

**GENETIC STRUCTURES OF NATURAL
POPULATION OF THREE NIPAH (*NYPA
FRUTICANS* (WURMB.) THUNB.) GENOTYPES
IN MALAYSIA REVEALED BY CHLOROPLAST
DNA**

ABDUL MUIN BIN MOHAMMAD

**MASTER OF SCIENCE
UNIVERSITI MALAYSIA TERENGGANU**

2014

**GENETIC STRUCTURES OF NATURAL
POPULATION OF THREE NIPAH (*NYPA
FRUTICANS* (WURMB.) THUNB.) GENOTYPES
IN MALAYSIA REVEALED BY CHLOROPLAST
DNA**

ABDUL MUIN BIN MOHAMMAD

**Thesis Submitted in Fulfilment of the Requirement
for the Degree of Master of Marine Science in the
Institute of Oceanography and Environment
Universiti Malaysia Terengganu**

February, 2014

I dedicated this thesis to my beloved mother, father and my family

&

to all my best friends and my beloved one, thank you for your understanding

and support.

Thank you

Abstract of thesis presented to the Senate of Universiti Malaysia Terengganu in fulfilment of the requirement for the degree of Master of Marine Science

**GENETIC STRUCTURES OF NATURAL POPULATION OF
THREE NIPAH (*NYPA FRUTICANS* (WURMB.) THUNB.)
GENOTYPES IN MALAYSIA REVEALED BY CHLOROPLAST
DNA**

ABDUL MUIN BIN MOHAMMAD

May, 2014

Main Supervisor : Associate Professor Sulong Ibrahim, M.Sc.

**Co-Supervisor : Associate Professor Dr. Zainudin Bachok, Ph.D
: Dr. Jasnizat Saidin, Ph.D**

Faculty : Institute Oceanography and Environment

Nypa fruticans (Wurmb. Thunb.) or known as nipah is one of the earliest and oldest angiosperms assignable to modern genus, dated back to Cretaceous period around 75 million years ago. Nipah is known to be one of the most versatile mangrove plants in term of its usage. However, its genetic diversity has been poorly documented. Hence, this study was carried out to evaluate nipah genetic diversity in Malaysia using DNA analysis via phylogenetic approach.

Ninety nipah leaflet samples were collected throughout Malaysia, where 6 samples (from 3 sampling sites) were collected in each state in Peninsular Malaysia and 12 samples (from 6 sampling sites) were collected in Sabah and Sarawak. The sampling sites were separated at least by 1 km away. There are two sampling point per sampling site separated by 3 m away to avoid sampling of similar genotype since nipah is capable to self-fertilize and propagate vegetatively. Middle parts of young nipah leaflets were

sampled. Nipah leaflet samples were preserved in CTAB and silica gel. Pre-assessment on nipah leaflet preservation showed that nipah leaflet samples preserved longer and better DNA quality obtained from CTAB preservation compared to silica gel preservation. DNA band from gel electrophoresis showed that CTAB preservation can preserve nipah leaflet up to 1 month with post-preservation, however it was recommended to isolate the DNA within 7 days, while leaflets preserved in silica gel were recommended to be extracted within 5 days.

Nipah isolated DNA was amplified using two chloroplast genetic markers; *matK* coding region and *trnL – trnF (trnLF)* non-coding region. *matK* gene exhibits unusually high nucleotide substitution, whereas *trnLF* gene is a rapidly evolved plastid, which makes both regions suitable with the need of this study. Multiple alignments pre-assessment showed that nipah genetic sequences were better aligned using MUSCLE compared to ClustalW. Best fit nucleotide substitution model analysis in Mega 5.10 showed that nipah samples were best analysed using T92 (Tamura 3-Parameter) model with uniform rates. Nipah phylogenetic trees were constructed using six different approaches; Maximum-Parsimony (MP), Minimum-Evolution (ME), Neighbour-Joining (NJ), Maximum-Likelihood (ML), Bayesian and BEAST approach. The reliability values and samples position in presented phylogenetic tree were cross-compared to develop nipah ultimate phylogenetic tree.

