BioTechnologia
vol. 95(1) • pp. 63-72 • 2014

Session 3
Long non-coding RNAs

Lecture 3.1

Genome-wide characterization of non-coding transcription by RNA Polymerase V

G. Bohmdorfer, M.J. Rowley, A.T. Wierzbicki

Department of Molecular, Cellular and Developmental Biology, University of Michigan, Ann Arbor, Michigan, USA

RNA-mediated transcriptional silencing prevents deleterious effects of transposon activity and controls the expression of protein-coding genes. It involves not only small interfering RNA (siRNA) but also long non-coding RNA (lncRNA). In *Arabidopsis thaliana* this lncRNA is produced by a specialized RNA Polymerase V (Pol V). The mechanism by which lncRNA is involved in the establishment of repressive chromatin states remains mostly unknown. We show evidence that lncRNA serves as a binding scaffold for several RNA-binding proteins. We further show that these proteins require both lncRNA and siRNA to recognize their genomic targets and are recruited to lncRNA in a stepwise manner. Finally, we show that genome-wide identification of lncRNAs produced by Pol V further supports their central role as scaffolds for protein binding and provides insights into the molecular mechanism of RNA-mediated transcriptional silencing.

Lecture 3.2

Dynamic changes of non-coding transcriptome of *Arabidopsis* during development

HUI-WEN WU, JUN LIU, HUAN WANG, NAM-HAI CHUA

Laboratory of Plant Molecular Biology, Rockefeller University, New York, New York, USA

Recent research on long non-coding RNAs (lncRNAs) has expanded our understanding of gene transcription regulation and the generation of cellular complexity. Depending on their genomic origins, lncRNAs can be transcribed from intergenic or intragenic regions or from introns of protein-coding genes. In *Arabidopsis*, we have identified 13,230 intergenic transcripts of which 6,480 can be classified as long intergenic non-coding RNAs (lincRNAs). Expression of 2,708 lincRNAs was detected by RNA sequencing experiments. Transcriptome profiling revealed that the majority of these lincRNAs are expressed at a level between those of mRNAs and precursors of miRNAs. A subset of lincRNA genes shows organ-specific expression, whereas many others are responsive to biotic and/or abiotic stresses (Liu et al, 2012). Next, we systematically identified long non-coding natural antisense transcripts (lncNATs), which are transcribed from the opposite DNA strand of coding or non-coding genes. We found a total of 37,238 sense-antisense transcript pairs and 70% of annotated mRNAs to be associated with antisense transcripts in *Arabidopsis*. These lncNATs could be reproducibly detected by different technical platforms. Moreover, we observed spatial and developmental-specific light effects on 626 concordant and 766 discordant NAT pairs. Genes for a large number of the light-responsive NAT pairs are associated with histone modification peaks, and histone acetylation is dynamically correlated with light-responsive expression changes of NATs (Wang et al., 2014). We have presented our data described above in a plant long non-coding RNA database (PLncDB), which also includes a list of epigenetic modifications and small RNA datasets (Jin et al., 2013). This database will be updated periodically for the research community. We have examined the function of several intronic ncRNAs in *Arabidopsis*. AGAMOUS (AG) is known to specify the development of the third and fourth whorl of wild type (WT) flowers. The AG genomic locus is marked by dispersed histone H3 lysine 27 trimethylations (H3K27me3) which are absent in mutants deficient in CURLY LEAF (CLF), the catalytic component of the polycomb repressive complex (PRC) 2. *Clf* mutants express AG transcripts in leaves and roots because of de-repression. We show that the AG second intron encodes several intronic non-coding RNAs (incRNAs), one which is important for CLF-mediated repression. Genetic interaction analysis demonstrated that the AG-incRNA associated with CLF to repress AG. Our results suggest that AG-incRNA is able to recruit CLF to the AG second intron to repress AG in leaf tissues through H3K27me3-mediated repression. The mechanism of AG-incRNA-mediated gene repression may serves as a general model for investigating tissue-specific expression of MADS-box genes in *Arabidopsis* flower development.

