

## EST-SSR Mining and Characterization from *Aquilaria malaccensis* Transcriptome Shotgun Assembly

Hasyati Shabrina<sup>1</sup>, Nurul Chaerani<sup>1\*</sup>, Dini Lestari<sup>1</sup>, Andrie Ridzki Prasetyo<sup>1</sup>, & Fauzan Fahrussiam<sup>1</sup>

<sup>1</sup>Program Studi Kehutanan, Fakultas pertanian, Universitas Mataram, Mataram, Nusa Tenggara Barat, Indonesia;

### Article History

Received : April 28<sup>th</sup>, 2024

Revised : May 10<sup>th</sup>, 2024

Accepted : June 14<sup>th</sup>, 2024

\*Corresponding Author:

**Nurul Chaerani**, Program Studi Kehutanan, Fakultas pertanian, Universitas Mataram, Mataram, Nusa Tenggara Barat, Indonesia; Email:

[nchaerani@unram.ac.id](mailto:nchaerani@unram.ac.id)

**Abstract:** *Aquilaria malaccensis* is a member of Thymelaceae family which is one of the main trees producing agarwood compounds in Indonesia and has high economic value. *A. malaccensis* has Appendix II conservation status due to high exploitation in the nature. One of the several efforts to maintain agarwood production while also maintaining agarwood population is efficient induction by manipulation of the pathway for agarwood formation and using individuals with higher production of agarwood compounds. Selection of superior individuals can be done using molecular markers that are linked to certain genes, including microsatellites or simple sequence repeats (SSR). This study aimed to identify SSRs in the agarwood transcriptome and characterized SSR sites that are linked to genes involved in the synthesis of agarwood compounds. The transcriptome data used was extracted from the NCBI database with accession number GHJS000000000.1. SSR mining was carried out using Krait v.1.4.0 software, then BLASTed with the protein database to analyze the linkage with the related genes. Perfect SSRs were found in 0.37% of the total sequences, and compound SSRs were 10.89%. The abundance of imperfect SSR (107,043) was greater than perfect SSR (31,262), indicating that SSR in the transcriptome had a low mutation rate and the SSR site was more conserved in *A. malaccensis*. The SSRs that linked to genes related to agarwood synthesis including HMGR, PAL, Mevalonate Kinase, MAPK, Delta-guaiene synthase, Farnesyl pyrophosphate synthase, and Phenylalanine ammonia-lyase. The discovery of SSRs linked to these genes can be used to create molecular markers that can be applied in agarwood producing trees selection.

**Keywords:** Agarwood synthesis related gene; *aquilaria malaccensis*; microsatellite; simple sequence repeat.

### Pendahuluan

Gaharu (*Aquilaria malaccensis*) termasuk anggota family Thymelaceae yang salah satu jenis penghasil utama gaharu di Indonesia. Pohon gaharu bernilai ekonomi tinggi karena menghasilkan senyawa gaharu, yang dalam bahasa dagang internasional disebut *oud* atau *agarwood*. Berdasarkan penelusuran (López-Sampson and Page, 2018) gaharu asal Indonesia merupakan gaharu dengan kualitas superior dan sudah terekam dalam perdagangan internasional sejak abad ke-13. Gaharu diperdagangkan dalam bentuk gubal, minyak ekstrak gubal, serpihan

atau chips, atau bubuk. Nilai jual gubal gaharu sendiri berada di kisaran Rp 500.000 per kg untuk gaharu kualitas *kamedangan* atau standar, sedangkan untuk kualitas super dapat mencapai Rp 5-10 juta per kg (Hidayat, Siburian and Indah Yuliana, 2020) Di Desa Fatmite Kabupaten Buru Selatan, gaharu berkontribusi terhadap pendapatan masyarakat pada kisaran Rp 4.1 juta per bulan (Kamaruddin, Titawael and Gawariah, 2022). Indonesia sendiri merupakan eksportir terbesar gaharu berdasarkan data tahun 1994-2015 (CITES, 2018).

