Name	Counterpart	Title
Ulfah J Siregar	**B14**	Increasing sequencing coverage in Petai and Sungkai using ONT technology

Research summary

Chloroplasts are photosynthetic organelles that have small, conservative genomes and semi-independent behavior toward the nuclear genome (Kim *et al.* 2005). Because of its size and conservatism, the chloroplast genome has been frequently utilized to study plant origin and evolution. Five tree species, namely Petai (*Parkia scpeciosa*) and Jengkol (*Archidendron pauciflorum*) which belong to Fabaceae, Sungkai (*Peronema canescens*) of Lamiaceae, Jelutung rawa (*Dyera polyphylla*) and Jelutung darat (*Dyera costulata*) of Apocynaceae were planted to enrich and improve the biodiversity in the oil palm plantation in Jambi Province. Although all those species are well known as favorite community trees and seemed able to grow well anywhere in Indonesia, recent observations in the field showed that not all species could adapt and grow well in the experiment. The objective of this research is to increase sequencing coverage in Petai (*Parkia speciosa*) and Sungkai (*Peronema canescens*) as the main species and three other species using Oxford Nanopore Technology (ONT).

Materials were leaves collected from the same trees used for short-read and long-read sequencing in the previous study. DNA samples were extracted using CTAB method (Cetyl Trimethyl Ammonium Bromide) with modification. The quantity of extracted DNA are measured using Qubit 1.0 Fluorometer Invitrogen (Qubit dsDNA BR) and quality was assessed using NanoPhotometer NP80 Implen. The usable DNA sample must have A260/A280 quality in the range of 1,7 – 2,0. Long-reads sequencing was done using MinION device connected to the computer of Intel Core i7,16 GB RAM, 1 TB hard drive, and operating system using Linux Mint 19.2, according to manufacture protocols SQK-NBD112.24 from Nanopore with the used flowcell type R.10.4.1. Sequencing result with Fast5 format then stored dan converted to Fastq format for further bioinformatics analysis, such as base calling.
Raw sequences data from a previous study (short-reads and long-reads) and new raw data (long-reads) with fastq format were assembled through the following process based on the pipeline from *https://github.com/asdcid/Chloroplast-genome-assembly*. For gene annotation, we used GeSeq (Tillich *et al.* 2017) under MPI-MP Chlorobox to annotate the assembly result. For the annotation reference, we use all chloroplast genome references of each family and visualized them using OGDraw. Phylogenetic analysis was constructed using MEGAX using a maximum likelihood algorithm with 1000 bootstrap values.

All the data from both previous and current experiments was used to assembly the chloroplast genome of Jengkol (*A. pauciflorum*) and Petai (*P. speciosa*) of (Fabaceae), Jelutung rawa (*D. polyphylla*) and jelutung darat (*D. costulata*) of (Apocynaceae), also Sungkai (*P. canescens*) of Lamiaceae as presented in Figure 1.

References

Kim KJ, Lee HL (2005) Widespread occurrence of small inversions in the chloroplast genomes of land plants. Mol Cells 9:104-113

Tillich M, Lehwark P, Pellizzer T, *et al.* (2017) GeSeq – versatile and accurate annotation of organelle genomes. Nucleic Acids Research 45: W6–W11

