Phylogenetic Analysis of Nuclear-Encoded RNA Maturases

Sunita Malik¹, KC Upadhyaya² and SM Paul Khurana¹

¹Amity Institute of Biotechnology, Amity University, Noida, India. ²School of Life Sciences, Jawaharlal Nehru University, New Delhi, India.

ABSTRACT: Posttranscriptional processes, such as splicing, play a crucial role in gene expression and are prevalent not only in nuclear genes but also in plant mitochondria where splicing of group II introns is catalyzed by a class of proteins termed maturases. In plant mitochondria, there are 22 mitochondrial group II introns. matR, nMAT1, nMAT2, nMAT3, and nMAT4 proteins have been shown to be required for efficient splicing of several group II introns in Arabidopsis thaliana. Nuclear maturases (nMATs) are necessary for splicing of mitochondrial genes, leading to normal oxidative phosphorylation. Sequence analysis through phylogenetic tree (including bootstrapping) revealed high homology with maturase sequences of A. thaliana and other plants. This study shows the phylogenetic relationship of nMAT proteins between A. thaliana and other nonredundant plant species taken from BLASTP analysis.

KEYWORDS: Maturases, nuclear maturases, splicing, homology, BLAST

Introduction

Mitochondria are responsible for aerobic energy production in eukaryotic cells. Although most of the proteins of mitochondrion are coded by nuclear DNA, a number of essential proteins are coded by mitochondrial DNA (mtDNA). Posttranscriptional processes, such as splicing, are vital for gene expression. In general, such processes are regulated by binding of specific proteins to structural domains present on some messenger RNA (mRNA) or pre-mRNAs but not others. Posttranscriptional regulatory processes are particularly prevalent in plant mitochondria. The removal of group II introns from plant mtDNA is essential for normal functioning of mitochondria, i.e., respiration. Intron-encoded maturases facilitate self-splicing of group II introns in bacteria and organellar genomes of several lower eukaryotes. Plant organelles also contain group II introns but have lost maturase open reading frame. Nuclear maturases (nMATs) are the proteins encoded by nuclear genes of angiosperms. The 4 nMAT proteins are classified as nMAT1, nMAT2, nMAT3, and nMAT4. Nuclear maturases are transported to mitochondria for specific and efficient splicing of group II introns of mitochondria-encoded genes. Three of these paralogs, nMAT1, nMAT2, and nMAT3, have been implicated in the splicing of mitochondrial introns in Arabidopsis thaliana. Also, nMAT mutants have been shown to result in defective embryos and other developmentally delayed phenotypes.

nMAT1 (also known as nMat1A) was reported to be involved in the splicing of mitochondrial nad4 intron2, but the specific intron target sequences were unknown. nMAT2-deficient plants show that this protein is required for the efficient splicing of at least 3 mitochondrial introns: a single intron within cox2 and intron 2 of nad1 and nad7 genes. nMAT1 is essential for trans-splicing of intron 1 of nad1 gene, cis-splicing of nad2 intron 1 and nad4 intron 2, and thus for the functional assembly of mitochondrial complex I. nMAT4 has been known for efficient splicing of introns 1, 3, and 4 of nad1, intron 1 of nad2, intron 1 of cox2, intron 3 of nad4, and introns 3 and 4 of nad7.

At least 5 nuclear-encoded factors are required for excision of the 4 introns within nad1 in Arabidopsis mitochondria, and in each case, the splicing of the nad1 introns involves at least 1 maturase protein. OTP43, nMAT1, and nMAT4 are required for trans-splicing of the nad1 intron 1. Splicing of the nad1 intron 2 involves at least 2 splicing factors: nMAT2 and nCSF1. Both nCSF1 and nMAT4 are required for trans-splicing of nad1 intron 3, and nMAT4 is required for splicing of nad1 intron 4. In addition, a DEAD box RNA helicase protein (PMH2) has been shown to influence the processing and stability of many mitochondrial pre-mRNAs in Arabidopsis, including nad1 introns 2 and 3. Above evidences suggest an important role for nMATs in splicing of group II introns in mitochondria and thus ensuring normal oxidative metabolism. On the basis of the type of maturase proteins, the organism found in, and the site of presence of the maturase-coding sequence, maturase proteins have been assigned different terms, namely, mat, matK, matR, ltrA, and nMat. These are found in bacteria, protists, algae, fungi, and plants. In general, besides nMAT, organellar maturases are also found, namely, matR (mitochondria) and matK (chloroplast maturases). Table 1 gives an overview of the different types of maturase protein–coding sequences across a wide range of organisms. As seen from the table, different types of maturase proteins have been found in various species.

Subsequently, nMATs have been found to have important functions, especially with respect to intron splicing in mitochondria of higher plants affecting growth and development,

Creative Commons Non Commercial CC BY-NC. This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (http://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
Table 1. List of organisms showing % identity (≥50%) match with nMAT1, nMAT2, nMAT3, and nMAT4 proteins of Arabidopsis thaliana.

