

Full Length Research Paper

***In silico* study in mitochondrial and chloroplast genomes of plants**

G. V. Padma Raju^{1*}, P. Srinivasa Rao², V. Chandra Sekhar¹ and C. Someswara Rao¹

¹SRKR Engineering College, Bhimavaram, AP, India.

²AU College of Engineering, Andhra University, Visakhapatnam, AP, India.

Received 18 June, 2015; Accepted 29 October, 2015

Simple sequence repeats (SSRs) or microsatellites constitute a countable portion of genomes. However, the significance of SSRs in organelle genomes has not been completely understood. The availability of organelle genome sequences allows us to understand the organization of SSRs in their genic and intergenic regions. In the current study, the mitochondrial and chloroplast genomes of different taxa of plants were surveyed. The present study only focused on different authors' investigations and conclusions made based on their results in relation to the different plants. This study helps the researchers to know the different structures of plant genomes, improves the understanding of existing genomes and helps to find newer genomes.

Key words: Chloroplast, mitochondria, simple sequence repeats (SSRs), microsatellites, plants.

INTRODUCTION

Simple sequence repeats (SSRs) or microsatellites are the genetic loci having tandem repeated tracts of DNA composed of 1 to 6 base pairs (bp). They are present in both prokaryotes and eukaryotes (Hancock et al., 1996). Microsatellite loci show extensive length polymorphism and are densely interspersed in eukaryotic genomes. The origin of SSRs can be equally represented in all regions of genomes including a coding portion that resulted in the appearance of single amino acid repeats or tandem oligopeptide repeats in protein sequences (Katti et al., 2000). Besides their usefulness as genetic markers, microsatellites are thought to play important role in genome evolution by creating genetic variability, DNA replication, and cell cycle (Li et al., 2002; Xue et al., 2010).

During the past decade, several genomes have been sequenced, leading to an increased interest in understanding the molecular mechanisms involved in the origin, evolution, and expansion/deletion of microsatellites. With the increase in number of sequenced genomes, several authors have reported the distribution of SSRs in exonic, intronic, and intergenic regions of several eukaryotic genomes (Shakyawar et al., 2009; Tóth et al., 2000). Such estimates reflected the basal level of SSR dynamics within different plant and animal species.

Most of the previous studies on microsatellite distribution are based on DNA sequence databases in which coding or gene-rich regions are overrepresented and are generally nuclear genomes based. On the other

*Corresponding author. E-mail: gvpadmaraju@gmail.com.

hand, the availability of complete organelle genome sequences now permits the determination of the frequencies of SSRs at the whole-genome level. The present article surveys the details of occurrences of variation in the microsatellite repeats in mitochondria and chloroplast.

LITERATURE REVIEW

Integrated view of researchers in the area of analyzing mitochondrial and chloroplast genomes is presented here. Researchers proposed different mechanism in both areas, in particular, to one category of kingdom of plant or kingdom of animal related to mitochondria as well as chloroplast.

Significant studies in microsatellites representations existing even among closely related species, suggested that microsatellites abundance may change relatively rapidly during evolution. The study is important in terms of identifying regions possessing microsatellites in mitochondrial and chloroplast genomes, but microsatellites distributions in these regions are non random and strongly biased. Thus, it may be useful to sketch the path involved in evolution of mitochondrial and chloroplast genomes from species to species and also the variations present between them after the sequencing of missing mitochondrial and chloroplast links, which are not sequenced to date.

Chloroplast genome analysis

A chloroplast is one of three types of plastids, characterized by its high concentration of chlorophyll. The other two types are the leucoplast and chromoplast, and they contain little chlorophyll and do not carry out photosynthesis (Cronn et al., 2008; Shaw et al., 2007).

Chloroplast represents an important autotrophic organelle, which involved in several important functions, most significantly photosynthesis. All plastids studied to date contain their own DNA, which is actually a reduced "genome" derived from a cyanobacterial ancestor that was captured early in the evolution of the eukaryotic cell. Land plant chloroplast genomes typically contain around 110 to 120 unique genes. Some algae have retained a large chloroplast genome with more than 200 genes, while the plastid genomes from non-photosynthetic organisms may retain only a few dozen genes. Several chloroplast genes (*rbcl*, *matK*, *ycf*, *psbA*) have been successfully used as barcodes to define the evolutionary biology among the plant clades (Cronn et al., 2008; Shaw et al., 2007).

Chloroplasts are highly dynamic, they circulate and are moved around within plant cells, and occasionally pinch in two to replicate. Their behavior is strongly influenced by environmental factors like light, colour, and intensity

(Cronn et al., 2008; Shaw et al., 2007).

Plant cells have several structures not found in other eukaryotes. In particular, organelles called chloroplasts allow plants to capture the energy of the sun in energy-rich molecules. Cell walls allow plants to have rigid structures as varied as wood trunks and supple leaves, and vacuoles allow plant cells to change size (Cronn et al., 2008; Shaw et al., 2007).

Chloroplasts originated from an ancient symbiosis, when a nucleated cell engulfed a photosynthetic prokaryote. Indeed, chloroplasts resemble modern cyano bacteria, which remain similar to the cyano bacteria of 3 million years ago. However, the evolution of photosynthesis goes back even more, to the earliest cells that evolved the ability to capture light energy and use it to produce energy-rich molecules. When these organisms developed the ability to split water molecules and use the electrons from these molecules, photosynthetic cells started generating oxygen, an event that had dramatic consequences in the evolution of all living things on earth (Cronn et al., 2008; Shaw et al., 2007).

Chloroplasts retain small and circular genomes that look a lot like cyan bacteria even though they are much smaller. Mitochondrial genomes are even smaller than the genomes of chloroplasts. Coding sequences for majority of chloroplast proteins have been lost, so these proteins are now encoded by the nuclear genome, synthesized in the cytoplasm then transformed from the cytoplasm into the chloroplast.

