Looping by RNA: Dynamic control of chromatin loop by long non-coding RNA in plants

Junghyun Kim and Sibum Sung

Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX 78712, USA

Correspondence should be addressed to Sibum Sung (sbsung@austin.utexas.edu)

In eukaryotes, the three-dimensional (3D) structure of genome plays a critical role in the transcriptional regulation. Even though the 3D structure of genome is mainly determined by DNA and protein components, increasing evidence has demonstrated that RNA also occupies a large portion of chromatin and influences genomic architecture (Ramirez-Colmenero et al., 2020; Wierzbicki et al., 2021). Long non-coding RNA (lncRNA) is a class of non-coding RNAs longer than 200 nucleotides but lack protein-coding potential. LncRNAs are key regulators of gene expression through their capacity to harbor regulatory proteins and to establish molecular interactions. In animals, lncRNAs play roles in dynamic chromatin remodeling through their contribution to the formation of chromatin loop, the maintenance of heterochromatin, and the function of chromatin-modifying complexes and affect various biological processes, including cellular developments, immunity, and cancer (Ramirez-Colmenero et al., 2020). Similarly, lncRNAs in plants have been implicated to function in the gene regulation by mediating chromatin modification and remodeling. Here, we highlight a novel lncRNA that modulates the recruitment of transcription factors and affects chromatin structure to control transcription in response to long-term cold (known as vernalization), in wheat. We discuss this finding in the context of similar studies in plants and animals.

A recent work in hexaploidy wheat (*Triticum aestivum*) has identified a lncRNA, *VAS* (*TaVRN1 Alternative Splicing*), that regulates the transcription of *TaVRN1* in part by affecting the chromatin structure (Figure 1a) (Xu et al., 2021). TaVRN1 promotes floral transition in response to vernalization. The *VAS* is an alternatively spliced transcript of *TaVRN1*, which contains a part of the 5' UTR, the entire first exon and a part of the first intron. The VAS transcript is induced by cold exposure, with the maximal expression observed at 21 days of cold in winter wheat accessions. However, it is not detectable in spring wheat accessions even after the vernalization treatment, possibly due to the deletion of critical sequence in the first intron of *TaVRN1*. Cold-induced *VAS* expression triggers vernalization-mediated activation of *TaVRN1* by recruiting a heterodimer of basic leucine zipper (bZIP) transcription factors (TaRF2a-TaRF2b) to the *TaVRN1* promoter. Chromatin conformation capture (3C) assay

showed that a chromatin loop between the promoter region near the TaRF2b binding sequence motif (Sp1 motif) and the first intron region, which overlaps with the *VAS* sequence, is gradually reduced during vernalization. The chromatin isolation by RNA purification associated with mass spectrometry (ChIRP-MS) assay to uncover potential protein factors associated with *VAS* has identified 11 proteins, including TaRF2b. Therefore, it appears that *VAS* provides a molecular structure to facilitate the recruitment of protein complexes to their target loci. Candidate *VAS*-associated proteins also include splicing factor-like protein, sister chromatid cohesion protein and kinase family proteins, implying that *VAS* may be involved in various chromatin remodeling processes. However, mechanistic details on how *VAS* and candidate VAS-associated proteins contribute to chromatin remodeling, including chromatin loop formation, at the *TaVRN1* locus remain to be understood.

In *Arabidopsis*, there are at least two examples in which lncRNAs contribute to the chromatin structures by forming chromatin loops, *APOLO* (Ariel et al., 2014) and *COLDWRAP* (Kim and Sung, 2017). The *APOLO* is an intergenic lncRNA transcribed from 5 kb upstream of the gene *PINOID* (*PID*). *PID* encodes a Ser/Thr protein kinase responsible for the polar localization of auxin transporter. Auxin-induced *APOLO* expression controls *PID* expression by affecting a chromatin loop formation between the *PID* promoter and *APOLO* locus, in cooperation with a component of Polycomb Repressive Complex 1 (PRC1), LHP1 (Figure 1b). *APOLO* interacts with LHP1 and LHP1 is associated with Histone H3 lysine 27 trimethylation (H3K27me3)-marked chromatin. Therefore, the expression of *APOLO* triggers dynamic formation of chromatin loop by *APOLO*-LHP1 (Ariel *et al.*, 2014). Auxin-induced *APOLO* expression is achieved by decreases in the levels of DNA methylation at *APOLO* and *PID* loci by auxin, although how auxin triggers the DNA demethylation remains to be understood. Interestingly, *APOLO* also contributes to chromatin looping at multiple *trans* target loci by targeting not only to spatially associated loci but also to distal loci through the sequence complementarity and DNA-RNA duplexes (R-loops) formation (Ariel et al., 2020). The R-loops driven by *APOLO* decoy LHP1, resulting in subsequent dynamic regulation of chromatin loop at target loci. Another example is *COLDWRAP*, a lncRNA transcribed from the proximal