*A. malaccensis* merupakan spesies dominan untuk ekspor dari Asia Tenggara

dengan konsumen utama negara-negara dari Timur Tengah (Liu *et al.*, 2017). Status gaharu, khususnya *A. malaccensis*, dalam konservasi berdasarkan CITES (*the Convention on International Trade in Endangered Species of Wild Fauna and Flora*) berada pada Appendix II yaitu terancam punah di alam. Keberadaan *A. malaccensis* di alam sendiri semakin berkurang akibat eksploitasi besar-besaran (Anwar, Wardati and Ardian, 2017). *A. malaccensis* sendiri lebih banyak dieksploitasi dibandingkan *Gyrinops verstigi* karena *A. malaccensis* menghasilkan rendemen gaharu yang lebih tinggi (Aqmarina, 2018). Belum adanya metode induksi gaharu yang efisien menyebabkan eksploitasi yang berlebihan tersebut. Salah satu upaya untuk menemukan metode induksi yang lebih efisien yaitu memetakan *pathway* proses pembentukan gaharu.

Salah satu senyawa utama yang terkandung dalam gaharu yaitu sesquiterpenoid. Senyawa tersebut dihasilkan melalui *pathway* mevalonate (Xu *et al.*, 2013; Aqmarina, 2018). Sintesis sesquiterpenoid difasilitasi beberapa enzim dan terjadi di plastid pada organ tanaman gaharu (Vranová, Coman and Gruissem, 2013). Sintesis sesquiterpenoid sendiri melibatkan beberapa gen seperti *3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGR)*, *1-deoxy-D-xylulose-5-phosphate synthase (DXS)*, *farnesyl diphosphate synthase (FPPS)* dan *sesquiterpene synthase (ASS)* (Xu *et al.*, 2023).

Penelitian sebelumnya yang dilakukan oleh Kadir *et al.* (2021) menggunakan metode RNA-Seq dari *A. malaccensis* yang diinduksi manusia maupun diinduksi secara alami menunjukkan terdapat perbedaan jumlah lokus gen unik yang terekspresi. Perbedaan jumlah transkrip yang terekspresi tersebut menunjukkan terdapat perbedaan metabolisme antara *A. malaccensis* yang diinduksi secara alami maupun induksi oleh manusia. Transkriptom yang didapatkan dari penelitian tersebut dapat digunakan untuk berbagai proses, seperti analisis *differentially expressed gene (DEG)* serta pembuatan berbagai macam penanda molekuler untuk analisis genetik.

Salah satu penanda molekuler yang umum dan mudah digunakan yaitu *simple sequence repeat (SSR)*. SSR juga dikenal dengan nama lain mikrosatelit. Penanda ini menggunakan

sekuens dengan repetisi singkat nukleotida pada asam nukleat. Penanda mikrosatelit dapat dirancang berdasarkan sekuens genomik maupun transkriptomik dari organisme. Keunggulan penanda ini yaitu bersifat kodominan, kelimpahannya di materi genetik cukup tinggi, serta *transferability* antar spesiesnya cukup baik. Selain itu, dalam analisis laboratorium, penanda ini memiliki kompleksitas kerja yang tidak terlalu rumit dan efisiensi biaya yang lebih terjangkau. Penanda mikrosatelit yang dirancang berdasarkan sekuens transkriptom disebut *Expressed Sequence Tag-SSR (EST-SSR)*. Penanda ini berada pada region yang terekspresi dan kemungkinan memiliki keterkaitan dengan ekspresi gen yang cukup erat.

