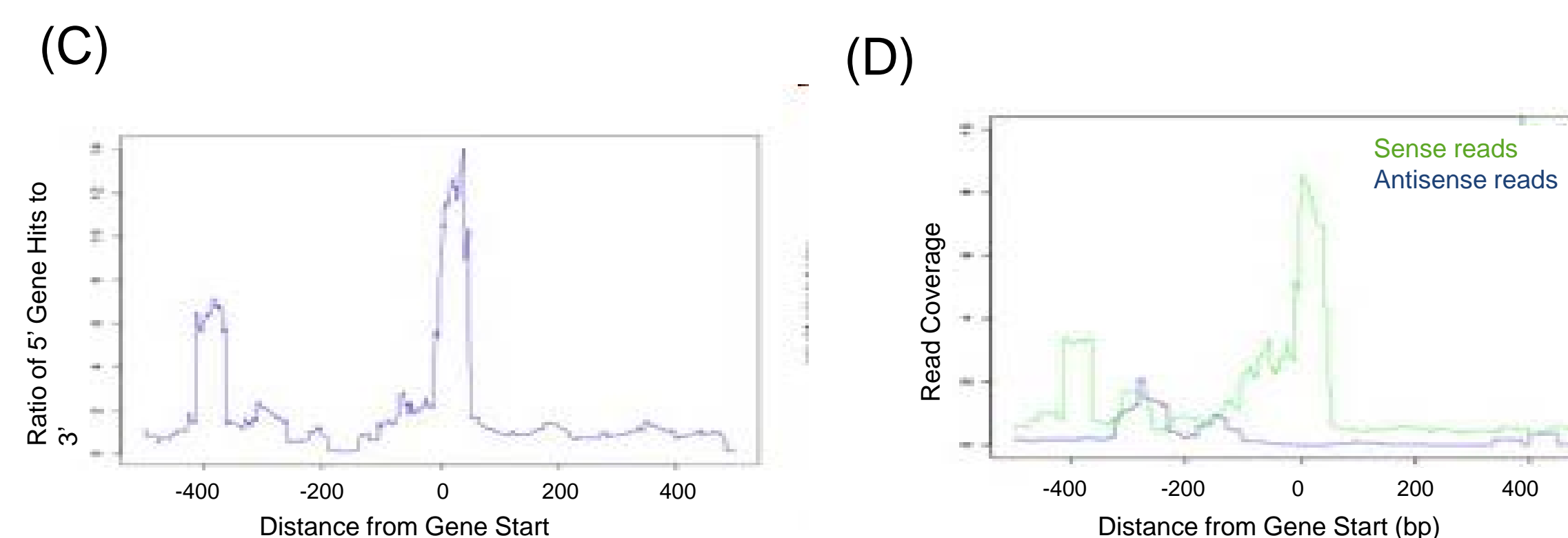
Different volume of hexmer-3'adapter(100 uM) were tested for reverse transcription (RT): (1)CNTB, 0.5 ul; (2)CNTC, 1.0 ul (data not shown) and (3) CNTF, 2 ul. Results shown using less adapter for RT can get more gene hit and reads. The pool library has been paired sequenced

## 5. Comparison of Different Methods for Removing rRNA



Different methods of removing rRNA and tRNA were compared: (1) COWN, non treatment; (2) COWO, MicroExpress alone; (3) COWP, MicroExpress plus Terminator Exonuclease; (4) COWS, Terminator Exonuclease alone; (5) COWT, Ribo-Zero alone and (6) COWU, Ribo-Zero plus Terminator Exonuclease.

Figure (A) shows some species of 16s and 23s rRNA can't be removed by MicroExpress and Terminator Exonuclease. Figure (B) shows the best results were obtained by using Ribo-Zero combined with Terminator Exonuclease.



C,D are all from COWU. (C) Most of the reads map to the 5' end of the gene, shown here by the ratio of read coverage at the gene start versus the middle of the gene. The upstream peak is likely from the start of the previous gene.

(D) Shows that the great majority of reads mapping at the transcription start site are sense stranded.

## Acknowledgements

We would like to thank Edward Kirton for Galaxy technical support