promoter region of *FLOWERING LOCUS C* (*FLC*), which is a major floral repressor in *Arabidopsis*. *COLDWRAP* interacts with a component of PRC2, CURLY LEAF (CLF) to form an intragenic chromatin loop between the promoter and 3' end of the first intron of *FLC* and confers repression of *FLC* expression by vernalization (Figure 1c) (Kim and Sung, 2017). In addition, another lncRNAs derived from the first intron of *FLC*, *COLDAIR* which is involved in the PRC2-mediated silencing of *FLC* by vernalization through its capacity in the recruitment of PRC2 to *FLC*, also contributes to the intragenic repressive chromatin loop accomplished by *COLDWRAP* and PRC2 (Figure 1c) (Kim and Sung, 2017).

Interestingly, both *APOLO* and *COLDWRAP* interact with proteins which belong to two major Polycomb group (PcG) repressive complex (PRCs; PRC1 and PRC2). Since PRCs are widely involved in the transcriptional repression by heterochromatin formation (Piunti and Shilatifard, 2021), those lncRNAs are expected to contribute to chromatin condensation at target loci by interacting with PcG proteins which affect heterochromatin formation. Indeed, in animals, many of lncRNAs have been identified to interact with PRCs including *Xist*, *Kcnq1ot1* and *HOTAIR* for their repressive chromatin loop formation (Ramirez-Colmenero *et al.*, 2020).

Studies using Hi-C, a genome-wide technique to determine chromatin interactions, have revealed that eukaryotic genome in the nucleus is highly organized. A basic unit of 3D organization of genome is "Topologically associating domains" (TADs) in many eukaryotes. TADs are compartments of chromatin folded into domains with preferential internal interactions among chromatin loops (Bonev and Cavalli, 2016). TADs are the architectural units of chromatin that define regulatory landscapes and the TAD-TAD boundaries contribute to the gene regulation by limiting interactions of *cis*-regulatory elements to their target genes. Consistent with the roles of lncRNAs in chromatin loop formation, there are examples of lncRNAs that contribute to the TAD formation. The ncRNAs, *Eleanor*, delineate TAD of the *ESR1* locus in the active nuclear compartment of LTED cells and activate the *ESR1* gene which encodes estrogen receptor-α (ER) (Abdalla et al., 2019). Another lncRNA, *Firre,* is located at the borders of TADs within CTCF binding regions and is necessary for super loop formation of inactive X-chromosome (Barutcu et al., 2018). Similarly, several studies defined

TADs in some plant species (Ouyang et al., 2020). *Arabidopsis* genome possesses relatively smaller range of chromatin loops but lacks apparent TADs. Interestingly, the repressive histone mark, H3K27me3 is favored for genes with promoter-promoter interactions over long ranges in *Arabidopsis* (Ouyang *et al.*, 2020), suggesting the role of chromatin modification in defining 3D chromatin structure in plants. However, our understanding of regulatory and structural elements controlling chromatin loops in plants is still at its infancy, although some chromatin regulators have been implicated in the formation of chromatin loop in *Arabidopsis* (Li et al., 2018; Zhao et al., 2021).

Increasing numbers of lncRNAs are being identified in plants, and their functional studies will undoubtfully unravel diverse functions of lncRNAs as regulatory elements. The identification of *VAS* in wheat and its role in the dynamic control of a chromatin loop by vernalization at *TaVRN1* locus provide an excellent example of lncRNAs that mediate changes in chromatin architecture although how this new lncRNA is directly involved in the chromatin loop remains to be elucidated. Given that lncRNAs can recruit, decoy, or scaffold protein complexes to mediate chromatin looping and the inter- or intra-chromosomal interactions, it is expected that more lncRNAs that participate in the regulation of gene expression and genomic stability will be uncovered in plants. Therefore, studying potential interactions between chromatin-remodeling factors and lncRNAs will be the promising first step to elucidate mechanisms underlying how lncRNAs contribute the 3D structure of plant genome.

