

Specifying an identified lncRNAs modulation in cellular processes during early embryonic development of *Nicotiana tabacum* seedlings

Elena Foust, Heather Garrett, and Dr. Michelle M. Barthet
Department of Biology
Coastal Carolina University

Abstract

Multiple forms of RNA exist in cells. The most common forms of RNA include tRNA, rRNA, and mRNA. LncRNAs are proposed to have multiple functions during seed development including gene silencing, organogenesis, and other molecular processes. LncRNAs are unknown or poorly understood because their importance has been questioned throughout early research. However, the lncRNA *NTAB_LNC014148.1* is presumed to function in seed development of *Nicotiana tabacum* based on global gene expression studies. Bioinformatics analyses predicted that lncRNA *NTAB_LNC014148.1* binds to at least two gene regions during early embryonic development (gene regions *Nitab4.5_0001672g0150.1* and *Nitab4.5_0001972g0020.1*). These two gene regions encode an Arf GTPase activating protein and oxidoreductase. ADP-ribosylation factor (Arf) is a type of small GTPase that acts as a carrier for vesicles during active transport and transports respective proteins through the Golgi complex. NADPH oxidoreductase are involved in the conduction of electron flow and generation, or prevention of reactive oxygen species (ROS) involved in cell signaling, cell death, and other aspects of cell development. Both the Arf GTPase and NADPH oxidoreductase are crucial proteins for seed germination. However, the exact role of lncRNA *NTAB_LNC014148.1* in seed development is unknown. We aim to discern specifically where and when lncRNA *NTAB_LNC014148.1* is expressed in *N. tabacum* seeds using RT-PCR as a first step in the characterization of this particular lncRNA. DNA: RNA interaction assays will be utilized to confirm interaction among lncRNA and predicted gene regions to further characterize lncRNA cellular function.

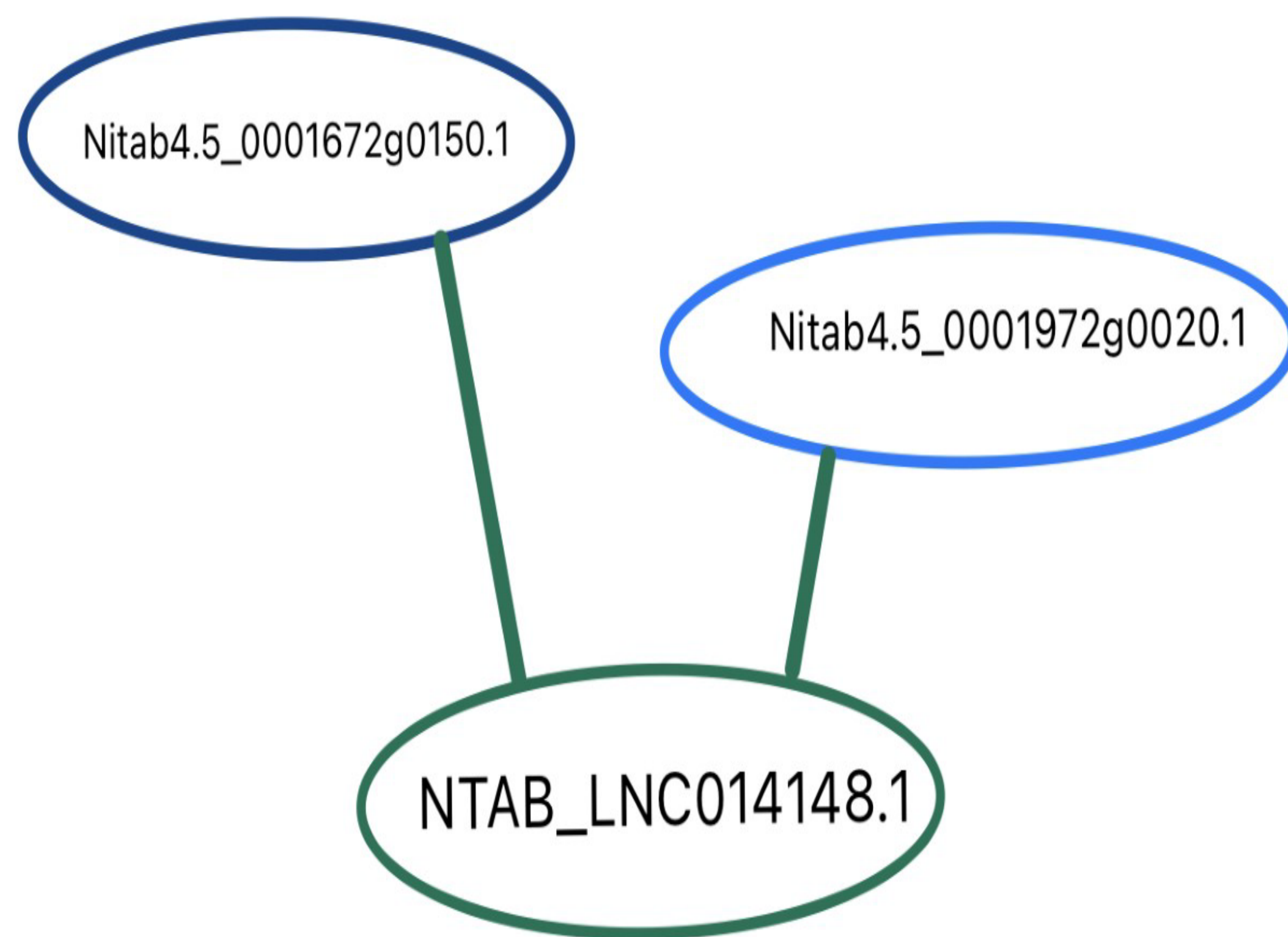


Figure 1. Bioinformatic knowledge map of the predicted genes associated with *NTAB_LNC014148.1*

Hypothesis

LncRNA (*NTAB_LNC014148.1*) binds to gene (*Nitab4.5_0001972g0020.1*) and is expressed in the early embryonic development of *Nicotiana tabacum* seedlings.

