

**Review Article**

# Multiplicity in Long Noncoding RNA in Living Cells

Riki Kurokawa<sup>1,\*</sup>, Reina Komiya<sup>2,3</sup>, Takanori Oyoshi<sup>4</sup>, Yoko Matsuno<sup>5</sup>, Hidenori Tani<sup>6</sup>, Masato Katahira<sup>7,8</sup>, Keisuke Hitachi<sup>9</sup>, Yuji Iwashita<sup>10</sup>, Takefumi Yamashita<sup>11</sup>, Keiko Kondo<sup>7</sup>, Ryoma Yoneda<sup>1</sup>, Yudai Yamaoki<sup>7</sup>, Naomi Ueda<sup>1</sup>, Tsukasa Mashima<sup>7,8</sup>, Naohiro Kobayashi<sup>12</sup>, Takashi Nagata<sup>7,8</sup>, Ayaka Kiyoshi<sup>8</sup>, Masayuki Miyake<sup>8</sup>, Fumi Kano<sup>13</sup>, Masayuki Murata<sup>14</sup>, Nesreen Hamad<sup>7,8</sup>, Kohei Sasaki<sup>11</sup>, Naoyuki Shoji<sup>11</sup>

<sup>1</sup>Division of Gene Structure and Function, Research Center for Genomic Medicine, Saitama Medical University, Hidaka, Japan

<sup>2</sup>Science and Technology Group, Okinawa Institute Science and Technology Graduate University, Tancha, Japan

<sup>3</sup>Japan Science and Technology Agency, PRESTO, Kawaguchi, Japan

<sup>4</sup>Department of Chemistry, Graduate School of Science, Shizuoka University, Suruga, Japan

<sup>5</sup>Division of Clinical Preventive Medicine, Niigata University, Niigata, Japan

<sup>6</sup>Environmental Management Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan

<sup>7</sup>Institute of Advanced Energy, Kyoto University, Kyoto, Japan

<sup>8</sup>Graduate School of Energy Science, Kyoto University, Kyoto, Japan

<sup>9</sup>Division for Therapies against Intractable Diseases, Fujita Health University, Toyoake, Japan

<sup>10</sup>Department of Tumor Pathology, Hamamatsu University School of Medicine, Hamamatsu, Japan

<sup>11</sup>Laboratory for Systems Biology and Medicine, Research Center for Advanced Science and Technology, the University of Tokyo, Tokyo, Japan

<sup>12</sup>Institute for Protein Research, Osaka University, Suita, Japan

<sup>13</sup>Cell Biology Center, Institute of Innovative Research, Tokyo Institute of Technology, Tokyo, Japan

<sup>14</sup>Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan

**Email address:**

[rkurokaw@saitama-med.ac.jp](mailto:rkurokaw@saitama-med.ac.jp) (R. Kurokawa)

\*Corresponding author

**To cite this article:**

Riki Kurokawa, Reina Komiya, Takanori Oyoshi, Yoko Matsuno, Hidenori Tani, Masato Katahira, Keisuke Hitachi, Yuji Iwashita, Takefumi Yamashita, Keiko Kondo, Ryoma Yoneda, Yudai Yamaoki, Naomi Ueda, Tsukasa Mashima, Naohiro Kobayashi, Takashi Nagata, Ayaka Kiyoshi, Masayuki Miyake, Fumi Kano, Masayuki Murata, Nesreen Hamad, Kohei Sasaki, Naoyuki Shoji. Multiplicity in Long Noncoding RNA in Living Cells. *Biomedical Sciences*. Vol. 4, No. 2, 2018, pp. 18-23. doi: 10.11648/j.bs.20180402.11