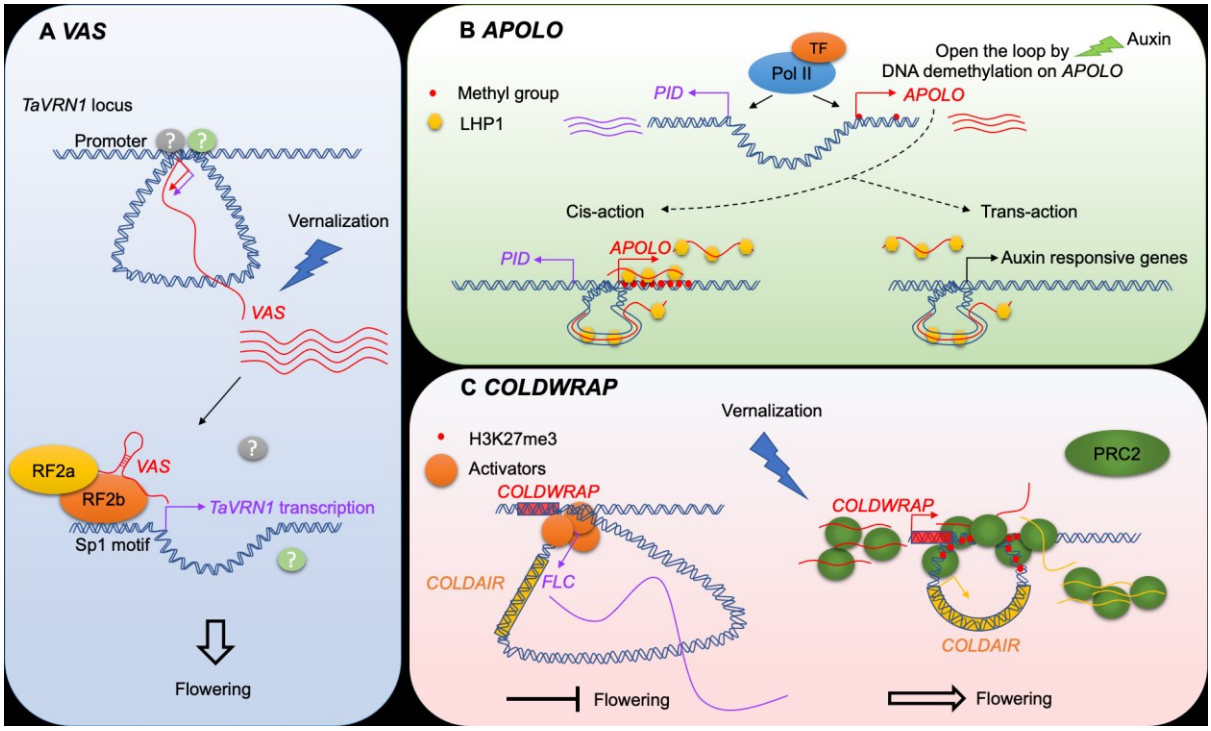

Figure 1. Dynamic chromatin loop formation by lncRNAs in plants.

(A) A chromatin loop formed at the *TaVRN1* locus in wheat. Upon vernalization, chromatin loop is lost, and more *VAS* transcripts are produced. Subsequently, *VAS* promotes the *TaVRN1* expression by recruiting the transcription factors, TaRF2b-TaRF2a, which bind to the Sp1 motif in the *TaVRN1* promoter to increase the transcription of *TaVRN1*. Activation of *TaVRN1* eventually promotes flowering in winter wheat.

(B) Chromatin loops mediated by *APOLO* in *Arabidopsis*. Auxin induces DNA demethylation at the *APOLO* locus and opens the loop encompassing the *PID* promoter region, resulting in increases in both *PID* and *APOLO* transcripts. *APOLO* lncRNA binds to its neighboring *PID* locus or distant loci, such as auxin responsive genes. *APOLO* decoys LHP1 and forms local chromatin loops to finetune the expression of target genes.