Methods

Bioinformatics analysis:

- Selected an intergenic lncRNA that is predicted to have an association with the genome of *Nicotiana tabacum* based on a plant-long non-coding RNA database[1]
- Analyzed a bioinformatics knowledge map that presented genes associated with the lncRNA that are suspected to have early embryonic development functionality[1]
- The (GOI) was selected because it was shown to have an association with the lncRNA of interest

Gene of Interest (GOI) : *Nitab4.5_0001972g0020.1*
lncRNA of interest : *NTAB_LNC014148.1*

Primer design:

- The (GOI) entire sequence was run against the lncRNA of interest in the clustal omega software[2]
- The local alignment tool WATER (EMBOSS) was used to determine where the lncRNA did not bind to the gene sequence[2]
- Analyzed the pairwise sequencing output and identified regions within the sequence where the lncRNA did not bind to the (GOI)
- Each applicable sequence where lncRNA did not bind was run against the whole genome shotgun sequence of *Nicotiana tabacum*[3]
- Primers for the (GOI) were designed by selecting a forward and a reverse strand in a 1,000 base-pair (bp) region

What are primers? Primers are short pieces of single-stranded DNA complementary to the target sequence. This helps to ensure specificity of the amplification reaction.
WATER (EMBOSS): pairwise sequence tool for local alignment

Forward:

5'GGCAGCTCTGCTGATGGAG

TTACATGCCGAATGATCTA 3'

3' AATGTACCGTTACTAGAT 5'

Reverse: 5'TAGATCATTGCCCATGTAA 3'

Figure 2. Forward and reverse primer sequences for the gene *Nitab4.5_0001972g0020.1*

