Published in final edited form as:

Methods Mol Biol. 2019 ; 1933: 49–65. doi:10.1007/978-1-4939-9045-0_3.

Identification of long noncoding RNAs in the developing endosperm of Maize

Eundeok Kim^{1,*}, Yuqing Xiong², Byung-Ho Kang^{2,3}, Sibum Sung^{1,4,#}

¹Department of Molecular Biosciences and Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, TX 78712, USA

²Microbiology and Cell Science, University of Florida, Gainesville, FL 32611, USA

³School of Life Sciences, State Key Laboratory for Agrobiotechnology, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong, China

⁴International Scholar, Kyung-Hee University, South Korea

Abstract

Maize endosperm consists of three distinct types of tissues, including the starchy endosperm (SE), the basal endosperm transfer cell layer (BETL), and the aleurone cell layer (AL).

Compartmentalization of these tissues during endosperm differentiation makes the endosperm development an excellent model to study changes in gene expression during development. By utilizing cryo-dissection of developing endosperm, morphologically distinct samples can be obtained for transcriptome and epigenome analysis. Here, we describe methods for the isolation of tissues from developing maize endosperm and for the transcriptome analysis to identify novel long noncoding RNAs. The transcriptome data can be further analyzed to illustrate spatiotemporal changes in both coding and non-coding transcripts during the endosperm development.

Keywords

Endosperm; long noncoding RNAs; transcriptome analysis; cryo-dissection; maize; *in situ* hybridization

1 Introduction

High-throughput RNA sequencing (RNA-Seq) has identified rare transcripts, including long noncoding RNAs (lncRNAs). Eukaryotic transcriptomes include a large number of long noncoding RNAs, some of which plays roles in various developmental programs [1]. A number of recent studies have shown that lncRNAs are essential *cis* and/or *trans* regulators of gene activity that can function as scaffolds for chromatin-modifying complexes and precursors for regulatory RNAs [2–5]. Although several lncRNAs have been known for decades, it has only been appreciated recently with the advent of high-throughput RNA sequencing (RNA-Seq) [6–8]. Despite growing numbers of identified lncRNAs in

#Corresponding Author: Sibum Sung, sbsung@austin.utexas.edu.

*Current address: Department of Biology, University of Washington, Seattle, WA 98195, USA

eukaryotes, functions and underlying mechanisms of lncRNA action are still poorly understood. As the first step to decipher the roles of lncRNAs, the establishment of robust bioinformatics analysis of transcriptome data is essential to correctly annotate lncRNAs. In addition, identification of differentially expressed lncRNAs during development would help identify potential regulatory lncRNAs.

Development of maize endosperm includes compartmentalization of differentiated tissues, making it feasible to micro-dissect endosperm to enrich tissue samples that are at different developmental stages of each specific cell types. Laser Capture Microdissection (LCM) [9] or cryo-dissection [10] can be used to isolate specific cell types. Cryo-dissection method is less invasive compared to the LCM, although it is rather labor-intensive. Despite of the large number of transcriptome data obtained from maize endosperm [11–16], spatiotemporal transcriptome data analysis using developing endosperm has been poorly reported. Here, we described methods of transcriptome analysis using cell-type specific tissues at different developmental stages, by which spatiotemporally expressed lncRNAs can be identified. We also described methods to validate identified lncRNAs by reverse transcription followed by quantitative PCR (RT-PCR) and by *in situ* hybridization in maize endosperm.