Short talk 3.1

Transcription, replication and recombination of genomic and subgenomic RNAs in brome mosaic virus, a model (+) strand RNA virus

J. SZTUBA-SOLINSKA[1], A. DZIANOTT[1], R. WIERZCHOSLAWSKI[1], J. HORN[2], J.J. BUJARSKI[1,3]

[1] Plant Molecular Biology Center and the Department of Biological Sciences, Northern Illinois University, DeKalb, Illinois, USA

[2]Department of Biological Sciences, Northern Illinois University, DeKalb, Illinois, USA

[3]Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland

The process of transcription from the brome mosaic virus (BMV) RNA3 segment generates two subgenomic (sg) RNAs, sgRNA3a and sgRNA4. By altering the elements of the intercistronic sg promoter (sgp) sequence that is responsible for sgRNA4, we observed both debilitating and stimulating effects on homologous RNA3-RNA3 recombination. This suggested that although linked, the mechanisms of RNA-RNA recombination and transcription follow separate molecular processes during BMV RNA replication. Other studies demonstrated that sequence alterations in the coat protein (CP)-binding cis-acting RNA motifs affected crossover frequencies between BMV RNA3 and sgRNA3a. This implied the role of CP in at least some of the BMV RNA recombination events by either facilitating RNA-RNA template interaction or by creating the obstacles for the viral replicase. Altogether, our results demonstrate a functional link of such viral functions as RNA replication, RNA transcription and RNA recombination not only in BMV, but likely in other RNA viruses.

Short talk 3.2

Non-coding transcription by alternative RNA polymerases dynamically regulates an auxin-driven chromatin loop

F. ARIEL[1], T. JEGU[2], D. LATRASSE[2], N. ROMERO-BARRIOS[1], A. CHRIST[1], M. BENHAMED[2], M. CRESPI[1]

[1] Institut des Sciences du Vegetal (ISV), Centre National de la Recherche Scientifique (CNRS), Gif-sur-Yvette, France

[2] Institut de Biologie des Plantes, Universite Paris-Sud XI, Orsay, France

The eukaryotic epigenome is shaped by the genome topology in three-dimensional space. Dynamic reversible variations in the epigenome structure direct the transcriptional responses to developmental cues. However, little is known about the control of epigenome dynamics. Here, we show that the *Arabidopsis thaliana* long intergenic non-coding RNA (lincRNA) *APOLO* is transcribed by RNA polymerase II (Pol II) and V (Pol V) complexes in response to auxin, a phytohormone controlling numerous facets of plant development. *APOLO* transcription facilitates the formation of an oscillating chromatin loop encompassing the promoter of its neighboring gene *PID* (or *PINOID*), a key regulator of auxin polar transport. Chromatin and RNA Immuno-precipitation (ChIP and RIP), together with Chromatin Isolation by RNA Purification (ChIRP) and Chromatin Conformation Capture (3C) served to decipher the ncRNA-mediated mechanisms controlling the chromatin loop opening and closing in response to auxin, modulating *PID* promoter activity. Components of the plant Polycomb Repressive Complexes as well as the transcriptional gene silencing and DNA demethylation machineries contribute to fine-tune chromatin loop dynamics. Altering *APOLO* expression or its 24nt siRNA-dependent DNA methylation affects loop formation and, consequently, *PID* expression. Hence, the active transcription of a lincRNA by alternative RNA polymerase complexes influences local chromatin topology and the expression of a neighboring locus, leading to far-reaching consequences on a variety of developmental outputs. Acknowledgements: FA is an EMBO postdoc fellow; NRB and TJ are Paris-Sud University PhD fellows. This work was funded by LABEX Saclay Plant Sciences.

Lecture 3.3

Novel antisense RNA regulation functions in plant abiotic stress responses

AKIHIRO MATSUI[1], KEI IIDA[2,5], KATSUSHI YAMAGUCHI[3], MAHO TANAKA[1], JUNKO ISHIDA[1], TAEKO MOROSAWA[1], SHUJI SHIGENOBU[3], KAZUO SHINOZAKI[4], TETSURO TOYODA[5], MOTOAKI SEKI[1,6,7]

[1] Plant Genomic Network Research Team, RIKEN Center for Sustainable Resource Science (CSRS), Yokohama, Japan

[2] Medical Research Support Center, Graduate School of Medicine, Kyoto University, Kyoto, Japan

[3] NIBB Core Research Facilities, National Institute for Basic Biology, Okazaki, Japan