**Received:** July 9, 2018; **Accepted:** July 26, 2018; **Published:** August 16, 2018

---

**Abstract:** This review article aims at demonstration of complexity or multiplicity of long noncoding RNA. The presentations of the meeting cover variety of biomedical sciences to indicate significance of long noncoding RNA in each field. The achievements of the meeting are to confirm a point of view of long noncoding RNAs in living cells. It should be in complexity. We will be able to analyze complicated phenomena of lncRNA biological actions and might have resolved the veiled rules in divergent biological programs. In the plant system, its long noncoding RNAs possesses unique property compared to mammalian systems, but tells us an unveiled fundamental principle behind all creatures in the globe. Then, experimental data from structural biology mainly by NMR analysis show that interaction of nucleic acids to RNA binding proteins focusing on TLS/FUS. Analyses of TLS are also performed by biochemistry and molecular biology to indicate precise figures of RNA binding specificity and also methylation effect on TLS. Biological meaning of the DNA local conformation like G-quadruplex is presented. XIST, one of the best known long noncoding RNA, related to X chromosome inactivation, has been linked to novel finding in epigenetic functional redundancy such as long noncoding RNA -chromatin associations. In the iPS cells, their long noncoding RNA s have been identified as a potential marker for chemical stress responses. Cellular

differentiation is also regulated by specific long noncoding RNA. Computational analysis is another approach to make progress. Transcriptome mining indicates an association between aging and sets of long noncoding RNAs with previously unidentified function. Molecular dynamics simulation of RNA presents unprecedented pathway to future of the field. The data in the manuscript are based upon the 2nd Annual Meeting of the long noncoding RNA study group. On Thursday, May 24 in 2018, 2nd Annual Meeting of the long noncoding RNA study group was held at Research Center for Advanced Science and Technology (RCAST), the University of Tokyo, Tokyo in Japan. The purpose of the meeting is to present recent data and have discussion regarding long noncoding RNA and related topics. We had thirteen sessions and exciting and fruitful debates there. The meeting has been successfully prorogued. We utilize the data to boost the activity of the field of long noncoding RNA.

**Keywords:** Long Noncoding RNA, RNA-binding Protein, Amyotrophic Lateral Sclerosis, TLS, FUS, X Chromosome Inactivation, LINES, Molecular Dynamics

## 1. Introduction

The divergent biological activity of long noncoding RNA has been focused owing to its contribution to biological programs including human genome. The intention of this review article is to emphasize divergent property of long noncoding RNA. Divergent researchers also contribute to this article. This implies that divergent readers should be attracted to the article.

The description of the body of the review article is based upon 2nd Annual Meeting of the long noncoding RNA study group. On Thursday, May 24 in 2018, 2nd Annual Meeting of the long noncoding RNA study group was held at Research center for advanced science and technology (RCAST), the University of Tokyo, Tokyo in Japan. This gave us a lot of beneficial discussion. We really appreciate every attendee there.

We have obtained a plan of the meeting at the chat of the dinner after the session at the 39th annual meeting of the Molecular Biology Society of Japan in Yokohama in December, 2016. Katahira, Yamashita, and Kurokawa voluntarily have started to take charge of the meeting. We decided to have a first meeting RCAST on Thursday, May 18 in 2017. We did not have any funding on it. This means that it should be a small meeting with less luxurious setting, but have the freedom to take time for discussion. We have prepared for the meeting with groping for a way to make better one. Fortunately, we have attained a great success on it. We have it as a closed one. Therefore, limited numbers of people attended. Then, we had a plenty of time to have fruitful discussion at the sessions. Official meeting of the big academic society has stringent regulation of the program because of time limitation. Then, we have appreciated the liberty of free discussion there.

We have determined to have the third meeting on Thursday, May 23 in 2019 near Tokyo station with service by Katahira. This will be another successful history of the Workshop.

## 2. Divergent Biological Activities in Long Noncoding RNA

Komiya demonstrates small RNAs, silencing elements, are divided into main two groups, microRNAs (miRNAs) and

small interfering RNAs (siRNAs). Phased small interfering RNAs (phasiRNAs), which are generated from long RNA precursors at intervals of 21- to 26-nucleotides (nt), have been identified in animals and plants. In the family Poaceae, numerous 21-nt phasiRNAs are expressed from the premeiotic to the early meiotic stage, and subsequently, 24-nt phasiRNAs are expressed during the meiotic stage. In the biogenesis of both types of reproductive phasiRNAs, precursors of long intergenic non-coding RNAs (lincRNAs), which contain consensus sequences complementary to 22-nt miR2118/miR2775, are cleaved within the miRNA targeting site. These miRNA cleavages induce dsRNA synthesis from cleaved RNAs and dsRNAs are processed by DICER-LIKE proteins (DCLs) into 21-/24-nt phasiRNAs [1, 2].

