Kethoxal-assisted single-stranded DNA sequencing (KAS-seq) for capturing transcription dynamics and enhancer activity

CURRENT STATUS: POSTED

Tong Wu
University of Chicago
ORCiD: https://orcid.org/0000-0001-9968-9909

Ruitu Lyu
University of Chicago

Chuan He
University of Chicago
chuanhe@uchicago.edu*Corresponding Author*
ORCiD: https://orcid.org/0000-0003-4319-7424

DOI:
10.21203/rs.3.pex-835/v2

SUBJECT AREAS
Biotechnology *Cell Biology*

KEYWORDS
single-stranded DNA, transcription, enhancer

Abstract

We describe a rapid and sensitive labeling of single-stranded DNA for sequencing (KAS-seq), based on a kethoxal-guanine reaction. KAS-seq procedure involves N_3-kethoxal labeling, DNA isolation, biotinylation, fragmentation and enrichment, library preparation and sequencing. The whole process can be finished in a day. KAS-seq enables rapid measurement of the dynamics of transcriptionally-engaged Pol II, transcribing enhancers, Pol I and Pol III activities, and non-canonical DNA structures involving ssDNA simultaneously *in situ*, by using as few as 1,000 cells or mice tissues. The robust and tissue-friendly nature coupled with low input material requirement make KAS-seq a method that can be broadly applied. This step-by-step protocol is related to the publication “Kethoxal-assisted single-stranded DNA sequencing captures global transcription dynamics and enhancer activity *in situ*” in *Nature Methods*.

Introduction

Transcription and its regulation determine cell fate and physiological functions, with dysfunctions in transcriptional regulation associated with various human diseases[1]. To understand global transcription regulation, genome-wide sequencing approaches have been developed to analyze the occupancy of RNA polymerases (ChIP-seq)[2], or detect the presence and level of nascent RNA. Nascent RNA analysis is usually based on run-on assays[3,4], metabolic labeling[5,6], and Pol II-associated or chromatin-associated RNA enrichment[7-11]. Although powerful, these methods also have limitations. Run-on-based methods and Pol II-associated RNA enrichment typically require millions of cells as starting materials. Pol II ChIP-seq could not distinguish whether RNA polymerases are simply bound or are actively engaged in transcription[3]. Metabolic labeling may not be able to accurately measure transient and low-abundant RNA species, such as enhancer RNAs (eRNAs), especially when using limited materials with modest sequencing depth. As most RNAs undergo post-transcriptional processing, their levels are indirect readouts that may not accurately reflect transcription dynamics *in situ*.

Transcriptionally engaged RNA polymerases resolve DNA double helices and generate single-stranded DNA bubbles. Mapping ssDNA throughout the genome provides a readout of the activity and dynamics

of transcriptionally engaged RNA polymerases.

Here we describe a rapid and sensitive labeling of single-stranded DNA for sequencing (KAS-seq, Figure 1), based on a kethoxal-guanine reaction[12]. We show that KAS-seq simultaneously measures the dynamics of transcriptionally-engaged Pol II, transcribing enhancers, Pol I and Pol III activities, and non-canonical DNA structures involving ssDNA *in situ*, by using as few as 1,000 cells or mice tissues.

Reagents

N_3-kethoxal (Synthesized as previously[12])

DMSO (Fisher Scientific, BP231-100)

DBCO-PEG_4-biotin (Sigma, 760749)

Boric acid (Sigma, 15663)

Potassium hydroxide (Sigma, P5958)

Tween-20 (Sigma, P9416)

Phosphate buffered saline (PBS, Thermo, 10010001)

Phosphate buffered saline, 10× (Thermo, AM9624)

1 M Tris-HCl buffer (Thermo, 15567027)

5 M sodium chloride solution (Thermo, AM9759)

0.5 M EDTA solution (Thermo, AM9206G)

RNase A (Thermo, 12091039)

Dynabeads Myone Streptavidin C1 (Thermo, 65001)

PureLink genomic DNA mini kit (Thermo, K182001)

