MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2016) **16**, 1279–1286 doi: 10.1111/1755-0998.12585

NEWS AND VIEWS

OPINION

The (in)complete organelle genome: exploring the use and nonuse of available technologies for characterizing mitochondrial and plastid chromosomes

MATHEUS SANITÁ LIMA,[1] LAURA C. WOODS,[1] MATTHEW W. CARTWRIGHT and DAVID ROY SMITH
Department of Biology, University of Western Ontario, London, Ontario, Canada, N6A 5B7

Abstract

Not long ago, scientists paid dearly in time, money and skill for every nucleotide that they sequenced. Today, DNA sequencing technologies epitomize the slogan 'faster, easier, cheaper and more', and in many ways, sequencing an entire genome has become routine, even for the smallest laboratory groups. This is especially true for mitochondrial and plastid genomes. Given their relatively small sizes and high copy numbers per cell, organelle DNAs are currently among the most highly sequenced kind of chromosome. But accurately characterizing an organelle genome and the information it encodes can require much more than DNA sequencing and bioinformatics analyses. Organelle genomes can be surprisingly complex and can exhibit convoluted and unconventional modes of gene expression. Unravelling this complexity can demand a wide assortment of experiments, from pulsed-field gel electrophoresis to Southern and Northern blots to RNA analyses. Here, we show that it is exactly these types of 'complementary' analyses that are often lacking from contemporary organelle genome papers, particularly short 'genome announcement' articles. Consequently, crucial and interesting features of organelle chromosomes are going undescribed, which could ultimately lead to a poor understanding and even a misrepresentation of these genomes and the genes they express. High-throughput sequencing and bioinformatics have made it easy to sequence and assemble entire chromosomes, but they should not be used as a substitute for or at the expense of other types of genomic characterization methods.

Keywords: genome announcement, genome report, next-generation sequencing, RNA-Seq, transcriptomics

Received 3 March 2016; revision received 23 June 2016; accepted 23 June 2016

Introduction

Sequencing an entire organelle genome was once a long and arduous task. Now it is commonplace (Smith 2016a). With the advent of next-generation sequencing (NGS) technologies and sophisticated user-friendly bioinformatics software, scientists of all stripes can sequence and assemble dozens of organelle genomes in a few days or less, and often for very little money (Gan *et al*. 2014; Mariac *et al*. 2014; Tang *et al*. 2014). This kind of progress is great. More sequences mean more data for comparative studies and a better understanding of organelle genome evolution. Organelle sequences are used in a wide range of disciplines and analyses (Smith 2016a), from medicine to anthropology to phylogenetics, and have helped resolve major scientific questions, including the origins and diversification of eukaryotic life (Gray 2012; Keeling 2013). But accurately characterizing a genome and the information it encodes requires much more than just DNA sequencing and bioinformatics analyses, and organelle genomes are no exception.

Mitochondria and plastids harbour some of the most complex genomes and gene expression systems of any genetic compartment (Smith & Keeling 2015). Take, for instance, the mitochondrial DNA (mtDNA) of the ichthyosporean *Amoebidium parasiticum*, which comprises several hundred small (0.3–8.3 kb) linear chromosomes (Burger *et al*. 2003), or the plastid DNAs (ptDNAs) of

Correspondence: David Roy Smith, Fax:519 661 3935;
Email: dsmit242@uwo.ca.

[1]Equal contribution.

peridinin dinoflagellate algae, such as *Symbiodinium minutum*, which are distributed across multiple minicircular (~2.5 kb) molecules that can differ in copy number throughout the life cycle (Mungpakdee *et al.* 2014; Dorrell & Howe 2015). Equally as impressive is the giant (>11 000 kb) multichromosomal mtDNA of the flowering plant *Silene conica* (Sloan *et al.* 2012) and the tiny 6-kb mtDNA of *Plasmodium falciparum* (Feagin 1992), which is organized as a linear concatemer (Wilson & Williamson 1997).

In addition to being structurally diverse, organelle genomes can undergo massive amounts of post-transcriptional processing (Smith & Keeling 2016). In the euglenozoan *Diplonema papillatum*, for example, *cox1* is transcribed from nine different mitochondrial chromosomes, giving nine partial transcripts that come together through trans-splicing to form a mature and intact mRNA (Vlcek *et al.* 2010). In the organelles of dinoflagellates, eleven of the twelve possible types of substitutional RNA editing (A-to-C, A-to-G, etc.) have been observed as well as a slew of other types of transcriptional modifications (Waller & Jackson 2009; Mungpakdee *et al.* 2014; Dorrell & Howe 2015). And this is to say nothing about nonstandard genetic codes (Knight *et al.* 2001), translational slippage (Masuda *et al.* 2010) and ribosomal jumping (Lang *et al.* 2014) within organelle systems.