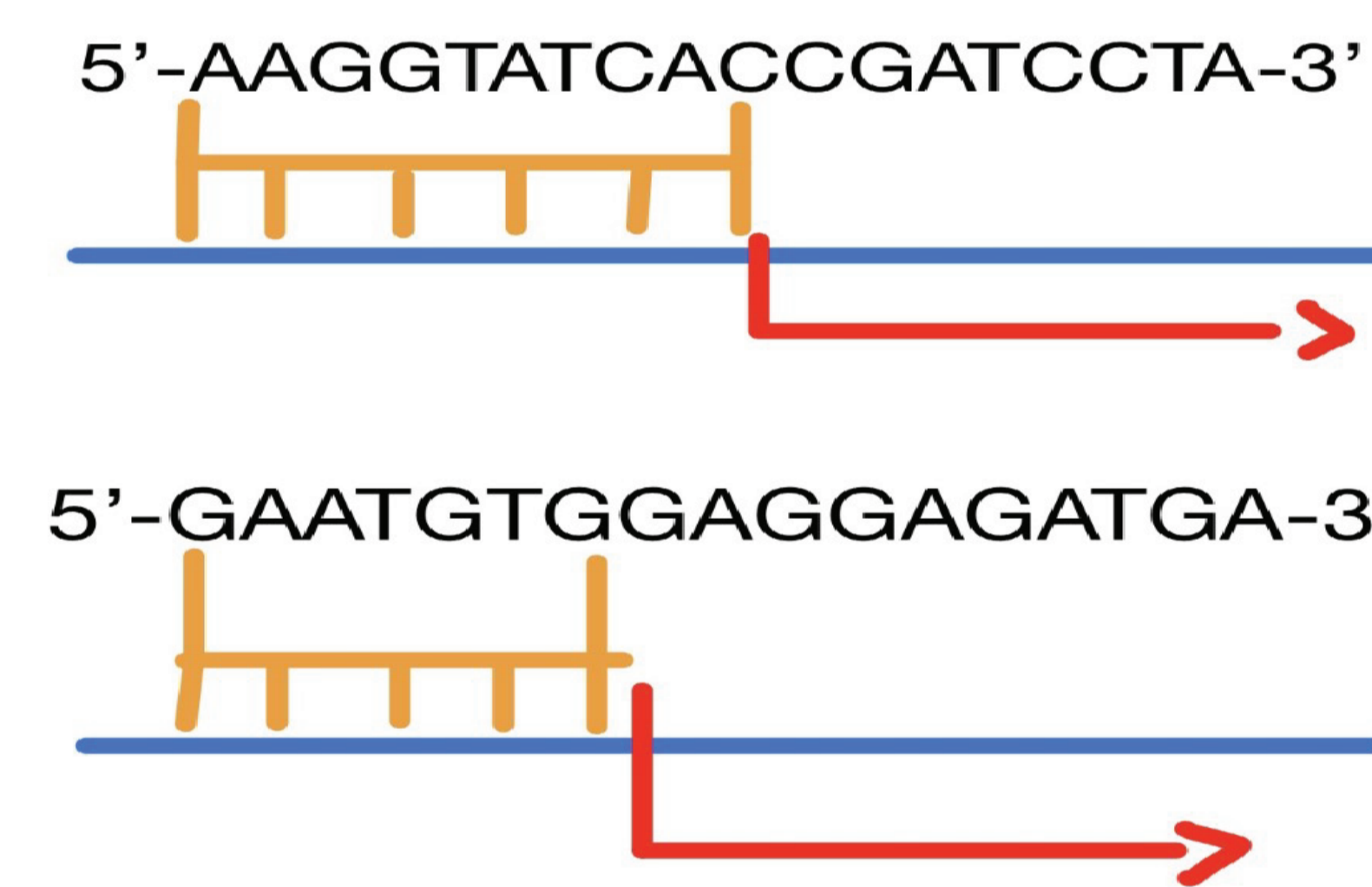


Figure 3. lncRNA primer sequences from two regions our lncRNA of interest *NTAB_LNC014148.1* does not bind to the gene sequence of *Nitab4.5_0001972g0020.1*

Plating/Tissue collection:

- Seeds were sterilized and stratified onto 3% agar plates
- Seeds were placed in the cold room for 4 days before plating
- Early embryonic tissue was collected 2-days and 4-days post-germination

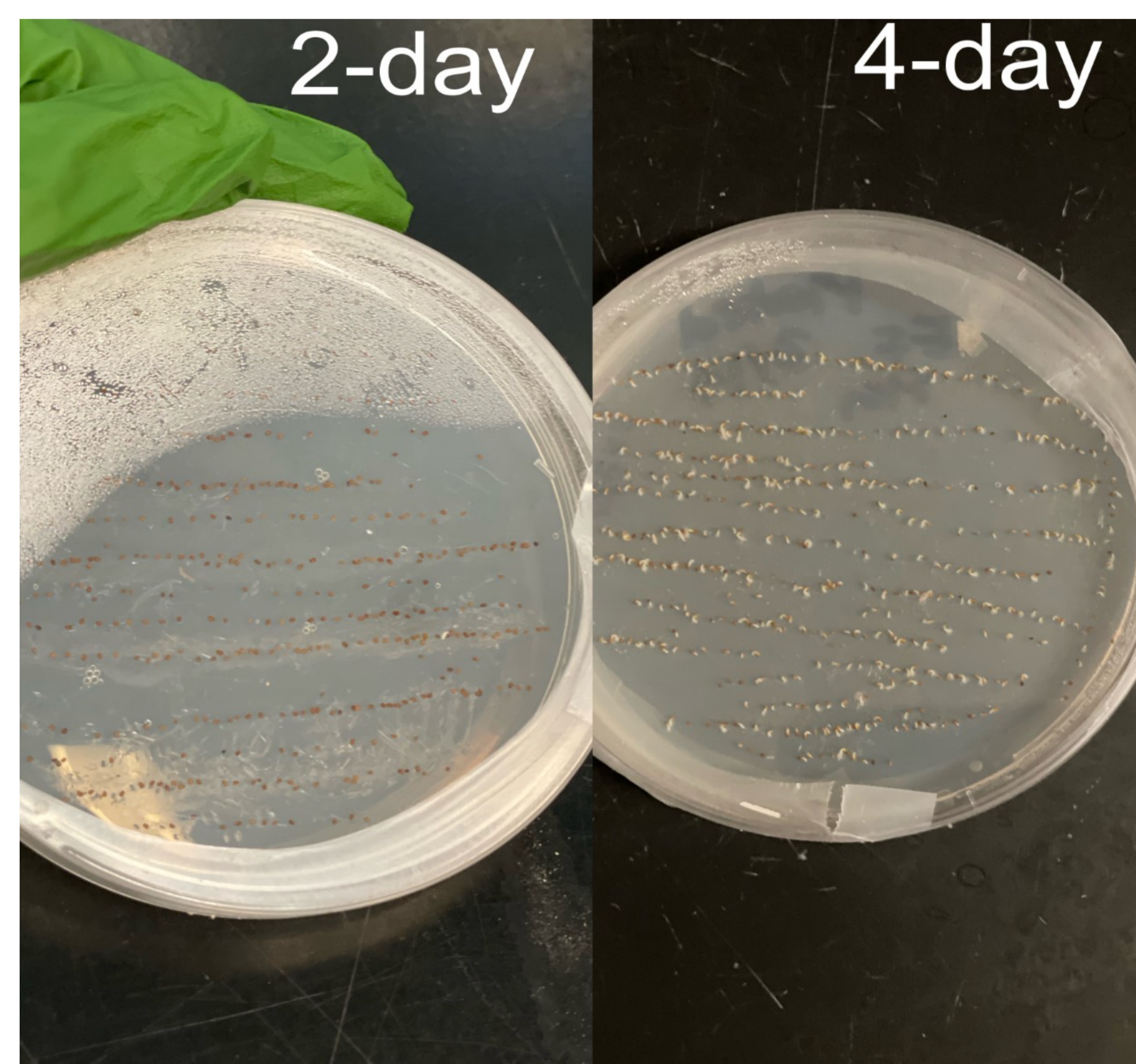


Figure 4. Day 2 and 4 embryonic tissue development in *Nicotiana Tabacum* seedlings

