Original article

Phylotranscriptomic analysis of *Dillenia indica* L. (Dilleniales, Dilleniaceae) and its systematics implication

Mohammad Ajmal Ali

Department of Botany and Microbiology, College of Science, King Saud University, Riyadh 11451, Saudi Arabia

ARTICLE INFO

ABSTRACT

The recent massive development in the next-generation sequencing platforms and bioinformatics tools including cloud based computing have proven extremely useful in understanding the deeper-level phylogenetic relationships of angiosperms. The present phylotranscriptomic analyses address the poorly known evolutionary relationships of the order Dilleniales to order of the other angiosperms using the minimum evolution method. The analyses revealed the nesting of the representative taxon of Dilleniales in the MPT but distinct from the representative of the order Santalales, Caryophyllales, Asterales, Cornales, Ericales, Lamiales, Saxifragales, Fabales, Malvales, Vitales and Berberidopsidales.

1. Introduction

The family Dilleniaceae (Parad. Lond. sub t. 73, 1806) previously placed in subclass Dilleniidae (Phytologia 74(3): 171, 1993) (Cronquist and Takhtadzhī, 1981), were unplaced in Angiosperm Phylogeny Group, APG (APG, 1998; APG II, 2003; APG III, 2009) to as an order; later APG IV (APG IV, 2016) recognized as the new monofamilial order Dilleniales (Prir. Rostlin 223, 1820) (characterized as mostly woody; lvs if veins strong, proceed to apex of teeth flw mostly K5, persisting, mostly Androceium ∞, Gynoceium mostly slightly connate seeds often with aril; fruits usually follicles) under clade core eudicots containing family Dilleniaceae only (-the family of trees, shrubs, and lianas) based on molecular phylogenetic studies (Soltis et al., 2011; Ruhfel et al., 2014). The APG system recognizes 12 genera [viz. 1. *Acrotrema* Jack (Malayan Misc. 1 (5): 36, 1820), 2. *Curatella* Loefl. (Iter Hispan. 229, 260, 1758), 3. *Davilla* Vand. (Fl. Lusit. Bras. Spec. 35, f. 14, 1788), 4. *Didesmandra* Stapf (Hooker's Icon. Pl. 27:, ad t. 2646, 1900), 5. *Dillenia* L. (Sp. Pl. 1: 535, 1753), 6. *Doliocarpus Rol.* (Kongl. Svenska Vetensk. Acad. Handl. 17: 260–261, 1756), 7. *Hibbertia* Andrews (Bot. Repos. 2:, t. 126, 1800), 8. *Neodillenia* Aymard (Harvard Pap. Bot. 10: 121, 1997), 9. *Pachynema* R. Br. ex DC. (Syst. Nat. 1: 397, 411, 1818), 10. *Pinzona* Mart. & Zucc. (Abh. Math.-Phys. Cl. Königl. Bayer. Akad. Wiss. 1: 371, 1832), 11. *Schumacheria* Vahl (Skr. Naturhist.-Selsk. 6: 122, 1810) and 12. *Tetracera* L. (Sp. Pl. 1: 533, 1753)] and c. 430 species of the family Dilleniaceae (http://www.mobot.org/MOBOT/research/APweb/welcome.html). Since, the evolutionary relationship of the order Dilleniales among the angiosperms phylogeny remains poorly known, the present study explores the relationship of the order Dilleniales in the angiosperm phylogeny based on phylotrasncriptomic analyses.

2. Materials and methods

2.1. Taxon selection

The RNA transcriptome SRA data of *Dillenia indica* L. (order Dilleniales, family Dilleniaceae) was available from the study of 'One Thousand Plant Transcriptomes Initiative' (Leebens-Mack, 2019) were retrieved, and analyzed together with representatives from the order Asterales, Berberidopsidales, Caryophyllales, Cornales, Ericales, Fabales, Gunnerales, Lamiales Malvales, Santalales, Saxifragales and Vitales. The RNA transcriptome data of *Gunnera manicata* (order Gunnerales, family Gunneraceae) was used as the outgroup in the phylotranscriptomic analysis (Table 1).

E-mail: ajmalpdrc@gmail.com, alimohammad@ksu.edu.sa
Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

Table 1
Taxon used in the phylotranscriptomic analyses to infer relationships of Dilleniales.