| ORGANISM            | NONANNOTATED/ANNOTATED ACCESSION NUMBERS | % OF IDENTITY MATCH | % OF IDENTITY MATCH | % OF IDENTITY MATCH | % OF IDENTITY MATCH | E VALUE |
|---------------------|------------------------------------------|---------------------|---------------------|---------------------|---------------------|---------|
|                     |                                           | nMAT1               | nMAT2               | nMAT3               | nMAT4               |         |
| Aegilops tauschii   | EMT06317.1                               | 66                  | 87                  | —                   | —                   | 0.0     |
|                     | EMT17620.1                               | 77                  | 79                  | —                   | —                   | 50      |
|                     | XP_020172013.1                           | 62                  | —                   | —                   | —                   |         |
| Arabidopsis thaliana| NP_174294.1 (nMAT1)                      | 100                 | —                   | —                   | —                   | 0.0     |
|                     | NP_199503.1 (nMAT2)                      | —                   | 100                 | —                   | —                   |         |
|                     | NP_001154695.1 (nMAT3)                   | —                   | —                   | 100                 | —                   |         |
|                     | NP_177575.1 (nMAT4)                      | —                   | —                   | —                   | 100                 |         |
|                     | CAB85525.1                               | —                   | —                   | —                   | 96                  |         |
|                     | OAO89517.1                               | —                   | —                   | —                   | 99                  |         |
|                     | NP_196025.5                              | —                   | —                   | —                   | 99                  |         |
|                     | OAP16468.1                               | —                   | —                   | —                   | 98                  |         |
|                     | OAP16469.1                               | —                   | —                   | —                   | 89                  |         |
| Amborella trichopoda| XP006844063.1                            | 63                  | —                   | —                   | —                   | 0.0     |
|                     | ERN17872.1                               | —                   | 73                  | —                   | —                   |         |
|                     | ERN07380.1                               | —                   | —                   | 50                  | —                   |         |
|                     | XP_006845705.2                           | —                   | —                   | 52                  | —                   |         |
| Ananas              | OAY76392.1                               | 64                  | —                   | —                   | —                   | 0.0     |
|                     | XP_020088875.1                           | 62                  | —                   | —                   | —                   |         |
|                     | OAY69737.1                               | —                   | 51                  | —                   | —                   |         |
|                     | XP_020109116.1                           | —                   | —                   | —                   | 53                  |         |
| Anthurium amnicola  | JAT50794.1                               | 67                  | —                   | —                   | —                   | 0.0     |
|                     | JAT40274.1                               | 55                  | 72                  | —                   | —                   |         |
|                     | JAT55260.1                               | —                   | —                   | —                   | 50                  |         |
| Arachis duranensis  | XP015948483                              | 71                  | 97                  | —                   | —                   | 0.0     |
|                     | XP_015959251.1                           | —                   | 74                  | —                   | —                   |         |
|                     | XP_015962934.1                           | —                   | —                   | 61                  | —                   |         |
|                     | XP_016163773.1                           | —                   | —                   | —                   | 59                  |         |
|                     | XP_015934970.1                           | —                   | —                   | —                   | 58                  |         |
| Arachis ipaensis    | XP_016203904.1                           | 74                  | —                   | —                   | —                   |         |
|                     | XP_016197721.1                           | —                   | 77                  | —                   | —                   |         |
|                     | XP_016194779.1                           | —                   | —                   | 61                  | —                   |         |
| Arabis alpina       | KFK44825.1                               | 90                  | —                   | —                   | —                   |         |
| Brachypodium distachyon| XP_014752427.1                          | 61                  | —                   | —                   | —                   | 0.0     |
|                     | XP_014757193.1                           | —                   | —                   | —                   | 65                  |         |
| Brassica sp.        | XP013587602.1                            | 87                  | 82                  | 77                  | —                   | 0.0     |
|                     | XP013662585.1                            | 81                  | 83                  | 76                  | —                   |         |
|                     | XP009115198.1                            | 88                  | 82                  | 76                  | —                   |         |
| ORGANISM       | NONANNOTATED/ANNOTATED ACCESSION NUMBERS | % OF IDENTITY MATCH | % OF IDENTITY MATCH | % OF IDENTITY MATCH | % OF IDENTITY MATCH | E VALUE |
|----------------|----------------------------------------|---------------------|---------------------|---------------------|---------------------|---------|
|                |                                        | NMAT1               | NMAT2               | NMAT3               | NMAT4               |         |
| XP_013598629.1 | —                                      | 91                  | —                   | —                   | —                   |         |
| XP_013698082.1 | —                                      | 85                  | —                   | —                   | —                   |         |
| XP_009101465.1 | —                                      | 86                  | —                   | —                   | —                   |         |
| XP_013668195.1 | —                                      | —                   | 80                  | —                   | —                   |         |
| XP_013606479.1 | —                                      | —                   | —                   | 79                  | —                   |         |
| XP_009122640.1 | —                                      | —                   | —                   | 78                  | —                   |         |
| XP_013668386.1 | —                                      | —                   | —                   | 78                  | —                   |         |
| XP_009106093.1 | —                                      | —                   | —                   | —                   | 87                  |         |
| XP_013649718.1 | —                                      | —                   | —                   | —                   | 87                  |         |
| XP_013649757.1 | —                                      | —                   | —                   | —                   | 87                  |         |
| XP_013587847.1 | —                                      | —                   | —                   | —                   | 87                  |         |
| Bovinia sp.    | BAC00897.1                             | —                   | —                   | —                   | —                   | 70      |
|                | BAC00896.1                             | —                   | 74                  | —                   | —                   | 70      |
| Beta vulgaris  | XP_019102903.1                         | 65                  | —                   | —                   | —                   |         |
|                | XP_019102999.1                         | 64                  | —                   | —                   | —                   |         |
|                | KMS96269.1                             | 68                  | —                   | —                   | —                   |         |
|                | KMS97692.1                             | —                   | 75                  | —                   | —                   |         |
|                | XP_010695416.1                         | —                   | 76                  | —                   | —                   |         |
|                | XP_010672964.2                         | —                   | —                   | 56                  | —                   |         |
|                | KMT09166.1                             | —                   | —                   | —                   | 57                  |         |
|                | XP_010680216.1                         | —                   | —                   | —                   | 57                  |         |
| Capsicum annuum| XP_016540023.1                         | 67                  | —                   | —                   | —                   |         |
|                | XP_016570924.1                         | —                   | 74                  | —                   | —                   |         |
|                | XP_016570996.1                         | —                   | 72                  | —                   | —                   |         |
|                | XP_016571240.1                         | —                   | 78                  | —                   | —                   |         |
|                | XP_016571839.1                         | —                   | —                   | 58                  | —                   |         |
|                | XP_016571840.1                         | —                   | —                   | 62                  | —                   |         |
| Citrus sinensis| XP_006467472.1                         | 66                  | —                   | —                   | —                   | 0.0     |
|                | XP_006465050.1                         | 76                  | 80                  | —                   | —                   |         |
|                | XP_006465053.1                         | 72                  | 78                  | —                   | —                   |         |
|                | JAV45571.1                             | 66                  | 80                  | —                   | —                   |         |
|                | XP_006485360.1                         | —                   | 82                  | —                   | —                   |         |
|                | JAV45493.1                             | —                   | 82                  | —                   | —                   |         |
|                | XP_006485361.1                         | —                   | —                   | 65                  | —                   |         |
|                | XP_006491615.1                         | —                   | —                   | 65                  | —                   |         |
|                | JAV45214.1                             | —                   | —                   | 65                  | —                   |         |
|                | XP_006465048.1                         | —                   | —                   | —                   | 61                  |         |
|                | XP_006465050.1                         | —                   | —                   | —                   | 61                  |         |
|                | XP_006465051.1                         | —                   | —                   | —                   | 59                  |         |
| ORGANISM         | ACCESSION NUMBERS | % OF IDENTITY MATCH | % OF IDENTITY MATCH | % OF IDENTITY MATCH | % OF IDENTITY MATCH | E VALUE |
|------------------|-------------------|---------------------|---------------------|---------------------|---------------------|---------|
| Camelina sativa  | XP019094631.1     | 93                  | —                   | —                   | —                   | 0.0     |
|                  | XP019090818.1     | 92                  | —                   | —                   | —                   |         |
|                  | XP019099276.1     | 94                  | —                   | —                   | —                   |         |
|                  | XP_019097174.1    | —                   | 96                  | —                   | —                   |         |
|                  | XP_019097173.1    | —                   | 95                  | —                   | —                   |         |
|                  | XP_010481395.1    | —                   | 95                  | —                   | —                   |         |
|                  | XP_010481396.1    | —                   | 95                  | —                   | —                   |         |
|                  | XP_019087551.1    | —                   | 96                  | —                   | —                   |         |
|                  | XP_010423639.1    | —                   | —                   | 91                  | —                   |         |
|                  | XP_019090510.1    | —                   | —                   | —                   | 90                  |         |