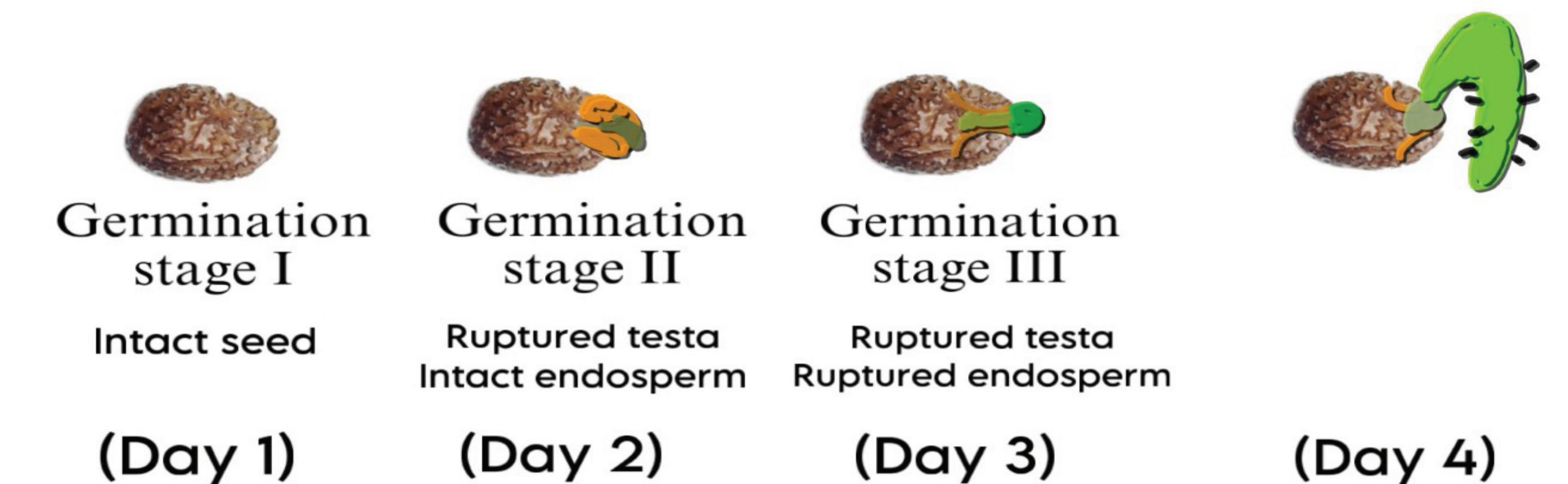


Figure 5. Day 1-4 germination stages and embryonic tissue development in *Nicotiana tabacum* seedlings

Expected Results:

After using RNase H and sequencing techniques, the original RNA will be present. This will support that the RNA of interest is bound to DNA.

WHEN is this lncRNA expressed?

Future Direction:

RNA extraction → RT-PCR → Gel electrophoresis → RNase protection assay with amplified target DNA → Validate the functionality of lncRNA from data other than RNA-Seq data → Publication