Penelitian tentang gen-gen yang terkait dengan senyawa pembentuk gaharu telah dilakukan oleh (Aqmarina, 2018) menggunakan RT-PCR untuk gen-gen yang terlibat dalam *metabolic pathway* sesquiterpen. Akan tetapi, penelitian terkait dengan deteksi atau analisis gen terkait senyawa pembentuk gaharu belum pernah dilakukan menggunakan penanda mikrosatelit yang disusun berdasarkan hasil sekuensing transkriptom gaharu. Dalam pembuatan primer berbasis mikrosatelit yang digenerate berdasarkan transkriptom, tahapan yang perlu dilakukan lebih dulu yaitu melakukan identifikasi situs-situs mikrosatelit. Tahapan selanjutnya yaitu mematchingkan situs-situs tersebut dengan sekuens gen yang terkait dengan sintesis senyawa gaharu. Oleh karena itu, dalam penelitian ini, kami bertujuan untuk mencari EST-SSR dari transkriptom gaharu (*Aquilaria malaccensis*) sebagai langkah awal dalam memahami regulasi gen dan keragaman genetik yang terlibat dalam produksi senyawa gaharu.

## Bahan dan Metode

Data transkriptom yang digunakan merupakan hasil *Transcriptome shotgun assembly (TSA)* dari penelitian (Abdul Kadir, Azizan and Othman, 2021), yang dapat diakses pada GenBank NCBI dengan nomor akses [GHJS000000000.1](https://www.ncbi.nlm.nih.gov/nuccore/GHJS000000000.1) atau pada link <https://www.ncbi.nlm.nih.gov/nuccore/GHJS000000000.1>. Analisis mikrosatelit dilakukan menggunakan software Krait v.1.4.0 (Du *et al.*, 2018). Kriteria untuk pencarian situs mikrosatelit yang digunakan yaitu minimum repetisi 10 untuk

mononukleotida, 6 untuk dinukleotida, dan 5 untuk trinukleotida, tetranukleotida, pentanukleotida, dan 4 untuk heksanukleotida. Pencarian situs mikrosatelit dilakukan untuk *perfect microsatellite*, *compound microsatellite*, dan *imperfect microsatellite*.

Data TSA gaharu kemudian digunakan untuk anotasi menggunakan *Basic Local Alignment Search Tools* (BLASTs) pada *cloud computing Galaxy v.23.0.0* (<https://usegalaxy.eu/>) (The Galaxy Community 2022). BLAST dilakukan dengan database protein untuk taksonomi Viridiplantae dari Swissprot ([https://www.uniprot.org/uniprotkb?facets=mod\\_el\\_organism%3A3702&query=viridiplantae](https://www.uniprot.org/uniprotkb?facets=mod_el_organism%3A3702&query=viridiplantae)) dengan *cut-off E-value* 1E-5 (Altschul *et al.*, 1990). Data mikrosatelit kemudian digabungkan dengan data anotasi menggunakan R statistical software version 4.1.2 (R Core Team, 2021)

## Hasil dan Pembahasan

### Ringkasan sekuens TSA *A. malaccensis*

Jumlah sekuens transkriptom yang terdapat pada TSA gaharu yaitu 127,993 sekuens dan kandungan GC 41.93%. Keterangan lengkap TSA gaharu disajikan pada Tabel 1. Total panjang sekuens yang dihasilkan yaitu 123,308,111 pasang basa (bp) dan tidak terdapat basa-basa yang tidak diketahui atau ambigu pada seluruh sekuens.

**Tabel 1.** Ringkasan Transcriptome Shotgun Assembly Gaharu GHJS00000000.1

Item	Deskripsi	Nilai
Total jumlah sekuens	Jumlah	127,993
Total panjang sekuens	A+T+C+G+N (bp)	123,308,111
Total panjang sekuens valid	A+T+C+G (bp)	123,308,111
Basa yang tidak diketahui (N) pada sekuens	Bp	0
Persentasi basa yang tidak diketahui (N)	Persen (%) (G+C)/(A+T+C+G) not include Ns (%)	0
Kandungan GC	(%)	41.93

Penelitian pada *A. sinensis*, yang juga menghasilkan senyawa gaharu, menunjukkan persentase basa G dan C pada transkriptom sebesar 46.55% (Lv *et al.*, 2022). Ikatan basa G dan C lebih stabil dibandingkan dengan basa A dan T (Borisova *et al.*, 1993). Penelitian Zhang *et al.*, (2011) menyatakan bahwa persentase G dan C yang tinggi pada ekson, ekson-ekson dengan *alternative splicing* dapat mempertahankan struktur yang lebih stabil. Rata-rata kandungan GC pada transkriptom tanaman eudikot sebesar 45-50% (Kotwal *et al.*, 2016).