S. No.	Taxon	Family	Order	Clade	GenBank
	Ingroup				
1.	*Dillenia indica* L.	Dilleniaceae	Dilleniales	Eudicots, Core eudicots	ERS1829252
2.	*Phoradendron serotinum* (Raf.) M.C. Johnst.	Santalaceae	Santalales	Eudicots, Superasterids	ERS1829342
3.	*Silene latifolia* Poir.	Caryophyllaceae	Caryophyllales	Superasterids	ERS1829296
4.	*Inula helenium* L.	Asteraceae	Asterales	Asterids, Campanulids	ERS368261
5.	*Cornus florida* L.	Cornaceae	Cornales	Eudicots, Asterids	ERS1829540
6.	*Cavendishia cuatrecasasii* A.C. Sm.	Ericaceae	Ericales	Eudicots, Asterids	ERS1829554
7.	*Lavandula angustifolia* Mill.	Lamiaceae	Lamiales	Asterids, Lamiids	ERS1829627
8.	*Astilbe chinensis* (Maxim.) Franch. & Sav.	Saxifragaceae	Saxifragales	Eudicots, Core eudicots	ERS1829362
9.	*Lupinus polyphyllus* C.E. Anderson	Fabaceae	Fabales	Rosids, Fabids	ERS631113
10.	*Hibiscus cannabinus* L.	Malvaceae	Malvales	Rosids, Malvids	ERS368256
11.	*Cissus quadrangularis* L.	Vitaceae	Vitales	Superrosids	ERS1829366
12.	*Berberidopsis beckleri* (F. Muell.) Veldkamp	Berberidopsidaceae	Berberidopsidales	Eudicots, Superasterids	ERS1829251
	Outgroup				
13.	*Gunnera manicata* Linden ex André	Gunneraceae	Gunnerales	Eudicots, Core eudicots	ERS1829249

2.2. Phylotranscriptomic analyses

The retrieved aligned data were then imported to MEGA X (Kumar et al., 2018) for pairwise sequence alignment. The aligned sequence data set saved into mega format and the evolutionary analyses (outgroup at *Gunnera manicata,* order Gunnerales, family Gunneraceae) were performed using Maximum Parsimony (MP) bootstrap methods (Felsenstein, 1985) in MEGA X (Kumar et al., 2018) in order to detect proximity of *Dillenia indica* (the representative of the order Dilleniales with the representatives of the order Asterales Link, Berberidopsidales Doweld, Caryophyllales Benth. & Hook. f., Cornales Link, Ericales Bercht. & J. Presl, Fabales Bromhead, Gunnerales Takht. ex Reveal, Lamiales Bromhead, Malvales Juss., Santalales R. Br. ex Bercht. & J. Presl, Saxifragales Bercht. & J. Presl and Vitales Juss. ex Bercht. & J. Presl.

3. Results and discussion

The present phylotranscriptomic analyses of 28,176 parsimony informative sites out of a total of 1,42,796 positions in the aligned (the aligned data set contains 64,202 conserved, 71,580 variable and 41,870 singleton sites, Fig. 1) transcriptome dataset of *D. indica, P. serotinum, S. latifolia, I. helenium, C. florida, C. cuatrecasasii, L. angustifolia, A. chinensis, L. polyphyllus, H. cannabinus, C. quadrangularis, B. beckleri* and *G. manicata* inferred using the MP method recovered the MPT with length of 1,72,735 [consistency index 0.795270 (0.647002), retention index 0.253898, composite index 0.201917]. The MPT revealed the nesting of the representative taxon of Dilleniales (*D. indica*) but distinct from the representative of Santalales (*P. serotinum*), Caryophyllales (*S. latifolia*), Asterales (*I. helenium*), Cornales (*C. florida*), Ericales (*C. cuatrecasasii*), Lamiales (*L. angustifolia*), Saxifragales (*A. chinensis*), Fabales (*L. polyphyllus*), Malvales (*H. cannabinus*), Vitales (*C. quadrangularis*) and Berberidopsidales (*B. beckleri*) (Fig. 2) which is also evident from the estimates of evolutionary divergence between the sequences (Zuckerkandl and Pauling, 1965) used to infer the phylogeny using MEGA X (Kumar et al., 2018) (Table 2).

