The search for the determinants of insertional RNA editing in *Physarum polycephalum* mitochondrial RNAs

Ralf Bundschuh
Department of Physics
Department of Chemistry&Biochemistry
Division of Hematology, Department of Internal Medicine
Center for RNA Biology
191 West Woodruff Av
Columbus, OH 43210, USA
bundschuh@mps.ohio-state.edu

ABSTRACT
The mitochondrial genome of the slime mold *Physarum polycephalum* is characterized by pervasive insertional RNA editing. Here, we provide an overview of the quest to identify the sequence determinants of the editing positions in this system. This work was presented at PhysNet 2015.

Categories and Subject Descriptors
J.3 [**Life and Medical Sciences**]: Biology and genetics

Keywords
Physarum polycephalum, RNA editing, computational biology

1. INTRODUCTION
The mitochondrion of *Physarum polycephalum* shows one of the highest rates of RNA editing discovered to date [14, 10]. While many types of RNA editing in its mitochondrially encoded genes have been reported, the by far most frequent editing event is the insertion of individual Cs. These insertions occur with extremely high accuracy at predetermined positions in the RNAs which have an average separation of about 25bp for protein coding RNAs and 40bp for structural RNAs [6]. This immediately raises the question about the mechanism by which *Physarum* identifies these positions. Biochemically it is known that the insertions occur co-transcriptionally [17] and that 9bp of DNA up- and downstream of the insertion sites are sufficient to ensure that editing takes place [4]. However, while there is some preference for C insertions to follow a purine-pyrimidine combination [11], this is neither an absolute rule nor would this signal provide enough specificity to determine the insertion positions. Additional sequence determinants are elusive. In fact, one of the problems of insertional editing is that it renders conventional gene finding, namely looking for long open reading frames, useless. As a result, only twelve protein coding genes were annotated when the mitochondrial genome was published [15] and for only six of those the editing sites were known (through sequencing the RNA), a number that should be compared to the 44 genes in the mitochondrial genome of the slime mold *Dictyostelium discoideum* [12] of similar length.

2. PROGRESS
Initially, we focused on extending the number of genes with known editing sites by developing computational methods that would allow the identification of protein coding regions within the mitochondrial genome as well as the prediction of the editing sites within these genes. We used alignment approaches to compare known mitochondrial protein sequences from other organisms to the mitochondrial genome of *Physarum polycephalum* while taking into account the existence of C insertions [2, 7] and also modified classical de novo gene finding Hidden Markov Models to allow for the existence of frame shifting C insertions [1]. The logic was that even a somewhat approximate prediction of the editing sites using computational methods could be used to design primers that then could be used in experiments to sequence the actual edited RNAs and thus determine the true editing sites in these genes. With continually improving computational approaches and wet lab experiments in Jonatha Gott's group, this resulted in seven additional genes for which the editing sites had been determined and ten additional genes, the genomic location of which had been identified and editing sites predicted but for which experimental verification was lacking. In the process, we also identified the first example of nucleotide deletion editing in *Physarum*.

At this time high throughput sequencing became available and we identified collaborators that could construct a library from *Physarum polycephalum* mitochondrial RNA extracted by Jonatha Gott. By aligning the sequencing reads from this library to the mitochondrial genome, every C insertion became immediately evident; with the missing Cs (and other bases including G and A insertions which had not been seen in *Physarum* before) inserted, gene finding became trivial. Thus, this experiment delivered the entire edited mitochondrial transcriptome of *Physarum* including 1255 C insertion bringing the total protein coding genes to 39, i.e., very comparable to *Dictyostelium* [3]. However, even with 1255 inser-

BICT 2015, December 03-05, New York City, United States

DOI 10.4108/eai.3-12-2015.2262393

tion sites, no sequence pattern in the vicinity of editing sites emerged. Even when comparing the *Physarum* mitochondrial genome with the recently determined closely related mitochondrial genome of *Didymium iridis* [16, 8] did not reveal any additional conservation in the vicinity of the editing sites as might be expected if the "editing signal" is encoded in these nucleotides in addition to the protein sequence [5].

3. CONCLUSIONS

While we have learned a lot about the mitochondrial genome of *Physarum polycephalum*, the crucial question of how it knows where to insert the added Cs remains open. Alignments of mitochondrial RNAs to the recently determined nuclear genome [13] of *Physarum polycephalum* reveal no hint of sequences coinciding with editing sites; neither did the RNA-Seq experiment show any sign of antisense (and thus possible guide) RNAs. Mapping of the mitochondrial RNA-Seq reads not alignable to the mitochondrial genome reveals a fairly large number of snoRNAs [13] but none of these appear to target editing sites based on their sequences.