Daniell et al. (2006) provides study on complete chloroplast genome sequences of *Solanum bulbocastanum* and *Solanum lycopersicum* (potato and tomato). Despite the agricultural importance of both potato and tomato, very little is known about their chloroplast genomes. Analysis of the complete sequences of tomato, potato, tobacco, and *Atropa* chloroplast genomes reveals significant insertions and deletions within certain coding regions or regulatory sequences.

Brysting et al. (2000) proposed the hybrid nature of *Poa jemtlandica*. The results show that some of these chromosome changes took place soon after hybridization in order to overcome the adverse interactions between the nuclear and the cytoplasmic genomes and to facilitate the successful establishment of the newly formed hybrid. The presence of intergenomic chromosome changes may play an important role in the evolution of natural hybrids and the establishment of new evolutionary lineages.

Logacheva et al. (2007) analyze the chloroplast genomes of flowering plants. This analysis show that the phylogenetic tree inferred from a combination of translated nucleotide sequences of genes encoding sub units of plastid RNA polymerase is closest to the tree constructed using all protein coding sites of the chloroplast genome.

Jansen et al. (2006) provides the study on

phylogenetic analyses of *Vitis* (Vitaceae). The Vitaceae (grape) is an economically important family of angiosperms. Recent phylogenetic analyses based on one to several genes have suggested several alternative placements of this family, including sister to Caryophyllales, asterids, Saxifragales, Dilleniaceae or to rest of rosids, though support for these different results has been weak. There has been a recent interest in using complete chloroplast genome sequences for resolving phylogenetic relationships among angiosperms. These studies have clarified relationships among several major lineages, but they have also emphasized the importance of taxon sampling and the effects of different phylogenetic methods for obtaining accurate phylogenies.

de Cambiaire et al. (2006) provides the study on chloroplast genome sequence of the chlorophycean green alga *Scenedesmus obliquus*. This study revealed that, although *Scenedesmus* and *Chlamydomonas* chloroplast DNA (cpDNA) display nearly identical gene repertoires and a high level of sidedness in the distribution of their genes on the two DNA strands, their gene orders are highly scrambled.

Gao et al. (2009) provides the study on complete chloroplast genome sequence of a tree fern *Alsophila spinulosa*. In this study, the authors sequenced the complete chloroplast genome of a scaly tree fern *A. spinulosa* (Cyatheaceae).

Rose et al. (2008) provides the study on chloroplast genome sequencing analysis of *Heterosigma akashiwo* CCMP452 (West Atlantic) and NIES293 (West Pacific) strains. They sequenced chloroplast genomes, using fosmids selected from a total cellular DNA library. This technique has been used to sequence chloroplast DNA of two *H. akashiwo* strains.

Brouard et al. (2008) provides the study on chloroplast DNA sequence of the green alga *Oedogonium cardiacum* (Chlorophyceae). The study provides a clear example that novel genes were acquired by the chloroplast genome through horizontal transfers, possibly from a mitochondrial genome donor.

Oliver et al. (2010) provides the study on chloroplast genome sequence of the moss *Tortula ruralis*. It presents the chloroplast genome sequence of *T. ruralis*, only the second published chloroplast genome for a moss, and the first for a vegetatively desiccation-tolerant plant.

Bausher et al. (2006) provides the study on complete chloroplast genome sequence of *Citrus sinensis* (L.) Osbeck var Ridge Pineapple. The production of *Citrus*, the largest fruit crop of international economic value, has recently been imperiled due to the introduction of the bacterial disease *Citrus canker*. The authors sequenced the *Citrus* chloroplast genome to facilitate genetic improvement of this crop and to assess phylogenetic relationships among major lineages of angiosperms.

Hirao et al. (2008) provides a study on complete nucleotide sequence of the *Cryptomeria japonica* D. Don. chloroplast genome. It provides a comparative analysis of

the gene content and genomic structure that illustrates the unique genomic features of gymnosperms. The results showed that differences in genomic structure between *C. japonica* and other land plants, including pines, strongly support the theory that the large inverted repeats stabilize the chloroplast genome.

Tsai et al. (2006) provides report on the use of the trnL intron and the trnL-trnF intergenic spacer (IGS) in the chloroplast genome and establishes a DNA sequence database for plant species identification. The sequence database described in this study can be used to identify plant species using DNA sequences of the trnL intron and trnL-trnF IGS of chloroplast genome and illustrates its value in plant species identification.

Masood et al. (2004) formed the complete nucleotide sequence of the chloroplast genome of wild rice, *Oryza nivara* and compared it with the corresponding published sequence of relative cultivated rice, *Oryza sativa*. This report describes comparative and genome wide chloroplast analysis between a wild and cultivated crop.

Hu et al. (2011) resolute the complete nucleotide sequence of the rapeseed (*Brassica napus* L.) chloroplast genome (cpDNA). The analysis of *B. napus* and *Brassica rapa* showed only 0.133% in the coding regions, 0.275% in the intron regions, and 0.348% in the intergenic spacer regions. This analysis also supported that *B. napus* is the closest species to *B. rapa* and *B. rapa* could be the maternal parent of *B. napus*.

King et al. (2002) provided the study on a variable minisatellite sequence in the chloroplast genome of *Sorbus* L. (Rosaceae: Maloideae), the chloroplast genome is now known to be more variable than was once thought. Reports of restriction fragment length polymorphism (RFLP) and sequence variation, as well as variation in chloroplast microsatellites, are common. Here, data are presented on the variability of a minisatellite sequence in the chloroplast genome of *Sorbus* species. Sequencing revealed the observed size polymorphism to be due to differences in the number of copies of an imperfect 9-bp motif.

Leseberg and Melvin (2009) investigate chloroplast genome (plastome) scale analyses. The methodological approach was to directly sequence overlapping amplicons from known plastome regions. Newly determined sequences were analyzed with published plastomes from representatives of Panicoideae, Ehrhartoideae, and Pooideae.