2 Materials

2.1 RNA extraction from three cell types of the maize endosperm

2.1.1. Equipment—1. Cryostat (Leica-Microsystem, CM1950)

2. Dissecting stereomicroscope (Meiji Techno, EMZ-200)

3. Staining jars

4. Scalpels

5. NanoDrop ND-1000 Spectrophotometer (Thermo Scientific)

2.1.2. Solutions and Chemicals—1. Liquid nitrogen

2. Tissue-Tek optimum cutting temperature-embedding compound (VWR)

3. HistoGene LCM Frozen Section Staining Kit (Applied Biosystems)

4. PEN Membrane Glass Slides (Applied Biosystems)

5. Picopure RNA isolation Kit (Applied Biosystems)

6. RapidOut DNA Removal Kit (Thermo Scientific)

7. RiboMinus Plant Kit for RNA-Seq (Invitrogen)

8. Glycogen (Sigma)

9. Sodium acetate

10. 100% ethanol

11. Xylene

2.2 Preparation of cDNA libraries for Illumina sequencing

2.2.1. Equipment—1. Agilent Bioanalyzer.

2.2.2. Solutions and Chemicals—1. Invitrogen Dynabeads®

2. Oligo(dT)₂₅ (Invitrogen)

3. Actinomycin D (Sigma)

4. AMPure XP (Agencourt)

5. RNAClean XP (Agencourt)

6. RNasin Plus RNase Inhibitor (Promega)

7. PicoGreen dsDNA Assay Kit (Invitrogen)

8. dNTP (10 mM each of dATP, dTTP, dGTP, dCTP)

9. Superscript III Reverse Transcriptase (Invitrogen)

10. DTT

11. DNA Polymerase I (NEB)

12. End-Repair Mix (Enzymatics)

13. T4 DNA Ligase (NEB)

14. Klenow 3'–5'exo– (NEB)

15. RNase H (NEB)

16. End-Repair Mix (NEB)

17. Phusion® High-Fidelity DNA Polymerase (NEB)

18. MgCl₂

19. DMSO

20. 100% ethanol

21. Nuclease-Free Water

2.2.3 Buffers—1. Binding Buffer: 20 mM Tris-HCl, pH 7.5, 1.0 M LiCl, 2 mM EDTA

2. Washing Buffer: 10 mM Tris-HCl, pH 7.5, 0.15 M LiCl, 1 mM EDTA

2.3. Equipment used to identification of long noncoding RNAs

2.3.1 Hardware—1. 64 bit computer with 1TB hard disk and 16GB of memory

2.3.2 Software—1. Unix/Linux platform (e.g., Mac OS or Ubuntu)

2. Python 2.7.12 with NumPy, SciPy and Pylab package

3. Bedtools

4. Cufflinks

5. Biopython

2.4 Quantitative RT-PCR

2.4.1 Equipment—1. Pellet pestle, or equivalent

2. Liquid nitrogen

3. Thermal cycler

4. Real-Time PCR system

5. DNase and RNase free Microfuge tubes, 1.5mL

6. DNase and RNase free PCR tubes, 0.2mL

2.4.2 Solutions and Chemicals—1. Nuclease-Free Water

2. Random primer

3. SYBR Green/ROX qPCR master MIX (Thermo Fisher)

4. M-MLV Reverse Transcriptase (Promega)

5. RNase H (Promega)

6. DNase I (Invitrogen)

7. Dithiothreitol (DTT)

8. Plant RNA extraction Kit (Qiagen)

9. PCR purification Kit (Qiagen)

2.5 *In Situ* hybridization

2.5.1 Equipment—1. NanoDrop ND-1000 Spectrophotometer (Thermo Fisher)

2.5.2 Solutions and Chemicals—1. AccuPrime™ *Pfx* DNA Polymerase (Invitrogen)

2. QIAquick PCR Purification Kit
3. Zero Blunt[®] TOPO[®] PCR Cloning Kit (Invitrogen)
4. M13-F (or M13-R) primers
5. DIG RNA Labeling Mix (Sigma-Aldrich)
6. Protector RNase Inhibitor (Sigma-Aldrich)
7. SP6 RNA Polymerase (Sigma-Aldrich)
8. T7 RNA Polymerase (or SP6 RNA Polymerase, Sigma-Aldrich)
9. Qiagen RNeasy MinElute Cleanup Kit (Qiagen)
10. Nitro blue tetrazolium chloride (Sigma-Aldrich)
11. BCIP (Sigma-Aldrich)
12. Acetic acid
13. Sodium acetate
14. 100% Ethanol
15. Diethyl pyrocarbonate (DEPC)
16. Acetone
17. Formamide
18. Polyvinyl alcohol,