Given this complexity, DNA sequencing data alone are often not sufficient to infer the true architecture and the resulting gene products of organelle genomes (Smith 2016a). Consequently, some of the most informative organelle genome analyses use a combination of different techniques, in addition to DNA sequencing and bioinformatics, to characterize the chromosome(s). For example, determining the mitochondrial genomic architecture of *D. papillatum* involved cloning, Sanger sequencing, high-throughput DNA and RNA sequencing, traditional and reverse transcription–PCR, DNA digestions, pulsed-field gel electrophoresis, and Southern and Northern blotting experiments, and still some of the chromosomes, coding regions and gene products remain undefined (Marande *et al.* 2005; Vlcek *et al.* 2010; Valach *et al.* 2014). A similar array of techniques was used to describe the mitochondrial and plastid genomes of dinoflagellates (Nash *et al.* 2007; Barbrook *et al.* 2012; Jackson *et al.* 2012), and new organelle genomic features and peculiarities are still being uncovered within this lineage (Mungpakdee *et al.* 2014; Dorrell & Howe 2015). Although the *P. falciparum* mtDNA was completely sequenced more than twenty years ago (Feagin 1992; Wilson & Williamson 1997), it has taken another twenty years of detailed RNA work to resolve the large and small subunit rRNA genes, which are fragmented and scrambled into ~25 distinct coding modules (Feagin *et al.* 2012).

Improvements to traditional molecular biology techniques and the development of new technologies have only made it easier to characterize complex organelle genomes and their modes of repair, replication and expression. State-of-the-art microscopes and cameras can now provide ultra-high-resolution images of organelles and their nucleoids, which in turn is giving new insights into mitochondrial and plastid DNA maintenance (Golczyk *et al.* 2014; Oldenburg & Bendich 2015). Advanced PCR, gel electrophoresis and blotting methods are exposing the dynamic and multifarious nature of organelle chromosomes (Lewis *et al.* 2015) and their resulting transcripts (Wende *et al.* 2014). High-throughput transcriptomics and proteomics are also helping to disentangle the genetic information within organelles (Jedelský *et al.* 2011; Marková *et al.* 2015), as are new methods for exploring DNA–protein interactions, such as chromatin immunoprecipitation (Yagi *et al.* 2012). But many of these methods are technically challenging, time-consuming and expensive, and unlike NGS, they cannot be easily outsourced. Nevertheless, as the rate of organelle genome sequencing increases, one might expect the use of 'complementary' characterization techniques, such as pulsed-field or two-dimensional gel electrophoresis (Slater *et al.* 1998), to also increase. However, this does not appear to be true. As described below, a scan of the recent literature reveals that apart from DNA sequencing and bioinformatics, there is a paucity of experimental data in many contemporary organelle genome studies, with some notable exceptions.

A snapshot of the experimental methods used in contemporary organelle genome papers

The first completely sequenced mitochondrial genomes (human and mouse) were published more than 30 years ago, using a Sanger sequencing approach (Anderson *et al.* 1981; Bibb *et al.* 1981). These feats were soon followed by the entire plastid genome sequencing of tobacco and the liverwort *Marchantia polymorpha* (Ohyama *et al.* 1986; Shinozaki *et al.* 1986). Over the ensuing years, organelle genome data steadily accumulated from diverse species, and by the turn of the millennium, which brought improvements to automated capillary Sanger sequencing, new organelle DNA sequences were being published every month or faster (Smith 2016a). Around 2010, following the advent of massively parallel high-throughput sequencing (NGS), the production and publication rate of organelle genome data skyrocketed, with hundreds—and more recently thousands—of sequences appearing annually (Smith 2016a).

Indeed, a PubMed search of scientific articles indexed in MEDLINE retrieved 2601 organelle genome papers