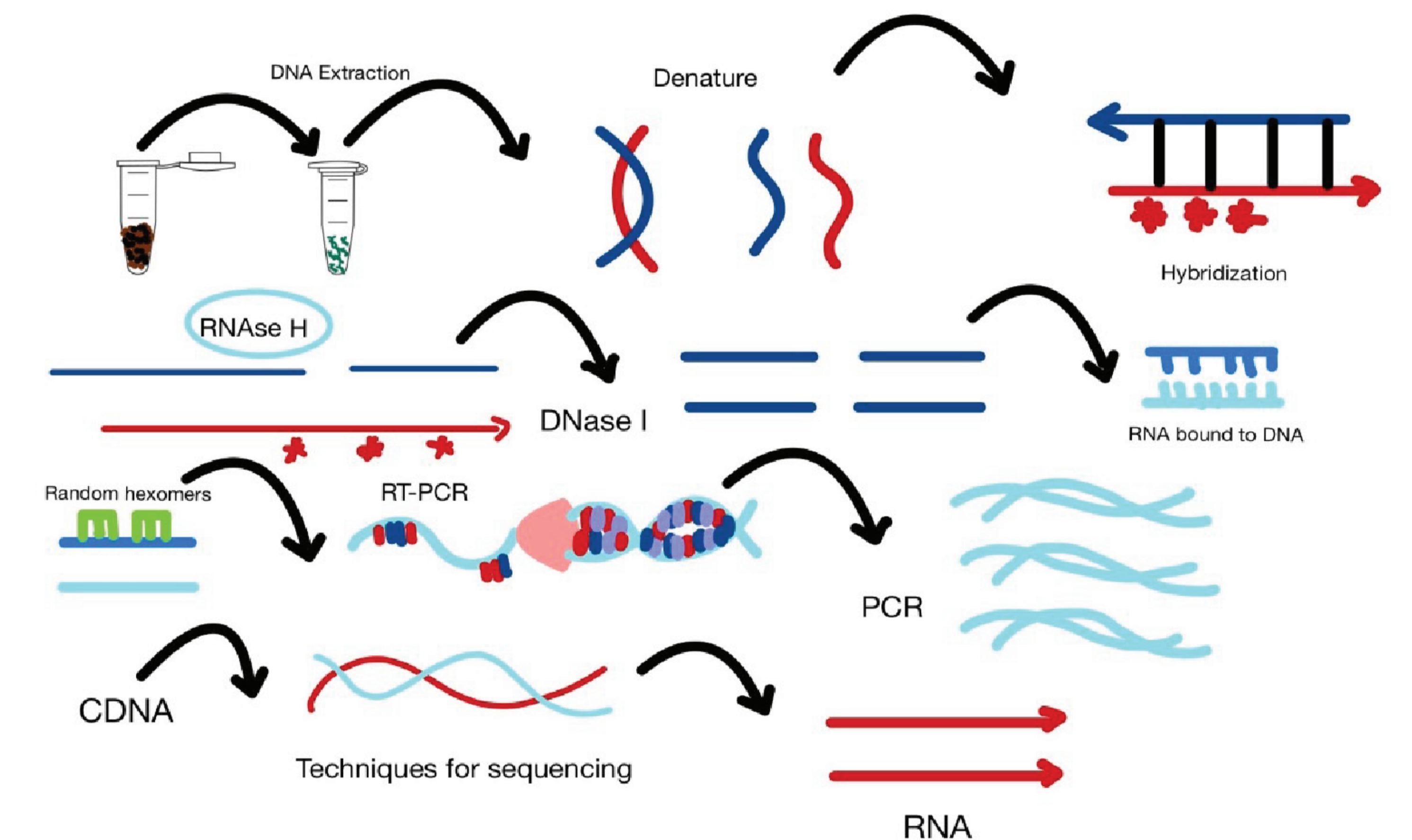


Figure 6. Protection assay to determine if the RNA binds to DNA

KEYWORDS:

Gene of interest (GOI), long non-coding RNA (lncRNA), Intergenic, Smith-Waterman algorithm, Basic Local Alignment Search Tool (BLAST), Vernalization, Reverse transcription polymerase chain reaction (RT-PCR), RNA-seq, Gel electrophoresis, Protection assay, RNase enzyme

Acknowledgments:

We would like to thank Coastal Carolina University, Department of Biology for funding this research project. Furthermore, we would like to thank Dr. Barthet for her contribution towards methodology and techniques specific to this project, as well as her guidance and support throughout the semester.

For further information:

Elena Foust (enfoust@coastal.edu)

Heather Garrett (hgarrett@coastal.edu)

Literature Cited:

- Jin JJ, Liu J, Wang H, Wong L, Chua NH. PLncDB: plant long non-coding RNA database. *Bioinformatics* 2013 APR 15; 29(8):1068-1071.
- Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Sievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins D. *Molecular Systems Biology* 7 Article number: 539 doi:10.1038/msb.2011.75
- Camacho C. et al. (2009) "BLAST+: architecture and applications"

Abstract

Multiple forms of RNA exist in cells. The most common forms of RNA include tRNA, rRNA, and mRNA. LncRNAs are proposed to have multiple functions during seed development including gene silencing, organogenesis, and other molecular processes. LncRNAs are unknown or poorly understood because their importance has been questioned throughout early research. However, the lncRNA *NTAB_LNC014148.1* is presumed to function in seed development of *Nicotiana tabacum* based on global gene expression studies. Bioinformatics analyses predicted that lncRNA *NTAB_LNC014148.1* binds to at least two gene regions during early embryonic development (gene regions *Nitab4.5_0001672g0150.1* and *Nitab4.5_0001972g0020.1*). These two gene regions encode an Arf GTPase activating protein and oxidoreductase. ADP-ribosylation factor (Arf) is a type of small GTPase that acts as a carrier for vesicles during active transport and transports respective proteins through the Golgi complex. NADPH oxidoreductase are involved in the conduction of electron flow and generation, or prevention of reactive oxygen species (ROS) involved in cell signaling, cell death, and other aspects of cell development. Both the Arf GTPase and NADPH oxidoreductase are crucial proteins for seed germination. However, the exact role of lncRNA *NTAB_LNC014148.1* in seed development is unknown. We aim to discern specifically where and when lncRNA *NTAB_LNC014148.1* is expressed in *N. tabacum* seeds using RT-PCR as a first step in the characterization of this particular lncRNA. DNA: RNA interaction assays will be utilized to confirm interaction among lncRNA and predicted gene regions to further characterize lncRNA cellular function.