The recent massive development in the next-generation sequencing platforms (Eid et al., 2009; Rothberg et al., 2011; Shendure and Aiden, 2012; Pattnaik et al., 2014; Jain et al., 2016), and bioinformatics tools (Mavromatis et al., 2007; Knudsen et al., 2010; Hu et al., 2012; Huang et al., 2012; McElroy et al., 2012; Shendure and Aiden, 2012; Yang and Rannala, 2012; Caboche et al., 2014; Shcherbina, 2014; Kwon et al., 2015; Escalona et al., 2016; Langmead and Nellore, 2018) and Cloud computing (Langmead and Nellore, 2018) have proven extremely useful in evaluating assembly, mapping, phasing or genotyping (Nielsen et al., 2011; Angly et al., 2012; Caboche et al., 2014; Shcherbina 2014; Li et al., 2014), and have led to public archives (Kodama et al., 2012), collaborations (Collins and Varmus, 2015; Melé et al., 2015; Gaziano et al., 2016; Lek et al., 2016; Trans-Omics for Precision Medicine (TOPMed) Program, 2017), and thus have entirely changed the face of plant systematics (Stuessy, 2020), and enhanced understanding of tree of life and plant genomics (Ali et al., 2020) and the transcriptomics based deeper-level phylogenetic relationships (Cannon et al., 2015; Smith et al., 2015; Yang et al., 2018).

The present phylotranscriptomic analyses revealed the nesting of the representative taxon of Dilleniales but distinct from the representative of Santalales, Caryophyllales, Asterales, Cornales, Ericales, Lamiales, Saxifragales, Fabales, Malvales, Vitales and Berberidopsidales. The monofamilial order Dilleniales (characterized as mostly woody; lvs if veins strong, proceed to apex of teeth flw mostly K5, persisting, mostly A∞, G mostly slightly connate seeds often with aril; fr usually follicles) under clade core eudicots containing family Dilleniaceae. The diagnostic taxonomic features of monophyletic Dilleniaceae (Hoot et al., 1999; Savolainen et al., 2000a,b; Ingrouille et al., 2002; Hilu et al., 2003) are: combination of scabrous leaves and primary stems, leaf venation with more or less straight, parallel secondaries and scalariform tertiaries, petioles with a broad insertion at the stem, orangish inner or outer bark that exfoliates thin plates or strips, uniformly persistent and typically accrescent calyces, apopetalous, caduceus corollas, polymerous, marcescent, or tardily deciduous androecia, free stylodia, arillate seeds (Horn, 2009). Moreover, The majority of molecular phylogenetic studies resolve the family Dilleniaceae as sister to
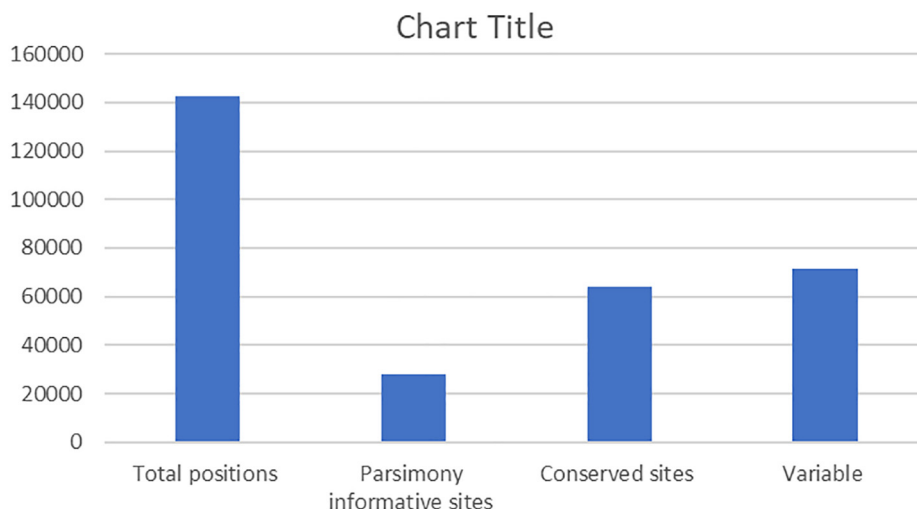
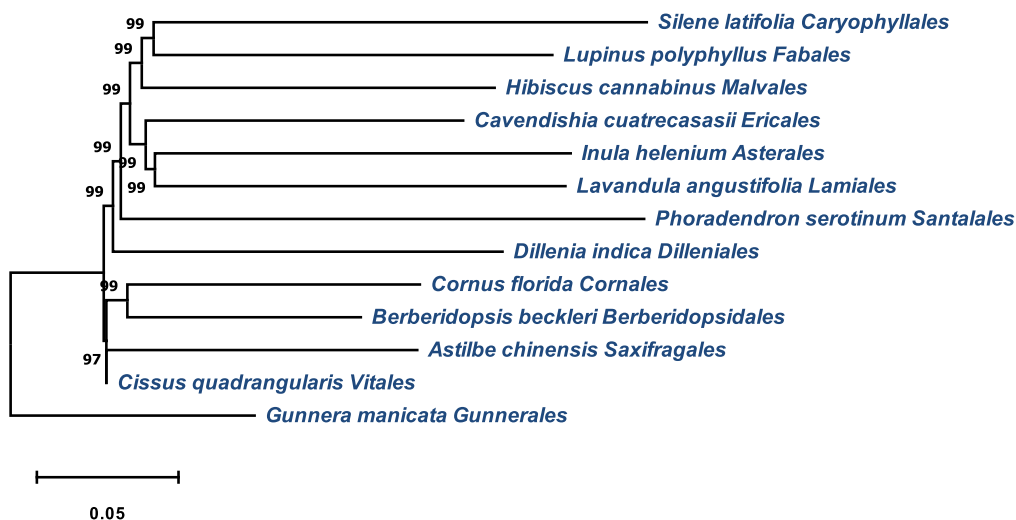**Fig. 1.** The final transcriptome data set showing number of sites.