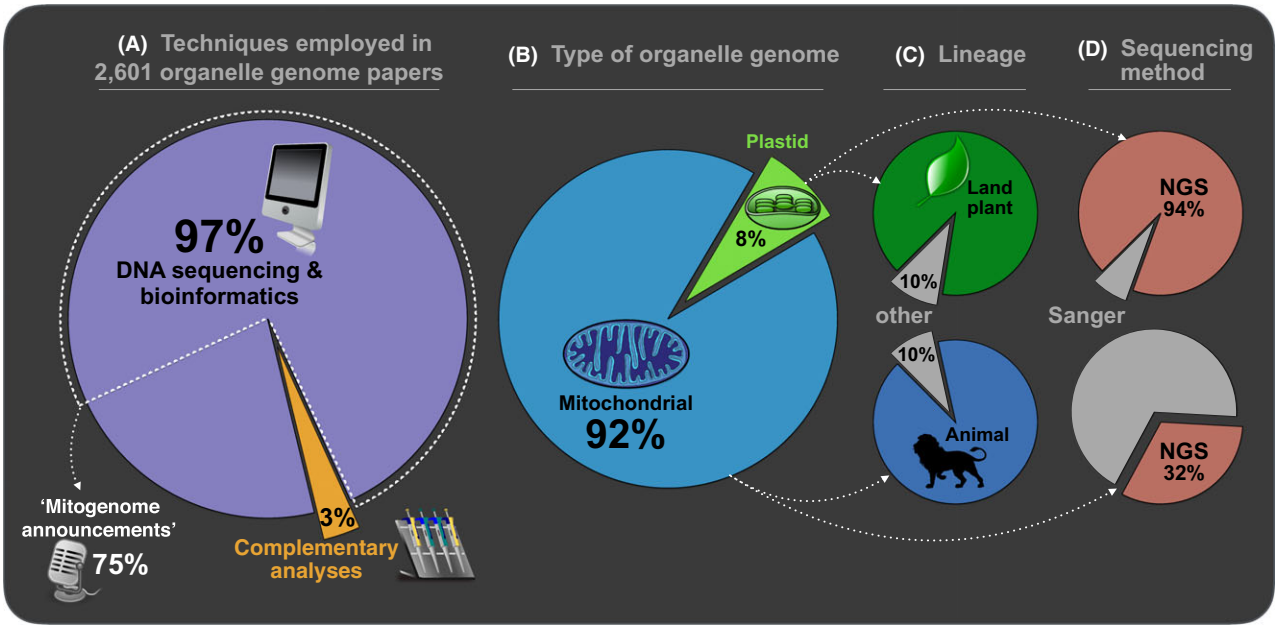

Fig. 1 A survey of organelle genome papers published in the last half-decade. Organelle genome papers indexed in MEDLINE were collected via the PubMed Advanced Search Builder at the National Center for Biotechnology Information website using the following keyword combinations: 'entire chloroplast/plastid/mitochondrial DNA/genome', 'complete chloroplast/plastid/mitochondrial DNA/genome', 'whole chloroplast/plastid/mitochondrial DNA/genome' and 'full chloroplast/plastid/mitochondrial DNA/genome'. We linked the different keyword combinations with OR (instead of AND), and did not use quotation marks, in order to retrieve as many hits as possible. We limited the search field to 'title/abstract', and the date range from 1 January 2010 to 1 November 2015. We scanned the results by eye, removing any obviously spurious hits. Altogether, we retrieved 2601 organelle genome papers (including 1781 Mitogenome Announcements), only 3% of which included complementary analyses (A). Approximately 92% and 8% of the collected articles were mitochondrial and plastid genome papers, respectively (B). The former comprised mostly animal mtDNAs, and the latter were primarily plant ptDNAs (C). Most of the ptDNAs were sequenced using next-generation sequencing (NGS) methods (or a combination of NGS and Sanger), whereas two-thirds of the mtDNAs were sequenced using a Sanger sequencing-only approach (D). Note: 'Lineage' (C) and 'Sequencing Method' (D) statistics do not include Mitogenome Announcements. See Appendix S1 (Supporting information) for further details.

published between 1 January 2010 and 1 November 2015 (Figure 1; Appendix S1, Supporting Information). About 92% of these papers describe mtDNAs, and 8% represent plastid genomes; these sequence data span a large breadth of eukaryotic diversity, but there is nonetheless an overrepresentation of metazoan mtDNAs and land plant ptDNAs, and a lack of data from many protist lineages (Figure 1; Appendix S1, Supporting information). Although some of these trends have been documented and discussed before (Smith & Keeling 2015; Smith 2016a), no one has yet surveyed the range of methods commonly employed in organelle genome studies.

We scanned the materials and methods from organelle genome papers published since 2010 (Fig. 1), recorded the techniques used to characterize the chromosomes and then placed these techniques into one of the following three broad categories. (i) 'DNA extraction, amplification and sequencing'. (ii) 'Bioinformatics', which includes, for example, genome assembly and annotation, molecular sequence alignments, phylogenetic analyses and estimations of genetic diversity. And (iii) 'complementary experiments', comprising any experiments not related to DNA sequencing or bioinformatics, such as restriction endonuclease digestion, gel electrophoresis, nucleotide blotting, real-time PCR, RNA analyses/sequencing or DNA imaging. Preparatory experiments for DNA sequencing, such as cloning or gel electrophoresis of PCR products prior to Sanger sequencing, were not considered complementary techniques.

Only a small fraction (3%) of organelle genome studies carried out over the past 5 years employed complementary experiments. In other words, most of the studies (97%) used only DNA sequencing and bioinformatics to characterize the chromosomes. Among the papers that did contain additional analyses, quantitative PCR was one of the most commonly employed experiments. Rarely did any of the papers include a detailed examination of organelle gene expression or chromosome structure. Instead, analyses relied upon bioinformatics software for RNA and protein predictions and for determining the size, conformation and number of chromosomes.

The compiled articles stem from an eclectic list of mostly life science journals, spanning an assortment of subdisciplines (e.g. genomics, evolution and molecular biology) and impact factors (Appendix S1, Supporting information). However, more than three quarters of the papers come from a single journal: *Mitochondrial DNA* (formerly called *DNA Sequence*, 1990–2008), which is published by Taylor & Francis and has a Thomson Reuters impact factor of 1.2 (2014). Most of the articles collected from *Mitochondrial DNA* are 'Mitogenome Announcements', short (~500 words) fast-tracked reports describing organelle genome sequences, which do not contain complementary analyses and mostly describe animal mtDNAs (Appendix S1, Supporting information). Other papers that we collected were similar to 'Mitogenome Announcements' in that they were brief reports highlighting a genome sequence and its GenBank accession, including papers from the journal *Genome Announcements*, published by the American Society for Microbiology, as well as Genome Reports from the journals *Genome Biology and Evolution*. Altogether, short genome announcement-type articles (<2000 words) represented ~75% of the papers that we surveyed.

