Faculty of Science
University of Copenhagen

Center for Computational and Applied Transcriptomics

Massive parallel sequencing based hydroxyl radical probing of RNA accessibility (**HRF-Seq**) using Fenton chemistry and synchrotron irradiation

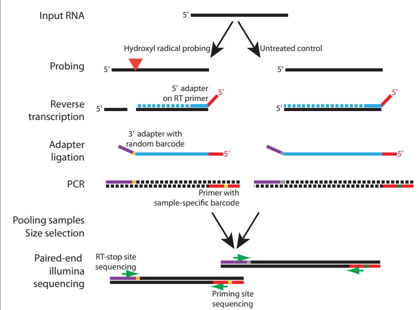
Łukasz Jan Kiełpiński and Jeppe Vinther
Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, DK-2200 Copenhagen N, Denmark
Email: Kielpinski@bio.ku.dk

Abstract

The hydroxyl radical footprinting (HRF) is a well established method for assaying nucleic acid backbone accessibility, which found its applications in studying the tertiary and quaternary structure of RNA and DNA. Traditionally, the signal is detected with either slab-gel or capillary electrophoresis, which considerably limits the throughput. Here we present a method of HRF signal detection utilizing the massive parallel sequencing, called HRF-Seq, which allows for a simultaneous analysis of multiple, long RNA molecules. The HRF-Seq workflow starts with the hydroxyl radical treatment of the RNA ensemble of interest, followed by a randomly primed reverse transcription, adapter ligation, PCR amplification and detection of cDNA 3' ends with Illumina sequencing. We describe a novel computational method of alleviating the PCR bias which uses the random barcodes introduced during the ligation. Moreover, we show the normalization procedure which unifies the signal over the regions of varying coverage and corrects for the background terminations. The HRF-Seq correlates well with the slab-gel electrophoresis and with the RNA backbone accessibility measured from the known crystal structures with the resolution that allows for the observation of differential reactivity of sides of helix. The results indicate that the method can generate useful constraints for the automated three dimensional RNA structure modeling in a high-throughput manner. As a next step we are applying the HRF-Seq method for in vivo RNA probing of mouse liver RNA. Samples were probed with synchrotron generated X-ray beams and the quality of RNA shows clear dose-response relationship with the irradiation time. Further analysis of the probed RNA is ongoing.

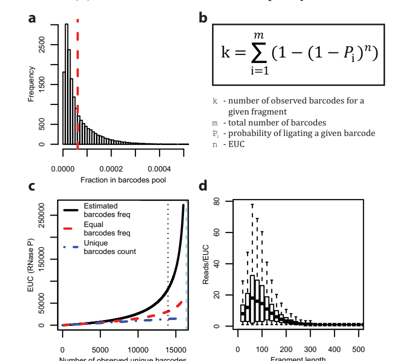
Experimental workflow

A complex mix of RNA is probed with hydroxyl radicals and reverse transcribed with random primers. The 3'end of cDNA is ligated to the adapter which is followed by PCR amplification and Illumina HiSeq sequencing.



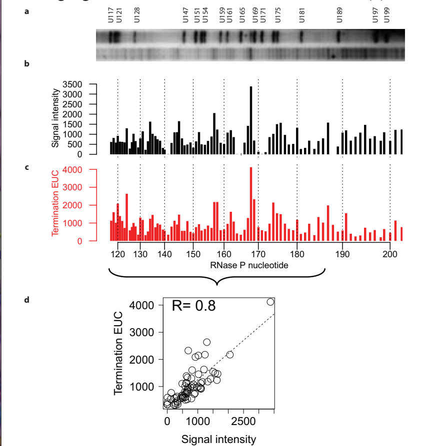
Estimated Unique Counts (EUC)

To alleviate the method biases we have decided to work with EUCs instead of with number of reads or unique barcodes. Fu *et al.* calculated EUC assuming equal barcode probability, which is not the case in our samples (**a**). We have generalized their method allowing different barcodes to have different probabilities (**b**). Our predictions are consistent with counting barcodes for lowly represented fragments but make a difference for highly abundant ones (**c**). Use of EUCs instead of raw reads decreased the weight of short fragments, likely correcting for the PCR bias (**d**), and increased the overall quality of our data.



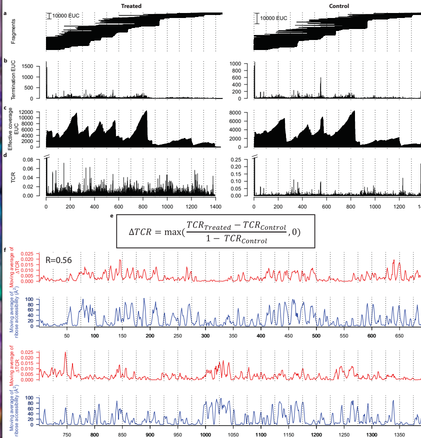
HRF-Seq EUCs correlate well with slab-gel detected intensities

The RNase P specificity domain was probed with hydroxyl radicals and the resulting degradation pattern was detected with either slab-gel electrophoresis (**a** and **b**) or HRF-Seq protocol (**c**) showing high level of correlation between the methods (**d**).