Over 700 lincRNAs, 21-nt reproductive phasiRNA precursors, are specifically expressed during reproduction. These reproductive lincRNAs 1. lack introns, 2. exhibit low expression, and 3. have ambiguous transcription start and end sites. Furthermore, comparative analysis indicates that these reproductive lincRNAs display low homology, even within the Poaceae (rice, Brachypodium, maize, and wheat, Komiya, unpublished). However, 21-nt reproductive phasiRNA biogenesis, including miR2118-cleavage triggering and DCL-processing, is conserved in monocots, dicots, and angiosperms [2], suggesting that there are diverse types of lincRNAs, phasiRNA precursors. The functions of lincRNAs, miR2118, and 21-nt phasiRNAs remain unknown. In this meeting, we introduced the role of miR2118 triggering the production of 21-nt phasiRNAs.

Kondo presents NMR analyses for the binary and ternary complexes formation of TLS/FUS protein with G-quadruplexes of telomeric DNA and TERRA, which induces telomere shortening. Translocated in liposarcoma protein (TLS/FUS) simultaneously interacts with G-quadruplexes of telomeric DNA and telomeric repeat-containing RNA (TERRA) to form ternary complex [3]. The complex of TLS and telomeric DNA and TERRA plays essential role in regulation of telomere lengthening. Phe and Tyr residues in the third Arg-Gly-Gly domain (RGG3) of TLS are reportedly responsible for the recognition of telomeric DNA and TERRA, respectively [4]. However, the structural basis for the G-quadruplexes recognition by RGG3 is yet to be elucidated. We identified the residues that are involved in the interaction between RGG3 and

G-quadruplexes by using NMR spectroscopy. In the ternary complex, Phe and Tyr residues preferentially bind to telomeric DNA and TERRA, respectively. While in the binary complexes, both Phe and Tyr residues interact with either telomeric DNA or TERRA. Based on these results, we propose the model of binary and ternary complexes, where Phe and Tyr residues switch their interaction modes depending on which G-quadruplex is a binding partner. [5]

Yoneda shows the regulation of CCND1 mRNA by TLS and long noncoding RNA transcribed from the promoter region of CCND1. Cyclin D1 (CCND1) is one of the factors which controls the G1/S phase checkpoint of the cell cycle. There are several long noncoding RNAs transcribed from the promoter region of CCND1 and are induced by irradiation. These long noncoding RNAs bind to RNA binding protein TLS (Translocated in LipoSarcoma, or FUS) and inhibits the HAT activity of CBP, subsequently inhibits the expression of CCND1. In this study, we focused on 602 nt long noncoding RNA, named promoter associated ncRNA-D (pncRNA-D), and examined the binding specificity between pncRNA-D and TLS [6]. First, we divided pncRNA-D into seven fragments, and RNA binding assay was performed. The results indicated that TLS bound strongly to the 5' and 3' ends of pncRNA-D [1]. TLS is expected to recognize GGUG sequences in RNAs, and in fact, the binding assay revealed that TLS preferentially binds to RNA fragments with GGUGs. Therefore, we mutated the GGUG sequence to CCUC, and the mutation dramatically reduced the binding between RNA and TLS. The data suggest the sequence specificity regulates the interaction between short RNA fragments and TLS.

Yamaoki presents in-cell NMR studies on the structures of nucleic acids introduced inside the living human cells. Intracellular environment is extremely crowded with large and small molecules. It has long been considered that the structures and functions of macromolecules, such as proteins and nucleic acids, inside the living cells are different from those in in vitro conditions. To elucidate this question, in-cell NMR spectroscopy is an attractive and robust methodology. So far, however, in-cell NMR spectroscopy has been applied mostly to proteins. There are only a few reports on nucleic acids [7]. In particular, the reported in-cell NMR studies on nucleic acids were performed almost entirely by using *Xenopus laevis* oocytes [8]. Here, we succeeded in recording the in-cell NMR studies of nucleic acids inside the living human cells. We used DNAs and RNAs, which form hairpin structures under in vitro conditions. They were introduced into HeLa cells by a method utilizing toxic protein, streptolysin O (SLO), which enables to form and reseal the pores on the cell membrane. After the SLO treatment, the introduction efficiency of nucleic acids and cell viability were examined by flow cytometry analysis. Confocal fluorescence microscopy showed that these DNAs and RNAs locate mainly inside the nuclei. The observation of the imino proton signals directly prove the formation of base-pairs and indicated the persistence of the hairpin structures inside the living human cells. [9]