Quick-DNA Microprep Plus kit (Zymo, D4074)

MinElute PCR purification kit (Qiagen, 28004)

DNA Clean & Concentrator kit (Zymo, D4013)

Tagmentation DNA Enzyme (Illumina, 15027865)

Tagmentation DNA Buffer (Illumina, 15027866)

i5 index primer (Illumina, 20027213)

i7 index primer (Illumina, 20027213)

NEBNext Ultra II Q5 Master Mix (NEB, M0544S)

Accel-NGS Methyl-seq DNA library kit (Swift 30024)

Equipment

1.5 mL tube

Nanodrop

Centrifuge

Themomixer

Thermo cycler

DynaMag-PCR Magnet

Bioruptor Pico (Diagenode)

0.65 mL Bioruptor Micotubes (Diagenode, C30010011)

Procedure

1. KAS-seq with mammalian cell cultures

Labeling and DNA isolation

1. Prepare 500 mM N_3-kethoxal stock solution using DMSO. Then prepare the labeling medium by diluting the N_3-kethoxal solution into pre-warmed (37 ˚C) cell culture medium to a final concentration of 5 mM. It is critical to pre-warm the medium to facilitate N_3-kethoxal dissolution.

2. Incubate 1-5 million cells in the labeling medium for 10 min at 37 ˚C. For adhesive cells, apply the labeling medium to the cell culture directly in dishes and incubate cells for 10 min at 37 ˚C. For suspension cells, suspend cells in the labeling medium and incubate them for 10 min at 37 ˚C.

3. Harvest cells after the incubation. Isolate total DNA from cells by using PureLink genomic DNA mini kit (Thermo K182001). Elute DNA by using 50 µL 25 mM K_3BO_3 (pH 7.0). **[Stop point 1]**

Biotinylation and purification

4. Prepare the click reaction mixture as the follows.

2 µg total DNA diluted in 85µL of 25 mM K_3BO_3 (pH 7.0)

10 µL 10× PBS

5 µL 20 mM DBCO-PEG_4-biotin (DMSO solution, Sigma 760749)

Incubate the mixture at 37 ˚C for 1.5 h with shaking at 500 rpm to facilitate the “click” reaction.

5. Add 5 µL RNase A (Thermo 12091039) to the reaction mixture. Incubate the mixture at 37 ˚C for 15 min with shaking at 500 rpm.

6. Purify DNA from the reaction mixture by DNA Clean & Concentrator kit (Zymo D4013). Elute DNA by using 50 µL 25 mM K_3BO_3 (pH 7.0). **[Stop point 2]**

Enrichment of N_3-kethoxal-modified DNA

7. Dilute 1 µg biotinylated DNA in 100 µL 25 mM K_3BO_3 (pH 7.0). Fragment the DNA by sonication using Bioruptor Pico under 30s-on /30s-off setting for 30 cycles. The size of fragmented DNA ranges from 150-350 bp.

8. Save 5 µL of the fragmented DNA as input. Use the remaining 95 µL for enrichment.

9. Wash 10 µL Dynabeads Myone Streptavidin C1 (Thermo 65001) 3 times with 50 µL 1× binding and wash buffer (5 mM Tris-HCl pH 7.4, 0.5 mM EDTA, 1 M NaCl, 0.05% Tween-20). Re-suspend pre-washed beads in 95 µL 2× binding and wash buffer (10 mM Tris-HCl pH 7.4, 1 mM EDTA, 2 M NaCl, 0.1% Tween-20).

10. Mix the beads with 95 µL fragmented DNA from step 8. Incubate the mixture at room temperature for 15 min with gentle rotation. After incubation, place the mixture on a magnetic rack to remove the supernatant and wash beads 5 times with 100 µL 1× binding and wash buffer.

11. Re-suspend the beads in 15 µL nuclease-free water and heat it at 95 ˚C for 10 min to elute enriched DNA. Then put the mixture on a magnetic rack, and transfer the supernatant to a new labeled tube.