### Perfect SSR

*Perfect microsatellite* atau *Perfect SSR* adalah penanda SSR yang memiliki pengulangan sempurna, seperti (AG)<sub>20</sub> atau (CCT)<sub>6</sub> dan tidak terganggu oleh basa-basa lain di sepanjang sekuensnya. Dari 127,993 sekuens, terdapat 31,262 *perfect SSR* dengan rata-rata panjang 14.27 bp. Panjang tersebut lebih pendek dibandingkan dengan rata-rata pada Angiospermae lainnya dalam penelitian (Ranade *et al.*, 2014). Jumlah SSR tersebut meliputi 0.37% dari total panjang sekuens. Penelitian pada Tabel ringkasan *perfect SSR* disajikan pada Tabel 2.

**Tabel 2.** Perfect microsatellite dari TSA Gaharu GHJS00000000

Item	Deskripsi	Nilai
Total jumlah <i>perfect SSR</i>	jumlah	31,262
Total panjang <i>perfect SSR</i>	Pasang basa (bp)	446,108
Rata-rata panjang <i>perfect SSR</i>	Total panjang SSR/total jumlah SSR (bp)	14.27
Persentase SSR dalam sekuens (%)	Total panjang SSR/total panjang sekuens (%)	0.37
Jumlah relatif	Total jumlah SSR/total panjang valid (lokus/Mb)	253.53
Kerapatan relatif	Total panjang SSR/total panjang valid (bp/Mb)	3,617.83

SSR yang dihasilkan terdiri dari mono hingga heksa nukleotida. SSR yang paling banyak ditemukan yaitu mononukleotida yang meliputi 64.48% dari total *perfect* SSR yang ditemukan. Hal tersebut merupakan hal yang umum pada EST-SSR karena pada ujung mRNA terdapat rantai poliadenin (A)<sub>n</sub> yang berfungsi melindungi RNA. Tipe terbanyak kedua yaitu

dinukleotida dan trinukleotida. Tipe terpanjang yang ditemukan pada *Perfect* SSR yaitu heksanukleotida. Tipe *Perfect* SSR yang ditemukan secara lengkap disajikan pada Tabel 3. Persentase tipe SSR yang ditemukan serupa dengan penelitian lain di tanaman karet (Feng *et al.*, 2009) dan ubi jalar (Hu *et al.*, 2011).

**Tabel 3.** Motif Perfect SSR

Tipe	Jumlah	Panjang(bp)	Persen (%)	Panjang rata-rata (bp)	Jumlah relatif (loci/Mb)	Kerapatan relatif (bp/Mb)
Mono	20,157	259,268	64.48	12.86	163.47	2102.6
Di	5,868	92,954	18.77	15.84	47.59	753.84
Tri	3,147	54,243	10.07	17.24	25.52	439.9
Tetra	1,376	23,800	4.4	17.3	11.16	193.01
Penta	412	8,475	1.32	20.57	3.34	68.73
Heksa	302	7,368	0.97	24.4	2.45	59.75

Pencarian *perfect* SSR berhasil menemukan 202 motif SSR. dari 202 motif tersebut, 10 motif dengan jumlah terbanyak yaitu A, AG, AC, AT, AAG, C, AAT, ATC, AAAT, dan AGG. Motif-motif tersebut paling banyak ditemukan yaitu dinukleotida dan trinukleotida. Mononukleotida dengan motif (A)<sub>n</sub> paling