The good, the bad and the ugly of organelle genomics

The publication of more than 2600 organelle genome articles over the past half-decade is an impressive achievement and a testament to how far and fast the field of genomics has progressed. (This number is likely even larger given that we could not feasibly capture every organelle genome paper using our PubMed search methods.) Together, these organelle genome data have helped to progress the field of genetics. For example, they have improved our understanding of genomic diversity and gene expression (Fitzgerald *et al.* 2011; Segovia *et al.* 2011), and yielded new insights into the mutational and population genetic processes impacting mtDNA and ptDNA (Hardouin & Tautz 2013). They have also advanced our understanding and/or treatment of human disease (Govindaraj *et al.* 2013), migration (Ning *et al.* 2016) and forensics (Just *et al.* 2015) and led to methodological advancements (Dong *et al.* 2013). But perhaps more than anything else, these data have provided the raw material for countless phylogenetic and population-level studies (Njuguna *et al.* 2013; Taylor *et al.* 2013), refining our view of the origins, evolution and diversity of eukaryotic life.

The efforts of the organelle research community to generate, annotate and describe these genomic data are laudable. And no matter what your opinion about the impact or level of detail to which the authors analysed these genomes, we are better off for having these data. There is no denying, however, that aside from bioinformatics analyses, many published organelle genomes have not been characterized in great detail, including some of those published by the corresponding author of this perspectives piece (e.g. Smith *et al.* 2012; Del Vasto *et al.* 2015). This lack of information about organelle DNA architecture is unfortunate given that some of the most interesting aspects of these genomes are found at the structural rather than the sequence level. The paucity of detailed data on organelle chromosome structure (as discussed further below) has also likely contributed to the popular misconception that mitochondrial and plastid genomes typically exist as intact circular molecules, which is known to be an oversimplification (Bendich 2004, 2010; Oldenburg & Bendich 2015).

What is driving the rapid growth in organelle genomics, and why are some researchers failing to include even the most straightforward experiments in their studies? NGS techniques have streamlined genomics (Gan *et al.* 2014; Mariac *et al.* 2014; Tang *et al.* 2014) and certainly contributed to the massive rise in organelle DNA sequencing and publishing over the past five years (Smith 2016a). But despite these advancements, the majority of the articles examined here (>65%), including many published in the past year, employed Sanger sequencing rather than next-generation methods (Fig. 1; Appendix S1, Supporting information). The continued popularity of Sanger sequencing can be partly explained by the fact that most newly sequenced organelle genomes are animal mtDNAs, which are generally small (<25 kb) and easily amplified using PCR, sometimes with a single set of primers (Cheng *et al.* 1994). In contrast, large organelle genomes (>50 kb), which are not amenable to PCR amplification, are now almost entirely sequenced using next-generation techniques or a combination of NGS and Sanger sequencing (Figure 1; Appendix S1, Supporting information).

Improved sequencing technologies may partly account for the large number of organelle DNAs being sequenced, but they cannot account for why so many investigators are ignoring traditional methods of genome characterization. One reason for the absence of additional analyses could be the growing popularity of 'genome announcement' articles, which serve to highlight a DNA sequence and little else, and by their very nature are too short to permit a thorough description of the sequence (Smith 2016b). These kinds of papers are also fast to prepare and are usually accepted within a few weeks or sooner after the initial submission, thereby catering to the increasing pressure within academia to publish more and publish often (Smith 2016b). In fact, from 2009 to 2015 the proportion of Mitogenome Announcements in the journal *Mitochondrial DNA* rose from 50% to 80% (DeSalle 2016a), leading to the creation in 2016 of a new open-access journal called *Mitochondrial*

DNA Part B: Resources, which is devoted almost entirely to short reports on whole mitochondrial genomes (DeSalle 2016b).

In defence of studies that do not include complementary analyses, many researchers who sequence and publish organelle genomes are not directly interested in or concerned with organelle genome structure or gene expression. Instead, their primary goal is to sequence organelle DNA for use in phylogenetic or population-level studies. In such cases, it might be unreasonable to expect the authors to perform a slew of complementary analyses unrelated to the questions that are being addressed—evolutionary relationships. Likewise, organelle genome sequences are sometimes generated as part of large studies, such as nuclear genome sequencing projects or broad-scale genetic diversity analyses. Again, in these instances it might be asking too much for the researchers to carry out additional analyses that are not directly connected to the project at hand. But whatever the reasons for the lack of complementary experiments in contemporary organelle genome papers, they could be negatively impacting the field of mitochondrial and plastid genomics. Soon, it might become increasingly important to incentivize more thorough analyses of organelle genomes in order to offset some of these potential negative effects.

Limitations and implications of a 'sequence-only' approach to organelle genomics

There are obvious limitations and drawbacks to characterizing an organelle genome using only DNA sequencing data. Yeast mitochondrial genomes, for example, typically assemble as genome-sized circular chromosomes, leading some to assume that these chromosomes have circular conformations in vivo. However, it is now well established that the mtDNAs of yeast, as well as those from other groups, can have much more complex and dynamic conformations than DNA assemblies may suggest, existing (at least in part) as complex multigenomic branched structures (Bendich 1996, 2010; Gerhold *et al.* 2010). Similar findings have come from the ptDNAs of land plants, which typically map as circles but in many instances are found in complex linear-branched forms larger than the size of the genome, similar to those of yeast mtDNAs (Bendich 2004; Oldenburg & Bendich 2016). And there is an assortment of protists that have linear mtDNAs with elaborate telomeres: for example, the linear mitochondrial genomes of the green algae *Chlamydomonas reinhardtii* and *Polytomella capuana* end in single-stranded 3′ overhangs and covalently closed hairpin loops, respectively (Vahrenholz *et al.* 1993; Smith & Lee 2008). The misrepresentation of organelle chromosome conformation is so widespread that some modern biology textbooks still describe mtDNAs and ptDNAs as unit-sized circular genomes (Hartwell *et al.* 2014). Moving forward, elucidating the dynamic structures of organelle chromosomes will require, in the very least, extensive gel electrophoresis work (Oldenburg & Bendich 2016).