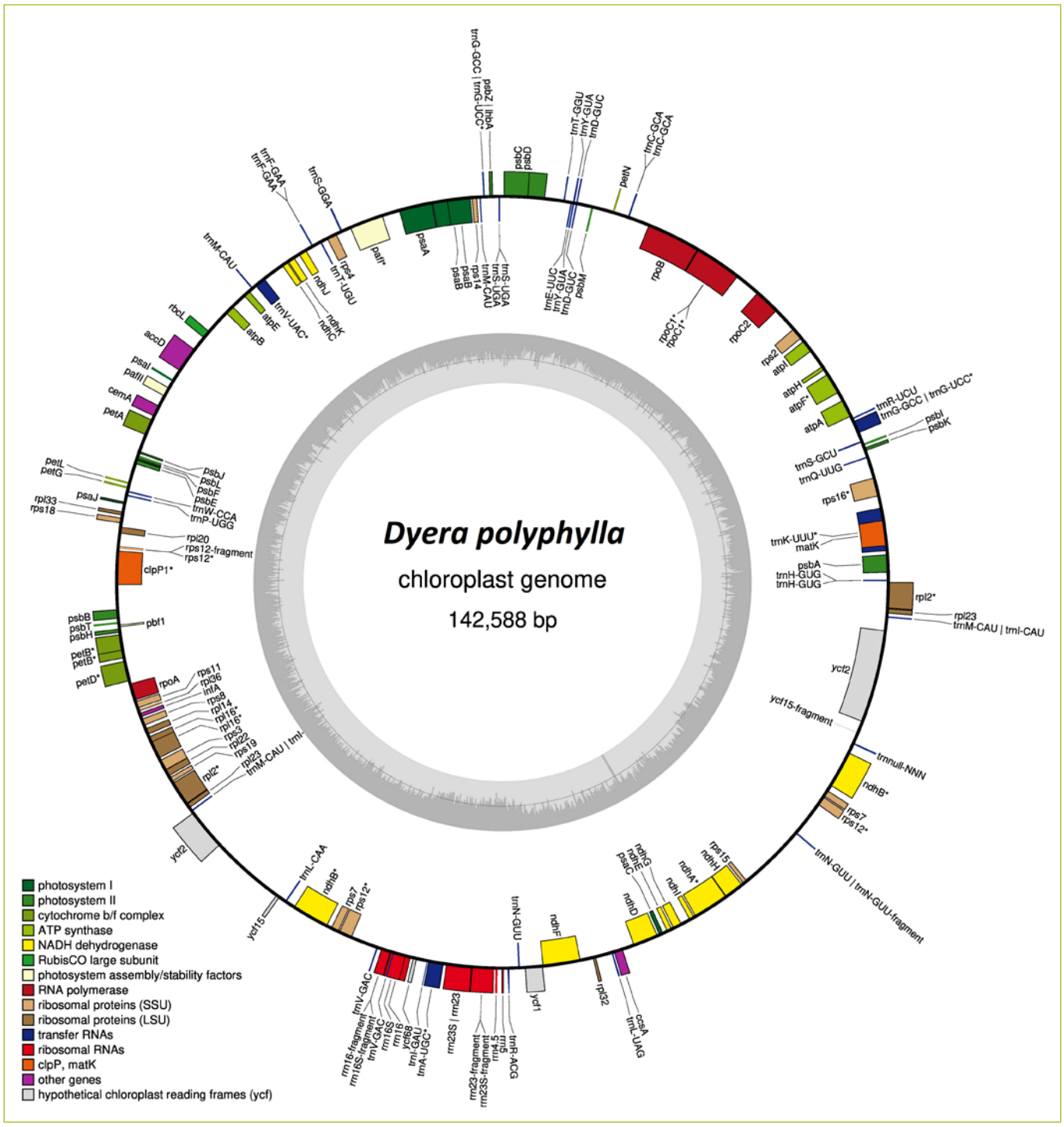
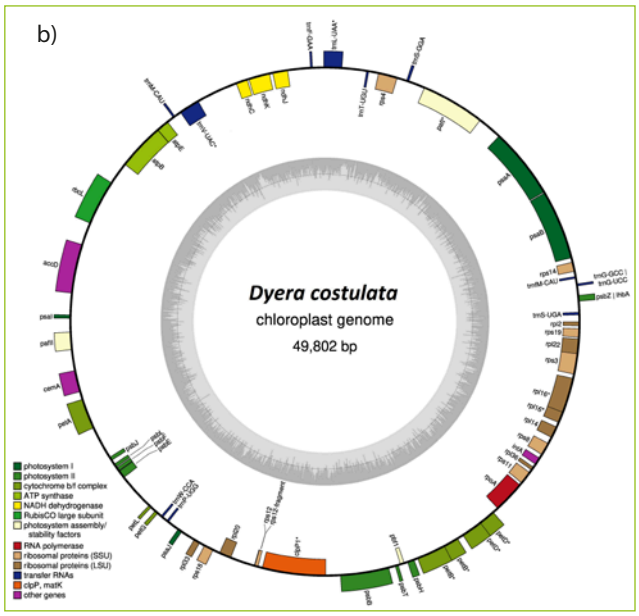
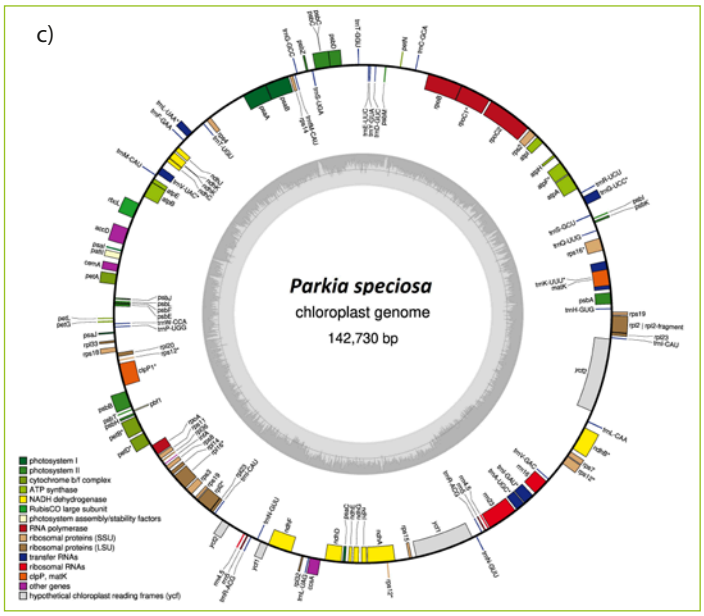