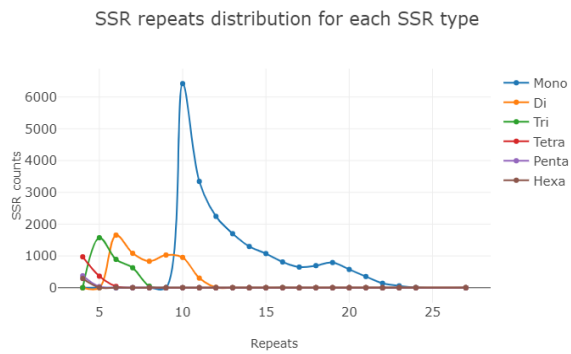
banyak ditemui karena rantai poliadenin adalah hal yang pasti ditemukan pada mRNA. Penelitian Gao *et al.*, (2022) menyatakan bahwa 91.74% dari total mononukleotida pada SSR adalah basa A/T. Rekapitulasi 10 motif dengan kelimpahan terbanyak disajikan pada Tabel 4.

**Tabel 4.** 10 Motif perfect microsatellite dengan kelimpahan terbanyak

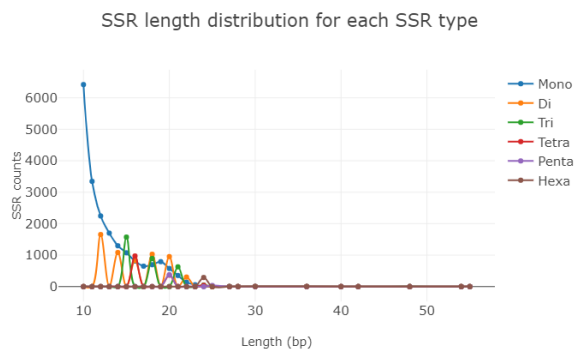
Motif	Tipe	Jumlah	Panjang (bp)	Persen (%)	Panjang Rata-Rata (bp)	Kelimpahan relatif (loci/Mb)	Kerapatan relatif (bp/Mb)
A	1	19,323	246,617	61.81	12.76	156.71	2000.01
AG	2	3,249	52,310	10.39	16.1	26.35	424.22
AC	2	1,366	20,912	4.37	15.31	11.08	169.59
AT	2	1,236	19,518	3.95	15.79	10.02	158.29
AAG	3	875	15,120	2.8	17.28	7.1	122.62
C	1	834	12,651	2.67	15.17	6.76	102.6
AAT	3	586	10,524	1.87	17.96	4.75	85.35
ATC	3	537	9,186	1.72	17.11	4.35	74.5
AAAT	4	452	7,668	1.45	16.96	3.67	62.19
AGG	3	329	5,637	1.05	17.13	2.67	45.71

Jumlah ulangan atau *repeat* untuk setiap motif ditemukan berbeda-beda yang disajikan pada Gambar 1. Jumlah *repeat* dengan kelimpahan terbanyak yaitu 10 *repeat* (6421 mononukleotida dan 953 dinukleotida), 6 *repeat* (1654 dinukleotida, 892 trinukleotida, 40 tetranukleotida, 5 heksanukleotida, dan 1 pentanukleotida), dan 5 *repeat* (1573 trinukleotida, 363 tetranukleotida, 34 pentanukleotida, dan 7 heksanukleotida). Hasil

dengan persentase *repeat* serupa ditemukan pada tanaman sainfoin (Shen *et al.*, 2019). Jumlah *repeat* akan berpengaruh terhadap panjang SSR yang dihasilkan. Tiga SSR terpanjang yang dihasilkan yaitu 55 nukleotida yang merupakan pentanukleotida dengan 11 *repeat*, 54 nukleotida dengan 27 *repeat* dinukleotida, dan 48 nukleotida dengan 16 *repeat* trinukleotida. Distribusi panjang *perfect* SSR disajikan pada Gambar 2.