Fig. 2. The evolutionary history of the order Dilleniales inferred from phylotranscriptomic analysis.

Table 2

The estimates of evolutionary divergence between the sequences (Zuckerkandl and Pauling, 1965) used to infer the phylogeny sung MEGA X (Kumar et al., 2018).

D. indica	0.256										
P. serotinum	0.309	0.326									
S. latifolia	0.309	0.322	0.359								
I. helenium	0.282	0.300	0.345	0.332							
C. florida	0.233	0.259	0.308	0.312	0.266						
C. cuatrecasasii	0.244	0.269	0.311	0.310	0.260	0.211					
L. angustifolia	0.283	0.292	0.344	0.327	0.292	0.270	0.262				
A. chinensis	0.232	0.252	0.302	0.302	0.277	0.226	0.235	0.281			
L. polyphyllus	0.283	0.289	0.337	0.315	0.303	0.273	0.274	0.290	0.270		
H. cannabinus	0.259	0.272	0.318	0.304	0.289	0.251	0.254	0.281	0.244	0.268	
C. quadrangularis	0.243	0.261	0.309	0.312	0.289	0.239	0.246	0.289	0.229	0.276	0.249
B. beckleri	0.202	0.228	0.278	0.294	0.264	0.186	0.208	0.269	0.192	0.255	0.227

Vitaceae (Chase et al., 1993; Savolainen et al., 2000b; Hilu et al., 2003), or Caryophyllales (Hoot et al., 1999; Soltis et al., 2000, 2003, 2007) or under within the core group of eudicots (APG, 2003, Moore et al., 2008), the present analyses supports the placement of Dilleniales as a new order in the angiosperm phylogeny.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

References

Ali, M.A., Rahman, M.O., Lee, J., Hemaid, F., Kambhar, S., Elangbam, M., Gurung, A., 2020. Dissecting molecular evolutionary relationship of Krameriaceae inferred from phylotranscriptomic analysis. Bangladesh J. Plant Tax. 27 (2), 427–433.

APG, 1998. An ordinal classification for the families of flowering plants. Ann. Mo. Bot. Gard. 85 (4), 531–553.

APG II, 2003. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG II. Bot. J. Linn. Soc. 141 (4), 399–436.