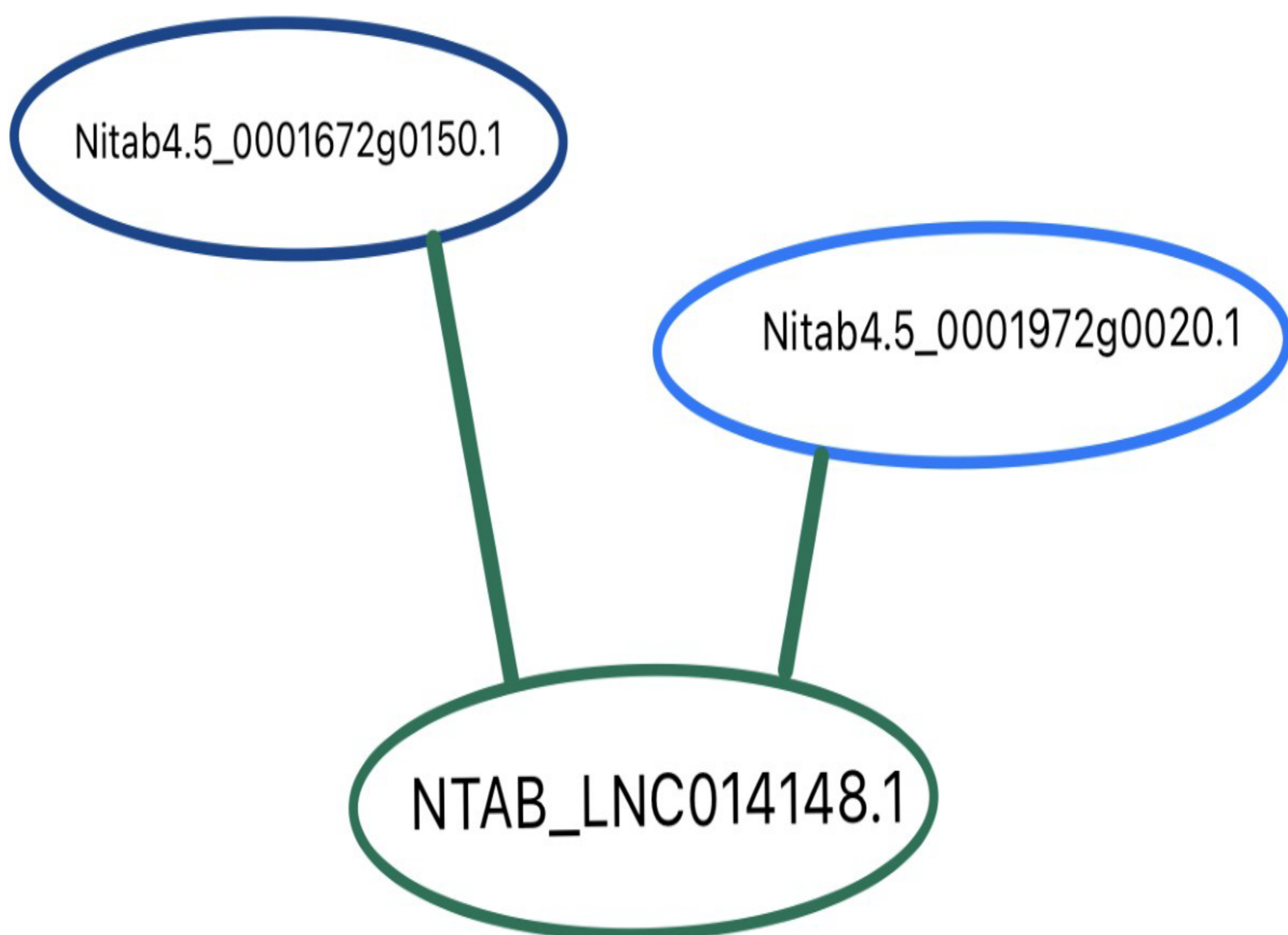


Figure 1. Bioinformatic knowledge map of the predicted genes associated with *NTAB_LNC014148.1*

Hypothesis

LncRNA (*NTAB_LNC014148.1*) binds to gene (*Nitab4.5_0001972g0020.1*) and is expressed in the early embryonic development of *Nicotiana tabacum* seedlings.

Methods

Bioinformatics analysis:

- Selected an intergenic lncRNA that is predicted to have an association with the genome of *Nicotiana tabacum* based on a plant-long non-coding RNA database[1]
- Analyzed a bioinformatics knowledge map that presented genes associated with the lncRNA that are suspected to have early embryonic development functionality[1]
- The (GOI) was selected because it was shown to have an association with the lncRNA of interest

Gene of Interest (GOI) : *Nitab4.5_0001972g0020.1*
LncRNA of interest : *NTAB_LNC014148.1*

Abstract

Multiple forms of RNA exist in cells. The most common forms of RNA include tRNA, rRNA, and mRNA. LncRNAs are proposed to have multiple functions during seed development including gene silencing, organogenesis, and other molecular processes. LncRNAs are unknown or poorly understood because their importance has been questioned throughout early research. However, the lncRNA *NTAB_LNC014148.1* is presumed to function in seed development of *Nicotiana tabacum* based on global gene expression studies. Bioinformatics analyses predicted that lncRNA *NTAB_LNC014148.1* binds to at least two gene regions during early embryonic development (gene regions *Nitab4.5_0001672g0150.1* and *Nitab4.5_0001972g0020.1*). These two gene regions encode an Arf GTPase activating protein and oxidoreductase. ADP-ribosylation factor (Arf) is a type of small GTPase that acts as a carrier for vesicles during active transport and transports respective proteins through the Golgi complex. NADPH oxidoreductase are involved in the conduction of electron flow and generation, or prevention of reactive oxygen species (ROS) involved in cell signaling, cell death, and other aspects of cell development. Both the Arf GTPase and NADPH oxidoreductase are crucial proteins for seed germination. However, the exact role of lncRNA *NTAB_LNC014148.1* in seed development is unknown. We aim to discern specifically where and when lncRNA *NTAB_LNC014148.1* is expressed in *N. tabacum* seeds using RT-PCR as a first step in the characterization of this particular lncRNA. DNA: RNA interaction assays will be utilized to confirm interaction among lncRNA and predicted gene regions to further characterize lncRNA cellular function.