|                  | XP_010490900.1    | —                   | —                   | —                   | 88                  |         |
|                  | XP_019083359.1    | —                   | —                   | —                   | 92                  |         |
|                  | XP_019084999.1    | —                   | —                   | —                   | 92                  |         |
| Cicer arietinum   | XP004486025.1     | 70                  | —                   | —                   | —                   | 0.0     |
|                  | XP_004486448.1    | —                   | 79                  | —                   | —                   |         |
|                  | XP_012575420.1    | —                   | —                   | 60                  | —                   |         |
|                  | XP_004485704.1    | —                   | —                   | —                   | 53                  |         |
| Cucumis sp.      | XP_008444296.1    | 66                  | —                   | —                   | —                   | 0.0     |
|                  | XP_011657261.1    | 71                  | —                   | —                   | —                   |         |
|                  | XP_008465920.1    | —                   | 76                  | —                   | —                   |         |
|                  | XP_004148504.1    | —                   | 79                  | 61                  | —                   |         |
|                  | XP_008463418.1    | —                   | —                   | 61                  | —                   |         |
|                  | XP_011650528.1    | —                   | —                   | 61                  | —                   |         |
|                  | XP_0084402019.1   | —                   | —                   | —                   | 60                  |         |
|                  | XP_011653460.1    | —                   | —                   | —                   | 59                  |         |
| Daucus carota    | XP_017239512.1    | 68                  | —                   | —                   | —                   |         |
| subsp. sativus   | KZN11262.1        | 68                  | —                   | —                   | —                   |         |
|                  | XP_017231682.1    | —                   | 76                  | —                   | —                   |         |
|                  | KZN04764.1        | —                   | 81                  | —                   | —                   |         |
|                  | KZM93721.1        | —                   | —                   | 58                  | —                   |         |
|                  | XP_017253957.1    | —                   | —                   | 55                  | —                   |         |
|                  | XP_017247780.1    | —                   | —                   | —                   | 57                  |         |
| Elaeis guineensis| XP_010913705.1    | 65                  | —                   | —                   | —                   |         |
|                  | XP_010933533.1    | —                   | 73                  | —                   | —                   |         |
|                  | XP_019705156.1    | —                   | —                   | 52                  | —                   |         |
|                  | XP_010942294.2    | —                   | —                   | —                   | 54                  |         |
| ORGANISM       | NONANNOTATED/ANNOTATED ACCESSION NUMBERS | % OF IDENTITY MATCH | % OF IDENTITY MATCH | % OF IDENTITY MATCH | % OF IDENTITY MATCH | E VALUE |
|----------------|----------------------------------------|---------------------|---------------------|---------------------|---------------------|---------|
|                |                                        | NMAT1               | NMAT2               | NMAT3               | NMAT4               |         |
| *Fragaria vesca* |                                        |                     |                     |                     |                     | 0.0     |
| XP_011460193.1  | 72                                     | —                   | —                   | —                   | —                   |         |
| XP_004300288.1  | —                                      | 78                  | —                   | —                   | —                   |         |
| XP_004303613.1  | —                                      | —                   | 58                  | —                   | —                   |         |
| XP_004307117.1  | —                                      | —                   | —                   | 63                  | —                   |         |
| *Glycine sp.*   |                                        |                     |                     |                     |                     | 0.0     |
| XP_006600812.1  | —                                      | 80                  | —                   | —                   | —                   |         |
| XP_003547207.1  | 70                                     | —                   | —                   | —                   | —                   |         |
| XP_003547325.1  | —                                      | 74                  | —                   | —                   | —                   |         |
| XP_003547691.1  | —                                      | 78                  | —                   | —                   | —                   |         |
| KRH09575.1      | —                                      | —                   | 61                  | —                   | —                   |         |
| XP_003547641.1  | —                                      | —                   | 61                  | —                   | —                   |         |
| XP_006599285.1  | —                                      | —                   | 59                  | —                   | —                   |         |
| XP_003548802.2  | —                                      | —                   | 59                  | —                   | —                   |         |
| XP_014624412.1  | —                                      | —                   | 59                  | —                   | —                   |         |
| KHN24575.1      | —                                      | —                   | 59                  | —                   | —                   |         |
| XP_006600812.1  | —                                      | —                   | —                   | 60                  | —                   |         |
| KHN09761.1      | —                                      | —                   | —                   | 60                  | —                   |         |
| *Gossypium sp.* |                                        |                     |                     |                     |                     | 0.0     |
| XP016722554.1   | 73                                     | —                   | —                   | —                   | —                   |         |
| XP012435002.1   | 73                                     | —                   | —                   | —                   | —                   |         |
| XP016708607.1   | 72                                     | —                   | —                   | —                   | —                   |         |
| XP_017628169.1  | 73                                     | —                   | —                   | —                   | —                   |         |
| XP_012472117.1  | —                                      | 82                  | —                   | —                   | —                   |         |
| XP_017638496.1  | —                                      | 82                  | —                   | —                   | —                   |         |
| XP_016721510.1  | —                                      | 82                  | —                   | —                   | —                   |         |
| XP_017608382.1  | —                                      | —                   | 66                  | —                   | —                   |         |
| XP_016669292.1  | —                                      | —                   | 66                  | —                   | —                   |         |
| XP_012485492.1  | —                                      | —                   | 66                  | —                   | —                   |         |
| XP_016673099.1  | —                                      | —                   | 66                  | —                   | —                   |         |
| XP_016715130.1  | —                                      | —                   | —                   | 61                  | —                   |         |
| XP_012487059.1  | —                                      | —                   | —                   | 61                  | —                   |         |
| XP_017608885.1  | —                                      | —                   | —                   | 61                  | —                   |         |
| *Jatropha curcas* |                                        |                     |                     |                     |                     | 0.0     |
| XP_012072095.1  | 71                                     | —                   | —                   | —                   | —                   |         |
| XP_012075086.1  | —                                      | 81                  | —                   | —                   | —                   |         |
| KDP45956.1      | —                                      | 81                  | —                   | —                   | —                   |         |
| XP_012085971.1  | —                                      | —                   | 64                  | —                   | —                   |         |
| XP_012087956.1  | —                                      | —                   | —                   | 60                  | —                   |         |
| XP_012087958.1  | —                                      | —                   | —                   | 60                  | —                   |         |
| *Malus*        |                                        |                     |                     |                     |                     | 0.0     |
| XP017189563.1   | 71                                     | —                   | —                   | —                   | —                   |         |
| XP_008373619.1  | —                                      | 79                  | —                   | —                   | —                   |         |
| XP_008385840.1  | —                                      | —                   | 63                  | —                   | —                   |         |
| XP_008348277.1  | —                                      | —                   | —                   | 64                  | —                   |         |
| ORGANISM          | NONANNOTATED/ANNOTATED ACCESSION NUMBERS | % OF IDENTITY MATCH | % OF IDENTITY MATCH | % OF IDENTITY MATCH | % OF IDENTITY MATCH | E VALUE |
|-------------------|-----------------------------------------|---------------------|---------------------|---------------------|---------------------|---------|
|                   |                                         | NMAT1   | NMAT2   | NMAT3   | NMAT4   |         |
| Morus notabilis   | XP010110974.1                           | 70      | —       | —       | —       | 0.0     |
|                   | XP_010094975.1                          | —       | 80      | —       | —       |         |
|                   | XP_010102735.1                          | —       | —       | 59      | —       |         |
| Noccaea sp.       | JAU99057.1                              | 92      | 96      | —       | —       | 0.0     |
|                   | JAU42137.1                              | 92      | 96      | —       | —       |         |
|                   | JAU86104.1                              | —       | 91      | —       | —       |         |
|                   | JAU33773.1                              | —       | 91      | —       | —       |         |
|                   | JAU09470.1                              | —       | 90      | —       | —       |         |
|                   | JAU85394.1                              | —       | —       | —       | 89      |         |
|                   | JAU07244.1                              | —       | —       | —       | 89      |         |
|                   | JAU58667.1                              | —       | —       | —       | 91      |         |
| Nicotiana sp.     | XP_016489538.1                          | 68      | —       | —       | —       | 0.0     |
|                   | XP_019234436.1                          | 68      | —       | —       | —       |         |
|                   | XP_009789087.1                          | 68      | —       | —       | —       |         |
|                   | XP_009608357.1                          | 68      | —       | —       | —       |         |
|                   | XP_016512170.1                          | 68      | —       | —       | —       |         |
|                   | XP_009776964.1                          | —       | 75      | —       | —       |         |
|                   | XP_009629279.1                          | —       | 75      | —       | —       |         |
|                   | XP_016437068.1                          | —       | 75      | —       | —       |         |
|                   | XP_019252521.1                          | —       | 75      | —       | —       |         |
|                   | XP_019236962.1                          | —       | —       | 60      | —       |         |
|                   | XP_016440529.1                          | —       | —       | 59      | —       |         |
|                   | XP_016478800.1                          | —       | —       | 60      | —       |         |
|                   | XP_009763877.1                          | —       | —       | 60      | —       |         |
|                   | XP_009622396.1                          | —       | —       | 59      | —       |         |
|                   | XP_009801671.1                          | —       | —       | —       | 60      |         |
|                   | XP_009801675.1                          | —       | —       | —       | 60      |         |
|                   | XP_016448890.1                          | —       | —       | —       | 60      |         |