Whittall et al. (2010) made a study on chloroplast genome sequence divergence in rare and widespread pines. Critical to conservation efforts and other investigations at low taxonomic levels, DNA sequence data offer important insights into the distinctiveness, biogeographic partitioning and evolutionary histories of the species. The resolving power of DNA sequences is often limited by insufficient variability at the intra specific level. This is particularly true of studies involving plant organelles, as the conservative mutation rate of

chloroplasts and mitochondria makes it difficult to detect polymorphisms necessary to track genealogical relationships among individuals, populations and closely related taxa, through space and time.

Nishikawa et al. (2005) investigated the phylogenetic analysis of *Oryza* species based on SSRs and their flanking nucleotide sequences from the mitochondrial and chloroplast genomes. In this investigation, they used a total of 50 accessions representing 21 different *Oryza* spp. A total of 50 accessions of *Oryza* that represented six different diploid genomes and three different allopolyploid genomes of *Oryza* spp. were analyzed. This is the first report of phylogenetic analysis among plant species, based on mitochondrial and chloroplast SSR and their flanking sequences.

Yuan et al. (2002) expressed DNA fiber-based fluorescence *in situ* hybridization (fiber-FISH) analyses of *O. sativa* spp. *japonica* nuclei to confirm that the insertion of organellar DNA was not a cloning artifact. The sequence of the chloroplast insertion is nearly identical (99.7% identity) to the corresponding regions in the published rice chloroplast genome sequence, suggesting that the transfer event occurred recently. Polymerase chain reaction (PCR) amplification and sequence analysis in two subspecies of rice, *O. sativa* spp. *japonica* and *indica*, indicate that the transfer event predated the divergence of these two subspecies. The chloroplast insertion is flanked by a 2.1-kb perfect direct repeat that is unique to this location in the rice genome.

Nock et al. (2011) provides study on chloroplast genome sequences from total DNA for plant identification. Chloroplast DNA sequence data are a versatile tool for plant identification or barcoding and establishing genetic relationships among plant species. This present chloroplast genome sequences of five grass species derived from massively parallel sequencing (MPS) of total DNA. These data accurately established the phylogenetic relationships between the species, correcting an apparent error in the published rice sequence.

Kim et al. (2006) made a study on complete sequence and organization of the cucumber (*Cucumis sativus* L. cv. Baekmibaekdadagi) chloroplast genome. The presence and relative positions of 113 genes (76 peptide-encoding genes, 30 tRNA genes, four rRNA genes, and three conserved open reading frames) were identified.

Chung et al. (2006) forms the complete nucleotide sequence of the chloroplast genome of potato *Solanum tuberosum* L. cv. Desiree. This provides comparison of chloroplast genomes of seven Solanaceae species and revealed that the gene content and their relative positions of *S. tuberosum* are similar to the other six Solanaceae species. However, undefined open reading frames (ORFs) in large single-copy (LSC) region were highly diverged in Solanaceae species except *Nicotiana sylvestris*. Detailed comparison was identified by numerous indels in the intergenic regions that were

mostly located in the LSC region.

This review demonstrates that researchers who have used primers developed in closely related species have had greater success in finding polymorphic loci. Thus, believed that future studies will benefit from a targeted approach using species-specific primers, rather than adapting published universal primers.

The abundance of chloroplast genome sequences should facilitate primer design for many taxa. Impediments to the use of SSRs include fragment size homoplasmy, introgression, and heteroplasmy. Fragment size homoplasmy can be overcome by careful use of these markers, such as targeted sequencing of unique alleles to evaluate potential for homoplasmy.

Mitochondrial genome analysis

The mitochondrion (plural mitochondria) is a membrane bound organelle found in most eukaryotic cells (the cells that make up plants, animals, fungi, and many other forms of life). Mitochondria range from 0.5 to 1.0 µm in diameter. These structures are sometimes described as "the powerhouse of the cell" because they generate most of the cell's supply of adenosine triphosphate (ATP), used as a source of chemical energy. In addition to supplying cellular energy, mitochondria are involved in other tasks such as signaling, cellular differentiation, cell death, as well as maintaining the control of the cell cycle and cell growth. Mitochondria have been implicated in several human diseases, including mitochondrial disorders and cardiac dysfunction, and may play a role in the aging process. More recent research indicates that autism, especially severe autism, is correlated with mitochondrial defects (Bartoli et al., 2004; Borecky et al., 2006; Calegario et al., 2003).

Several characteristics make mitochondria unique. The number of mitochondria in a cell can vary widely by organism, tissue, and cell type. For instance, red blood cells have no mitochondria, whereas liver cells can have more than 2000. The organelle is composed of compartments that carry out specialized functions. These compartments or regions include the outer membrane, the intermembrane space, the inner membrane, and the cristae and matrix. Mitochondrial proteins vary depending on the tissue and the species. In humans, 615 distinct types of proteins have been identified from cardiac mitochondria, whereas in rats, 940 proteins have been reported. The mitochondrial proteome is thought to be dynamically regulated. Although most of a cell's DNA is contained in the cell nucleus, the mitochondrion has its own independent genome. Further, its DNA shows substantial similarity to bacterial genomes (Bartoli et al., 2004; Borecky et al., 2006; Calegario et al., 2003).

Mitochondria are membrane-enclosed organelles that occur in most eukaryotic cells. Mitochondria have inner and outer membranes composed of phospholipid bilayers

and proteins. Mitochondria, which have been called “cellular power plants”, produce most of the cell’s supply of adenosine triphosphate (ATP), which is used as a source of energy. Mitochondria are not necessarily inherited solely through the maternal line; they can be inherited from both the parents (Bartoli et al., 2004; Borecky et al., 2006; Calegario et al., 2003).

Xiong et al. (2008) summarizes the gene duplication and transfer events in the plant mitochondrial genome. The plant mitochondrial genome is a circular double-stranded DNA molecule that encodes tRNAs, rRNAs, ribosomal proteins, and a portion of the enzymes used in respiration

Mollier et al. (2002) demonstrated that the nuclear gene encoding the plastid S13 has been partially duplicated in *Arabidopsis thaliana*, such that the copy has lost the exon encoding the plastid transit peptide and has acquired a sequence capable of encoding a mitochondrial targeting sequence. The mitochondrial S13 ribosomal protein has probably been replaced by its homologue from plastids in *A. thaliana*.