In MP approach, nipah samples exhibited to be parsimonious with CI = 0.995822, RI = 0.999856 and RCI = 0.995678 for *matK* amplified samples, whereas CI = 0.994819, RI = 0.999675 and RCI = 0.994495 for *trnLF* amplified samples. All twelve originally obtained phylogenetic trees indicated that nipah samples divided into three major clades; Clade A, Clade B and Clade A1. Clade A and Clade B were highly supported by bootstrap and probability value, while Clade A1 moderately supported by both values. BEAST and Bayesian outcome evaluated via TRACER, supported the analysis with ESS value >100. Nonetheless, in nipah ultimate phylogenetic tree, the obtained clades were supported by mean diversity, divergence over sequence pairs within clades, coefficient of evolution differentiation, branch length and evolution divergence pairs between clades. Clade B showed highest mean evolution diversity between clades (0.050) followed by Clade A and Clade A1.

On the other hand, consensus phylogenetic tree branch length of *matK* gene stretched longer for Clade B compared to Clade A and Clade A1. *matK* gene evolutionary divergence over sequence pairs within clades also showed similar output, where Clade B exhibited highest divergence with 0.336 followed by Clade A and Clade A1. Furthermore, nucleotide base substitution supported nipah evolution and diversity with distinctive differences between clades value. *matK* amplified samples exhibited nucleotide base substitution for Clade A ranging from 0.006655 to 0.055583, Clade B ranging from 0.380949 to 0.705165 and Clade A1 ranging from 0.002272 to 0.062037, while *trnLF* amplified samples

exhibited nucleotide base substitution for Clade A ranging from 0.285758 to 0.705140, Clade B ranging from 0.056329 to 0.152249 and Clade A1 ranging from 0.287545 to 0.706491.

According to the obtained output of branch length, nucleotide substitution, mean diversity between clades, divergence over sequence pairs within groups and evolution divergence pairs between groups, Clade B exhibited higher evolution rate, which suggest the early evolution occurred, followed by Clade A and A1. Using these information and phylogenetic outcome, phylogeographical distribution and possible morphology evolution of nipah in Malaysia was estimated. Phylogeographical distribution suggested that nipah was firstly settled in Peninsular Malaysia and Sarawak before scattered around Sabah, where Clade A1 can only be observed from Sabah samples. Divergence time estimation suggested that nipah in Malaysia were isolated and started to evolve around 30 million years ago. Tajima test statistic (D) suggested that nipah *matK* region ($D = -2.323 = 99\%$ confident level) evolution is affected by neutral mutation including bottleneck effect, while *trnLF* region ($D = -0.19838 = <90\%$ confident level) is slightly affected by neutral mutation.

Nevertheless, *matK* amplified samples suggested that nipah morphology may evolved from being big in size and length but produced less seed cluster to being small in size and length but produced more seed cluster, while *trnLF* amplified samples proposed the other way around. The outcome of this study exhibited that *Nypa fruticans* in Malaysia consist

three varieties that are genetically different. *matK* coding region was proven to exhibit unusually high rate of evolution compared to *trnLF* non-coding region, which showed more conserved than in reality. However, both regions had proven to be useful in indicating varieties in nipah samples with recommendation to use *matK* gene is more preferred. Further study should be conducted to verify nipah evolution pattern using mitochondrial and nuclear region and/or the combination of regions to conclude the overall evolution pattern in nipah.

Abstrak tesis yang dikemukakan kepada Senat Universiti Malaysia Terengganu sebagai memenuhi keperluan untuk Ijazah Master Sains Marin

**KAJIAN STRUKTUR GENETIK TIGA GENOTIP NIPAH (*NYP*
FRUTICANS (WURMB.) THUNB.) DALAM POPULASI
SEMULAJADI DI MALAYSIA DIDEHAHKAN OLEH DNA
KLOOROPLAS**

ABDUL MUIN BIN MOHAMMAD

Februari, 2014

Penyelia Utama : Profesor Madya Sulong Ibrahim, M.Sc.

**Penyelia Bersama : Profesor Madya Dr. Zainudin Bachok, Ph.D
: Dr. Jasnizat Saidin, Ph.D**

Fakulti : Institut Oseanografi dan Sekitaran

Nypa fruticans atau lebih dikenali sebagai nipah adalah salah satu angiosperma yang terawal dan tertua dihubungkaitkan dengan genus yang hidup pada hari ini. Penemuan fosil nipah menunjukkan bahawa nipah telah wujud sejak zaman Cretaceous, yang berusia lebih kurang 75 juta tahun dahulu. Nipah dikenali sebagai salah satu tumbuhan bakau yang banyak kegunaannya. Namun begitu, kepelbagaian genetik dalam nipah masih lagi kurang didokumentasikan. Oleh itu, kajian ini dijalankan bagi menilai kepelbagaian genetik nipah di Malaysia menggunakan analisis DNA dengan pendekatan filogenetik.

