

Unravelling organelle genome transcription using publicly available RNA-sequencing data

Matheus Sanita Lima¹, David Roy Smith¹

1. Department of Biology, University of Western Ontario, London, Ontario, Canada

Abstract:

Background: Organelle genomes are the most sequenced type of chromosome, given their relative compactness and importance on molecular research. Additionally, next generation sequencing (NGS) has generated unprecedented amounts of genomic and transcriptomic data that can be used for organelle genome evolution studies. Indeed, NGS has been influencing the organelle genomics field by being a major facilitator in the current sequencing efforts. However, most of the NGS data, particularly the RNA sequencing (RNA-seq) data, is still used only for the study of cell nucleus, being the organelle reads even discarded in such investigations. Here, we used this untapped data source to investigate the transcription of publicly available organelle genomes in plastid-bearing protists. We mapped raw transcripts over genomes of 116 protist species to verify the utility of RNA-seq data in unraveling the organelle genome transcription architecture.

Results: 77 species had their organelle genomes nearly entirely (> 85%) mapped by transcripts despite of the amount of non-coding DNA (intergenic regions and introns) they had. These genomes come from various protist groups – including the volvocine algae – that exhibit diverse genome structure and content. Yet, all of them appear to go full transcription.

Conclusions: Publically available (and unused) RNA-seq data can be employed to investigate organelle genome transcription. Even though organelle genomes can exhibit significant portions of non-coding DNA, these regions are still actively transcribed. That allowed us to show that nearly entire (when not entire) organelle genomes can be recovered from RNA-seq data. This represents a valuable alternative to determine organelle genomes. Beyond that, organelles seem to be producing a plethora of non-coding RNAs that can regulate fundamental processes such as organelle differentiation in land plants and trophic mode choice in mixotrophs.