On top of providing minimal details about genome architecture, DNA sequencing data give limited insights into organelle transcription and translation. Mitochondria and plastids are veritable circus acts of gene expression (Smith & Keeling 2016). The mtDNAs of most metazoans, fungi and protists have undergone one or more changes to the standard genetic code (Knight *et al.* 2001). Many groups undergo organelle RNA editing, whereby nucleotides are substituted, inserted and/or deleted from transcripts. In the mitochondria of kinetoplastids, such as *Trypanosoma brucei*, uracil insertion/deletion editing can affect up to 90% of the codons in a single protein-coding transcript (Simpson & Shaw 1989). Post-transcriptional editing can be nearly as extreme in the mitochondria and plastids of various land plants and dinoflagellates where nucleotide substitution editing is often rampant (Waller & Jackson 2009; Mungpakdee *et al.* 2014; Dorrell & Howe 2015). Other elaborate types of post-transcriptional processing, such as trans-splicing, transcriptional cleavage and polyadenylation, are also widespread in mitochondria and plastids, and new idiosyncrasies are continually being uncovered (Masuda *et al.* 2010; Lang *et al.* 2014). Sometimes the levels of post-transcriptional editing and processing are so severe that given the DNA sequence alone, it is not possible to distinguish coding from noncoding DNA. In such cases, data at the RNA and/or protein level are crucial to understanding the information encoded in the organelle DNA.

With notable exceptions (e.g. Mercer *et al.* 2011), we still have a poor understanding of organelle gene expression, especially in nonmodel species. But this is poised to change in the near future. There are now thousands of eukaryotic RNA sequencing projects in GenBank's Sequence Read Archive. These publically available data abound with mitochondrial- and plastid-derived reads, most of which are unanalysed and represent an excellent untapped resource for exploring organelle transcription (Smith 2013). Already, scientists have started publishing organelle transcriptome papers (Bundschuh *et al.* 2011; Kolondra *et al.* 2015; Wu *et al.* 2015; Tian & Smith 2016) or begun to include next-generation RNA sequencing data alongside whole organelle genome analyses (Margam *et al.* 2011; Fang *et al.* 2012; Jackson *et al.* 2012). RNA sequencing data may not be a substitute for more sophisticated transcript detection technologies, but they certainly add an additional layer of understanding and well-needed depth to any organelle genome paper.

Moving forward, organelle genome studies need to combine high-throughput sequencing with molecular-biology-focused methods. This combined with information on population genetics and mutation rates, as well as a more unified understanding of cytonuclear interactions will result in some very exciting analyses. And even if these additional data are not of immediate interest to all researchers who sequence organelle genomes, then perhaps a central resource database linking the different types of experimental information for each genome would be useful.

Concluding remarks

The last thing we want to do is discourage scientists from sequencing and publishing organelle genomes, even if they are in the form of a genome announcement. Rather, we want to encourage authors to include more in-depth information about those genomes. And, again, we support the view that more genome sequence data, even if the genomes from which they are derived are not characterized in great detail, are still a scientific asset and better than no data at all. The editor in chief of the journal *Mitochondrial DNA*, Rob DeSalle, recently took such a stance in an eloquent commentary article defending mitochondrial genome papers:

'Publications announcing mtDNA genomes serve an important purpose in science. Access to information should be enhanced whenever we can [sic] and it seems to me that having the information about a newly sequenced mtDNA genome in the literature is an enhancing element. More importantly, an announcement can link the specimen's archival data to a sequence and clarify the provenance of a sequence. In addition, if phylogenetic analysis of the generated sequence is required (as the journal *mtDNA* requires) then the validity of the sequence can be determined by its phylogenetic placement with other known sequences' (DeSalle 2016a). These are all valid points. DeSalle (2016a) ultimately concludes: 'If the incentive of publishing the findings from a novel mtDNA genome is removed ... I fear that the generation of these genomes will be severely slowed and in essence a reachable goal of a mitochondrial/chloroplast DNA genomic database for all organisms on the planet with these genomes will not be realized'.

A database of organelle genome sequences for all eukaryotes is an admirable goal and one that would undoubtedly contribute to the barcoding and resolution of life on Earth. Future innovations in DNA sequencing and bioinformatics will only make it easier to achieve such a goal. But these innovations should not be used as a substitute for or come at the expense of other types of genomic characterization methods.

It is important to remember that most of the greatest contributions from the field of organelle genetics have not necessarily come from the raw genome sequence data themselves but from the complete picture of the organelle, its genome and chromosome(s), and mode of expression, including knowledge of mutation rates, population genetic landscapes and nuclear-encoded organelle-targeted proteins. If researchers had not been striving towards this 'complete' understanding, we may not have seen the development of leading evolutionary theories, such as constructive neutral evolution, which was based largely on studies of organelle post-transcriptional editing and processing (Covello & Gray 1993; Stoltzfus 1999).