APG III, 2009. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG III. Bot. J. Linn. Soc. 161 (2), 105–121.

APG IV, 2016. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG IV. Bot. J. Linn. Soc. 181 (1), 1–20.

Caboche, S., Audebert, C., Lemoine, Y., Hot, D., 2014. Comparison of mapping algorithms used in high-throughput sequencing: application to Ion Torrent data. BMC Genomics 15, 264.

Angly, F.E., Willner, D., Rohwer, F., Hugenholtz, P., Tyson, G.W., 2012. Grinder: a versatile amplicon and shotgun sequence simulator. Nucleic Acids Res. 40, e94.

Cannon, S.B., McKain, M.R., Harkess, A., Nelson, M.N., Dash, S., Deyholos, M.K., Peng, Y., Joyce, B., Stewart Jr, C.N., Rolf, M., Kutchan, T., Tan, X., Chen, C., Zhang, Y., Carpenter, E., Wong, G.K., Doyle, J.J., Leebens-Mack, J., 2015. Multiple polyploidy events in the early radiation of nodulating and nonnodulating legumes. Mol. Biol. Evol. 32, 193–210.

Chase, M.W., Douglas, E.S., Richard, G.O., David, M., Donald, H.L., Brent, D., Mishler, M.R., Duvall, R.A., Price, H.G., Hills, Y.-L.-Q., Kathleen, A.K., Jeffrey, H.R., Elena, C., Jeffrey, D.P., James, R.M., Kenneth, J.S., Helen, J.M., John Kress, W., Kenneth, G., Karol, W., Dennis, C., Mikael, H., Brandon, S.G., Robert, K.J., Ki-Joong, K., Charles, F.W., James, F.S., Glenn, R.F., Steven, H.S., Qui-Yun, X., Gregory, M.P., Pamela, S.S., Susan, M.S., Stephen, E.W., Paul, A.G., Christopher, J.Q., Luis, E.E., Edward, G., Gerald, H.L., Sean, W.G., Spencer, C.H.B., Selvadurai, D., Victor, A.A., 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene rbcL. Ann. Missouri Bot. Gard., 528–580.

Collins, F.S., Varmus, H., 2015. A new initiative on precision medicine. N. Engl. J. Med. 372, 793–795.

Cronquist, A., Takhtadzhian, A.L., 1981. An integrated system of classification of flowering plants. Columbia University Press.

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., et al., 2009. Real-time DNA sequencing from single polymerase molecules. Science 323, 133–138.

Escalona, M., Rocha, S., Posada, D., 2016. A comparison of tools for the simulation of genomic next-generation sequencing data. Nat. Rev. Genet. 17, 459.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Gaziano, J.M., Concato, J., Brophy, M., Fiore, L., Pyarajan, S., Breeling, J., Whitbourne, S., Deen, J., Shannon, C., Humphries, D., Guarino, P., Aslan, M., Anderson, D., LaFleur, R., Hammond, T., Schaa, K., Moser, J., Huang, G., Muralidhar, S., Przygodzki, R., O'Leary, T.J., 2016. Million Veteran Program: A mega-biobank to study genetic influences on health and disease. J. Clin. Epidemiol. 70, 214–223.

Hilu, K.W., Borsch, T., Müller, K., Soltis, D.E., Soltis, P.S., Savolainen, V., Chase, M.W., Powell, M.P., Alice, L.A., Evans, R., Sauquet, H., Neinhuis, C., Slotta, T.A., Rohwer, J.G., Campbell, C.S., Chatrou, L.W., 2003. Angiosperm phylogeny based on matK sequence information. Am. J. Bot. 90, 1758–1776.

Hoot, S.B., Magallon, S., Crane, P.R., 1999. Phylogeny of basal eudicots based on three molecular data sets: atpB, rbcL, and 18S nuclear ribosomal DNA sequences. Ann. Missouri Bot. Gard., 1–32

Horn, J.W., 2009. Phylogenetics of Dilleniaceae using sequence data from four plastid loci (rbcL, infA, rps4, rpl16 intron). Int. J. Plant Sci. 170, 794–813.