Kurokawa shows that translocated in liposarcoma (TLS) is

an RNA-binding protein and a transcription-regulatory sensor of DNA damage. TLS binds promoter-associated noncoding RNA (pncRNA) and inhibits histone acetyltransferase (HAT) activity of CREB-binding protein (CBP)/E1A-binding protein P300 (p300) on the cyclin D1 (CCND1) gene. Although post-translational modifications of TLS, such as arginine methylation, are known to regulate TLS's nucleocytoplasmic shuttling and assembly in stress granules, its interactions with RNAs remain poorly characterized. Herein, using various biochemical assays, we show confirmed the earlier observations that TLS is methylated by protein arginine methyltransferase 1 (PRMT1) in vitro [10]. The arginine methylation of TLS disrupted binding to pncRNA and also prevented binding of TLS to and inhibition of CBP/p300. This result indicated that arginine methylation of TLS abrogates both binding to pncRNA and TLS-mediated inhibition of CBP/p300 HAT activities. After extensive experiments, we conclude propose that arginine methylation of TLS regulates both TLS–nucleic acid and TLS–protein interactions and thereby participates in transcriptional regulation.

Oyoshi demonstrates G-quadruplex binding ability of RGG domain in TLS/FUS. The Arg-Gly-Gly repeat (RGG) domain is an evolutionarily conserved amino acids sequence in various DNA and RNA binding proteins [11]. These proteins are responsible for DNA damage signaling, histone modification, transcription, pre-mRNA splicing, RNA transport and translation [11, 12]. Translocated in liposarcoma (TLS), which is also called fused in sarcoma (FUS), is such a protein with RGG domain involved in the regulation of gene expression and telomere maintenance [3, 13]. Previously, we reported that RGG domain in C-terminal region of TLS/FUS, which was expressed in *Escherichia coli*, and purified with low concentration of KCl, binds G-quadruplex telomere DNA and telomeric repeat-containing RNA and TLS/FUS *in vivo* regulates histone modification by associating with histone methyl transferase in the telomere region [3]. On the other hand, other group reported that TLS/FUS and the RGG domain purified with high concentration of urea and KCl bind weakly and non-specifically to DNA and RNA [14]. In order to investigate why the DNA and RNA binding activities of TLS/FUS are different, we analyzed the effect of urea and KCl for DNA and RNA binding activities and second structure of RGG domain. These results indicated that high concentration of urea and KCl decreases the G-quadruplex binding activities and specificities of TLS/FUS and the RGG domain and induces denaturation of the  $\beta$ -spiral structure of the RGG domain. Moreover, the Arg-Gly-Gly repeat region in RGG domain of TLS/FUS by itself alone cannot form a  $\beta$ -spiral and bind to G-quadruplex specifically. The RGG domain consists of Arg-Gly-Gly repeat region and the proline- and arginine-rich regions in N-terminal and C-terminal domain, respectively. These regions promoted  $\beta$ -spiral formation and G-quadruplex binding activities and specificities of TLS/FUS and the RGG domain. Our findings suggest that G-quadruplex binding activities and specificities

of TLS/FUS depend on  $\beta$ -spiral structure of RGG domain.