We will have to wait and see whether the next 5 years bring as many new mtDNA papers as the previous five, and whether those studies are short genome reports or detailed investigations. Whatever the outcome, the choice to include or not include complementary experiments will likely have a major impact on where the study ultimately gets published. Of the small fraction of papers in our survey that included additional techniques, three quarters were published in a journal with an impact factor >3. Conversely, the vast majority (>80%) of papers that contained only DNA sequencing and bioinformatics data were published in a journal with an impact factor <2. So if you are planning to write an organelle genome paper, there is a lot to think about—or not.

Acknowledgements

DRS is supported by a Discovery Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Anderson S, Bankier AT, Barrell BG *et al*. (1981) Sequence and organization of the human mitochondrial genome. *Nature*, **290**, 457–465.

Barbrook AC, Dorrell RG, Burrows J, Plenderleith LJ, Nisbet RER, Howe CJ (2012) Polyuridylylation and processing of transcripts from multiple gene minicircles in chloroplasts of the dinoflagellate *Amphidinium carterae*. *Plant Molecular Biology*, **79**, 347–357.

Bendich AJ (1996) Structural analysis of mitochondrial DNA molecules from fungi and plants using moving pictures and pulsed-field gel electrophoresis. *Journal of Molecular Biology*, **255**, 564–588.

Bendich AJ (2004) Circular chloroplast chromosomes: the grand illusion. *Plant Cell*, **16**, 1661–1666.

Bendich AJ (2010) The end of the circle for yeast mitochondrial DNA. *Molecular Cell*, **39**, 831–832.

Bibb MJ, Van Etten RA, Wright CT, Walberg MW, Clayton DA (1981) Sequence and gene organization of mouse mitochondrial DNA. *Cell*, **26**, 167–180.

Bundschuh R, Altmüller J, Becker C, Nürnberg P, Gott JM (2011) Complete characterization of the edited transcriptome of the mitochondrion of *Physarum polycephalum* using deep sequencing of RNA. *Nucleic Acids Research*, **39**, 6044–6055.

Burger G, Forget L, Zhu Y, Gray MW, Lang BF (2003) Unique mitochondrial genome architecture in unicellular relatives of animals. *Proceedings of National Academy of Sciences of the USA*, **100**, 892–897.

Cheng S, Higuchi R, Stoneking M (1994) Complete mitochondrial genome amplification. *Nature Genetics*, **7**, 350–351.

Covello PS, Gray MW (1993) On the evolution of RNA editing. *Trends in Genetics*, **9**, 265–268.

Del Vasto M, Figueroa-Martinez F, Featherston J *et al.* (2015) Massive and widespread organelle genomic expansion in the green algal genus *Dunaliella*. *Genome Biology and Evolution*, **7**, 656–663.

DeSalle R (2016a) Comments on Smith (2015)—The past, present and future of mitochondrial genomics: have we sequenced enough mtDNAs. *Briefings in Functional Genomics*, doi: 10.1093/bfgp/elv052. (in press).

DeSalle R (2016b) To new authors and readers of *Mitochondrial DNA Part B: Resources*. *Mitochondrial DNA Part B*, **1**, 1.

Dong W, Xu C, Cheng T, Lin K, Zhou S (2013) Sequencing angiosperm plastid genomes made easy: a complete set of universal primers and a case study on the phylogeny of Saxifragales. *Genome Biology and Evolution*, **5**, 989–997.

Dorrell RG, Howe CJ (2015) Integration of plastids with their hosts: lessons learned from dinoflagellates. *Proceedings of the National Academy of Sciences of the USA*, **112**, 10247–10254.

Fang Y, Wu H, Zhang T *et al.* (2012) A complete sequence and transcriptomic analyses of date palm (*Phoenix dactylifera* L.) mitochondrial genome. *PLoS ONE*, **7**, e37164.

Feagin JE (1992) The 6 kb element of *Plasmodium falciparum* encodes mitochondrial cytochrome genes. *Molecular and Biochemical Parasitology*, **52**, 145–148.

Feagin JE, Harrell MI, Lee JC *et al.* (2012) The fragmented mitochondrial ribosomal RNAs of Plasmodium falciparum. *PLoS ONE*, **7**, e38320.

Fitzgerald TL, Shapter FM, McDonald S *et al.* (2011) Genome diversity in wild grasses under environmental stress. *Proceedings of the National Academy of Sciences*, **108**, 21140–21145.

Gan HM, Schultz MB, Austin CM (2014) Integrated shotgun sequencing and bioinformatics pipeline allows ultra-fast mitogenome recovery and confirms substantial gene rearrangements in Australian freshwater crayfishes. *BMC Evolutionary Biology*, **14**, 19.

Gerhold JM, Aun A, Sedman T, Jõers P, Sedman J (2010) Strand invasion structures in the inverted repeat of *Candida albicans* mitochondrial DNA reveal a role for homologous recombination in replication. *Molecular Cell*, **39**, 851–861.

Golczyk H, Greiner S, Wanner G *et al.* (2014) Chloroplast DNA in mature and senescing leaves: a reappraisal. *Plant Cell*, **26**, 847–854.