An avenue to address the challenge of finding the sequence determinants of RNA editing further with computational tools might be to sequence the mitochondrial genome of another editing slime mold, such as, e.g., *Stemonitis flavogenita*, for which the cox1 gene has been sequenced, shows editing, but also a lot less conservation with *Physarum* [9], which might make it easier to identify additional conservation driven by the need to specify the editing site positions than in the comparison of the relatively closely related organisms *Physarum polycephalum* and *Didymium iridis*. Even more useful would of course be an experimental system that would allow probing editing of RNAs transcribed from DNA sequences with site-directed mutations.

4. ACKNOWLEDGMENTS

The quest for understanding insertional RNA editing in *Physarum polycephalum* has involved many collaborators and students (C. Beargie, C, Chen, D. Frankhouser, H.Y. Lee, T. Liu, and P. Nürnberg). However, the person who introduced us to this fascinating problem at a coffee break at the 2001 Rustbelt RNA meeting and has been our collaborator throughout the entire time is Jonatha Gott. This work has been supported by the National Science Foundation through grants DMR-0404615 and DMR-0706002 and through a Pelotonia fellowship to Cai Chen.

5. REFERENCES

[1] C. Beargie, T. Liu, M. Corriveau, H. Y. Lee, J. Gott, and R. Bundschuh. Genome annotation in the presence of insertional RNA editing. *Bioinformatics*, 24(22):2571–2578, Nov 2008.

[2] R. Bundschuh. Computational prediction of RNA editing sites. *Bioinformatics*, 20(17):3214–3220, Nov 2004.

[3] R. Bundschuh, J. Altmuller, C. Becker, P. Nurnberg, and J. M. Gott. Complete characterization of the edited transcriptome of the mitochondrion of Physarum polycephalum using deep sequencing of RNA. *Nucleic Acids Res.*, 39(14):6044–6055, Aug 2011.

[4] E. M. Byrne and J. M. Gott. Cotranscriptional editing of Physarum mitochondrial RNA requires local features of the native template. *RNA*, 8(9):1174–1185, Sep 2002.

[5] C. Chen, D. Frankhouser, and R. Bundschuh. Comparison of insertional RNA editing in Myxomycetes. *PLoS Comput. Biol.*, 8(2):e1002400, 2012.

[6] J. M. Gott. RNA editing in *Physarum polycephalum*. In B. L. Bass, editor, *RNA Editing*, pages 20–37. Oxford University Press, Oxford, UK, 2001.

[7] J. M. Gott, N. Parimi, and R. Bundschuh. Discovery of new genes and deletion editing in Physarum mitochondria enabled by a novel algorithm for finding edited mRNAs. *Nucleic Acids Res*, 33(16):5063–5072, 2005.

[8] P. G. Hendrickson and M. E. Silliker. RNA editing in six mitochondrial ribosomal protein genes of Didymium iridis. *Curr. Genet.*, 56(3):203–213, Jun 2010.

[9] T. L. Horton and L. F. Landweber. Evolution of four types of RNA editing in myxomycetes. *RNA*, 6(10):1339–1346, Oct 2000.

[10] T. L. Horton and L. F. Landweber. Rewriting the information in DNA: RNA editing in kinetoplastids and myxomycetes. *Curr Opin Microbiol*, 5(6):620–626, Dec 2002.

[11] D. Miller, R. Mahendran, M. Spottswood, H. Costandy, S. Wang, M. L. Ling, and N. Yang. Insertional editing in mitochondria of Physarum. *Semin Cell Biol*, 4(4):261–266, Aug 1993.

[12] S. Ogawa, R. Yoshino, K. Angata, M. Iwamoto, M. Pi, K. Kuroe, K. Matsuo, T. Morio, H. Urushihara, K. Yanagisawa, and Y. Tanaka. The mitochondrial DNA of Dictyostelium discoideum: complete sequence, gene content and genome organization. *Mol Gen Genet*, 263(3):514–519, Apr 2000.

[13] P. Schaap and et al. *in review*.

[14] H. C. Smith, J. M. Gott, and M. R. Hanson. A guide to RNA editing. *RNA*, 3(10):1105–1123, Oct 1997.

[15] H. Takano, T. Abe, R. Sakurai, Y. Moriyama, Y. Miyazawa, H. Nozaki, S. Kawano, N. Sasaki, and T. Kuroiwa. The complete DNA sequence of the mitochondrial genome of Physarum polycephalum. *Mol Gen Genet*, 264(5):539–545, Jan 2001.

[16] S. J. Traphagen, M. J. Dimarco, and M. E. Silliker. RNA editing of 10 Didymium iridis mitochondrial genes and comparison with the homologous genes in Physarum polycephalum. *RNA*, 16(4):828–838, Apr 2010.

[17] L. M. Visomirski-Robic and J. M. Gott. Insertional editing of nascent mitochondrial RNAs in Physarum. *Proc Natl Acad Sci U S A*, 94(9):4324–4329, Apr 1997.