12. In parallel with the previous step, heat the saved input at 95 ˚C for 10 min. Then, put both the input and the eluted samples on ice immediately. **[Stop point 3]**

Library preparation and sequencing

13. Perform library construction by using the Accel-NGS Methyl-seq DNA library kit (Swift 30024). Perform PCR amplification for 7-8 cycles for the input samples and 12-14 cycles for the enriched samples. The library size should range from 200-500 bp.

14. Sequence libraries on Illumina platforms by using single-end mode, aiming to get 30 million reads per sample

DNAs can be stored at -20 ˚C at the noted stop points if needed.

2. KAS-seq with mouse liver

Labeling and DNA isolation

1. Homogenize mouse liver tissue to a cell suspension in ice-cold PBS by using a dounce homogenizer or a pellet pestle.

2. Spin the cell suspension at 100 g for 15 seconds to sediment and remove potential large tissue pieces at the bottom of the tube.

3. Spin the cell suspension at 800 g for 5 min. Remove the supernatant and save the cell pellet at the bottom of the tube for labeling.

4. Suspend 1-5 million cells in the labeling medium for 10 min at 37 ˚C (refer to step 1 in the protocol for mammalian cell cultures for labeling medium preparation).

5. Isolate total DNA from cells by using PureLink genomic DNA mini kit (Thermo K182001). Elute DNA by using 50 µL 25 mM K_3BO_3 (pH 7.0). **[Stop point 1]**

Perform Biotinylation and purification, Enrichment of N_3-kethoxal-modified DNA, Library preparation and sequencing according to the protocol for mammalian cell cultures.

3. Low-input KAS-seq

Labeling and DNA isolation

1. Label 1,000, 5,000, or 10,000 cells according to the protocol for bulk mammalian cell cultures.

2. Isolate DNA from cells by using the Quick-DNA Microprep Plus kit (Zymo D4074). Elute DNA by using 43 µL 25 mM K_3BO_3 (pH 7.0).

Biotinylation and purification

3. Prepare the click reaction mixture as follows.

42.5 µL DNA in 25 mM K_3BO_3 (pH 7.0)

5 μL 10× PBS

2.5 μL 20 mM DBCO-PEG$_4$-biotin (DMSO solution, Sigma 760749)

Incubate the mixture at 37 ˚C for 1.5 h with shaking at 500 rpm to facilitate the “click” reaction.

4. Add 2.5 μL RNase A (Thermo 12091039) to the reaction mixture. Incubate the mixture at 37 ˚C for 5 min.

5. Purify DNA from the reaction mixture by DNA Clean & Concentrator kit (Zymo D4013). Elute DNA by using 25 μL 25 mM K_3BO_3 (pH 7.0).

Tagmentation and enrichment

6. Perform tagmentation on biotinylated DNA by using Tagmentation DNA Enzyme (Illumina 15027865). For 1,000 cells, 23.5 μL DNA was mixed with 1.5 μL enzyme and 25 μL Tagmentation DNA Buffer (Illumina 15027866). For 5,000 cells, 23 μL DNA was mixed with 2 μL enzyme and 25 μL Tagmentation DNA Buffer. For 10,000 cells, 20 μL DNA was mixed with 5 μL enzyme and 25 μL Tagmentation DNA Buffer.

7. Incubate the tagmentation mixture at 37 ˚C for 30 min with shaking at 500 rpm.

8. Purify DNA from the reaction mixture by DNA Clean & Concentrator kit (Zymo D4013). Elute DNA by using 55 μL 25 mM K_3BO_3 (pH 7.0).

9. Save 5 μL of the tagmented DNA as input. Use the remaining 50 μL for enrichment.

Enrichment of N_3-kethoxal-modified DNA

10. Wash 5 μL Dynabeads Myone Streptavidin C1 (Thermo 65001) 3 times with 25 μL 1× binding and wash buffer (5 mM Tris-HCl pH 7.4, 0.5 mM EDTA, 1 M NaCl, 0.05% Tween-20). Re-suspend pre-washed beads in 50 μL 2× binding and wash buffer (10 mM Tris-HCl pH 7.4, 1 mM EDTA, 2 M NaCl, 0.1% Tween-20).