Govindaraj P, Nalini A, Krishna N *et al.* (2013) Mitochondrial DNA variations in Madras motor neuron disease. *Mitochondrion*, **13**, 721–728.

Gray MW (2012) Mitochondrial evolution. *Cold Spring Harbor Perspectives in Biology*, **4**, a011403.

Hardouin EA, Tautz D (2013) Increased mitochondrial mutation frequency after an island colonization: positive selection or accumulation of slightly deleterious mutations? *Biology Letters*, **9**, 20121123.

Hartwell LH, Hood L, Goldberg ML *et al.* (2014) *Genetics: From Genes to Genomes*, 1st Canadian edn. McGraw-Hill, Toronto, Canada.

Jackson CJ, Gornik SG, Waller RF (2012) The mitochondrial genome and transcriptome of the basal dinoflagellate *Hematodinium* sp.: character evolution within the highly derived mitochondrial genomes of dinoflagellates. *Genome Biology and Evolution*, **4**, 59–72.

Jedelský PL, Doležal P, Rada P *et al.* (2011) The minimal proteome in the reduced mitochondrion of the parasitic protist *Giardia intestinalis*. *PLoS ONE*, **6**, e17285.

Just RS, Scheible MK, Fast SA *et al.* (2015) Full mtGenome reference data: development and characterization of 588 forensic-quality haplotypes representing three US populations. *Forensic Science International: Genetics*, **14**, 141–155.

Keeling PJ (2013) The number, speed, and impact of plastid endosymbioses in eukaryotic evolution. *Annual Review of Plant Biology*, **64**, 583–607.

Knight RD, Freeland SJ, Landweber LF (2001) Rewiring the keyboard: evolvability of the genetic code. *Nature Reviews Genetics*, **2**, 49–58.

Kolondra A, Labedzka-Dmoch K, Wenda JM, Drzewicka K, Golik P (2015) The transcriptome of *Candida albicans* mitochondria and the evolution of organellar transcription units in yeasts. *BMC Genomics*, **16**, 827.

Lang BF, Jakubkova M, Hegedusova E *et al.* (2014) Massive programmed translational jumping in mitochondria. *Proceedings of National Academy of Sciences*, **111**, 5926–5931.

Lewis SC, Joers P, Willcox S, Griffith JD, Jacobs HT, Hyman BC (2015) A rolling circle replication mechanism produces multimeric lariats of mitochondrial DNA in *Caenorhabditis elegans*. *PLoS Genetics*, **11**, e1004985.

Marande W, Lukeš J, Burger G (2005) Unique mitochondrial genome structure in diplonemids, the sister group of kinetoplastids. *Eukaryotic Cell*, **4**, 1137–1146.

Margam VM, Coates BS, Hellmich RL *et al.* (2011) Mitochondrial genome sequence and expression profiling for the legume pod borer *Maruca vitrata* (Lepidoptera: Crambidae). *PLoS ONE*, **6**, e16444.

Mariac C, Scarcelli N, Pouzadou J *et al.* (2014) Cost-effective enrichment hybridization capture of chloroplast genomes at deep multiplexing levels for population genetics and phylogeography studies. *Molecular Ecology Resources*, **14**, 1103–1113.

Marková S, Filipi K, Searle JB, Kotlík P (2015) Mapping 3′ transcript ends in the bank vole (*Clethrionomys glareolus*) mitochondrial genome with RNA-Seq. *BMC Genomics*, **16**, 870.

Masuda I, Matsuzaki M, Kita K (2010) Extensive frameshift at all AGG and CCC codons in the mitochondrial cytochrome c oxidase subunit 1 gene of *Perkinsus marinus* (Alveolata; Dinoflagellata). *Nucleic Acids Research*, **38**, 6186–6194.

Mercer TR, Neph S, Dinger ME *et al.* (2011) The human mitochondrial transcriptome. *Cell*, **146**, 645–658.

Mungpakdee S, Shinzato C, Takeuchi T *et al.* (2014) Massive gene transfer and extensive RNA editing of a symbiotic dinoflagellate plastid genome. *Genome Biology and Evolution*, **6**, 1408–1422.

Nash EA, Barbrook AC, Edwards-Stuart RK, Bernhardt K, Howe CJ, Nisbet RER (2007) Organization of the mitochondrial genome in the dinoflagellate *Amphidinium carterae*. *Molecular Biology and Evolution*, **24**, 1528–1536.

Ning C, Gao S, Deng B *et al.* (2016) Ancient mitochondrial genome reveals trace of prehistoric migration in the east Pamir by pastoralists. *Journal of Human Genetics*, **61**, 103–108.

Njuguna W, Liston A, Cronn R, Ashman TL, Bassil N (2013) Insights into phylogeny, sex function and age of *Fragaria* based on whole chloroplast genome sequencing. *Molecular Phylogenetics and Evolution*, **66**, 17–29.

Ohyama K, Fukuzawa H, Kohchi T *et al.* (1986) Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature*, **322**, 572–574.

Oldenburg DJ, Bendich AJ (2015) DNA maintenance in plastids and mitochondria of plants. *Frontiers in Plant Science*, **6**, 883.

Oldenburg DJ, Bendich AJ (2016) The linear plastid chromosomes of maize: terminal sequences, structures, and implications for DNA replication. *Current Genetics*, **62**, 431–442.