|                   | XP_018631962.1                          | —       | —       | —       | 60      |         |
|                   | XP_009620643.1                          | —       | —       | —       | 60      |         |
|                   | XP_019259655.1                          | —       | —       | —       | 60      |         |
|                   | XP_019259660.1                          | —       | —       | —       | 60      |         |
|                   | XP_019259661.1                          | —       | —       | —       | 57      |         |
| Oryza brachyantha | XP_015698849.1                          | 63      | —       | —       | —       | 0.0     |
|                   | EEC69135.1                              | 62      | —       | —       | —       |         |
|                   | XP_015619809.1                          | 62      | —       | —       | —       |         |
|                   | ABA97559.1                              | 59      | —       | —       | —       |         |
|                   | BAF29652.1                              | 63      | —       | —       | —       |         |
|                   | XP_006662081.1                          | —       | 66      | —       | —       |         |
|                   | XP_015614743.1                          | —       | 65      | —       | —       |         |
| ORGANISM       | NONANNOTATED/ANNOTATED ACCESSION NUMBERS | % OF IDENTITY MATCH | % OF IDENTITY MATCH | % OF IDENTITY MATCH | % OF IDENTITY MATCH | E VALUE |
|---------------|----------------------------------------|---------------------|---------------------|---------------------|---------------------|---------|
|               |                                        | NMAT1               | NMAT2               | NMAT3               | NMAT4               |         |
| Phaseolus vulgaris | ESW27588.1                             | 75                  | 79                  | —                   | —                   | 0.0     |
|                | ESW27589.1                             | 75                  | 79                  | —                   | —                   |         |
|                | ESW27590.1                             | 75                  | 79                  | —                   | —                   |         |
|                | XP_007147490.1                         | —                   | 73                  | —                   | —                   |         |
|                | XP_007155594.1                         | —                   | —                   | 59                  | —                   |         |
| Populus sp.   | AAU03342.1                             | —                   | —                   | —                   | —                   | 0.0     |
|                | XP_011009930.1                         | 73                  | —                   | —                   | —                   |         |
|                | XP_006368468.1                         | —                   | 82                  | —                   | —                   |         |
|                | XP_011027297.1                         | —                   | —                   | 60                  | —                   |         |
|                | XP_011031175.1                         | —                   | —                   | —                   | 60                  |         |
| Prunus mume   | XP_008225479.1                         | 79                  | —                   | —                   | —                   | 0.0     |
|                | XP_008229517.1                         | —                   | 78                  | —                   | —                   |         |
|                | XP_007208922.1                         | —                   | 77                  | —                   | —                   |         |
|                | XP_016651634.1                         | —                   | —                   | 64                  | —                   |         |
|                | ONH95023.1                             | —                   | —                   | 63                  | —                   |         |
|                | ONI20797.1                             | —                   | —                   | —                   | 62                  |         |
|                | ONI20796.1                             | —                   | —                   | —                   | 62                  |         |
| Ricinus communis | XP_002527885.2                     | —                   | —                   | —                   | —                   | 0.0     |
|                | XP_002519696.1                         | 70                  | —                   | —                   | —                   |         |
|                | XP_002528785.1                         | —                   | 77                  | —                   | —                   |         |
|                | XP_002533678.1                         | —                   | —                   | 66                  | —                   |         |
|                | XP_002527885.2                         | —                   | —                   | —                   | 59                  |         |
|                | EEF34516.1                             | —                   | —                   | —                   | 60                  |         |
| Raphanus sativus | XP018434021.1                        | 87                  | —                   | —                   | —                   | 0.0     |
|                | XP_018485322.1                         | —                   | 86                  | —                   | —                   |         |
|                | XP_018442287.1                         | —                   | —                   | 84                  | —                   |         |
|                | XP_018442286.1                         | —                   | —                   | 79                  | —                   |         |
|                | XP_018446864.1                         | —                   | —                   | 85                  |         |         |
| Setaria italica | XP_014660530.1                     | 62                  | —                   | —                   | —                   | 0.0     |
|                | XP_004983602.1                         | —                   | 66                  | —                   | —                   |         |
| Sorghum bicolor | XP_002468369.1                     | 63                  | —                   | —                   | —                   | 0.0     |
|                | XP_002465818.1                         | —                   | 65                  | —                   | —                   |         |
|                | XP_002438709.1                         | —                   | —                   | —                   | 62                  |         |
| Solanum sp.   | XP_004231081.1                         | 69                  | —                   | —                   | —                   | 0.0     |
|                | XP_006364540.1                         | 69                  | —                   | —                   | —                   |         |
|                | XP_006358607.1                         | —                   | 75                  | —                   | —                   |         |
|                | XP_015083496.1                         | —                   | 75                  | —                   | —                   |         |
|                | XP_004246177.1                         | —                   | 75                  | —                   | —                   |         |
| ORGANISM                  | NONANNOTATED/ANNOTATED ACCESSION NUMBERS | % OF IDENTITY MATCH | % OF IDENTITY MATCH | % OF IDENTITY MATCH | % OF IDENTITY MATCH | E VALUE |
|---------------------------|-----------------------------------------|---------------------|---------------------|---------------------|---------------------|---------|
|                           |                                         | NMAT1              | NMAT2              | NMAT3              | NMAT4              |         |
| XP_015067974.1            | —                                       | —                  | 58                 | —                  | —                   |         |
| XP_006343366.1            | —                                       | —                  | 58                 | —                  | —                   |         |
| XP_004234535.1            | —                                       | —                  | 58                 | —                  | —                   |         |
| XP_006341072.1            | —                                       | —                  | —                  | 59                 | —                   |         |
| XP_004246478.1            | —                                       | —                  | —                  | 58                 | —                   |         |
| XP_015086285.1            | —                                       | —                  | —                  | 58                 | —                   |         |
| Sesamum                   | XP011096554.1                           | 70                 | —                  | —                  | —                   | 0.0     |
|                           | XP_011100130.1                          | —                  | 79                 | —                  | —                   |         |
|                           | XP_011078929.1                          | —                  | —                  | 60                 | —                   |         |
|                           | XP_011077836.1                          | —                  | —                  | —                  | 58                 |         |
| Spinacia oleracea         | KNA15840.1                              | 76                 | —                  | —                  | —                   |         |
|                           | KNA24686.1                              | —                  | 77                 | —                  | —                   |         |
|                           | KNA14637.1                              | —                  | —                  | 54                 | —                   |         |
| Tarenaya hassleriana      | XP010549486.1                           | 80                 | —                  | —                  | —                   | 0.0     |
|                           | XP_010518672.1                          | —                  | 85                 | —                  | —                   |         |
|                           | XP_010552974.1                          | —                  | —                  | 76                 | —                   |         |
|                           | XP_019058746.1                          | —                  | —                  | 74                 | —                   |         |
| Theobroma cacao           | XP_007029012.2                          | 73                 | —                  | —                  | —                   | 0.0     |
|                           | EOY09514.1                              | 75                 | —                  | —                  | —                   |         |
|                           | XP_007039600.2                          | —                  | 80                 | —                  | —                   |         |
|                           | EOY24103.1                              | —                  | 80                 | —                  | —                   |         |
|                           | EOY24100.1                              | —                  | 80                 | —                  | —                   |         |
|                           | XP_017977755.1                          | —                  | —                  | 76                 | —                   |         |
|                           | EOY14249.1                              | —                  | —                  | 67                 | —                   |         |
|                           | EOX92573.1                              | —                  | —                  | 63                 | —                   |         |
|                           | XP_017980919.1                          | —                  | —                  | —                  | 63                 |         |
| Vitis vinifera            | XP_002269568.1                          | 72                 | —                  | —                  | —                   | 0.0     |
|                           | XP_002269047.1                          | —                  | 80                 | —                  | —                   |         |
|                           | CBI30076.3                              | —                  | —                  | 59                 | —                   |         |
|                           | XP_002274379.2                          | —                  | —                  | 66                 | —                   |         |
|                           | XP_010646090.1                          | —                  | —                  | —                  | 62                 | —       |
| Vigna radiata             | XP_017434120.1                          | 70                 | —                  | —                  | —                   | 0.0     |
|                           | XP_014517011.1                          | 70                 | —                  | —                  | —                   |         |
|                           | XP_017437103.1                          | —                  | 76                 | —                  | —                   |         |
|                           | XP_014519153.1                          | —                  | 76                 | —                  | —                   |         |
|                           | XP_014517263.1                          | —                  | —                  | 63                 | —                   |         |
|                           | XP_017423160.1                          | —                  | —                  | 60                 | —                   |         |
|                           | KOM43061.1                              | —                  | —                  | 60                 | —                   |         |
controlling reactive oxygen species in roots and leaves, seed germination, and establishment and development of seedlings. In this study, the phylogenetic relationship of the 4 nMAT sequences of *A. thaliana* is established, especially with respect to other plant species.