11. Mix the beads with 50 μL fragmented DNA from step 9. Incubated the mixture at room temperature for 15 min with gentle rotation. After incubation, place the mixture on a magnetic rack to remove the supernatant and wash beads 5 times with 100 μL 1× binding and wash buffer.

PCR amplification and library sequencing

12. Adjust the volume of input to 20 µL by adding 15 µL nuclease-free water to the 5 µL input saved in step 9. Re-suspend the washed beads in 20 µL nuclease-free water.

13. Prepare the PCR reaction mix as follows:

20 µL input DNA or beads suspension

2.5 µL i5 index primer (Illumina 20027213)

2.5 µL i7 index primer (Illumina 20027213)

25 µL NEBNext Ultra II Q5 Master Mix (NEB M0544S)

14. Perform PCR reaction in a thermocycler as follows:

1 cycle: 5 min 72 ˚C

10 min 95 ˚C

15 cycles: 10 sec 98 ˚C

30 sec 60 ˚C

1 min 72 ˚C

15. Remove the beads from the enriched samples by using a magnetic rack. Purify the library by using MinElute PCR purification kit (Qiagen 28004). The library size should range from 200-600 bp.

16. Sequence libraries on Illumina platforms by using single-end mode with 30 million reads per sample.

Time Taken

Labeling, cell collection, and DNA isolation: ~1 h

Biotinylation: ~2 h

Fragmentation and enrichment: ~1 h

Library preparation: ~3 h

Anticipated Results

1. Sequencing results show traditional KAS-seq features. Signals are enriched at gene-coding regions, with sharp and strong peaks on TSS, broad peaks on the gene body and terminal regions (Figure 3).

2. KAS-seq signals show up in some enhancers.

3. KAS-seq signals show up in some predicted non-B form ssDNA-containing regions.

References

1. Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473**, 337-342 (2011).

2. Kim, T.-K. *et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465**, 182-187 (2010).

3. Core, L.J., Waterfall, J.J. & Lis, J.T. Nascent RNA Sequencing Reveals Widespread Pausing and Divergent Initiation at Human Promoters. *Science* **322**, 1845-1848 (2008).

4. Kwak, H., Fuda, N.J., Core, L.J. & Lis, J.T. Precise Maps of RNA Polymerase Reveal How Promoters Direct Initiation and Pausing. *Science* **339**, 950-953 (2013).

5. Fuchs, G. *et al.* 4sUDRB-seq: measuring genomewide transcriptional elongation rates and initiation frequencies within cells. *Genome Biology* **15**, R69 (2014).

6. Schwalb, B. *et al.* TT-seq maps the human transient transcriptome. *Science* **352**, 1225-1228 (2016).

7. Churchman, L.S. & Weissman, J.S. Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* **469**, 368-373 (2011).

8. Christopher, Ramachandran, S. & Henikoff, S. Nucleosomes Are Context-Specific, H2A.Z-Modulated Barriers to RNA Polymerase. *Molecular Cell* **53**, 819-830 (2014).

9. Nojima, T. *et al.* Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. *Cell* **161**, 526-540 (2015).

10. Mayer, A. *et al.* Native Elongating Transcript Sequencing Reveals Human Transcriptional Activity at Nucleotide Resolution. *Cell* **161**, 541-554 (2015).

11. Hirabayashi, S. *et al.* NET-CAGE characterizes the dynamics and topology of human transcribed cis-regulatory elements. *Nature Genetics* 51, 1369-1379 (2019).

12. Weng, X. *et al.* Keth-seq for transcriptome-wide RNA structure mapping. *Nature Chemical Biology* (2020).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

ProtocolExchangeFig3.png
ProtocolExchangeFig1.png
ProtocolExchangeFig2.jpg

Kethoxal-assisted single-stranded DNA sequencing captures global transcription dynamics and enhancer activity in situ

by Tong Wu, Ruitu Lyu, Qiancheng You, +1

Nature Methods (06 April, 2020)