2.5.3 Buffers—1. Sodium carbonate buffer: 200mM sodium carbonate, 200,

2. Maleic acid buffer: 100 mM maleic acid, 150 mM NaCl, pH 7.5
3. Developing solution: 100 mM Tris-HCl pH 9.5, 150 mM NaCl, 50 mM MgCl₂
4. 20X SSC buffer: 3M NaCl, 0.3M sodium citrate, pH7.0
5. 20X SSPE buffer: 0.02M EDTA, 2.98M NaCl, 0.2M phosphate buffer (pH 7.4)
6. Washing buffer: 50% formamide, 2× SSC
7. NTE buffer: 500 mM NaCl, 10 mM Tris pH 7.5, 1 mM EDTA
8. 10X Phosphate buffer saline: 1.27 M NaCl, 27 mM KCl, 100 mM Na₂HPO₄, 18 mM KH₂PO₄
9. Detection buffer 1: 100 mM Tris pH 7.5, 150 mM NaCl

10. Detection buffer 2: 100 mM Tris pH 9.5, 100 mM NaCl, 50 mM MgCl₂

3 Methods

3.1 RNA extraction from three cell types in the maize endosperm

3.1.1 cryosectioning and isolating total RNA samples from maize endosperm

- 1. Collect maize kernels at the field and freeze them immediately with liquid nitrogen.
2. Mount frozen kernels on cryostat stubs and prepare 30 micrometer sections in cryostat (−22°C).
3. Collect sections on PEN membrane glass slides. After thawing the frozen sections, quickly stain the sections with HistoGene LCM Frozen Section Staining Solution.
4. Dehydrate the stained sections in ethanol (75%, 90%, and 100%) and finally in 100% xylene.
5. After drying 5 min in the fume hood, dissect out starchy endosperm, basal endosperm cell layer, and aleuron cell layer with scalpel. The PEN membrane facilitate separation of tissue specimens from the glass slide.
6. Isolate total RNA samples from dissected tissue specimens with a Picopure RNA isolation Kit according to the manufacturer's direction and remove DNA from the total RNA samples with a RapidOut DNA removal kit. RNA samples should be more than 1 µg for Illumina library preparation.

3.1.2 Removal of ribosomal RNA—1. To a sterile, RNase-free 1.5 mL microcentrifuge tube, add the following:

Total RNA (1–5 µg):	4.8 µL
Spike-in control RNA (50 pg/µL):	0.2µL
RiboMinus™ Probe (15 pmol/µL):	2.5 µL
Hybridization Buffer:	22.5 µL

2. Incubate the tube at 70–75°C for 5 min to denature RNA.
3. Allow the sample to cool to 37°C slowly over a period of 30 minutes by placing the tube in a 37°C water bath. It is important to allow slow cooling for sequence-specific hybridization.
4. After the 37°C incubation step of the hybridized sample (above), briefly centrifuge the tube to collect the sample to the bottom of the tube.
5. Transfer the sample to the RiboMinus Magnetic beads in the kit. Mix by pipetting up and down or low-speed vortexing.

6. Incubate the tube at 37°C for 15 minutes. During incubation, gently mix the contents occasionally. Briefly centrifuge the tube to collect the sample to the bottom.
7. Place the tube on a magnetic separator for 1 minute to pellet the rRNA-probe complex. Collect the supernatant containing RiboMinus RNA.
8. Repeat steps 2–4 to further remove rRNA-probe complex.
9. Transfer the RiboMinus RNA sample into a clean, nuclease-free 1.5 mL microcentrifuge tube and add the following components to precipitate RNA.

1 μ L glycogen (20 mg/mL)
1/10th sample (eluted RNA) volume of 3M sodium acetate
2.5X sample volumes of 100% ethanol

10. Centrifuge the tube for 30 minutes at the highest speed at 4°C. Carefully discard the supernatant without disturbing the pellet.
11. Wash the pellet with 70% cold ethanol twice.
12. Air-dry the pellet for ~5 minutes. Resuspend the RNA pellet in 10 μ L DEPC-treated water.

3.2 Preparation of cDNA libraries for Illumina sequencing

- ### 3.2.1 RNA isolation and fragmentation—
1. Add 10 μ L of Binding Buffer to RNA samples from 3.1 and heat up to 65°C to disrupt secondary structures. After 2 min at 65°C, place on ice immediately.
 2. To the 20 μ L of RNA samples, add 10 μ L Dynabeads. Mix the beads and solution thoroughly and anneal by rotating continuously on a mixer for 5 min at room temperature.
 3. Place the tube on the magnet for 1–2 min and carefully remove all the supernatant.
 4. Remove the tube from the magnet and add 20 μ L Washing Buffer. Mix by pipetting carefully a couple of times. Use the magnet to pull the beads to the side of the tube. Carefully remove all the supernatant.
 5. Repeat the washing step as described in step 4.
 6. Remove the tube from the magnet and resuspend the beads using 10 μ L 2 $\times$ SuperScript III buffer with 10 mM DTT.
 7. Incubate at 94°C for 10 min to fragment the mRNA. The incubation time, however, must be optimized for your samples.
 8. Briefly spin the tube and place it on the magnet. Transfer the supernatant quickly to a new 0.5 ml micro centrifuge tube.