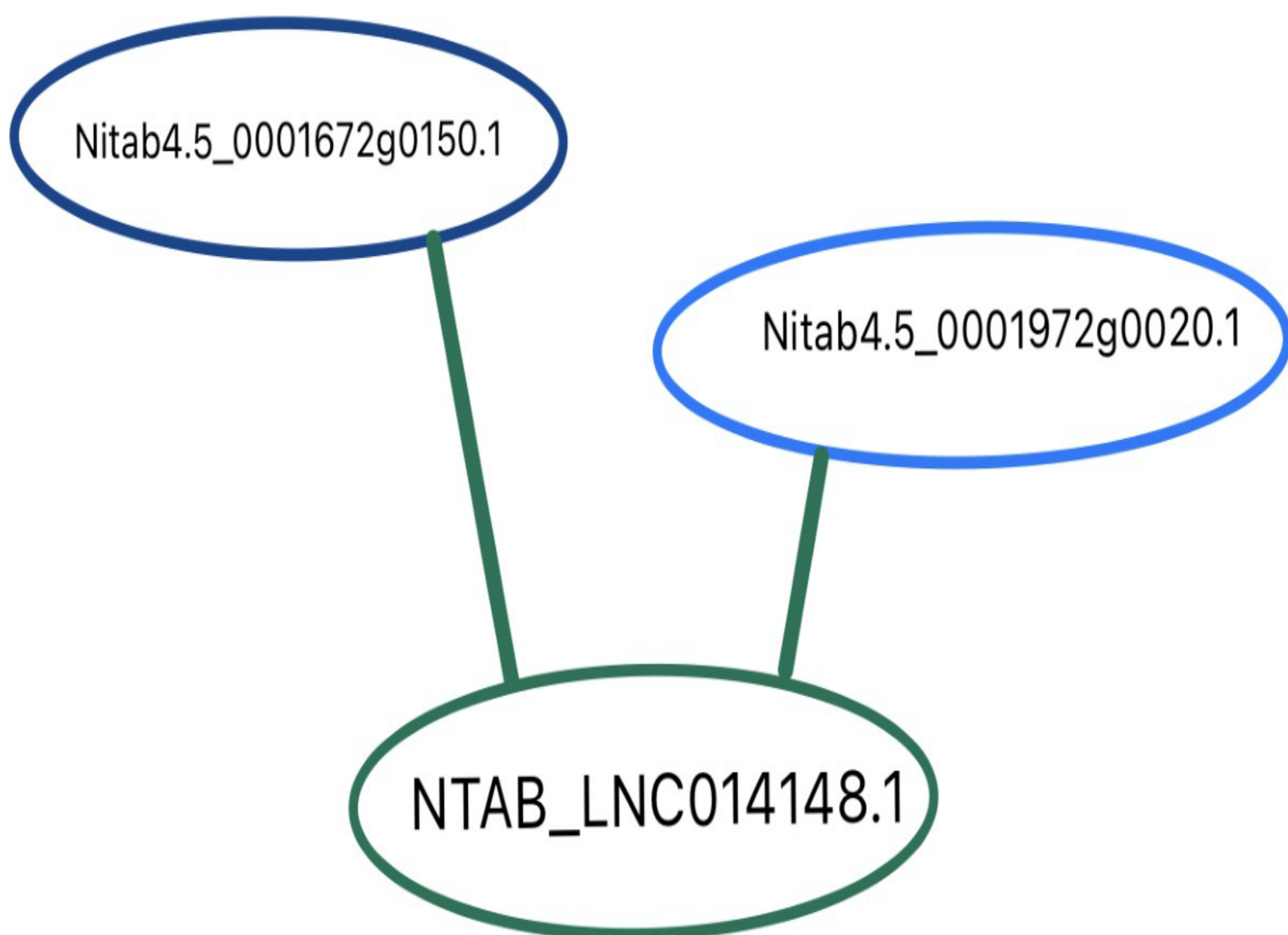


Figure 1. Bioinformatic knowledge map of the predicted genes associated with *NTAB_LNC014148.1*

Hypothesis

lncRNA (*NTAB_LNC014148.1*) binds to gene (*Nitab4.5_0001972g0020.1*) and is expressed in the early embryonic development of *Nicotiana tabacum* seedlings.

Methods

Bioinformatics analysis:

- Selected an intergenic lncRNA that is predicted to have an association with the genome of *Nicotiana tabacum* based on a plant-long non-coding RNA database[1]
- Analyzed a bioinformatics knowledge map that presented genes associated with the lncRNA that are suspected to have early embryonic development functionality[1]
- The (GOI) was selected because it was shown to have an association with the lncRNA of interest

Gene of Interest (GOI) : *Nitab4.5_0001972g0020.1*
LncRNA of interest : *NTAB_LNC014148.1*

Abstract

Multiple forms of RNA exist in cells. The most common forms of RNA include tRNA, rRNA, and mRNA. LncRNAs are proposed to have multiple functions during seed development including gene silencing, organogenesis, and other molecular processes. LncRNAs are unknown or poorly understood because their importance has been questioned throughout early research. However, the lncRNA *NTAB_LNC014148.1* is presumed to function in seed development of *Nicotiana tabacum* based on global gene expression studies. Bioinformatics analyses predicted that lncRNA *NTAB_LNC014148.1* binds to at least two gene regions during early embryonic development (gene regions *Nitab4.5_0001672g0150.1* and *Nitab4.5_0001972g0020.1*). These two gene regions encode an Arf GTPase activating protein and oxidoreductase. ADP-ribosylation factor (Arf) is a type of small GTPase that acts as a carrier for vesicles during active transport and transports respective proteins through the Golgi complex. NADPH oxidoreductase are involved in the conduction of electron flow and generation, or prevention of reactive oxygen species (ROS) involved in cell signaling, cell death, and other aspects of cell development. Both the Arf GTPase and NADPH oxidoreductase are crucial proteins for seed germination. However, the exact role of lncRNA *NTAB_LNC014148.1* in seed development is unknown. We aim to discern specifically where and when lncRNA *NTAB_LNC014148.1* is expressed in *N. tabacum* seeds using RT-PCR as a first step in the characterization of this particular lncRNA. DNA: RNA interaction assays will be utilized to confirm interaction among lncRNA and predicted gene regions to further characterize lncRNA cellular function.