It appears that mitochondrial splicing factors in plants may represent a step in gradual evolutionary transformation from specific maturase-facilitated self-splicing introns toward the complex spliceosomal machinery found in eukaryotes. This hypothesis has received credence by homology of the core splicing factor Prp8 with maturases. We propose that such an important process must be conserved in all plants. To investigate further the functions and mechanisms of these proposed spliceosomal “descendants” within the nucleus and their conservation, we adopted a bioinformatics-based approach to analyze all available sequences for the presence or absence of nMAT genes or nMATs.

**Materials and Methods**

**nMAT homologs of Arabidopsis using BLASTP**

Each of 4 nMAT full-length protein sequences from *Arabidopsis* was screened on National Center for Biotechnology Information (NCBI) BLAST, against existing sequences of nonredundant GenBank CDS translations, PDB database, Swiss-Prot database, and PIR and PRF database, excluding environmental samples from Whole Genome Shotgun (WGS) projects, keeping the default settings, after which the organisms with an E value of 0.0 and identity match ≥50% were considered for the test. NR database was individually extracted for every nMAT protein. A sample search date for the BLASTP is query ID: lcl|Query_115880 for nMAT1, lcl|Query_12347 for nMAT2, lcl|Query_75782 for nMAT3, and lcl|Query_27575 for nMAT4 on March 11, 2017.

**Dendrogram of nMATs and their homologs**

Phylogenetic trees with bootstrapping were derived using the Molecular Evolutionary Genetic Analysis version 6.0 (MEGA) software. Clustal Omega Multiple Sequence Alignment (MSA) output files for all the 4 nMAT sequences were converted to MEG files in the MEGA software which were further used to build the phylogenetic trees.

Phylogenetic trees were constructed using the Neighbour-Joining method, keeping the bootstraping values at 1000 replicates (the Poisson model was used for the amino acid substitution). Phylogenetic trees when constructed need to be established with respect to certainty of the estimations, which represent the end result confidence of the tree topology. Bootstrapping results in a tree with bootstrap support values, signifying the number of times that particular clade was formed during bootstrapping.

The tree consisted of sequences of all the high-homology proteins with nMATs 1 to 4. Finally, BLAST tree view (Fast minimum method) generated from each BLAST (Figures S1, S2, S3, S4) was also analyzed as distance trees (RID numbers: nMAT1: RIDC6UYKM6P016; nMAT2: RIDC7F5J9CS016; nMAT3: RIDC7FUZES016; nMAT4: RIDC7W8SBED013).
Results and Discussion

Genome-wide analysis for nMAT homologs of Arabidopsis

Analysis of protein databases led to the identification of several protein sequences of the nMAT family in several plants. The nMAT protein sequences were found in a wide range of plants. In this study, the proteins which are nonannotated (annotated as hypothetical, predicted, or unknown protein) but found to have high identity with nMAT proteins were included for further analysis (see Table 1). Table 2 is a compiled representation of all the proteins in different plants which share high identity with nMATs, and they are annotated as nMAT-like. The accession with an "*" are already annotated Intron maturase, type II family protein in NCBI database.

BLAST search for all 4 nMAT sequences resulted in several similar sequences with high identity score. Nuclear maturase sequences were found to have high-identity match with 282 protein sequences distributed in various plant species enlisted in Table 1. The organisms in which nMATs 1 to 4 have identity matches are Agriolaps tauschii, A thaliana, Amborella trichopoda, Arabidopsis thaliana, Ananas, Anthurium amnicola, Arachis sp., Arabis alpina, Brachypodium distachyon, Brassica sp., Bowenia sp., Beta vulgaris, Capsicum annuum, Citrus sp., Camelina sativa, Cicer arietinum, Cucumis sp., Daucus carota subsp. Sativus, Elaeis guineensis, Fragaria vesca, Glycine sp., Gossypium sp., Jatropha curcas, Malus, Morus nobilis, Noccaea sp., Nicotiana sp., Oryza brachyantha, Phascolus vulgaris, Populus sp., Prunus mume, Ricinus communis, Raphanus sativus, Setaria italic, Sorghum bicolor, Solanum sp., Sesamum, Spinacia oleracea, Tarenaya haustoriana, Theobroma cacao, Vitis vinifera, Vigna radiate, Zea mays, and Ziziphus.

Phylogenetic analysis reveals conservation of nMATs across plants

The phylogenetic trees/dendrograms with bootstrap values (for 1000 replicates) were generated in MEGA using Clustal Omega MSA outputs. These trees show high bootstrap values for all 4 nMAT sequences.

MEGA tree view revealed that all the 4 nMATs share homology with the members of eudicots, monocots, gymnosperms, angiosperms, pteridophytes, and bryophytes. Tree view of nMATs 1 to 4 shows that all 4 nMATs are present in bacteria, fungi, and all plant groups (Figures S1, S2, S3, S4).

The MEGA phylogenetic tree of nMAT proteins (Figure 1) shows the presence of multiple groups of closely related branches with high bootstrap values, even with 100% bootstrap value. However, a few bootstrap values are as low as 11%, showing that they have very low homology between the nMAT sequences in A thaliana and others.

A dendrogram was generated using the homologs to nMATs 1 to 4 along with other maturase sequences of A thaliana (Figure 1). Among the highly homologous sequences were XP017638496.1, XP007039600.2, XP004148504.1, XP00430288.1, XP019087551.1, XP006485360.1, XP011657261.1, XP013662585.1, XP006358607.1, XP010549486.1, JAU09470.1, OAY76392.1, OAP16468.1, EOY92573.1, KZJ11262.1, BAF29652.1, EOY14249.1, and many other protein sequences. Analysis of evolutionary relationships showed that the maturase protein had 100 bootstrap value, signifying that this sequence is most probably a maturase-coding sequence.