Xu et al. (2009) provides the study of the evolution and function of lignin biosynthesis genes, thus has two-fold implications. It analyzed lignin biosynthesis genes from fourteen plant species and one symbiotic fungal species. In addition, it also carried out the biomass composition analysis of nine *Arabidopsis* mutants. The results were analyzed together with the genomics analysis. The research revealed that, among the species analyzed, the complete lignin biosynthesis pathway first appeared in moss; the pathway is absent in green algae.

Jackson et al. (2007) investigated the structure, content and expression of dinoflagellate mitochondrial genomes. This analysis of cDNAs suggests several novel aspects of dinoflagellate mitochondrial gene expression. It also exhibits several unique characteristics, most notable are the expansion of gene copy numbers and their arrangements within the genome, RNA editing, loss of stop codons, and use of trans-splicing.

Behura et al. (2011) presented complete sequences of the mitochondrial genomes for two important mosquitoes, *Aedes aegypti* and *Culex quinquefasciatus*, that are major vectors of dengue virus and lymphatic filariasis, respectively. It also analyzed insertions of nuclear copies of mtDNA-like sequences (NUMTs) in a comparative manner between the two mosquitoes. It also provides new insights on understanding the roles of nuclear mtDNA sequences in genome complexities of these mosquitoes.

Zhou and Xia (2009) analyzed mitochondrial genome in six different plants to find general patterns of codon usage in plant mitochondrial genomes. These findings suggest that natural selection is likely to be playing a large role in codon usage bias in plant mitochondrial genomes, not only natural selection, but also other several factors are likely to be involved in determining the selective constraints on codon bias in plant mitochondrial

genomes.

Sperisen et al. (2001) studied the utility of two polymorphic mitochondrial tandem repeats located in the second intron of the *nad1* gene of Norway spruce. Their study provides the way to further tandem repeat markers in a wide range of taxa which should greatly improve the understanding of population genetic processes, structure, and history of plant species.

Sugiyama et al. (2005) specified the valuable model system for investigating the origin of mitochondrial DNA (mtDNA) in amphidiploid plants and studying the genetic interaction between mitochondria and chloroplasts in various functions of the plant cell. It provides the complete sequencing of tobacco mtDNA that has identified the *rps1* and *wrps14* genes, which had previously been thought to be absent from tobacco mtDNA based on southern analysis, and provided the basis for identifying RNA editing sites in tobacco mitochondrial transcripts.

Lee et al. (2011) performed the comparative analysis of the root mitochondrial proteome with mitochondria isolated from photosynthetic shoots to define the role of protein abundance with these differences. The major differences observed were in the abundance and/or activities of enzymes in the tricarboxylic acid (TCA) cycle and the mitochondrial enzymes involved in photorespiration. The metabolic pathways that relied on the supply of intermediates from TCA cycle and photorespiration were also altered, namely cysteine, formate, and one-carbon metabolism, as well as amino acid metabolism focused on 2-oxoglutarate generation.

According to Taylor and Doug (2005), human mitochondrial genome is extremely small when compared with the nuclear genome, and mitochondrial genetics presents unique clinical and experimental challenges. Despite the diminutive size of the mitochondrial genome, mitochondrial DNA (mtDNA) mutations are an important cause of inherited disease.

Rohou et al. (2001) reported the identification and characterization of a new mutation (ts9) in the *Saccharomyces cerevisiae* mitochondrial genome, which was first genetically mapped in the tRNA_{gly} region and further identified by means of sequencing as consisting of a G to A transition at position 30 in the tRNA. The mutation causes an almost complete disappearance of mature tRNA_{gly}, while a second mitochondrial mutation with a compensatory C to T change restores it in normal quantities.

Jianping (2010) proposed that the mutation of mitochondria genome is the trigger of somatic cell transforming to cancer cell. From his hypothesis, mitochondrial genomes lack introns, this is contrary to chromosome genomes, so missed introns result in mitochondrial genome vulnerable to be affected by malignant environments, such as free radicals (reactive oxygen species). These results showed that mutations in mtDNA have been identified in various types of human

cancer, such as breast cancer, ovary cancer, and colon cancer.

Desplanque et al. (2000) proposed that linkage disequilibrium between chloroplast DNA and mitochondrial DNA haplotypes in *Beta vulgaris* subsp. *maritima* (L.). This observed genetic association between chloroplasts and mitochondria (that is, two maternally inherited cytoplasmic genomes) may indicate whether or not homoplasmy occurs in the mitochondrial genome. Four-hundred and fourteen individuals sampled in wild populations of beets from France and Spain was screened for their mitochondrial and chloroplast polymorphisms. Mitochondrial DNA (mtDNA) polymorphism was investigated with restriction fragment length polymorphism (RFLP) and cpDNA polymorphism was investigated with PCR-RFLP, using universal primers for the amplification. Identical mitochondrial variants found in populations of different regions probably occurred as a result of migration. Final conclusion from this study that mtDNA is a tool as valuable as cpDNA when a maternal marker is needed for population genetics analyses in beet on a large regional scale.

Nakazono et al. (1995) says that small repeated sequence contains the transcription initiation sites for both *trnfM* and *rrn26* in rice mitochondria. These analysis sequences that include promoter sequences are used for the initiation of transcription of mitochondrial genes. These results are an evidence for the functional significance of small repeated sequences in plant mtDNA.

Tang et al. (2014) developed a novel multiplex sequencing and assembly pipeline allowing for simultaneous acquisition of full mitogenomes from pooled animals without DNA enrichment or amplification. By concatenating assemblies from three *de novo* assemblers, obtained high-quality mitogenomes for all pooled taxa. The *in silico* simulation showed that by recruiting multiple mito-loci, taxon detection was improved at a fixed sequencing depth.