Sembilan puluh sampel daun nipah telah dikumpulkan dari seluruh Malaysia, di mana 6 sampel (dari 3 tempat penyampelan) diperolehi dari setiap negeri di Semenanjung Malaysia dan 12 sampel (dari 6 tempat penyampelan) dari Sabah dan Sarawak. Jarak pemisahan di antara dua tempat penyampelan adalah sekurang-kurangnya 1 km. Terdapat dua lokasi

penyampelan per tempat penyampelan dengan jarak pemisahan sekurang-kurangnya 3 m bagi mengelakkan penyampelan dari genotip yang sama, memandangkan nipah berupaya menjalankan persenyawaan sendiri dan boleh juga disebarkan secara vegetatif. Penyampelan dibuat dengan mengambil bahagian tengah dari daun nipah yang muda. Sampel daun nipah disimpan dalam CTAB dan gel silika. Prapenilaian kebolehsimpanan daun nipah menunjukkan bahawa daun nipah yang disimpan dalam CTAB tahan lebih lama dan kualiti DNA yang diekstrak adalah lebih baik berbanding menggunakan gel silika. Jalur DNA dari gel elektroforesis menunjukkan bahawa daun nipah yang diawet di dalam pengawet CTAB tahan sehingga sebulan, namun begitu adalah disarankan agar pemencilan DNA dilakukan dalam tempoh 7 hari, manakala saranan jangka waktu pemencilan DNA bagi daun yang disimpan di dalam gel silika adalah dalam tempoh 5 hari.

DNA nipah yang dipencilkan digandakan menggunakan dua penanda genetik; gen pengkodan *matK* dan gen bukan pengkodan *trnL – trnF* (*trnLF*). Gen *matK* mempamerkan penggantian nukleotida yang tinggi, manakala gen *trnLF* merupakan plastid yang cepat berevolusi, menjadikan kedua-dua gen ini sesuai dengan keperluan kajian ini. Prapenilaian penjajaran berganda menunjukkan bahawa jujukan genetik nipah dijajarkan dengan lebih baik menggunakan MUSCLE berbanding ClustalW. Analisis penggantian nukleotida dalam Mega 5.10 menunjukkan bahawa model T92 (Tamura 3-Parameter) adalah model yang terbaik untuk menganalisis sampel nipah pada kadar yang sekata. Pokok filogenetik nipah dibina menggunakan enam kaedah, iaitu kaedah ‘Maximum-Parsimony’ (MP),

‘Minimum-Evolution’ (ME), ‘Neighbour-Joining’ (NJ), ‘Maximum-Likelihood’ (ML), ‘Bayesian’ dan ‘BEAST’. Nilai kebolehpercayaan dan kedudukan sampel dalam pokok filogenetik dibandingkan di antara satu sama lain bagi menghasilkan pokok filogenetik nipah yang utama.

Dalam kaedah MP, sampel nipah telah menunjukkan ciri-ciri kesamaan dengan $CI = 0.995822$, $RI = 0.999856$ dan $RCI = 0.995678$ diperolehi daripada sampel yang digandakan menggunakan *matK*, manakala $CI = 0.994819$, $RI = 0.999675$ dan $RCI = 0.994495$ diperolehi dari sampel yang digandakan menggunakan *tmLF*. Kesemua dua belas pokok filogenetik asal menunjukkan bahawa sampel nipah boleh dibahagikan kepada tiga klad utama, iaitu Klad A, B dan A1. Klad A dan B sangat disokong oleh nilai kebarangkalian, manakala Klad A1 disokong secara sederhana oleh nilai kebarangkalian. Melalui TRACER, nilai ESS >100 menyokong hasil daripada analisis BEAST dan Bayesian. Hasil dari pembahagian klad yang diperolehi menerusi pokok filogenetik nipah yang utama, didapati bahawa klad disokong oleh analisis min kepelbagaian, pencapahan pasangan jujukan dalam klad, pekali pembeza evolusi, panjang ranting dalam pokok filogenetik dan pencapahan evolusi di antara pasangan klad. Klad B menunjukkan nilai min kepelbagaian evolusi di antara klad yang tertinggi iaitu pada 0.050, diikuti dengan Klad A dan Klad A1.