Segovia R, Pett W, Trewick S, Lavrov DV (2011) Extensive and evolutionarily persistent mitochondrial tRNA editing in velvet worms (phylum Onychophora). *Molecular Biology and Evolution*, **28**, 2873–2881.

Shinozaki K, Ohme M, Tanaka M *et al.* (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *EMBO Journal*, **5**, 2043–2049.

Simpson L, Shaw J (1989) RNA editing and the mitochondrial cryptogenes of kinetoplastid protozoa. *Cell*, **57**, 355–366.

Slater GW, Kist TB, Ren H, Drouin G (1998) Recent developments in DNA electrophoretic separations. *Electrophoresis*, **19**, 1525–1541.

Sloan DB, Alverson AJ, Chuckalovcak JP *et al*. (2012) Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PloS Biology*, **10**, 53.

Smith DR (2013) RNA-Seq data: a goldmine for organelle research. *Briefings in Functional Genomics*, **12**, 454–456.

Smith DR (2016a) The past, present and future of mitochondrial genomics: Have we sequenced enough mtDNAs? *Briefings in Functional Genomics*, **15**, 47–54.

Smith DR (2016b) Goodbye genome paper, hello genome report: the increasing popularity of "genome announcements" and their impact on science. *Briefings in Functional Genomics*, doi: 10.1093/bfgp/elw026. (in press).

Smith DR, Keeling PJ (2015) Mitochondrial and plastid genome architecture: Reoccurring themes, but significant differences at the extremes. *Proceedings of National Academy of Sciences of the USA*, **112**, 10177–10184.

Smith DR, Keeling PJ (2016) Protists and the 'Wild West' of gene expression: new frontiers, lawlessness, and misfits. *Annual Review of Microbiology*, **70**, 161–178.

Smith DR, Lee RW (2008) Mitochondrial genome of the colorless green alga *Polytomella capuana*: a linear molecule with an unprecedented GC content. *Molecular Biology Evolutionary*, **25**, 487–496.

Smith DR, Kayal E, Yanagihara AA, Collins AG, Pirro S, Keeling PJ (2012) First complete mitochondrial genome sequence from a box jellyfish reveals a highly fragmented linear architecture and insights into telomere evolution. *Genome Biology and Evolution*, **4**, 52–58.

Stoltzfus A (1999) On the possibility of constructive neutral evolution. *Journal of Molecular of Evolution*, **49**, 169–181.

Tang M, Tan M, Meng G *et al*. (2014) Multiplex sequencing of pooled mitochondrial genomes—a crucial step toward biodiversity analysis using mito-metagenomics. *Nucleic Acids Research*, **42**, e166.

Taylor JE, Pacheco MA, Bacon DJ *et al*. (2013) The evolutionary history of *Plasmodium vivax* as inferred from mitochondrial genomes: parasite genetic diversity in the Americas. *Molecular Biology and Evolution*, **30**, 2050–2064.

Tian Y, Smith DR (2016) Recovering complete mitochondrial genome sequences from RNA-Seq: a case study of *Polytomella* non-photosynthetic green algae. *Molecular Phylogenetics and Evolution*, **98**, 57–62.

Vahrenholz C, Riemen G, Pratje E, Dujon B, Michaelis G (1993) Mitochondrial DNA of *Chlamydomonas reinhardtii*: the structure of the ends of the linear 15.8-kb genome suggests mechanisms for DNA replication. *Current Genetics*, **24**, 241–247.

Valach M, Moreira S, Kiethega GN, Burger G (2014) Trans-splicing and RNA editing of LSU rRNA in *Diplonema* mitochondria. *Nucleic Acids Research*, **42**, 2660–2672.

Vlcek C, Marande W, Teijeiro S, Lukeš J, Burger G (2010) Systematically fragmented genes in a multipartite mitochondrial genome. *Nucleic Acids Research*, **39**, 979–988.

Waller RF, Jackson CJ (2009) Dinoflagellate mitochondrial genomes: stretching the rules of molecular biology. *BioEssays*, **31**, 237–245.

Wende S, Platzer EG, Jühling F *et al*. (2014) Biological evidence for the world's smallest tRNAs. *Biochimie*, **100**, 151–158.

Wilson RJ, Williamson DH (1997) Extrachromosomal DNA in the Apicomplexa. *Microbiology and Molecular Biology Reviews*, **61**, 1–16.

Wu Z, Stone JD, Štorchová H, Sloan DB (2015) High transcript abundance, RNA editing, and small RNAs in intergenic regions within the massive mitochondrial genome of the angiosperm *Silene noctiflora*. *BMC Genomics*, **16**, 938.

Yagi Y, Ishizaki Y, Nakahira Y, Tozawa Y, Shiina T (2012) Eukaryotic-type plastid nucleoid protein pTAC3 is essential for transcription by the bacterial-type plastid RNA polymerase. *Proceedings of National Academy of Sciences of the USA*, **109**, 7541–7546.

DRS conceived and designed the study. MSL, LCW and MWC collected and analyzed the data. DRS, MSL, LCW and MWC wrote the paper.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Methodological survey of organelle genome papers indexed in MEDLINE from 1 January 2010 to 1 November 2015 [Excel file].