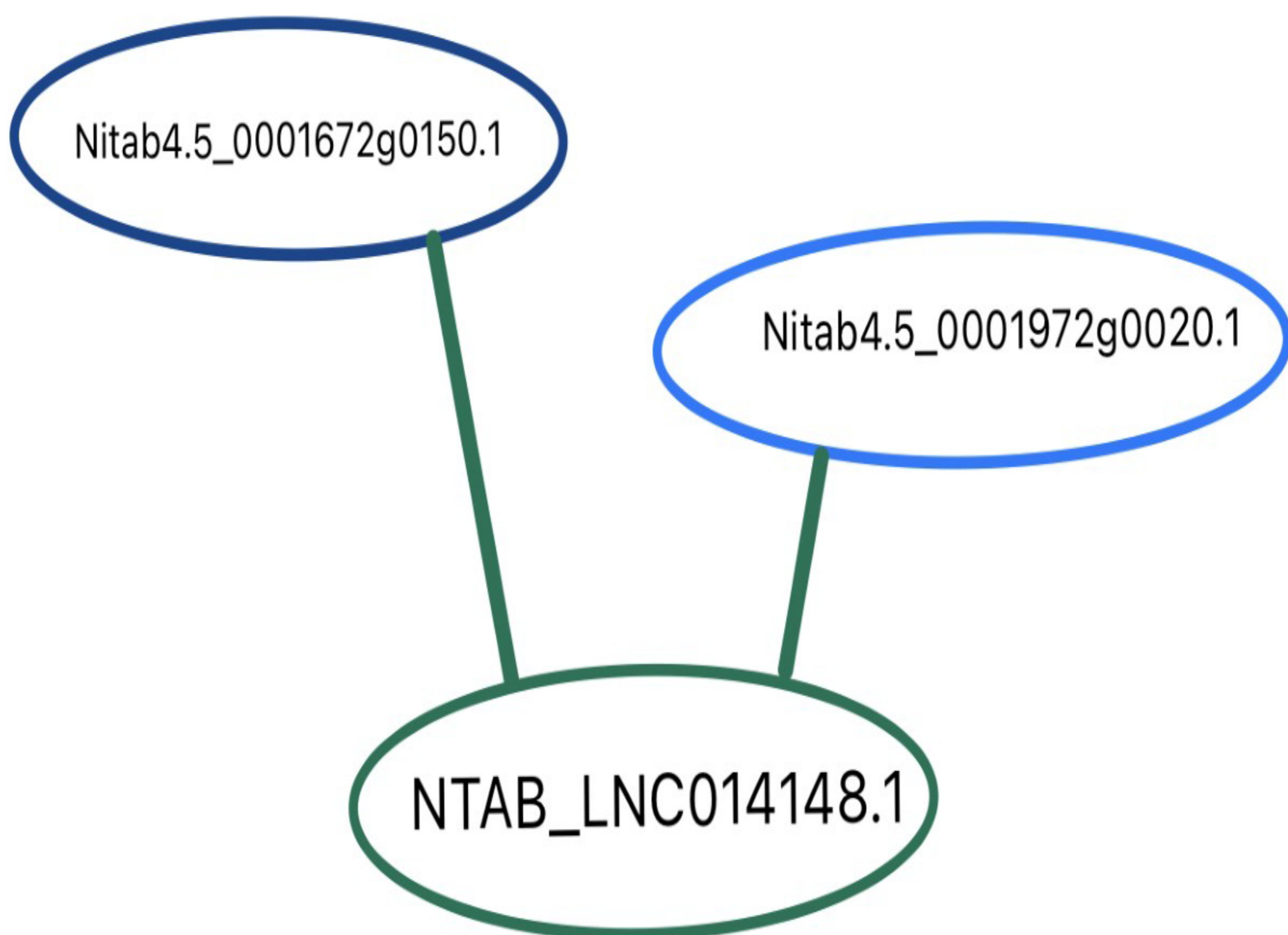


Figure 1. Bioinformatic knowledge map of the predicted genes associated with *NTAB_LNC014148.1*

Hypothesis

LncRNA (*NTAB_LNC014148.1*) binds to gene (*Nitab4.5_0001972g0020.1*) and is expressed in the early embryonic development of *Nicotiana tabacum* seedlings.

Methods

Bioinformatics analysis:

- Selected an intergenic lncRNA that is predicted to have an association with the genome of *Nicotiana tabacum* based on a plant-long non-coding RNA database[1]
- Analyzed a bioinformatics knowledge map that presented genes associated with the lncRNA that are suspected to have early embryonic development functionality[1]
- The (GOI) was selected because it was shown to have an association with the lncRNA of interest

Gene of Interest (GOI) : *Nitab4.5_0001972g0020.1*
LncRNA of interest : *NTAB_LNC014148.1*