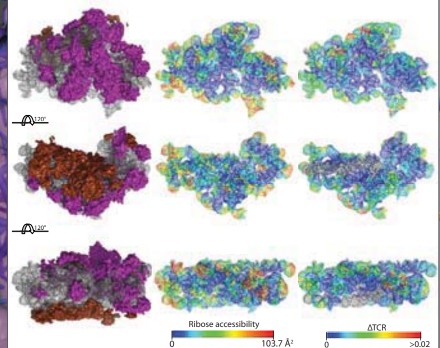
HRF-Seq ΔTCR correlates with RNA backbone solvent accessibility

Sequenced fragments from the treated and control samples were mapped to the *E. coli* 16S ribosomal RNA (**a**) and used to count the termination EUC (**b**) and the effective coverage EUC (**c**). Those values were used to calculate the termination-coverage ratio (TCR, **d**). Using a formula modified from the QuSHAPE procedure (**e**) we obtained the ΔTCR, which were smoothened with the 3-nt moving average and compared with the ribose solvent accessibility as measured from the crystal structure (**f**).

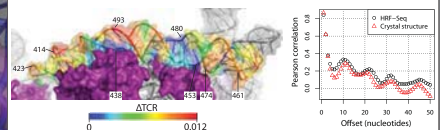


16S rRNA surface accessibility

Three views of the crystal structure of 16S rRNA (pdb: 3OFC) with the bound proteins (**left**, grey - RNA, purple - 16S proteins, brown - fragments of the large subunit), colored with moving average of the backbone accessibility as measured from the crystal structure (**center**) and colored with the smoothened HRF-Seq ΔTCR values (**right**). The HRF-Seq clearly recapitulates the footprints of proteins and of the large subunit.

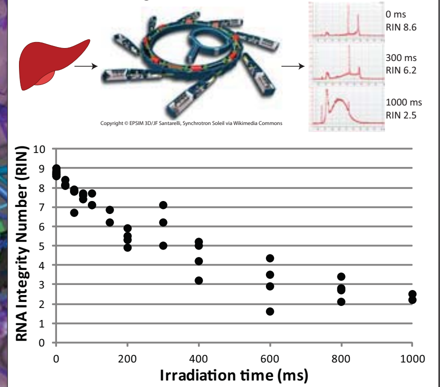


The resolution of HRF-Seq allows for discrimination between sides of a helix exposed on the ribosomal surface and for the observation of a molecule-wide signal correlation between positions offset by multiples of a helical turn.



RNA tertiary structure *in vivo*

Our current work focuses on applying the HRF-Seq to detect the hydroxyl radical cleavages induced with synchrotron generated X-rays in the frozen fragments of mouse liver. We observe that the extent of breakages correlates with the irradiation time.



Summary

We present a new method for HRF signal detection with massive parallel sequencing. The HRF-Seq largely increases the experimental throughput, allowing the analysis of multiple, long molecules simultaneously and produces data that agree well with the structures determined with the X-ray crystallography. The HRF-Seq incorporates the novel way of tackling the biases with the estimated unique count estimation based on the incorporated random barcodes and uses a new method of data normalization which utilizes the priming site information. We expected the method and the computational improvements to contribute to the expanding field of global-scale RNA structure analysis.

References

1. Kielpinski, L.J. and Vinther, J. (2014) Massive parallel-sequencing-based hydroxyl radical probing of RNA accessibility. *Nucleic Acids Res.*
2. Fu, G.K., Hu, J., Wang, P.H. and Fodor, S.P. (2011) Counting individual DNA molecules by the stochastic attachment of diverse labels. *Proc Natl Acad Sci U S A*,
3. Karabiber, F., McGinnis, J.L., Favorov, O.V. and Weeks, K.M. (2013) QuShape: rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis. *RNA*, 19, 63-73. 108, 9026-9031.
4. Dunkle, J.A., Xiong, L., Mankin, A.S. and Cate, J.H. (2010) Structures of the Escherichia coli ribosome with antibiotics bound near the peptidyl transferase center explain spectra of drug action. *Proc Natl Acad Sci U S A*, 107, 17152-17157.
5. Adilakshmi, T. (2006) Hydroxyl radical footprinting in vivo: mapping macromolecular structures with synchrotron radiation. *Nucleic Acids Research*, 34, e64-e64.

Acknowledgments

We are grateful to Jan Christiansen who helped purify *E. coli* ribosomes, to Anders Krogh for advice on calculation of EUCs and to Sayan Gupta from The National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory for performing the tissue irradiation. We thank the Danish National DNA Sequencing Center for performing sequencing. Research was funded by Danish Council for Strategic Research and Department of Biology, University of Copenhagen. Figures 2, 3, 4 and 5 reproduced here from reference 1 with modifications (CC BY 3.0 license).

Download the HRF-Seq paper

http://dx.doi.org/10.1093/nar/gku167

Download this poster

http://people.binf.ku.dk/lukasz/HRF-Seq_poster.pdf

Recorded poster presentation

http://lkielpinski.wordpress.com/2014/03/10/rna-meeting-poster/

Copyright © 2014 The Authors