3.2.2 First-strand cDNA synthesis—1. Assemble the following RT reaction:

Fragmented mRNA	10 μ L
Random hexamer	0.5 μ L
Rnasin Plus	0.5 μ L

2. Heat at 50°C for 1 min. Immediately place on ice. Add 9 μ L of the following master mix to the fragmented mRNA

H ₂ O	6.88 μ L
Actinomycin D (1 μ g/ μ L)	0.12 μ L
DTT (100 mM)	1 μ L
dNTP (25 mM)	0.5 μ L
SuperScript III	0.5 μ L

3. Perform the RT reaction (25°C for 10 min and 50°C for 50 min)

4. Add 36 μ L of RNAClean XP and incubate the mixture on ice for 15 min. (The solution is viscous; mix thoroughly by pipetting up and down.)

5. Collect RNA/cDNA hybrid molecules with solid phase reversible immobilization (SPRI) beads and wash the beads twice with 75% ethanol gently.

6. Air-dry the beads for 5 min. Elute the RNA/cDNA hybrid with 10 μ L distilled water.

3.2.3 Second-strand synthesis and end-repair—1. Prepare the second strand reaction master mix on ice as follows:

NEBuffer 2	1.5 μ L
dNTP mix (10 mM each)	1 μ L
RNase H (5 U/ μ L)	0.2 μ L
DNA polymerase I (10 U/ μ L)	1 μ L
H ₂ O	1.3 μ L

2. Add 5 μ L the master mix to each 10 μ L of RNA/cDNA hybrid samples. Incubate at 16°C for 2.5h.

3. Purify double-stranded DNA (dsDNA) using 1.8 volumes of AMPure XP beads. Elute with 10 μ L of H₂O.

4. Prepare an appropriate amount of the end-repair master mix on ice as follows:

End-Repair Buffer (10×)	1.5 µL
dNTP mix (1 mM)	3 µL
End-Repair Mix	0.5 µL

5. Add 5 µL of the master mix to each 10 µL of dsDNA. Incubate at 20°C for 30 min.

6. Purify using 1.8 volumes of AMPure XP beads. Elute with 10 µL of H₂O

3.2.3 A tailing and Y-Shape Adapter Ligation—1. Prepare A-Tailing master mix on ice:

NEBuffer 2	1.5 µL
dATP (10 mM)	0.5 µL
H ₂ O	2.5 µL
Klenow 3′–5′ exo	0.5 µL

2. Add 5 µL of the master mix to each sample and incubate at 37°C for 30 min.

3. Purify using 1.8 volumes of AMPure XP beads. Elute with 8 µL of H₂O

4. Add 1 µL of the desired barcode adapter (5 µM) to each sample

5. Prepare the master mix on ice as follows:

Rapid Ligation Buffer (2×)	8.5 µL
T4 DNA Ligase (600 U/µL)	0.5 µL

6. Add 9 µL of the master mix to each sample and incubate at room temperature for 15 min.

7. Purify DNA using one volume (18 µL) of AMPure XP beads. Elute with 10 µL of TE buffer.

8. Purify the library again using the precisely same amount of AMPure XP beads (10 µL) and elute with 10 µL of TE buffer.

3.2.3 PCR enrichment—1. Set up PCR reactions as follows:

ds DNA Sample	5 µL
primer A (10 µM)	0.2 µL
primer B (10 µM)	0.2 µL
Phusion HF Buffer (5×)	4 µL

dNTP (10 mM)	0.2 µL
H ₂ O	10 µL
Phusion II	0.5 µL

2. Perform initial denaturation at 94°C for 2 min, followed by 10–12 cycles of amplification (98°C for 10 sec, 65 °C for 30 sec, 72 °C for 30 sec). (This is a critical step. To avoid over-amplification, the PCR cycle should be optimized for each cell/tissue type. For many plant RNA samples: use 14–15 cycles for 0.5–1 µg total RNA, 10–12 cycles for 1–5 µg RNA, 8–10 cycles if >5 µg of RNA has been used)

3. If needed, remove adapter dimers (~100 bp) from PCR reactions with SPRI clean up or isolate 200–500 bp PCR products using Qiagen QIAquick Gel Extraction Kit.