Matsuno shows characteristic converging variations in LINE nucleotides can benefit forming triplexes with lncRNA in X chromosome inactivation. Long noncoding Xist RNA-chromatin associations are essential intermolecular interactions in the process of X chromosome inactivation. Using bioinformatic analyses, we defined short-length motifs ( $\geq 5$  nucleotides), termed redundant UC/TC (r-UC/TC) or AG (r-AG) motifs, which may help mediate triplex formation between Xist RNA and dsDNA, and were confirmed with CD spectra and quartz-crystal microbalance (QCM). r-UC motifs are dispersed throughout the involved long noncoding RNAs of three species: 255 sites, 9.0% of human 19-kb XIST RNA, 210 sites, 8.4% of mouse 17-kb Xist RNA; 287 sites, 7.7% of opossum 24-kb Rsx RNA, while the complementary r-AG motifs are dispersed throughout LINES with an even higher prevalence than the r-UC motifs in the corresponding lncRNAs: the average prevalence of r-AG motifs in 50 randomly selected LINES ( $>6000$  bp) are 13.6% (human), 14.1% (mouse), and 19.2% (opossum). Furthermore, r-AG and r-TC motifs in LINES have a similar length-distribution pattern across the three species (Kolmogorov-Smirnov test,  $z$ -values: human 6.965; mouse 2.303; opossum 7.690,  $p < 0.001$ ). Regardless of such similarities, sequence alignments of LINES in each species revealed high frequency variations in r-AG motif sequences probably due to transition, transversion, or deletion/insertion mutations. These characteristic converging variations in mammalian LINES may exist to retain the ability to form triplexes with the involved long noncoding RNA as a common process of X chromosome inactivation.

Tani demonstrates short-lived long non-coding RNAs as surrogate indicators for chemical exposure. Because of the limitations of whole animal testing approaches for toxicological assessment, new cell-based assay systems have been widely studied. However, appropriate methods have not been developed. In this study, we focused on two biological products for toxicological assessment: human induced pluripotent stem cells (human iPS cells) and long noncoding RNAs [15]. Human iPS cells possess the abilities of self-renewal and differentiation into multiple cell types. Long noncoding RNAs are an important class of pervasive non-protein-coding transcripts involved in the molecular mechanisms associated with responses to chemicals. Here, we identified four long noncoding RNAs (TUG1, GAS5, FAM222-AS1, and SNHG15) that were significantly increased in response to typical chemical stresses (oxidative stress, heavy metal stress, and protein synthesis stress) in neural stem cells derived from human iPS cells. We propose that these long noncoding RNAs have the potential to be surrogate indicators of chemical stress responses in human iPS cells.

Katahira shows the interaction of TLS with noncoding RNA. Translocated in liposarcoma (TLS), also known as fused in liposarcoma (FUS), is a multifunctional DNA/RNA-binding protein that is implicated in numerous fatal neurodegenerative diseases. It was also revealed that the

TLS binds to the noncoding RNA transcribed from the upstream region of the cyclin D1 gene and represses transcription of the gene [13]. The sequence of the full-length noncoding RNA was determined and the RNA regions responsible for binding of TLS were identified [6]. The interaction of TLS with the noncoding RNA was analyzed by NMR. The interaction was also analyzed by fluorescence spectroscopy. These studies would provide the basis how TLS exerts the activity of the repression of the transcription.

Hitachi demonstrates tight spatiotemporal regulation of both specification of cell differentiation lineage and cell cycle withdrawal by long noncoding RNA in skeletal muscle cells. Recent findings have revealed that thousands of long non-coding RNAs are produced from the genome including cis-regulatory regions. Here, we identified and characterized a novel long noncoding RNA derived from the 5'-promoter region of myogenin gene, one of critical regulators of differentiation of skeletal muscle cells (called myogenesis), and designated it as Myoparr (myogenin promoter associated myogenic regulatory long noncoding RNA). We showed that Myoparr coordinately regulates both specification of cell differentiation lineage and cell cycle withdrawal. Taken together, our findings indicate that promoter-associated long noncoding RNA, Myoparr, is a novel key regulator of myogenic cell differentiation by determining the timing of important events during myogenesis.

Iwashita presents identification of long noncoding RNAs associated with aging by microarray analysis. Alteration of gene expression is one of the important aspects of aging and can become a molecular cause of irreversible decline of tissue function. Previous research demonstrated that a large number of protein-coding genes were differentially-expressed in various aging contexts. However, expression dynamism of long noncoding RNAs that are not translated into proteins are not fully elucidated. To identify aging-associated long noncoding RNAs from published mouse transcriptome data, we used aging-related keywords as a query and retrieved five microarray data sets from NCBI Gene Expression Omnibus (GEO) after quality control of transcriptome data. Comparison of the expression before and after aging revealed that long noncoding RNAs detected by five probes were differentially expressed in more than three data sets and sequences of these probes were clustered in the 700 kb region of Chromosome 1. These results suggest that our method for transcriptome mining is efficient to give a novel annotation of expression association to long noncoding RNAs with unknown function.