Dari sudut yang lain, panjang ranting pada pokok filogenetik konsensus gen *matK* adalah lebih panjang pada Klad B berbanding Klad A dan A1. Pencapahan pasangan jujukan dalam klad daripada gen *matK* juga

menunjukkan hasil yang serupa, iaitu Klad B menunjukkan pencapaian tertinggi dengan nilai sebanyak 0.336 diikuti dengan Klad A dan A1. Analisa penggantian tapak nukleotida menyokong evolusi dan kepelbagaian dalam nipah dengan nilai yang berbeza di antara Klad A, B dan A1. Sampel yang digandakan menggunakan *matK* menunjukkan penggantian tapak nukleotida bagi Klad A adalah di antara 0.006655 hingga 0.055583, Klad B pula di antara 0.380949 hingga 0.705165 dan Klad A1 adalah di antara 0.002272 hingga 0.062037, manakala sampel yang digandakan menggunakan *trnLF* menunjukkan penggantian tapak nukleotida bagi Klad A adalah di antara 0.285758 hingga 0.705140, Klad B di antara 0.056329 hingga 0.152249 dan Klad A1 di antara 0.287545 hingga 0.706491.

Dengan menilai hasil yang diperolehi dari panjang ranting pokok filogenetik, penggantian nukleotida, min kepelbagaian di antara klad, pencapaian pasangan jujukan dalam klad dan pencapaian evolusi di antara pasangan klad, Klad B mempamerkan kadar evolusi yang lebih tinggi berbanding Klad A dan A1, di mana fenomena ini mencadangkan bahawa Klad B berevolusi lebih awal berbanding klad-klad yang lain. Dengan menggunakan informasi ini dan hasil daripada filogenetik, taburan filogeografi dan evolusi morfologi yang mungkin bagi nipah di Malaysia adalah dianggarkan. Taburan filogeografi mencadangkan bahawa nipah tersebar di Semenanjung Malaysia dan Sarawak terlebih dahulu sebelum meluaskan taburannya ke Sabah, di mana hanya sampel yang diperolehi dari Sabah membentuk Klad A1. Anggaran masa pencapaian mencadangkan bahawa nipah di Malaysia telah terencil dan mula berevolusi sejak lebih

kurang 30 juta tahun dahulu. Statistik ujian Tajima (D) mencadangkan bahawa evolusi dalam gen *matK* nipah ($D = -2.323 = 99\%$ tahap kebolehpercayaan) dipengaruhi oleh mutasi neutral termasuklah kesan 'bottleneck', sementara evolusi dalam gen *trnLF* ($D = -0.19838 = <90\%$ tahap kebolehpercayaan) dipengaruhi oleh mutasi neutral pada kadar yang rendah.

Sampel yang digandakan dengan *matK* menunjukkan bahawa morfologi nipah mungkin berevolusi dari menjadi besar dan panjang tetapi menghasilkan gugusan buah yang sedikit kepada menjadi kecil dan rendah tetapi menghasilkan gugusan buah yang banyak, manakala sampel yang digandakan dengan *trnLF* menunjukkan sebaliknya. Hasil dari kajian ini menunjukkan bahawa *Nypa fruticans* di Malaysia terbahagi kepada tiga jenis yang berbeza dari segi genetik. Gen pengkodan *matK* dalam nipah di Malaysia telah terbukti mempamerkan kadar evolusi yang tinggi berbanding gen bukan pengkodan *trnLF*, di mana pada realitinya menunjukkan gen ini lebih terpelihara. Namun begitu, kedua-dua gen ini telah terbukti boleh digunakan bagi membuktikan kepelbagaian dalam sampel nipah, dengan penggunaan gen *matK* lebih disarankan. Kajian lanjutan perlu dijalankan bagi mengesahkan corak evolusi bagi nipah dengan menggunakan gen dari mitokondria dan nuklear dan/atau gabungan gen bagi mendapatkan gambaran menyeluruh corak evolusi dalam nipah.