4. Purify using one volume of AMPure XP beads. Elute with 20 µL of TE buffer

3.3 Identification of long noncoding RNAs

3.3.1 Identify long noncoding transcripts—1. Generate fasta file using bed file which includes the coordinate information of unannotated transcript from an intergenic region by using Bedtools (<http://bedtools.readthedocs.io/en/latest/index.html>)

```
$ bedtools getfasta [OPTIONS] -fi <input FASTA file> - bed <BED/GFF/VCF>
```

2. Identify coding regions within the transcripts sequence generated by de novo transcript assembly by using coding potential program (TransDecoder tool). TransDecoder tool will identify coding sequences that have open reading frame and homology to known proteins through blast or pfam searches.

```
$ TransDecoder.LongOrfs -t <input FASTA file>
```

```
$ TransDecoder.Predict -t <input FASTA file>
```

3. Remove the transcripts that are identified from step 2.

4. Remove putative precursors of all small RNA using the python script from [17]— also see Supplementary File 1 for the python script.

5. Remove transposable element (TE) transcripts using a modified python script (Supplementary File 2, python script 2). (See Note 1)

3.3.2 Cluster analysis by using the self-organizing map (SOM) analysis tool

—1. Transform the expression data by shift FPKM values so that the minimum value to be zero and normalize the maximum values to be 10. (See Note 2)

2. Upload the processed value to <https://genepattern.broadinstitute.org> to use SOM module embedded in the GenePattern.

3. Choose SOM module and perform clustering analysis with 4,200,000 repetitions of the process. (See Note 3)

3.3.3 Identification of *cis*-regulatory elements—1. Extract the genomic sequences from 1kb upstream and downstream of long noncoding RNA transcripts from each clusters from step 3.3.2.

2. Upload the extracted sequences to <http://meme-suite.org/> to identify *de novo* motifs in those regions [18].

3. Only those with *q*-values lower than 0.01 from step 2 were taken for further analysis to identify motifs with the significant similarity between coding RNA and lncRNA clusters by using TOMTOM program [19] (See Note 4)

3.4 Validate expression patterns of lncRNAs by quantitative Real-Time PCR

3.4.1 Total RNA isolation and First-strand cDNA synthesis for quantitative RT-PCR—1. Grind approximately 150 cryoselected samples (from 3.2) to a fine powder in liquid nitrogen using pellet pestle. (See Note 5)

2. Isolate total RNA using RNA isolation kit according to the manufacturer's instructions. (See Note 6)

3. Remove genomic DNA using DNase I according to the manufacturer's instructions.

4. Purify DNase I treated total RNA using RNA purification kit according to the manufacturer's instructions.

5. Synthesize First-strand cDNA with 1ug of total RNA from step 4 by using random primer according to the manufacturer's instructions using thermal cycler.

6. Purify the reaction from step5 using PCR purification kit according to the manufacturer's instructions and add Nuclease-free water to make a total 100ul volume.

3.4.2 Measure the relative expression of lncRNA by quantitative Real-Time PCR—1. Design the primer set targeting the identified lncRNA transcripts using a Primer3 program including ZmTXN as an internal reference (See Note 7)

2. Prepare the qPCR amplification mix in a 0.2mL tube on ice.

PCR amplification reaction mix	
Component	Volume (uL)
2X SYBR master mix	7.5
Water	4.5
10 uM Primer F	0.5
10 uM Primer R	0.5

Purified cDNA	2
Total	15

3. Mix well and perform PCR cycling as described Table X using a real-time machine.

Cycle number	Temperature (°C)	Time
1	95	10 min
40 cycles	95	15 sec
	60	30sec
	72	30sec
1 cycles	4	Hold

4. Calculate the relative transcript levels of lncRNAs and analyze the data according to manufacturer's protocol.

3.5 *In Situ* hybridization

3.5.1 *in vitro* transcription and probe shortening—1. Set up *in vitro* transcription reactions on ice as follows.