[4] Gene Discovery Research Group, RIKEN Center for Sustainable Resource Science (CSRS), Yokohama, Japan

[5] Integrated Database Unit, RIKEN Advanced Center for Computing and Communication (ACCC), Wako, Japan

[6] Kihara Institute for Biological Research, Yokohama City University, Yokohama, Japan

[7] Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology (JST), Kawaguchi, Japan

Arabidopsis whole genome transcriptome analysis under drought, cold, high-salinity stress and ABA treatment conditions, using tiling array, showed that approximately 6,000 novel non-codding transcripts existed in the antisense strand of the AGI code genes and their expression was induced under abiotic stress (Matsui et al. (2008) Plant Cell Physiol.). To study the biogenesis mechanism of the antisense RNAs, accumulation of *RD29A* antisense RNA was analyzed, using RNA-dependent RNA polymerase (RDR) mutant series. Finally, we found that *rdr1/2/6* mutation decreased accumulation of *RD29A* antisense RNA. And *rdr1/2/6* mutants showed arrest of root growth after drought stress, suggesting that the antisense RNAs have novel biological function in response to abiotic stress. RNase protection and RNA decay analyses showed that *RD29A* antisense RNA formed double stranded RNA and promoted degradation of *RD29A* mRNA. However, RNA-seq analysis did not detect siRNA accumulation in this locus. We propose a novel RNA regulation mechanism that RDR-mediated synthesis of antisense RNAs functions in the turnover of sense mRNAs under the stress conditions in plants.

Lecture 3.4

Global analysis of RNA-protein interaction sites and RNA secondary structure in the *Arabidopsis* nucleus

S.J. Gosai[1,#], S.W. Foley[1,2,#], Dongxue Wang[3], I.M. Silverman[1,2], A.D.L. Nelson[4], N. Selamoglu[1], M.A. Beilstein[4], F. Daldal[1], R.B. Deal[3], B.D. Gregory[1,2]

[1] Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[2] Cell and Molecular Biology Graduate Group, University of Pennsylvania, Philadelphia, Pennsylvania USA

[3] Department of Biology, Emory University, Atlanta, Georgia, USA

[4] School of Plant Sciences, University of Arizona, Tucson, Arizona, USA

[#] equal contributions

At the heart of post-transcriptional regulatory pathways in eukaryotes are *cis*- and *trans*-acting features and factors including RNA secondary structure as well as RNA-binding proteins (RBPs) and their binding sites on target RNAs. However, to date a comprehensive view of the structural and RBP interaction landscape of RNAs in the nucleus has not been accomplished for any organism. Here, we use our novel protein interaction profile sequencing (PIP-seq) approach on mixed nuclei from *Arabidopsis* seedlings to globally profile both RNA secondary structure and RNA-protein interaction sites in this cellular compartment. From our analysis, we reveal opposing patterns of secondary structure and RBP binding levels throughout nascent messenger RNAs (mRNAs). These patterns underlie the regulation of alternative splicing and polyadenylation. We also uncover a class of protein bound, highly conserved, nuclear, long non-coding RNAs (lncRNAs), some of which interact with chromatin. Finally, we uncover a large collection of protein bound sequence motifs, and identify their interacting proteins and structural contexts. In total, we provide the first simultaneous, transcriptome-wide view of RNA secondary structure and RNA-protein interaction sites in a plant nucleus.

Short talk 3.3

The exosome and its cofactors contribute to multiple steps of 18S rRNA maturation in *Arabidopsis thaliana*

P.J. SIKORSKI[1,2], **D. GAGLIARDI**[2], **J. KUFEL**[1], **H. LANGE**[2]

[1] Institute of Genetics and Biotechnology, University of Warsaw, Warsaw, Poland

[2] Institut de Biologie Moleculaire des Plantes du Centre National de la Recherche Scientifique, Universite de Strasbourg, Strasbourg, France