Yamashita shows current status of molecular dynamics simulations of RNA. Due to recent advances in computer hardware, all-atom molecular dynamics (MD) simulation (with an explicit water model) has become a useful tool in providing atomic-level dynamic information of biopolymers (especially proteins). For example, we successfully applied MD simulations to antibody systems and characterized antigen-antibody interactions from the perspective of structure and energy [16].

Although the MD simulation of RNA has more difficult

factors than that of protein, the MD simulation can also provide a detailed insight into RNA. For short single-stranded RNAs, we showed that the stabilities of the stacking structures were largely dependent on the size of the base by using MD simulations [17]; the stacking of purine RNA was shown to be much more stable than that of pyrimidine RNA because purine moieties have larger  $\pi$ - $\pi$  interactions than pyrimidine moieties. In our on-going studies, MD simulations of larger RNA systems (including RNA-protein complexes) supplement additional valuable data. The accuracy of MD simulations is essentially dependent on the force field [18, 19]. Recently, the force field of RNA has been improved significantly [20]. By such a fundamental effort, MD simulations can be widely used to clarify atomic-level dynamics and to complement biochemical and

structural experiments of RNAs (and RNA-protein complexes).

### 3. Conclusion

The presentations of the meeting cover variety of biomedical sciences to indicate significance of long noncoding RNAs in each field (Figure 1). The achievements of the meeting is to make sure a point of view of long noncoding RNA in living cells. It should be in complexity. We will be able to analyze complicated phenomena of lncRNA biological actions and might have resolved the veiled rules. We have already made appointment of the meeting coming May in 2018. It is certain that we will have more fruitful activity there in next year meeting.

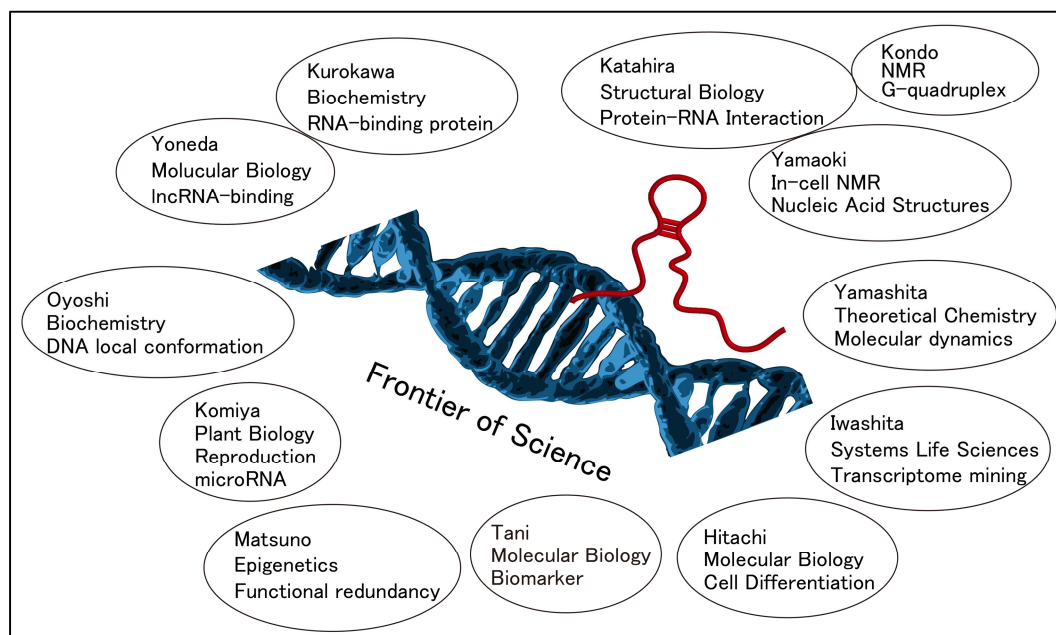


Figure 1. Summary of the Talks.