Figure 1a-e. Circular map of Chloroplast genome **(a)** *D. polyphylla*; **(b)** *D. costulata*; **(c)** *P. speciosa*; **(d)** *A. pauciflorum*; **(e)** *P. canescens*.

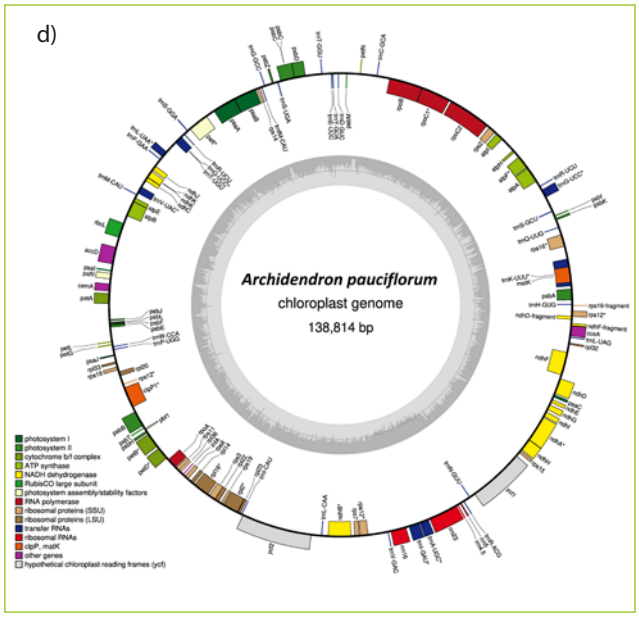
b)


Dyera costulata
chloroplast genome
49,802 bp

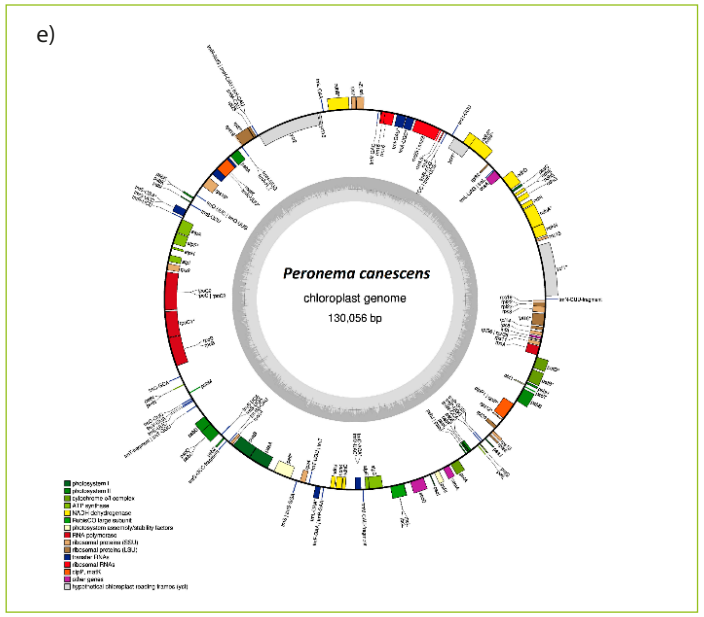
c)


Parkia speciosa
chloroplast genome
142,730 bp

d)


Archidendron pauciflorum
chloroplast genome
138,814 bp

e)


Peronema canescens
chloroplast genome
130,056 bp