Maturation of 18S rRNA from the polycistronic rRNA precursor involves a number of conserved endonucleolytic cleavages. Recent studies in yeasts and human have now revealed that exoribonucleolytic degradation by RRP6, a cofactor of the nuclear exosome, can also contribute to processing or degradation of 18S precursors. Here, we investigate the impact of the exosome and its cofactors on 18S rRNA maturation in *Arabidopsis thaliana*. We show that plants deficient in the core exosome, the exoribonuclease RRP44, or the nucleolar RNA helicase AtMTR4 accumulate different forms of both 3′- and 5′- extended 18S rRNA precursors. Downregulation of the core exosome or RRP44 is associated with accumulation of a 2.5 kb 18S precursor spanning from P to A3 processing sites. In *rrp41* mutants, these P-A3 precursors are frequently polyadenylated with long poly-A tails up to 200 nucleotides. By contrast, loss of the RNA helicase AtMTR4 results in decreased levels of P-A3 and to increased levels of 5′ shortened P′- A3 fragments, which have no or only short oligo(A) tails. Preliminary results indicate that more than one poly(A) polymerase is involved in the addition of poly(A) tails to rRNA precursors prior to their degradation by the exosome. Moreover, loss of the exoribonuclease RRP6L2 is specifically associated with the accumulation of precursors that have mature 5′ ends but carry a 3′ extension of 20 nucleotides (18S-A2), indicating that 18S precursors undergo exoribonucleolytic trimming by RRP6L2. Interestingly, the 18S-A2 fragments that accumulate in *rrp6l2* mutants often carry short oligo(U) tails and can be detected in the cytoplasm. Taken together, our data suggest that in addition to conserved endonucleolytic cleavages, 3′-5′ exoribonucleolytic degradation by the exosome and its cofactors contribute to different steps of 18S maturation in plants.

Short talk 3.4

Genome-wide annotation of functional regulatory elements in plant snoRNA/scaRNA genes

GE QU[1], K, KRUSZKA[2], Z. SZWEYKOWSKA-KULINSKA[2], M. ECHEVERRIA[3], W. KARLOWSKI[1]

[1] Laboratory of Computational Genomics / Bioinformatics Laboratory, Faculty of Biology, Adam Mickiewicz University in Poznan, Poland

[2] Department of Gene Expression, Faculty of Biology, Adam Mickiewicz University in Poznan, Poland

[3] Universite de Perpignan Via Domitia, Perpignan, France

The small nucleolar RNAs (snoRNA) and the related scaRNAs represent an abundant class of small trans-acting RNAs with very diverse and important functions in all eukaryotes. The snoRNAs fall into two classes that direct modification of ribosomal RNAs (rRNA): the C/D snoRNAs that direct 2′-O-ribose methylation and H/ACA snoRNAs that direct pseudouridylation - these modifications are essential to produce functional ribosomes. The major targets of scaRNAs are the spliceosomal snRNAs where introduced modifications are essential for splicing. The snoRNAs and related scaRNAs have additional targets – for example they direct methylation of a brain specific mRNA in vertebrates. Moreover, many so-called 'orphan' snoRNAs, for which no target could be predicted, have been identified in all species, suggesting that these may have additional functions. Several hundreds of genes encoding snoRNAs, scaRNAs and orphan snoRNAs have been identified already in *Arabidopsis* and rice. The studies show that while the structure of snoRNAs and scaRNAs is relatively conserved, the plant snoRNA genes exhibit unique genomic organization which is rarely found in other eukaryotes. We are interested in the factors controlling snoRNA and scaRNA biogenesis in plants. A preliminary analysis of snoRNA gene promoters in *Arabidopsis* revealed that all polycistronic snoRNA genes encoding snoRNAs targeting rRNA modifications have conserved promoter motifs (Telo-box and GCCCR-elements). Moreover, these elements are also found in the promoter of genes encoding ribosomal proteins and other proteins controlling ribosome biogenesis. Most interestingly, we found that scaRNAs genes have also a distinct type of conserved elements. This altogether may suggest that the snoRNA/scaRNA gene promoter contains elements which, in addition to basic transcriptional regulation, play a significant role in the snoRNA or scaRNA function and localization. The main object of the project is to address the question of the role of the promoter of snoRNA/scaRNA genes by combining *in silico* and laboratory experimental approaches.

Short talk 3.5

Novel players in plant stress response: micropeptides and/or long non-coding RNAs?