DNA template	5 µL (~0.5 µg)
H ₂ O	8 µL
10× NTP mix	2 µL
10× transcription buffer	2 µL
SP6 (or T7) polymerase	2 µL
RNase inhibitor	1 µL

2. Mix gently, centrifuge briefly, and incubate at 37 °C, 2 hrs.

3. Check 1 µl reaction on 1% agarose gel.

4. Purify the RNA using the Qiagen RNeasy MinElute Cleanup Kit. Elute with 50 µl of RNase free water.

5. Add 50 µl 200 mM carbonate buffer pH 10.2 and incubate at 60 °C for calculated times to prepare particular sizes of probe (Time = [size (kb) of the probe – 0.15] / [0.11 × size (kb) of the probe × 0.15]) .

6. Transfer tubes to ice and terminate probe hydrolysis reaction by adding 10 µl 10% acetic acid, and 12 µl 3 M NaOAc. (Gas bubbles will appear.)

7. Add 312 µl ethanol and incubate at –20 °C for 60 min.

8. Spin down at 4 °C for 10 min and wash the pellet with 100% ethanol.

9. Air-dry the pellet about 10 min and resuspend in 50 μ l DEPC-water. This solution can be stored at -20°C .

3.5.2 Hybridization—1. Prepare hybridization buffer by mixing the following (for 12 slides)

100 μ l	3 M NaCl, 0.1 M Tris-HCl pH 6.8, 0.1 M NaPO ₄ -buffer, 50 mM EDTA
400 μ l	Deionized formamide
200 μ l	50% Dextran sulfate
10 μ l	100 mg/ml tRNA
20 μ l	50 $\times$ Denhardt's solution
70 μ l	H ₂ O

2. Add 16 μ l probe mix to 64 μ l hybridization buffer, leave in 60C water bath.

3. Spread 80 μ l of mix evenly on the slides. Cover with a clean cover-slip.

4. Incubate over-night at 50 $^{\circ}\text{C}$ in an air-tight, moist box lined with tissue soaked in 50% formamide + 2 $\times$ SSC (See Note 8). Seal the box with tape and cover with foil.

3.5.3 Washing—For washing, incubate in Washing buffer, NTE buffer, and PBS according to the following order and conditions.

1. Washing buffer	10–15 min	50 $^{\circ}\text{C}$
2. Washing buffer	30–60 min	50 $^{\circ}\text{C}$
3. Washing buffer	30–60 min	50 $^{\circ}\text{C}$
4. NTE	5 min	37 $^{\circ}\text{C}$ with slowly shaking in incubator
5. NTE	5 min	37 $^{\circ}\text{C}$ with slowly shaking in incubator
6. RNase A	30 min	37 $^{\circ}\text{C}$ with slowly shaking in incubator
7. NTE	5 min	room temperature shaking
8. NTE	5 min	room temperature shaking
9. Washing buffer	30–60 min	50 $^{\circ}\text{C}$
10. PBS	5 min	room temperature shaking

3.5.4 Detection—Detection can be performed under the following order. Each incubation can be done in a small tray containing 100 ml of buffer at RT with gentle shaking. For color reaction, add 1.5 μ l of NBT and BCIP per ml of 10% polyvinylalcohol prepared in Detection buffer 2.

1. Detection Buffer	5 min
2. Detection Buffer 1 with 0.5% blocking Reagent	60 min

3. Detection Buffer 1 with 1% BSA, 0.3% Triton	30–60 min
4. Anti-DIG-AP (diluted 1:3K in Buffer 1 with 1% BSA, 0.3% Triton)	60 min
5. Detection Buffer 1 with 0.3% Triton	10–20 min
6. Detection Buffer 1 with 0.3% Triton	10–20 min
7. Detection Buffer 1 with 0.3% Triton	10–20 min
8. Detection Buffer 1 with 0.3% Triton	10–20 min
9. Detection Buffer	5 min
10. Detection Buffer	5 min
11. Color reaction (under darkness for overnight up to 3 days at room temperature)	1–3 days

4. Notes

1. To execute this script, you need to create the blastable database of transposable elements from fasta file including transposable element sequence information. (<http://maizetdb.org/~maize/>). Use makeblastdb command to create blastable database from fasta file from http://maizetdb.org/~maize/TE_12-Feb-2015_15–35.fa. And all database should be the same directory.