### Funding

Reina Komiya: This work was supported by JST, PRESTO grant number JPMJPR17Q3 and by JSPS KAKENHI grant number 17H05608; Keiko Kondo: This work was supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan [#18K14636 and #18H04550]; Ryoma Yoneda: This work was supported in part by KAKENHI from JSPS to MEXT of Japan (17K18065); Yudai Yamaoki: This research was supported by a JSPS grant (#17K14515, #17H05878 and #18H04550) and a JSPS Research Fellow Program grant; Riki Kurokawa

This work was supported in part by a Grant-in-aid for "Support Project of Strategic Research Center in Private Universities" from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) to Saitama Medical University Research Center for Genomic Medicine, and by KAKENHI from JSPS to MEXT of Japan (KibanB nos22390057; 25293073, KibanC 18K06939); Takanori Oyoshi:

This work was supported by a JGC-S Scholarship Foundation, a Grant-in-Aid for Scientific Research (C) (No. 17K05930 to T.O.) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan; Hidenori Tani: This work was supported by a JSPS Grant-in-Aid for Young Scientists (B), grant number 26810100; Masato Katahira: This research was supported by MEXT grants (#18H04550 and #18K19397); Keisuke Hitachi: This work was supported by JSPS KAKENHI (25860151, 16K08599, and 17K08646); Takefumi Yamashita: This research used computational resources of the TSUBAME through the HPCI System Research Project (hp170130, hp180044), and of the Research Center for Computational Science, Okazaki, Japan. T.Y. acknowledges support from the Council for Science, Technology and Innovation (CSTI), the Cross-ministerial Strategic Innovation Promotion (SIP) program, "Structural Materials for Innovation" (Funding agency: JST), Scientific Research (JSPS KAKENHI (S)15H05752, (C)15KT0103, (C)18K05025), and Innovative Drug Discovery Infrastructure through Functional Control of Bio-molecular Systems (hp180191).

## Acknowledgements

Ryota Yagi at the Department of Chemistry, Graduate School of Science, Shizuoka University, has experimental contribution to the project; Sayaka Okuda<sup>1,2</sup>, Kaoru Nakamura<sup>1</sup>, Motohide Aoki<sup>2</sup>, Tomonari Umemura<sup>2</sup> (<sup>1</sup>Environmental Management Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 16-1, Onogawa, Tsukuba, Ibaraki, Japan, <sup>2</sup>Department of Molecular Sciences, School of Life Sciences, Tokyo University of Pharmacy and Life Sciences, 1432-1 Horinouchi, Hachioji, Tokyo, Japan), have significant experimental contribution to the project; Kunihiro Tsuchida, at the Division for Therapies against Intractable Diseases, Institute for Comprehensive Medical Science, Fujita Health University, has significant experimental contribution to the project; Haruhiko Sugimura, at the Department of Tumor Pathology, Hamamatsu University School of Medicine, Japan, has significant experimental contribution to the project.