Hu, X., Yuan, J., Shi, Y., Lu, J., Liu, B., Li, Z., Chen, Y., Mu, D., Zhang, H., Li, N., Yue, Z., Bai, F., Li, H., Fan, W., 2012. pIRS: Profile-based Illumina pair-end reads simulator. Bioinformatics 28, 1533–1535.

Huang, W., Li, L., Myers, J.R., Marth, G.T., 2012. ART: a next-generation sequencing read simulator. Bioinformatics 28, 593–594.

Ingrouille, M.J., Chase, M.W., Fay, M.F., Bowman, D., Van Der Bank, M., Bruijn, A.D.E., 2002. Systematics of Vitaceae from the viewpoint of plastid rbcL DNA sequence data. Bot. J. Linn. Soc. 138, 421–432.

Jain, M., Olsen, H.E., Paten, B., Akeson, M., 2016. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biol. 17, 239.

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35 (6), 1547–1549.

Knudsen, B., Forsberg, R., Miyamoto, M.M., 2010. A computer simulator for assessing different challenges and strategies of de novo sequence assembly. Genes (Basel). 1, 263–282.

Kodama, Y., Shumway, M., Leinonen, R., 2012. The sequence read archive: explosive growth of sequencing data. Nucleic Acids Res. 40, D54–D56.

Kwon, T., Yoo, W.G., Lee, W.J., Kim, D.W., 2015. Next-generation sequencing data analysis on cloud computing. Genes Genom 37, 489–501.

Langmead, B., Nellore, A., 2018. Cloud computing for genomic data analysis and collaboration. Nat. Rev. Genet. 19, 208.

Lek, M., Karczewski, K., Minikel, E., et al., 2016. Analysis of protein-coding genetic variation in 60,706 humans. Nature 536, 285–291.

Li, B., Fillmore, N., Bai, Y., Collins, M., Thomson, J.A., Stewart, R., Dewey, C.N., 2014. Evaluation of de novo transcriptome assemblies from RNA-Seq data. Genome Biol. 15, 553.

Leebens-Mack, J.H., Barker, M.S., et al., 2019. One thousand plant transcriptomes and the phylogenomics of green plants. Nature 574, 679–685.

Mavromatis, K., Ivanova, N., Barry, K., et al., 2007. Use of simulated data sets to evaluate the fidelity of metagenomic processing methods. Nat Methods 4, 495–500.

McElroy, K.E., Luciani, F., Thomas, T., 2012. GemSIM: general, error-model based simulator of next-generation sequencing data. BMC Genomics 13, 74.

Melé, M., Ferreira, P.G., Reverter, F., DeLuca, D.S., Monlong, J., Sammeth, M., Young, T.R., Goldmann, J.M., Pervouchine, D.D., Sullivan, T.J., Johnson, R., Segrè, A.V., Djebali, S., Niarchou, A.; GTEx Consortium, Wright, F.A., Lappalainen, T., Calvo, M., Getz, G., Dermitzakis, E.T., Ardlie, K.G., Guigó, R., 2015. Human genomics. The human transcriptome across tissues and individuals. Science. 2015 348 (6235), 660-665.

Moore, M., Bell, C., Soltis, P.S., Soltis, D.E., 2008. Analysis of an 83-gene, 86-taxon plastid data set resolves relationships among several recalcitrant deep-level eudicot lineages. Botany 2008 Conference Abstracts.

Nielsen, R., Paul, J.S., Albrechtsen, A., Song, Y.S., 2011. Genotype and SNP calling from next-generation sequencing data. Nat. Rev. Genet. 12, 443–451.

Pattnaik, S., Gupta, S., Rao, A.A., Panda, B., 2014. SInC: an accurate and fast error-model based simulator for SNPs, Indels and CNVs coupled with a read generator for short-read sequence data. BMC Bioinf. 15, 1–9.

Rothberg, J., Hinz, W., Rearick, T., et al., 2011. An integrated semiconductor device enabling non-optical genome sequencing. Nature 475, 348–352.

Ruhfel, B.R., Gitzendanner, M.A., Soltis, P.S., Soltis, D.E., Burleigh, J.G., 2014. From algae to angiosperms-inferring the phylogeny of green plants (Viridiplantae) from 360 plastid genomes. BMC Evol. Biol. 14, 23.