However, there were many other sequences which had high bootstrap values more than 50 and less than 100: XP012472117.1 (78), XP019097173.1 (99), XP015083496.1 (64), XP009776964.1 (53), XP017235912.1 (55), XP006491615.1 (99), XP015891101.1 (94), XP006465505.1 (72), XP006600812.1 (81), XP016197721 (97), XP017603828.1 (94), XP002269047 (83), JAV44975.1 (57), NP177575.1 (65), EMT17620.1 (81), and others.

Apart from these, XP016571839.1, XP011078929.1, XP003534769.1, XP016673099.1, XP002519696.1, CABB5525.1, ESW 27588.1, AAU 03342.1, CBI30076.3, KMS9769.1, KZM39721.1, YP740389.1, and a few others represented very low bootstrap values. However, CABB5525.1 has been annotated as maturase-like protein (https://www.ncbi.nlm.nih.gov/protein/cab85525.1). The highest bootstrap values were noted for A tauschii, Glycerina sp., Brassica sp., J curcas, C sativa, Noccaea sp., C arietinum, S bicolor, C sativus, R communis, Gossypium sp., A duranensis, T cacao, V vinifera, P nume, Zea sp., Ziziphus sp., Nicotiana sp., Oryza sp., Vigna sp., Ananas sp., Phascolus, and a few other species for maturase protein against nMAT of A thaliana. A few sequences of these plants, however, showed low bootstrap values, meaning that they have very low evolutionary relation.

Moradian et al.19 on complete mitochondrial genome sequencing of 3 Tetrahymena species revealed that there were mutation hot spots and accelerated nonsynonymous substitutions in Ymf genes. These could be mitochondrial maturases, but mitochondrial molecular characterization will have to be done for relevance of the assumption. In this study, 72 accesses were found to be ymf-like. The matK gene encodes a maturase that is involved in splicing type II introns, and the matK sequence has often been used in phylogenetic and evolutionary studies.20 On maturase phylogeny, recently Guo and Mower21 provided an insight of evolution of maturases in different land plants and showed that nMATs 3 and 4 are closely related to chloroplast matK Maturase. In this study, matK of Bowenia sp. (BAC00897.1 and BAC00896.1) shows high homology with nMAT2 and nMAT4 proteins (Table 1). It indicates possibility of nucleus encoding some maturase proteins which may perform splicing in both mitochondria and chloroplast.

Thus, the MEGA tree view analysis shows that all the Arabidopsis nMAT proteins have phylogenetic relations among all the plant groups. As nMATs are crucial for aerobic respiration, it was expected that they would be present in all plant genomes. In silico analysis has led to the identification of
Table 2. List of all plants considered in the test (% identity \(\geq 50\%) and E=0) and the phylogenetic relation with each of the nMAT proteins.

| ORGANISM               | nMAT1            | nMAT2            | nMAT3            | nMAT4            |
|------------------------|------------------|------------------|------------------|------------------|
| Aegilops tauschii      | EMT06317.1       | EMT1620.1        | EMT06317.1       | EMT1620.1        |
|                        | XP_02172013.1    |                  |                  |                  |
| Arabidopsis thaliana   | NP_174294.1* (nMAT1) | NP_199503.1* (nMAT2) | NP_001154695.1* (nMAT3) | NP_177575.1* (nMAT4) |
|                        | CAB85525.1*      | OAO85171.1       | OAP16468.1       | OAP16469.1       |
| Amborella trichopoda   | XP_006844063     | ERN17872.1       | ERN07380.1       | XP_006845705.2   |
| Ananas                 | OAY76392.1       |                  | OAY69737.1       | XP_020109116.1   |
|                        | XP_020088875.1   |                  |                  |                  |
| Anthurium amnicola     | JAT50794.1       | JAT40274.1       |                  |                  |
| Arachis duranensis     | XP_015948483     | XP_01599251.1    | XP_015962934.1   | XP_015962934.1   |
| Arachis ipaensis       | XP_016203904.1   | XP_016197721.1   | XP_016194779.1   |                  |
| Arabis alpina          | KFK44825.1       |                  |                  |                  |
| Brachypodium distachyon| XP_014752427.1   | XP_014757193.1   |                  |                  |
| Brassica sp.           | XP_013587602.1   | XP_013662585.1   | XP_009115198.1   | XP_009106093.1   |
|                        | XP_013598629.1   | XP_013698082.1   | XP_013668195.1   | XP_013649757.1   |
|                        | XP_013668386.1   |                  | XP_013668386.1   |                  |
| Bowenia sp.            | BAC00896.1       |                  |                  |                  |
|                        | BAC00897.1       |                  |                  |                  |
|                        | BAC00896.1       |                  |                  |                  |
| Beta vulgaris          | XP_019102903.1   | XP_019102999.1   | KMS96269.1       |                  |
|                        | XP_010672964.2   |                  |                  |                  |
|                        | KMT09166.1       |                  |                  |                  |
| Capsicum annuum        | XP_016540023.1   | XP_016570924.1   | XP_016571839.1   |                  |
|                        | XP_016570961.1   |                  | XP_016571840.1   |                  |
|                        |                    |                  |                  |                  |
| Citrus sp.             | XP_006467472.1   | XP_006465050.1   | XP_006485360.1   |                  |
|                        | XP_006450481.1   |                  |                  |                  |
|                        | XP_006465501.1   |                  |                  |                  |
|                        | XP_004148504.1   |                  |                  |                  |
| Camelina sativa        | XP_019094631.1   | XP_019090081.1   | XP_019083359.1   |                  |
|                        | XP_01909099276.1 |                  | XP_019084999.1   |                  |
|                        |                  |                  |                  |                  |
| Cicer arietinum        | XP_004486025.1   | XP_004486448.1   | XP_012575420.1   | XP_004485704.1   |
|                        | XP_004014850.1   |                  |                  |                  |
|                        | XP_004014850.1   |                  |                  |                  |
| Cucumis sp.            | XP_008444296.1   | XP_008465924.1   | XP_008442019.1   |                  |
|                        | XP_001653460.1   |                  |                  |                  |
| Daucus carota subsp.   | XP_017239512.1   | XP_017231682.1   | KZM03721.1       |                  |
| sativus                | KZN11262.1       | KZN04764.1       | XP_017253957.1   |                  |
| Elaeis guineensis      | XP_010913705.1   | XP_010933533.1   | XP_019705166.1   |                  |
| Fragaria vesca         | XP_011460193.1   | XP_004303613.1   | XP_004307117.1   |                  |
(Continued)
| ORGANISM          | NMAT1               | NMAT2               | NMAT3               | NMAT4               |
|-------------------|---------------------|---------------------|---------------------|---------------------|
| Glycine sp.       | XP_003547207.1      | XP_003547325.1      | KRH09575.1          | XP_006600812.1      |
|                   |                     | XP_0035434769.1     |                     | KHN09761.1          |
|                   |                     |                     |                     |                     |
| Gossypium sp.     | XP_016722554.1      | XP_012472117.1      | XP_017608382.1      | XP_016715130.1      |
|                   | XP_012435002.1      | XP_017638496.1      | XP_016669292.1      | XP_012487059.1      |
|                   | XP_016708607.1      | XP_016721510.1      | XP_012485492.1      | XP_017608885.1      |
|                   | XP_017628169.1      |                     |                     |                     |
| Jatropha curcas   | XP_012072095.1      | XP_012075086.1      | XP_012085971.1      | XP_012087956.1      |
|                   |                     |                     |                     | XP_012087958.1      |
| Malus             | XP_017189563.1      | XP_008373619.1      | XP_008385840.1      | XP_008348277.1      |
| Morus notabilis   | XP_010110974.1      | XP_010094975.1      |                     |                     |
|                   |                     |                     |                     |                     |
| Noccaea sp.       | JAU99057.1          | JAU99057.1          | JAU85394.1          | JAU58667.1          |
|                   | JAU42137.1          | JAU42137.1          |                     |                     |
| Nicotiana tabacum | XP_016489538.1      | XP_009776964.1      | XP_019236962.1      | XP_009860171.1      |
|                   | XP_019234436.1      | XP_009629279.1      | XP_016440529.1      | XP_00986175.1       |
|                   | XP_009789087.1      | XP_016437068.1      | XP_016478800.1      | XP_016448890.1      |
|                   | XP_016512170.1      | XP_019252521.1      | XP_009763877.1      | XP_018631962.1      |
|                   |                     |                     | XP_009622396.1      |                     |
| Oryza brachyantha | XP_015668849.1      | XP_006662081.1      |                     |                     |
|                   | EEC69135.1          |                     |                     |                     |
|                   | XP_015619809.1      |                     |                     |                     |
|                   | ABA97559.1*         |                      |                     |                     |
|                   | BAF29652.1          |                      |                     |                     |
| Phaseolus vulgaris| ESW27588.1          | ESW27589.1          | ESW27590.1          |                     |
|                   | ESW27589.1          | ESW27590.1          | XP_007147490.1      |                     |
|                   |                     |                     |                     |                     |
| Populus sp.       | XP_011009930.1      | XP_006368468.1      |                     |                     |
|                   |                     |                     |                     |                     |
| Prunus mume       | XP_008225479.1      | XP_008229517.1      |                     |                     |
|                   |                     |                     |                     |                     |
| Ricinus communis  | XP_002519696.1      |                     |                     |                     |
|                   | XP_002528785.1      |                     |                     |                     |
| Raphanus sativus  | XP_018434021.1      | XP_018443287.1      |                     |                     |
|                   |                     |                      |                     |                     |
| Setaria italica   | XP_014660530.1      |                     |                     |                     |
|                   |                     |                     |                     |                     |
| Sorghum bicolor   | XP_002468369.1      |                     |                     |                     |
|                   |                     |                     |                     |                     |
| Solanum sp.       | XP_004231081.1      |                     |                     |                     |
|                   | XP_006364540.1      |                     |                     |                     |
|                   |                     |                     |                     |                     |
| Sesamum           | XP_011096554.1      |                     |                     |                     |
|                   |                     |                     |                     |                     |
| Spinacia oleracea | KNA15840.1          |                     |                     |                     |
|                   | KNA24686.1          |                     |                     |                     |
|                   |                     |                     |                     |                     |
| Tarenaya hassleriana| XP010549486.1      |                     |                     |                     |
|                   |                     |                     |                     |                     |
| Theobroma cacao   | XP_007029012.2      |                     |                     |                     |
|                   | EOY09514.1          |                     |                     |                     |
|                   |                     |                     |                     |                     |
several protein sequences in numerous plants which may have splicing role in plants.