P. TAVORMINA, J. NEUKERMANS, C. VOS, B.P.A. CAMMUE, B. DE CONINCK

Centre of Microbial and Plant Genetics, KU Leuven, Belgium
Department of Plant Systems Biology, VIB, Belgium

Plants as sessile organisms have to resist abiotic and biotic stresses without the option to escape. To cope with theses stresses, plants rely on a complex array of fine-tuned networks. Plant peptides are part of these networks and are known to exert their functions directly as antimicrobial peptides or indirectly via interaction with signaling cascades. However, the diversity and the total amount of plant peptides is strongly underestimated. As a consequence, recent genome annotations of the model plant *Arabidopsis thaliana* are far from complete. Moreover, peptides that are produced *de novo* from a small open reading frame (sORF, <100 aa), without the involvement of a precursor, are rarely being studied in plants. In other eukaryotes, such sORF-encoded peptides are being studied more intensively as important bioactive molecules, but still appear to be an unexplored part of the peptidome. However, big controversy arises in this research domain because transcripts (>200 bp) containing no ORFs longer than 100 aa are being considered as long non-coding RNAs (lncRNAs). The need emerges to study the dual character of these transcripts as functional small peptides 'micropeptides' and/or as RNA. In the present study we aimed at i) identifying novel stress-induced genes, and ii) unravelling the biological function of the gene products, as peptide(s) and/or as RNA, in the stress response of *Arabidopsis thaliana*. We therefore performed a tiling array-based transcriptome study on leaf material under reactive oxygen species (ROS)-inducing biotic and abiotic stress conditions. As such we identified 195 and 176 transcriptional active regions (TARs) that were differentially expressed after infection with the fungal pathogen *Botrytis cinerea* and the herbicide paraquat, respectively (De Coninck et al., J. Exp. Bot. 2013). Functional analysis of these TARs is performed at two levels. At one hand, we study novel peptides that are possibly encoded by small ORFs in these TARs, more specifically via a peptidomics approach. At the other hand, after determining the exact transcripts of several TARs by rapid amplification of cDNA ends (RACE), we also study the TARs as potential long non-coding RNAs (lncRNAs). Further characterization of the identified TARs was done via phenotypical analyses on a selection of corresponding knock-out mutants. For example, *Tar1* was more susceptible to infection with the fungal pathogen *Fusarium oxysporum*. *Tar2* showed interveinal chlorosis and was also more susceptible to infection with *Fusarium oxysporum*. In conclusion, these TARs function in the stress response of plants. It is currently being assessed by complementation assays whether this involvement is mediated via peptide(s) and/or RNA.

Short talk 3.6

Long non-coding RNA modulates alternative splicing regulators in *Arabidopsis*

F. BARDOU[1], F. ARIEL[1], C.G. SIMPSON[2], N. ROMERO-BARRIOS[1], S. BALZERGUE[3], J.W.S. BROWN[2,4], M. CRESPI[1]

[1] Institut des Sciences du Végétal, CNRS, Saclay Plant Sciences, Gif-sur-Yvette, France

[2] Cell and Molecular Sciences, the James Hutton Institute, Invergowrie, Dundee, Scotland, United Kingdom

[3] Genomique Fonctionnelle d'Arabidopsis, Unite de Recherche en Genomique Vegetale, UMR INRA 1165 – Universite d'Evry Val d'Essonne – ERL CNRS 8196, Evry, France

[4] Plant Sciences Division, College of Life Sciences, University of Dundee at the James Hutton Institute, Invergowrie, Dundee, Scotland, United Kingdom

Alternative splicing (AS) of pre-mRNA represents a major mechanism that underlies increased transcriptome and proteome complexity. Here we show that the nuclear speckle RNA-binding protein (NSR) and the alternative splicing competitor long non-coding RNA (ASCO-lncRNA) constitute an AS regulatory module. AtNSR-GFP translational fusions are expressed in primary and lateral root (LR) meristems. Double *Atnsr* mutants and ASCO overexpressors exhibit an altered ability to form LRs after auxin treatment. Interestingly, auxin induces a major change in AS patterns of many genes, a response largely dependent on NSRs. RNA-immunoprecipitation assays demonstrate that AtNSRs not only interact with their alternatively-spliced mRNA targets but also with the ASCO-RNA *in vivo*. The ASCO-RNA displaces an AS target from an NSR-containing complex in vitro. The expression of ASCO-RNA in *Arabidopsis* affects the splicing patterns of several NSR-regulated mRNA targets. Hence, lncRNA can hijack nuclear AS regulators to modulate AS patterns during development.