Cui et al. (2013) investigated the performance characteristics and to gain new insights into the analysis of the mitochondrial genome. The entire mitochondrial genome was analyzed as a single amplicon using a long-range PCR-based enrichment approach coupled with MPS. Results demonstrated the uniform coverage of the entire mitochondrial genome. MPS of the single amplicon revealed the presence of single-nucleotide polymorphisms and nuclear homologs of mtDNA sequences that cause the erroneous and inaccurate variant calls when PCR/Sanger sequencing approach was used.

Rodríguez-Moreno et al. (2011) determined the nucleotide sequences of the melon chloroplast and mitochondrial genomes. Comparison of the cucumber and melon chloroplast genomes showed differences in only approximately 5% of nucleotides, mainly due to short indels and SNPs. Additionally, 2.74 Mb of mitochondrial

sequence, accounting for 95% of the estimated mitochondrial genome size, were assembled into five scaffolds and four additional unscaffolded contigs out of 84% of the mitochondrial genome are contained in a single scaffold.

Chang et al. (2011) sequenced the mitotypes of cam (*B. rapa*), ole (*Brassica oleracea*), Jun (*Brassica juncea*), and car (*Brassica carinata*). The sequence relationship analysis showed that there has been genome compaction and inheritance in the course of *Brassica* mitotype evolution. They have sequenced four *Brassica* mitotypes, compared six *Brassica* mitotypes and suggested a mechanism for mitochondrial genome formation in *Brassica*, including evolutionary events such as inheritance, duplication, rearrangement, genome compaction, and mutation.

Tian et al. (2006) divided mitochondrial sequence variations into two basic categories, intravarietal and intersubspecific. Intravarietal polymorphisms are variations within mitochondrial genomes of an individual variety. Intersub specific polymorphisms are variations between subspecies among their major genotypes. In this study, identified 96 single nucleotide polymorphisms (SNPs), 25 indels, and three segmental sequence variations as intersubspecific polymorphisms.

Hahn et al. (2013) presented an *in silico* approach for the reconstruction of complete mitochondrial genomes of non-model organisms directly from next-generation sequencing (NGS) data-mitochondrial baiting and iterative mapping (MITObim). The method is straightforward even if only distantly related mitochondrial genomes or mitochondrial barcode sequences are available as starting-reference sequences or seeds, respectively. This approach overcomes the limitations of traditional strategies for obtaining mitochondrial genomes for species with little or no mitochondrial sequence information at hand and represents a fast and highly efficient *in silico* alternative to laborious conventional strategies relying on initial long-range PCR.

Handa (2003) provides the complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*B. napus* L.). The entire mitochondrial genome of rapeseed (*B. napus* L.) was sequenced and compared with that of *A. thaliana*. This gene content is almost identical to that of *Arabidopsis*; however, the rps14 gene, which is a pseudo-gene in *Arabidopsis*, is intact in rapeseed. On the other hand, five tRNA genes are missing in rapeseed compared to *Arabidopsis*, although the set of mitochondrially encoded tRNA species is identical in the two Cruciferae. The results suggest that higher plant mitochondria are extremely conservative with respect to coding sequences and somewhat conservative with respect to RNA editing, but non-coding parts of plant mitochondrial DNA are extraordinarily dynamic with respect to structural changes, sequence acquisition and/or sequence loss.

Turmel et al. (2003) determined the mtDNA sequence

of *Chara vulgaris* (Charophyceae), a green alga belonging to the charophycean order (Charales) that is thought to be the most closely related alga to land plants. Overall, comparisons provide unequivocal support for a sister-group relationship between the Charales and the land plants. Only four introns in land plant mtDNAs appear to have been inherited vertically from a charalean algal ancestor.

Kubo et al. (2000) determined the complete nucleotide sequence of the mitochondrial genome of an angiosperm, sugar beet (*Beta vulgaris* cv TK81-O). They also identified a novel tRNA^{Cys} gene (trnC2-GCA) which shows no sequence homology with any tRNA^{Cys} genes reported so far in higher plants. Intriguingly, this tRNA gene is actually transcribed into a mature tRNA, whereas the native tRNA^{Cys} gene (trnC1-GCA) is most likely a pseudogene.

Ma et al. (2012) determined a largely completed mt genome from a bamboo, *Ferocalamus rimosivaginus* (Poaceae), through Illumina sequencing of total DNA. With combination of *de novo* and reference-guided assembly, 39.5-fold coverage Illumina reads were finally assembled into scaffolds totaling 432,839 bp. The assembled genome contains nearly the same genes as the completed mt genomes in Poaceae. Results demonstrate that it is a rapid and efficient approach to obtain angiosperm mt genome sequences using Illumina sequencing technology.

Dettai et al. (2012) proposes a novel approach for the isolation and sequencing of a universal, useful and popular marker across distant, non-model metazoans. It relies on the properties of metazoan mito genomes for enrichment, on careful choice of the organisms to multiplex, as well as on the wide collection of accumulated mitochondrial reference datasets.

Groenenberg et al. (2012) describes how the complete mitogenome of a terrestrial snail, *Cylindrus obtusus* (Draparnaud, 1805) was sequenced without PCRs from a collection specimen. The mitogenome was obtained with Illumina GALLx shot gun sequencing. Although the used specimen was collected relatively recently and kept in a DNA-friendly preservative, believe that the exclusion of PCRs as facilitated by next generation sequencing (NGS) removes a great obstacle in DNA sequencing of collection specimens. A brief comparison is made between our Illumina GALLx approach and a similar study that made use of the Roche 454-FLX platform.

Iannelli et al. (2007) sequenced the complete mt genome of two congeneric ascidian species, *Phallusia mammillata* and *Phallusia fumigata*. The two mtDNAs are surprisingly rearranged, both with respect to one another and relative to those of other tunicates and chordates, with gene rearrangements affecting both protein-coding and tRNA genes. The new data highlight the extraordinary variability of ascidian mt genome in base composition, tRNA secondary structure, tRNA gene content, and non-coding regions.