(e.g. \$ makeblastdb -in TE.fasta -dbtype nucl)

2. In many cases, normalizing the values improves the clustering since all FPKM values are weighted equally by shifting and normalizing. This ensures that individual expression values are from the shape of their expression patterns, not from their absolute levels.

3. The parameter below was used for SOM analysis as follow. For coding gene clusters, the generated clusters were tested by GO enrichment analysis to adjust a parameter to get a better fit to the researched model.

Parameter	
output stub	output name
cluster range	cluster range you predict
seed range	42
iterations	4200000
cluster by	rows
som rows	0
som cols	0
initialization	Random_Datapoints
neighborhood	Bubble
alpha initial	0.1
alpha final	0.005
sigma initial	5.0

sigma final 0.5

4. TOMTOM program allows you to identify shared or unique *cis*-regulatory elements between each cluster.
5. Grind tissue using pellet pestle in e-tube because the amount of tissue is about 20–30ul. If you grind the tissue using mortar and pestle, you will lose the most of sample.
6. Ensure that all work surfaces pipettes, and reagents needed for isolation and preparation of first-strand cDNA synthesis are free of RNase contamination by wiping down the work surfaces and pipettes with a nucleic acid decontamination detergent before starting work.
7. *ZmTXN* was used as an internal reference gene because it shows the least level of variation among different tissues. And all primer pairs need to be tested for the specificity and efficiency by visualizing bands on 3% agarose gel and blast primer sequences back to Maize genome.
8. 50% formamide + 2× SSC can be replaced with 1× SSPE

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

B.-H Kang and S. Sung is supported by USDA NIFA Award AFRI grant (2011–67013-30119). B.-H. Kang is also supported by the grants from the Research Grants Council (RGC) of Hong Kong (GRF14126116, C4011–14R, and AoE/M-05/12), Cooperative Research Program for Agriculture Science and Technology Development (Project No. 10953092018) Rural Development Administration, Republic of Korea, and S. Sung is also supported by NIH (GM100108) and NSF (1656764).

References

1. Fatica A, Bozzoni I (2014) Long non-coding RNAs: new players in cell differentiation and development. *Nat Rev Genet* 15 (1):7–21. doi:10.1038/nrg3606 [PubMed: 24296535]
2. Kim DH, Sung S (2017) Vernalization-Triggered Intragenic Chromatin Loop Formation by Long Noncoding RNAs. *Dev Cell* 40 (3):302–312 e304. doi:10.1016/j.devcel.2016.12.021 [PubMed: 28132848]
3. Heo JB, Sung S (2011) Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science* 331 (6013):76–79. doi:science.1197349[pil]10.1126/science.1197349 [PubMed: 21127216]
4. Quinn JJ, Chang HY (2016) Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet* 17 (1):47–62. doi:10.1038/nrg.2015.10 [PubMed: 26666209]
5. Chu C, Zhang QC, da Rocha ST, Flynn RA, Bharadwaj M, Calabrese JM, Magnuson T, Heard E, Chang HY (2015) Systematic discovery of Xist RNA binding proteins. *Cell* 161 (2):404–416. doi:10.1016/j.cell.2015.03.025 [PubMed: 25843628]
6. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, Lagarde J, Veeravalli L, Ruan X, Ruan Y, Lassmann T, Carninci P, Brown JB, Lipovich L, Gonzalez JM, Thomas M, Davis CA, Shiekhattar R, Gingeras TR, Hubbard TJ, Notredame C, Harrow J, Guigo R (2012) The GENCODE v7 catalog of human long noncoding