However, this study revealed $\geq 50\%$ identity match of nMAT proteins among *A. thaliana* and most other species, even after usage of predicted and unknown accession numbers. As seen from the results, among plants, nMAT1 sequence has high-identity match with sequences in XP006844063.1, OAY76392.1, XP_020088875.1, JAT50794.1, XP016203904.1, KFK44825.1, XP013587602.1, XP009115198.1, XP_011657261.1, XP019094631.1, XP019090818.1, XP0190909276.1, XP004486025.1, XP_017239512.1, XP_011460193.1, XP_003547207.1, XP_017628169.1, XP_012072095.1, JAT541237.1, XP_009789087.1, XP_011009930.1, XP_014660530.1, XP_00634540.1, XP011096554.1, XP010549486.1, EOY95141.1, XP_002269568.1, XP_017434120.1, XP_008651784.1, XP_015898038.1, and others.

In nMAT2, the sequence was found to have high-identity match with sequences in EMT06317.1, EMT17620.1, ERN17872.1, JAT50794.1, XP015948483, XP_016197721.1, XP_00634540.1, XP011096554.1, XP010549486.1, EOY95141.1, XP_002269568.1, XP_017434120.1, XP_008651784.1, XP_015877305.1, EEC69135.1, XP_015619809.1, XP_020172013.1, XP_019102999.1, KMS96269.1, and others.

In nMAT3, the sequence was found to have high-identity match with sequences in CAB85525.1, OAY69737.1, XP_013598629.1, XP_013668386.1, XP_016721510.1, KDP45956.1, XP_008348277.1, JAU58667.1, JAU85394.1, JAU07244.1, XP_019259655.1, ONI20796.1, XP_018446864.1, XP_019058746.1, XP_015891101.1, KOM32641.1, and others.

Finally, the nMAT4 sequence was shown to have high-identity match with sequences in EMT17620.1, OAP16469.1, OAP16469.1, XP_020109116.1, XP_015934970.1, XP_013587847.1, XP_010680216.1, JAV49751.1, XP_006465048.1, XP_019083359.1, XP_019084999.1, XP_008442019.1, XP_004307117.1, XP_017247780.1, KHN09761.1, XP_017608885.1, XP_008348277.1, JAU58667.1, JAU85394.1, JAU07244.1, XP_019259655.1, ONI20796.1, XP_018446864.1, XP_019058746.1, XP_015891101.1, KOM32641.1, and others.

It is implied from the study that nMAT1, nMAT2, nMAT3, and nMAT4 of *A. thaliana* had a high number of identity matches with that of other plants (Table 1). As the identity observed with plant proteome is significantly high, the corresponding proteins which are currently described as predicted proteins, hypothetical proteins, or uncharacterized proteins may now forth be characterized for their function in splicing mitochondrial pre-mRNA. This significance was calculated using homology modeling where the protein structures were compared against each other. Table S1 enlists the present annotations of these putative “nMAT-like proteins.”

Analysis of dendrogram and the bootstrap scores showed that the proteins predicted for different nMATs are highly potent candidates to be annotated as nMAT1-, nMAT2-, nMAT3-, or nMAT4-like proteins due to consistent high homology with previously annotated *A. thaliana* sequences. The homology between selected candidates is so high that the candidates having high homology with nMAT1 are more closely related to nMAT1 of different genera than to nMAT2 proteins of *A. thaliana*. Moreover, it was also found from the dendrograms that some of the protein sequences showed high-identity match, but on bootstrapping it was found that they had low evolutionary relations.

Multiple Sequence Alignments showed high degree of conservation of matraverse domains among nMATs and their
respective high-identity score protein sequences. NP_196025.5 shows nonalignment with the X-domain but aligns with En and RT domains (domain information from Mohr and Lambowitz). It might have retained reverse transcriptase role but not that of maturase, which needs to be confirmed experimentally to further understand the role of different maturase domains in splicing.

Conclusions
Plants have been found to have mitochondrial and plastidial group II introns which require proteins called maturases that are important for splicing. The maturases though considered to be evolutionarily linked to other microorganisms, this study was confined to the maturase sequences of Arabidopsis with the available sequences of plant genomes at NCBI. Several plant loci have been detected which are capable of producing a protein with maturase function or maturase-like function. As nMATs are present in all plant proteomes, it possibly suggests the essential nature of these proteins in the regulation of oxidative phosphorylation in mitochondria by splicing of several vital genes.