Sperisen et al. (2001) demonstrated the utility of two polymorphic mitochondrial tandem repeats located in the second intron of the nad1 gene of Norway spruce. Their study provides the way to further tandem repeat markers in a wide range of taxa that should greatly improve the understanding of population genetic processes, structure, and history of plant species.

Kuntal and Vinay (2011) surveyed the patterns of SSRs in mitochondrial genomes of different taxa of plants. In this study, authors use the 16 species of plants including algae, marchantiophyte, bryophyte, anthocerotophyte, gymnosperm, monocots and dicots, belonging to streptophyta were downloaded from NCBI's Genome data bank. The accession numbers of mitochondrial genome sequences were used in the study. These results show that the abundance of repeat types varies with the genomic region and distribution is the characteristic of the taxonomic group examined. The repeat motifs are not uniformly distributed across the genomes, but mostly confined to intergenic regions and non-coding segments of mitochondrial DNA in contrast to the coding regions, and with that considerable variation in numbers of SSRs is observed in non coding regions of mitochondrial genomes across different species.

CONCLUSIONS

The complete genome sequences, including both the mitochondrial and chloroplast genomes, of various organisms are becoming available, and this can be considered a major step forward toward exploiting the usefulness of bio informatics engineering technology. The earlier work showed that many gene analyses have occurred during mitochondrial and chloroplast evolution, but there is no reliable estimate of the total number of genes that have been transferred. And also this analysis helps the researchers to develop the better techniques in this era of bio-informatics especially analyzing the genome sequence of mitochondrial and chloroplast genomes of various species.

Conflict of Interests

The authors have not declared any conflict of interests.

REFERENCES

- Bartoli CG, Facundo G, Dana EM, Juan JG (2004). Mitochondria are the main target for oxidative damage in leaves of wheat (*Triticum aestivum* L.). J. Experimental Botany pp.1663-1669.
- Bausher MG, Nameirakpam DS, Seung-Bum L, Robert KJ, Henry Daniell (2006). The complete chloroplast genome sequence of *Citrus sinensis* (L.) Osbeck var'Ridge Pineapple': organization and phylogenetic relationships to other angiosperms. BMC Plant. Biol. 6:1-21.
- Behura SK, Neil FL, Brian H, Diane DL, Martin FS, Daniela P, Jeanne R-S, Vishvanath N, David WS (2011). Complete sequences of