## References

- [1] Komiya R, Ohyanagi H, Niihama M, Watanabe T, Nakano M, Kurata N & Nonomura K (2014) Rice germline-specific Argonaute MEL1 protein binds to phasiRNAs generated from more than 700 lincRNAs. *Plant J* 78, 385-397, doi: 10.1111/tj.12483.
- [2] Komiya R (2017) Biogenesis of diverse plant phasiRNAs involves an miRNA-trigger and Dicer-processing. *J Plant Res* 130, 17-23, doi: 10.1007/s10265-016-0878-0.
- [3] Takahama K, Takada A, Tada S, Shimizu M, Sayama K, Kurokawa R & Oyoshi T (2013) Regulation of telomere length by G-quadruplex telomere DNA- and TERRA-binding protein TLS/FUS. *Chemistry & biology* 20, 341-350, doi: 10.1016/j.chembiol.2013.02.013.
- [4] Takahama K & Oyoshi T (2013) Specific binding of modified RGG domain in TLS/FUS to G-quadruplex RNA: tyrosines in RGG domain recognize 2'-OH of the riboses of loops in G-quadruplex. *J Am Chem Soc* 135, 18016-18019, doi: 10.1021/ja4086929.
- [5] Kondo K, Mashima T, Oyoshi T, Yagi R, Kurokawa R, Kobayashi N, Nagata T & Katahira M (2018) Plastic roles of phenylalanine and tyrosine residues of an RGG motif of TLS/FUS in binary and ternary complex formation with the G-quadruplex structures of telomeric DNA and TERRA. *Sci Rep* 8, 2864, doi: 10.1038/s41598-018-21142-1.
- [6] Yoneda R, Suzuki S, Mashima T, Kondo K, Nagata T, Katahira M & Kurokawa R (2016) The binding specificity of Translocated in LipoSarcoma/FUsed in Sarcoma with lncRNA transcribed from the promoter region of cyclin D1. *Cell & bioscience* 6, 4, doi: 10.1186/s13578-016-0068-8.
- [7] Hansel R, Luh LM, Corbeski I, Trantirek L & Dotsch V (2014) In-cell NMR and EPR spectroscopy of biomacromolecules. *Angew Chem Int Ed Engl* 53, 10300-10314, doi: 10.1002/anie.201311320.
- [8] Giassa IC, Rynes J, Fessl T, Foldynova-Trantirkova S & Trantirek L (2018) Advances in the cellular structural biology of nucleic acids. *FEBS letters*, doi: 10.1002/1873-3468.13054.
- [9] Yamaoki Y, Kiyoishi A, Miyake M, Kano F, Murata M, Nagata T & Katahira M (2018) The first successful observation of in-cell NMR signals of DNA and RNA in living human cells. *Phys Chem Chem Phys* 20, 3109-3117, doi: 10.1039/c7cp05188c.
- [10] Cui W, Yoneda R, Ueda N & Kurokawa R (2018) Arginine methylation of translocated in liposarcoma (TLS) inhibits its binding to long noncoding RNA, abrogating TLS-mediated repression of CBP/p300 activity. *The Journal of biological chemistry*, doi: 10.1074/jbc.RA117.000598
- [11] Thandapani P, O'Connor TR, Bailey TL & Richard S (2013) Defining the RGG/RG motif. *Molecular Cell* 50, 613-623, doi: 10.1016/j.molcel.2013.05.021.
- [12] Sun X, Haider Ali MSS & Moran M (2017) The role of interactions of long non-coding RNAs and heterogeneous nuclear ribonucleoproteins in regulating cellular functions. *The Biochemical journal* 474, 2925-2935, doi: 10.1042/BCJ20170280.
- [13] Wang X, Arai S, Song X, Reichart D, Du K, Pascual G, Tempst P, Rosenfeld MG, Glass CK & Kurokawa R (2008) Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature* 454, 126-130, doi: nature06992 [pii]10.1038/nature06992.
- [14] Wang X, Schwartz JC & Cech TR (2015) Nucleic acid-binding specificity of human FUS protein. *Nucleic Acids Research* 43, 7535-7543, doi: 10.1093/nar/gkv679.
- [15] Tani H, Okuda S, Nakamura K, Aoki M & Umemura T (2017) Short-lived long non-coding RNAs as surrogate indicators for chemical exposure and LINC00152 and MALAT1 modulate their neighboring genes. *PLoS One* 12, e0181628, doi: 10.1371/journal.pone.0181628.
- [16] Yamashita T (2018) Toward rational antibody design: recent advancements in molecular dynamics simulations. *Int Immunol* 30, 133-140, doi: 10.1093/intimm/dxx077.
- [17] Yamashita T (2016) Towards Physical Understanding of Molecular Recognition in the Cell: Recent Evolution of Molecular Dynamics Techniques and Free Energy Theories. *Biomedical Sciences* 2, 34-47 doi: 10.11648/j.bs.20160205.11.
- [18] Yamashita T, Ueda A, Mitsui T, Tomonaga A, Matsumoto S, Kodama T & Fujitani H (2015) The Feasibility of an Efficient Drug Design Method with High-Performance Computers. *Chemical and Pharmaceutical Bulletin* 63, 147-155, doi: 10.1248/cpb.c14-00596.
- [19] Yamashita T (2015) Improvement in Empirical Potential Functions for Increasing the Utility of Molecular Dynamics Simulations, *JPS Conf Proc*, 5, 010003, doi:10.7566/JPSCP.5.010003
- [20] Tan D, Piana S, Dirks RM & Shaw DE (2018) RNA force field with accuracy comparable to state-of-the-art protein force fields. *Proceedings of the National Academy of Sciences of the United States of America* 115, E1346-E1355, doi: 10.1073/pnas.1713027115.