In this study, the 4 nMAT proteins, namely, nMAT1, nMAT2, nMAT3 and nMAT4, on conducting BLASTP represented values for various plants, and it was found that they had expect values very near to zero, representing that the proteins had high-identity match among sequences of Arabidopsis and other plants. When phylogenetic inspection with bootstrapping was done in MEGA, it showed that the nMAT sequences had evolutionary relationship with Arabidopsis nMAT sequences. Some had very high evolutionary relation with Arabidopsis, whereas some showed low phylogenetic relationship. Thus, the phylogenetic relationship between Arabidopsis and other plants has clearly been established. nMATs 1 to 4 are essential for floral development and seed set. This role of nMATs in floral development may be employed in Jatropha seeds. Jatropha seeds are used to generate biodiesel, and unstable and poor flowering cause low seed yield in Jatropha. It has been established during this research that the sequences of J curcas plant showed high-identity match as well as the phylogenetic relations of maturase proteins; further research focus should be made into enhancement of seed production by overexpression of nMATs for increased production of biodiesel from Jatropha seeds and seeds of other potential biofuel-producing plants. Similarly, this approach may be employed to enhance oil yield of edible oil–yielding plants such as Brassica.

Acknowledgements
The authors acknowledge the support provided by Amity Institute of Biotechnology, Amity University, Haryana. They also thank Dr Suneyna Bansal and Dr Ravi Dutta who provided insight and expertise that greatly assisted in research analysis.
Author Contributions
SM, KCU, and SMPK conceived and designed the experiments. SM, RD, and KCU analyzed the data. SM wrote the first draft of the manuscript. SM, SB, and KCU contributed to the writing of the manuscript. SM, KCU, SMPK, and RD agree with manuscript results and conclusions. SM and KCU jointly developed the structure and arguments for the paper. SM and KCU made critical revisions and approved final version. All authors reviewed and approved of the final manuscript.

Disclosures and Ethics
As a requirement of publication, author(s) have provided to the publisher signed confirmation of compliance with legal and ethical obligations including, but not limited to, the following: authorship and contributorship, conflicts of interest, privacy and confidentiality, and (where applicable) protection of human and animal research subjects. The authors have read and confirmed their agreement with the ICMJE authorship and conflict of interest criteria. The authors have also confirmed that this article is unique and not under consideration or published in any other publication, and that they have permission from rights holders to reproduce any copyrighted material. Any disclosure made in this section. The external blind peer reviewers report no conflicts of interest.

REFERENCES
1. Krebs H, Johnson WA. The role of citric acid in intermediate metabolism in animal tissues. Enzymologia. 1937;4:148–156.
2. Mitchell P. Coupling of phosphorylation to electron and hydrogen transfer by a chemiosmotic type of mechanism. Nature. 1961;191:144–148.
3. Matsuura M, Saldanha R, Ma H, et al. A bacterial group II intron encoding reverse transcriptase, maturase, and DNA endonuclease activities: biochemical demonstration of maturase activity and insertion of new genetic information within the intron. Gene Dev. 1997;11:2910–2924.
4. Lambowitz AM, Zimmerly S. Group II introns: mobile ribozymes that invade DNA. Cold Spring Harb Perspect Biol. 2011;3:a003616.
5. Mohr G, Lambowitz A. Putative proteins related to group II intron reverse transcriptase/maturases are encoded by nuclear genes in higher plants. Nucleic Acids Res. 2003;31:647–652.
6. Nakagawa N, Sakurai N. A mutation in At-nMat1a, which encodes a nuclear gene having high similarity to group II intron maturase, causes impaired splicing of mitochondrial NAD4 transcript and altered carbon metabolism in A. thaliana. Plant Cell Physiol. 2006;47:772–783.
7. Keren I, Bazawark-Geleta A, Kolton M, et al. AtMat2, a nuclear-encoded maturase required for splicing of group-II introns in Arabidopsis mitochondria. RNA. 2009;5:2289–2311.
8. Tarres I, Gal L, Colas des Frances-Small C, et al. nMAT1, a nuclear-encoded maturase involved in the trans-splicing of nad1 intron 1, is essential for mitochondrial complex I assembly and function. Plant J. 2012;71:413–426.
9. Cohen S, Zmudjak M, Colas des Frances-Small C, et al. nMAT4, a maturase factor required for nad1 pre-mRNA processing and maturation, is essential for helicase I biogenesis in Arabidopsis mitochondria. Plant J. 2014;78:253–268.
10. Falcon de Longevialle A, Meyer EH, Andres C, et al. The pentatricopeptide repeat gene OTP43 is required for trans-splicing of the mitochondrial nad1 intron 1 in Arabidopsis thaliana. Plant Cell. 2007;19:3256–3265.
11. Zmudjak M, Colas des Frances-Small C, Keren I, et al. nCSF1, a nucleus-encoded CRM protein required for the processing of many mitochondrial introns, is involved in the biogenesis of respiratory complexes I and IV in Arabidopsis. New Phytol. 2013;199:379–394.
12. Köhler D, Schmidt-Gattung S, Binder S. The DEAD-box protein PMH2 is required for efficient group II intron splicing in mitochondria of Arabidopsis thaliana. Plant Mol Biol. 2010;72:459–467.
13. Brown G, des Frances-Small CC, Östersetzer-Biran O. Group II intron splicing factors in plant mitochondria. Front Plant Sci. 2014;5:1–13.
14. Dlakic M, Mouhebian A. Pip1, the pirvotal protein of the spliceosomal catalytic center, evolved from a retroelement-encoded reverse transcriptase. RNA. 2011;17:799–808.
15. Autschal S, Warren G, Miller W, Myers E, Lipman D. Basic local alignment search tool. J Mol Biol. 1990;215:403–410.
16. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013;30:2725–2729.
17. Sievers F, Wilm A, Dineen D, et al. Fast, scalable generation of high-quality protein sequence multiple sequence alignments using Clustal Omega. Mol Syt Biol. 2014;2:539–559.
18. Yang Z, Rannala B. Molecular phylogenetics: principles and practice. Nat Rev Genet. 2012;13:303–314.
19. Moradian MM, Beglayan D, Skoyslay JM, Kerkorian V. Complete mitochondrial genome sequence of three Tetrahymena species reveals mutation hot spots and accelerated nonsynonymous substitutions in Ymf genes. PLoS ONE. 2007;2:e650.
20. Carbonell-Caballeró J, Alonso R, Ibáñez V, Terol J, Talón M, Dopazo J. A phylogenetic analysis of 34 chloroplast genomes elucidates the relationships between wild and domestic species within the genus citrus. Mol Biol Evol. 2015;32:2725–2729.
21. Sievers F, Wilm A, Dineen D, et al. Fast, scalable generation of high-quality protein sequence multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2014;2:539–559.
22. Guo W, Mower JP. Evolution of plant mitochondrial intron-encoded maturases: frequent lineage-specific loss and recurrent intracellular transfer to the nucleus. J Mol Evol. 2013;77:43–54.
23. Nahar K, Oostes-Hampton M. Jatropha: An Alternative Substitute to Fossil Fuel (HS1193). Gainesville, FL: IFAS Publication, Institute of Food and Agricultural Sciences, University of Florida; 2011:12–17.
24. Ghosh A. Paclobutrazol arrests vegetative growth and unveils unexpressed yield potential of Jatropha curcas. J Plant Growth Regul. 2010;29:307–315.