- mitochondria genomes of *Aedes aegypti* and *Culex quinquefasciatus* and comparative analysis of mitochondrial DNA fragments inserted in the nuclear genomes. *Insect Biochem. Mol. Biol.* 41(10): 770-777.
- Zhou M, Xia L (2009). Analysis of synonymous codon usage patterns in different plant Mitochondrial genomes. *Mol. Biol. Rep.* 36(8): 2039-2046.
- Borecky J, Nogueira FTS, de Oliveira KAP (2006). The plant energy-dissipating mitochondrial systems: depicting the genomic structure and the expression profiles of the gene families of uncoupling protein and alternative oxidase in monocots and dicots. *J. Exp. Bot.* 57(4):849-864.
- Brouard J-S, Christian O, Claude L, Monique T (2008). Chloroplast DNA sequence of the green alga *Oedogonium cardiacum* (Chlorophyceae): unique genome architecture, derived characters shared with the Chaetophorales and novel genes acquired through horizontal transfer. *BMC genomics* 9:1-290.
- Brysting AK, Holst-Jensen A, Leitch I (2000). Genomic origin and organization of the hybrid *Poa jemtlandica* (Poaceae) verified by genomic in situ hybridization and chloroplast DNA sequences. *Annals of botany.* 85(4):439-445.
- Calegario FF, Ricardo GC, Marcia MF, Fernanda VA, Wilson FJ, Petr J, Paulo A, Anibal EV (2003). Stimulation of potato tuber respiration by cold stress is associated with an increased capacity of both plant uncoupling mitochondrial protein (PUMP) and alternative oxidase. *J. Bioenerg. Biomembr.* 35(3):211-220.
- Chang S, Tiantian Yang, Tongqing Du, Yongjuan Huang, Jianmei Chen, Jiyong Yan, Jianbo He, and Rongzhan Guan (2011). Mitochondrial genome sequencing helps show the evolutionary mechanism of Mitochondrial genome formation in *Brassica*. *BMC genomics* 12:1-497.
- Chung HJ, Jong DJ, Hyun-Woo P, Joo-Hwan K, Hyun WC, Sung RM, Won-Joong J, Jang RL (2006). The complete chloroplast genome sequences of *Solanum tuberosum* and comparative analysis with Solanaceae species identified the presence of a 241-bp deletion in cultivated potato chloroplast DNA sequence. *Plant cell reports.* 25(12):1369-1379.
- Cui H, Fangyuan L, David C, Guoli W, Cavatina KT, Gregory ME, Brett G (2013). Comprehensive next-generation sequence analyses of the entire Mitochondrial genome reveal new insights into the molecular diagnosis of Mitochondrial DNA disorders. *Genet. Med.* 15(5):388-394.
- Cronn R, Aaron L, Matthew P, David SG, Rongkun S, Todd M (2008). Multiplex sequencing of plant chloroplast genomes using Solexa sequencing-by-synthesis technology. *Nucleic Acids Res.* 36(19):e122-e122.
- Daniell H, Seung-Bum L, Justin G, Christopher S, Tania Q-V, Chittibabu G, Jeffrey T, Robert KJ (2006). Complete Chloroplast genome sequences of *Solanum bulbocastanum*, *Solanum lycopersicum* and comparative analyses with other Solanaceae genomes. *Theor. Appl. Genet.* 112(8): 1503-1518.
- de Cambiaire JC, Christian O, Claude L, Monique T (2006). The complete chloroplast genome sequence of the chlorophycean green alga *Scenedesmus obliquus* reveals a compact gene organization and a biased distribution of genes on the two DNA strands. *BMC Evol. Biol.* 6:1-37.
- Dettai A, Cyril Gallut, Sophie Brouillet, Joel Pothier, Guillaume Lecointre, Régis Debruyne (2012). "Conveniently pre-tagged and pre-packaged: extended molecular identification and metagenomics using complete metazoan Mitochondrial genomes. *PloS one* 7:12-e51263.
- Desplanque B, F. Viard, J. Bernard, D. Forcioli, P. SauMitou - Laprade, J. Cuguen, and H. Van Dijk (2000). The linkage disequilibrium between Chloroplast DNA and Mitochondrial DNA haplotypes in *Beta vulgaris* ssp. *maritima* (L.), the usefulness of both genomes for population genetic studies." *Mol. Ecol.* 9(2):141-154.
- Gao L, Xuan Y, Yong-Xia Y, Ying-Juan S, Ting W (2009). "Complete chloroplast genome sequence of a tree fern *Alsophila spinulosa*: insights into evolutionary changes in fern chloroplast genomes. *BMC Evol. Biol.* 9:1-130.
- Groenenberg DSJ, Walter Pirovano, Edmund Gittenberger, and Menno Schilthuizen (2012). The complete Mitogenome of *Cylindrus Obtusus* (Helicidae, Ariantinae) using Illumina next generation sequencing. *BMC genomics.* 13:1-114.
- Hancock JM (1996). Simple sequences in a 'minimal' genome. *Nature genet.* 14(1):14-15.
- Hahn C, Lutz B., Bastien C. (2013). Reconstructing Mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Res.* 41(13):e129-e129.
- Handa H (2003). The complete nucleotide sequence and RNA editing content of the Mitochondrial genome of rapeseed (*Brassica napus* L.) comparative analysis of the Mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. *Nucleic Acids Res.* 31(20):5907-5916.
- Hirao T, Atsushi W, Manabu K, Teiji K, Katsuhiko T (2008). "Complete nucleotide sequence of the *Cryptomeria japonica* D. Don. chloroplast genome and comparative chloroplast genomics: diversified genomic structure of coniferous species." *BMC plant biol.* 8:1-70.
- Hu ZY, Wei H, Shun-Mou H, Han-Zhong W (2011). Complete chloroplast genome sequence of rapeseed (*Brassica napus* L.) and its evolutionary implications." *Genetic Resources and Crop Evolution* 58(6): 875-887.
- Jackson CJ, John EN, Murray NS, Michael WG, Patrick JK, Ross FW (2007). "Broad genomic and transcriptional analysis reveals a highly derived genome in dinoflagellate mitochondria." *BMC biology* 5:1-41.
- Jansen RK, Charalambos K, Christopher S, Seung-Bum L, Jeffrey T, Andrew JA, Henry D (2006). "Phylogenetic analyses of *Vitis* (Vitaceae) based on complete chloroplast genome sequences: effects of taxon sampling and phylogenetic methods on resolving relationships among rosids." *BMC Evol. Biol.* 6:1-32.
- Iannelli F, Francesca G, Graziano P, Carmela G (2007). The Mitochondrial genome of *Phallusia mammillata* and *Phallusia fumigata* (Tunicata, Ascidiacea): high genome plasticity at intra-genus level. *BMC Evol. Biol.* 7:1-155.
- Jianping D (2010). Mutation of Mitochondria genome: trigger of somatic cell transforming to cancer cell. *Int. Arch. Med.* 3:1-4.
- Katti MV, Sami-Subbu R, Prabhakar KR, Vidya SG (2000). Amino acid repeat patterns in protein sequences: their diversity and structural-functional implications. *Prot. Sci.* 9(06):1203-1209.
- Kim J-S, Jong DJ, Jung-Ae L, Hyun-Woo P, Kwang-Hoon O, Won-Joong J, Dong-Woog C, Jang RL, Kwang Yun Cho (2006). "Complete sequence and organization of the cucumber (*Cucumis sativus* L. cv. Baekmibaekdadagi) chloroplast genome." *Plant cell reports* 25(4): 334-340.
- King RA, Colin F (2002). A variable minisatellite sequence in the chloroplast genome of *Sorbus* L. (Rosaceae: Maloideae). *Genome.* 45(3):570-576.
- Kubo T, Satsuki Nishizawa, Akira Sugawara, Noriko Itchoda, Amy Estiati, and Tetsuo Mikami. "The complete nucleotide sequence of the Mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for tRNACys (GCA). *Nucleic Acids Res.* 28(13):2571-2576.
- Kuntal H, Vinay S (2011). *In silico* analysis of SSRs in Mitochondrial genomes of plants. *Omics: a J. Integr. Biol.* 15(11):783-789.
- Lee CP, Holger E, Nicholas O'T, Harvey AM (2011). "Combining proteomics of root and shoot mitochondria and transcript analysis to define constitutive and variable components in plant mitochondria." *Phytochemistry.* 72(10):1092-1108.
- Li Y-C Abraham BK, Tzion F, Avigdor B, Eviatar N (2002). Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Mol. Ecol.* 11(12):2453-2465.
- Leseberg CH, Melvin RD (2009). The complete chloroplast genome of *Coix lacryma-jobi* and a comparative molecular evolutionary analysis of plastomes in cereals. *J. Mol. Evol.* 69(4):311-318.
- Logacheva MD, Penin AA, Samigullin TH, Vallejo-Roman CM, Antonov AS (2007). Phylogeny of flowering plants by the chloroplast genome sequences: in search of a "lucky gene". *Biochemistry (Moscow).* 72(12):1324-1330.
- Masood MS Shahid, Tomotaro N, Shu-ichi F, Peter KN, Takahiko T, Koh-ichi K (2004). The complete nucleotide sequence of wild rice (*Oryza nivara*) chloroplast genome: first genome wide comparative sequence analysis of wild and cultivated rice. *Gene* 340(1):133-139.
- Ma PF, Zhen-Hua G, De-Zhu L (2012). Rapid sequencing of the bamboo Mitochondrial genome using Illumina technology and parallel episodic evolution of organelle genomes in grasses. *PLoS*