Savolainen, V., Chase, M.W., Hoot, S.B., Morton, C.M., Soltis, D.E., Bayer, C., Fay, M.F., De Bruijn, A.Y., Sullivan, S., Qiu, Y.-L., 2000a. Phylogenetics of flowering plants based on combined analysis of plastid atpB and rbcL gene sequences. Syst. Biol. 49, 306–362.

Savolainen, V., Fay, M.F., Albach, D.C., Backlund, A., van der Bank, M., Cameron, K.M., Johnson, S.A., Lledó, M.D., Pintaud, J.C., Powell, M., Sheahan, M.C., Soltis, D.E., Soltis, P.S., Weston, P., Whitten, W.M., Wurdack, K.J., Chase, M.W., 2000b. Phylogeny of the eudicots: a nearly complete familial analysis based on rbcL gene sequences. Kew Bull., 257–309

Shcherbina, A., 2014. FASTQSim: platform-independent data characterization and in silico read generation for NGS datasets. BMC Res. Notes 7, 533.

Shendure, J., Aiden, E.L., 2012. The expanding scope of DNA sequencing. Nat. Biotechnol. 30, 1084.

Smith, S.A., Moore, M.J., Brown, J.W., Yang, Y., 2015. Analysis of phylogenomic datasets reveals conflict, concordance, and gene duplications with examples from animals and plants. BMC Evol. Biol. 15, 150.

Soltis, D.E., Gitzendanner, M.A., Soltis, P.S., 2007. A 567-taxon data set for angiosperms: the challenges posed by Bayesian analyses of large data sets. Int. J. Plant Sci. 168, 137–157.

Soltis, D.E., Senters, A.E., Zanis, M.J., Kim, S., Thompson, J.D., Soltis, P.S., Ronse De Craene, L.P., Endress, P.K., Farris, J.S., 2003. Gunnerales are sister to other core eudicots: implications for the evolution of pentamery. Am. J. Bot. 90, 461–470.

Soltis, D.E., Smith, S.A., Cellinese, N., Wurdack, K.J., Tank, D.C., Brockington, S.F., Refulio-Rodriguez, N.F., Walker, J.B., Moore, M.J., Carlsward, B.S., Bell, C.D., Latvis, M., Crawley, S., Black, C., Diouf, D., Xi, Z., Rushworth, C.A., Gitzendanner, M.A., Sytsma, K.J., Qiu, Y.L., Hilu, K.W., Davis, C.C., Sanderson, M.J., Beaman, R.S., Olmstead, R.G., Judd, W.S., Donoghue, M.J., Soltis, P.S., 2011. Angiosperm phylogeny: 17 genes, 640 taxa. Am. J. Bot. 98, 704–730.

Soltis, D.E., Soltis, P.S., Chase, M.W., Mort, M.E., Albach, D.C., Zanis, M., Savolainen, V., Hahn, W.H., Hoot, S.B., Fay, M.F., Axtell, M., Swensen, S.M., Prince, L.M., John Kress, W., Nixon, K.C., Farris, J.S., 2000. Angiosperm phylogeny inferred from 18S rDNA, rbcL, and atpB sequences. Bot. J. Linn. Soc. 133, 381–461.

Stuessy, T.F., 2020. Challenges facing systematic biology. Taxon 69 (4), 655–667.

Trans-Omics for Precision Medicine (TOPMed) Program 2017. National Heart, Lung, and Blood Institute https://www.nhlbi.nih.gov/science/trans-omicsprecision-medicine-topmed-program (2017).

Yang, Z., Rannala, B., 2012. Molecular phylogenetics: principles and practice. Nat. Rev. Genet. 13, 303–314.

Yang, Y., Moore, M.J., Brockington, S.F., Mikenas, J., Olivieri, J., Walker, J.F., Smith, S. A., 2018. Improved transcriptome sampling pinpoints 26 ancient and more recent polyploidy events in Caryophyllales, including two allopolyploidy events. New Phytol. 217, 855–870.

Zuckerkandl, E., Pauling, L, 1965. Evolutionary divergence and convergence in proteins. Edited in Evolving Genes and Proteins by V. Bryson and H.J. Vogel. 97–166. Academic Press, New York.