- RNAs: analysis of their gene structure, evolution, and expression. *Genome Res* 22 (9):1775–1789. doi:10.1101/gr.132159.111 [PubMed: 22955988]
7. Rutenberg-Schoenberg M, Sexton AN, Simon MD (2016) The Properties of Long Noncoding RNAs That Regulate Chromatin. *Annu Rev Genomics Hum Genet* 17:69–94. doi:10.1146/annurev-genom-090314-024939 [PubMed: 27147088]
 8. Goff LA, Rinn JL (2015) Linking RNA biology to lncRNAs. *Genome Res* 25 (10):1456–1465. doi:10.1101/gr.191122.115 [PubMed: 26430155]
 9. Zhang S, Thakare D, Yadegari R (2018) Laser-Capture Microdissection of Maize Kernel Compartments for RNA-Seq-Based Expression Analysis. *Methods in molecular biology* 1676:153–163. doi:10.1007/978-1-4939-7315-6_9 [PubMed: 28986909]
 10. Kim ED, Xiong Y, Pyo Y, Kim DH, Kang BH, Sung S (2017) Spatio-temporal analysis of coding and long noncoding transcripts during maize endosperm development. *Sci Rep* 7 (1):3838. doi:10.1038/s41598-017-03878-4 [PubMed: 28630499]
 11. Kiyota E, Pena IA, Arruda P (2015) The saccharopine pathway in seed development and stress response of maize. *Plant Cell Environ*. doi:10.1111/pce.12563
 12. Zhang Z, Yang J, Wu Y (2015) Transcriptional Regulation of Zein Gene Expression in Maize through the Additive and Synergistic Action of opaque2, Prolamine-Box Binding Factor, and O2 Heterodimerizing Proteins. *Plant Cell* 27 (4):1162–1172. doi:10.1105/tpc.15.00035 [PubMed: 25901087]
 13. Zhan J, Thakare D, Ma C, Lloyd A, Nixon NM, Arakaki AM, Burnett WJ, Logan KO, Wang D, Wang X, Drews GN, Yadegari R (2015) RNA sequencing of laser-capture microdissected compartments of the maize kernel identifies regulatory modules associated with endosperm cell differentiation. *Plant Cell* 27 (3):513–531. doi:10.1105/tpc.114.135657 [PubMed: 25783031]
 14. Zhang W, Yan H, Chen W, Liu J, Jiang C, Jiang H, Zhu S, Cheng B (2014) Genome-wide identification and characterization of maize expansin genes expressed in endosperm. *Mol Genet Genomics* 289 (6):1061–1074. doi:10.1007/s00438-014-0867-8 [PubMed: 25213600]
 15. Xiong Y, Mei W, Kim ED, Mukherjee K, Hassanein H, Barbazuk WB, Sung S, Kolaczowski B, Kang BH (2014) Adaptive expansion of the maize maternally expressed gene (Meg) family involves changes in expression patterns and protein secondary structures of its members. *BMC Plant Biol* 14:204. doi:10.1186/s12870-014-0204-8 [PubMed: 25084677]
 16. Xiong YQ, Li QB, Kang BH, Chourey PS (2011) Discovery of Genes Expressed in Basal Endosperm Transfer Cells in Maize Using 454 Transcriptome Sequencing. *Plant Molecular Biology Reporter* 29 (4):835–847. doi:10.1007/s11105-011-0291-8
 17. Boerner S, McGinnis KM (2012) Computational identification and functional predictions of long noncoding RNA in *Zea mays*. *PLoS One* 7 (8):e43047. doi:10.1371/journal.pone.0043047PONE-D-12-12079[pii] [PubMed: 22916204]
 18. Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings / International Conference on Intelligent Systems for Molecular Biology ; ISMB International Conference on Intelligent Systems for Molecular Biology* 2:28–36
 19. Gupta M (2007) Generalized hierarchical markov models for the discovery of length-constrained sequence features from genome tiling arrays. *Biometrics* 63 (3):797–805. doi:10.1111/j.1541-0420.2007.00760.x [PubMed: 17825011]

Table 1.

Bioinformatics Tools and Databases

Resource	URL
Primer3	http://frodo.wi.mit.edu/primer3/
NCBI Blast	http://blast.ncbi.nlm.nih.gov
MEME	http://meme-suite.org/
Maize Genetics and Genomics Database	https://www.maizegdb.org/
Maize transposable element (TE) database	http://maizetdb.org/~maize/
Transdecoder program	https://github.com/TransDecoder/TransDecoder/wiki
Biopython	http://biopython.org/
Blast software and database	https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastDocs&DOC_TYPE=Download
Genepattern	https://genepattern.broadinstitute.org/gp/pages/login.jsf
Bedtools	http://bedtools.readthedocs.io/en/latest/index.html