- One. 7(1):1-13.
- Mollier P, Beate H, Cedrig D, Ian S (2002). The gene encoding *Arabidopsis thaliana* mitochondrial ribosomal protein S13 is a recent duplication of the gene encoding plastid S13. *Current genetics*. 40(6):405-409.
- Nakazono M, Nobuhiro T, Masahiro S, Atsushi H (1995). A small repeated sequence contains the transcription initiation sites for both *trnFM* and *rrn26* in rice Mitochondria. *Plant Mol. Biol.* 28(2):343-346.
- Nishikawa T, Duncan A. Vaughan, and Koh-ichi Kadowaki (2005). Phylogenetic analysis of *Oryza* species, based on simple sequencerepeats and their flanking nucleotide sequences from the Mitochondrial and Chloroplast genomes. *Theor. Appl. Genet.* 110(4):696-705.
- Nock CJ, Nock CJ, Waters DL, Edwards MA, Bowen SG, Rice N, Cordeiro GM, Henry RJ (2011). Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnol J.* 9(3):328-33.
- Oliver MJ, Andrew GM, Brent DM, Jennifer VK, Jeffrey LB, Dina FM, Karin DE, Paul GW, Aaron MD, Kenneth GK (2010). Chloroplast genome sequence of the moss *Tortula ruralis*: gene content, polymorphism, and structural arrangement relative to other green plant chloroplast genomes. *BMC genomics* 11:1-143.
- Rodríguez-Moreno L, Rodríguez-Moreno, L, Víctor MG, Andrej B, Carmen M, Pere P, Miguel AA, Jordi G (2011). "Determination of the melon Chloroplast and Mitochondrial genome sequences reveals that the largest reported Mitochondrial genome in plants contains a significant amount of DNA having a nuclear origin." *BMC genomics*. 12:1- 424.
- Rohou H, Francisci S, Rinaldi T, Frontali L, Bolotin-Fukuhara M (2001). Reintroduction of a characterized Mit tRNA glycine mutation into yeast Mitochondria provides a new tool for the study of human neurodegenerative diseases. *Yeast* 18(3):219-227.
- Rose C, Michael AJ, Yang Z, Jean C, Melinda D, Terry L, John M, Han CO, Elizabeth S, Gabrielle R (2008). Chloroplast genome sequencing analysis of *Heterosigma akashiwo* CCMP452 (West Atlantic) and NIES293 (West Pacific) strains.
- Shakyawar SK, Balwinder KJ, Dinesh K (2009). SSR repeat dynamics in mitochondrial genomes of five domestic animal species. *Bioinformation*. 4: 4-158.
- Shaw J, Edgar BL, Edward ES, Randall LS (2007). Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *Ame J. Bot.* 94(3):275-288.
- Sperisen C, Büchler U, Gugerli F, Mátyás G, Geburek Th, Vendramin GG (2001). Tandem repeats in plant mitochondrial genomes: application to the analysis of population differentiation in the conifer Norway spruce. *Mol. Ecol.* 10(1):257-263.
- Sugiyama Y, Watase Y, Nagase M, Makita N, Yagura S, Hirai A, Sugiura M (2005). The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants. *Mol Genet. Genomics*. 272(6): 603-615.
- Taylor RW, Doug MT (2005). Mitochondrial DNA mutations in human disease", *Nature Reviews Genetics* 6(5): 389-402.
- Tang M, Meihua Tan, Guanliang Meng, Shenzhou Yang, Xu Su, Shanlin Liu, Wenhui Song (2014). "Multiplex sequencing of pooled Mitochondrial genomes-a crucial step toward biodiversity analysis using Mito-metagenomics. *Nucleic Acids Res.* 42(22):e166-e166.
- Tian X, Jing Zheng, Songnian Hu, and Jun Yu (2006). The rice Mitochondrial genomes and their variations. *Plant Physiol.* 140(2):401-410.
- Tóth G, Zoltán G, Jerzy J (2000). Microsatellites in different eukaryotic genomes: survey and analysis. *Genome research* 10(7):967-981.
- Tsai LC, Yung-Chien Y, Hsing-Mei H, Jenn-Che W, Adrian L, James C-I (2006). Species identification using sequences of the *trnL* intron and the *trnL-trnF* IGS of chloroplast genome among popular plants in Taiwan. *Forensic Sci. Int.* 164(2):193-200.
- Turmel M, Christian Otis, and Claude Lemieux (2003). The Mitochondrial genome of *Chara vulgaris*: insights into the Mitochondrial DNA architecture of the last common ancestor of green algae and land plants. *Plant Cell.* 15(8):1888-1903.
- Whittall JB, John S, Matthew P, Jason B, Dick C, Aaron L, Richard C (2010). Finding a (pine) needle in a haystack: chloroplast genome sequence divergence in rare and widespread pines. *Mol. Ecol.* 19(1):100-114.
- Xiong A-S, Ri-He Peng, Jing Zhuang, Feng Gao, Bo Zhu, Xiao-Yan Fu, Yong Xue (2008). Gene duplication and transfer events in plant mitochondria genome. *Biochem. Biophys. Res. Commun.* 376(1):1-4.
- Xue J-Y, Yang L, Libo L, Bin W, Yin-Long Q (2010). The complete mitochondrial genome sequence of the hornwort *Phaeoceros laevis*: retention of many ancient pseudogenes and conservative evolution of mitochondrial genomes in hornworts. *Current genetics*. 56(1):53-61.
- Xu Z, Dandan Z, Jun Hu, Xin Z, Xia Ye, Kristen LR, Nathan RS (2009). "Comparative genome analysis of lignin biosynthesis gene families across the plant kingdom. *BMC bioinformatics* 10:11-3.
- Yuan Q, Hill J, Hsiao J, Moffat K, Ouyang S, Cheng Z, Jiang J, Buell C (2002). Genome sequencing of a 239-kb region of rice chromosome 10L reveals a high frequency of gene duplication and a large chloroplast DNA insertion. *Mol. Genet. Genomics* 267(6):713-720.