(C) An intragenic chromatin loop by *COLDWRAP* and PRC2 in *Arabidopsis*. *COLDAIR* and *COLDWRAP* are gradually induced by vernalization and supports PRC2 occupation to the promoter region to repress *FLC* expression by H3K27me3 enrichments.

FUNDING

We acknowledge supports from NIH (R01GM100108) and NSF (IOS1656764) to S. S.

ACKNOWLEDGEMENTS

No conflict of interest declared.

REFERENCES

Abdalla, M.O.A., Yamamoto, T., Maehara, K., Nogami, J., Ohkawa, Y., Miura, H., Poonperm, R., Hiratani, I., Nakayama, H., Nakao, M., et al. (2019). The Eleanor ncRNAs activate the topological domain of the ESR1 locus to balance against apoptosis. Nat Commun **10**:3778. 10.1038/s41467-019-11378-4.

Ariel, F., Jegu, T., Latrasse, D., Romero-Barrios, N., Christ, A., Benhamed, M., and Crespi, M. (2014). Noncoding transcription by alternative RNA polymerases dynamically regulates an auxin-driven chromatin loop. Mol Cell **55**:383-396. 10.1016/j.molcel.2014.06.011.

Ariel, F., Lucero, L., Christ, A., Mammarella, M.F., Jegu, T., Veluchamy, A., Mariappan, K., Latrasse, D., Blein, T., Liu, C., et al. (2020). R-Loop Mediated trans Action of the APOLO Long Noncoding RNA. Mol Cell **77**:1055-1065 e1054. 10.1016/j.molcel.2019.12.015.

Barutcu, A.R., Maass, P.G., Lewandowski, J.P., Weiner, C.L., and Rinn, J.L. (2018). A TAD boundary is preserved upon deletion of the CTCF-rich Firre locus. Nature Communications **9**ARTN 1444

10.1038/s41467-018-03614-0.

Bonev, B., and Cavalli, G. (2016). Organization and function of the 3D genome. Nat Rev Genet **17**:661-678. 10.1038/nrg.2016.112.

Kim, D.H., and Sung, S. (2017). Vernalization-Triggered Intragenic Chromatin Loop Formation by Long Noncoding RNAs. Dev Cell **40**:302-312 e304. 10.1016/j.devcel.2016.12.021.

Li, Z., Jiang, D., and He, Y. (2018). FRIGIDA establishes a local chromosomal environment for FLOWERING LOCUS C mRNA production. Nat Plants **4**:836-846. 10.1038/s41477-018-0250-6.

Ouyang, W., Xiong, D., Li, G., and Li, X. (2020). Unraveling the 3D Genome Architecture in Plants: Present and Future. Mol Plant **13**:1676-1693. 10.1016/j.molp.2020.10.002.

Piunti, A., and Shilatifard, A. (2021). The roles of Polycomb repressive complexes in mammalian development and cancer. Nat Rev Mol Cell Biol **22**:326-345. 10.1038/s41580-021-00341-1.

Ramirez-Colmenero, A., Oktaba, K., and Fernandez-Valverde, S.L. (2020). Evolution of Genome-Organizing Long Non-coding RNAs in Metazoans. Front Genet **11**:589697. 10.3389/fgene.2020.589697.

Wierzbicki, A.T., Blevins, T., and Swiezewski, S. (2021). Long Noncoding RNAs in Plants. Annu Rev Plant Biol **72**:245-271. 10.1146/annurev-arplant-093020-035446.

Xu, S., Dong, Q., Deng, M., Lin, D., Xiao, J., Cheng, P., Xing, L., Niu, Y., Gao, C., Zhang, W., et al. (2021). The vernalization-induced long non-coding RNA VAS functions with the transcription factor TaRF2b to promote TaVRN1 expression for flowering in hexaploid wheat. Mol Plant 10.1016/j.molp.2021.05.026.

Zhao, B., Xi, Y., Kim, J., and Sung, S. (2021). Chromatin architectural proteins regulate flowering time by precluding gene looping. Sci Adv **7**10.1126/sciadv.abg3097.