**Gambar 1** Distribusi repetisi *perfect* SSR untuk setiap tipe mikrosatelit



**Gambar 2.** Distribusi panjang *perfect* SSR

### Compound SSR

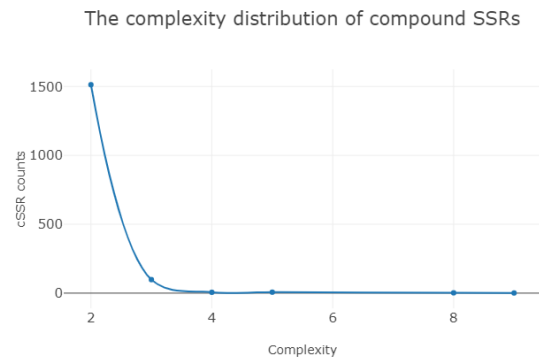
Beberapa situs *perfect* SSR dapat terletak pada transkrip yang sama pada RNA yang diselingi basa nukleotida non-SSR. *Compound* microsatellite atau *compound* SSR merupakan gabungan 2 atau lebih *perfect* SSR dengan motif yang berbeda maupun sama, dengan jarak interupsi non-repeat maksimum 100 nukleotida. Ringkasan *compound* SSR disajikan pada Tabel 5. Panjang repeat pada masing-masing SSR dalam *compound* SSR akan menentukan panjang *compound* SSR. *Compound* SSR paling kompleks terdiri dari 9 SSR yang tergabung menjadi 1 *compound*. Persentase *compound* SSR pada *A. malaccensis* dalam penelitian ini sebesar 5.2%. Penelitian pada tanaman *Psathyrostachys juncea* menunjukkan persentase *Compound* SSR sebesar 4.3% (Li *et al.*, 2022), sedangkan pada tanaman *Morinda officinalis* sebesar 5.67% (Liao *et al.*, 2019).

*Compound* SSR dapat terbentuk dari dua atau lebih *perfect* SSR. Semakin banyak *perfect* SSR yang menyusun *compound* SSR maka *compound* SSR semakin kompleks. Kompleksitas *compound* microsatellite disajikan

pada gambar 3. *Compound* SSR terpanjang yaitu sepanjang 68 nukleotida, sementara *compound* SSR terbanyak yaitu 22 nukleotida dengan kelimpahan 162 SSR. adapun jarak antara dua SSR membentuk *compound* SSR paling rendah 2 basa dan paling banyak sejauh 39 basa.

**Tabel 5.** Ringkasan *compound* SSR

Item	Deskripsi	Nilai
Jumlah total <i>compound</i> SSRs	Jumlah cSSR: Jumlah SSR tunggal penyusun	1627
Jumlah SSR penyusun <i>compound</i> SSRs	<i>Compound</i> SSRs total jumlah cSSR /total jumlah SSR	3404
Persentase <i>compound</i> SSRs	total jumlah cSSRs/total panjang valid transkriptom	10.89
Kelimpahan relatif	(loci/Mb) total panjang cSSRs/total panjang valid transkriptom	13.19
Kerapatan relatif	(bp/Mb)	382.53



**Gambar 3.** Kompleksitas *compound* SSR berdasarkan jumlah SSR penyusun

### Imperfect SSR

*Imperfect* SSR atau interrupted SSR merupakan situs mikrosatelit yang dianggap tidak sempurna karena dalam sekuensnya terdapat sisipan 1-3 nukleotida di antara repetisi yang sama. Jumlah *imperfect* SSR ditemukan

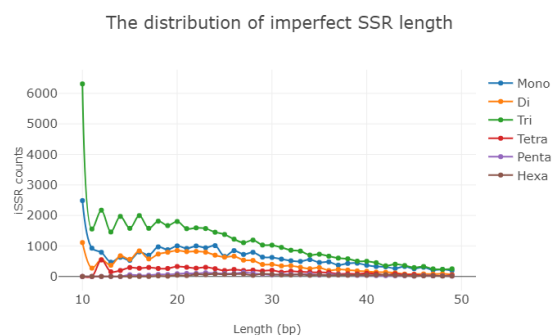


sebanyak 107,043 SSR, tiga kali lebih banyak dibandingkan *perfect* SSR. Kerapatan dan kerapatan relatif *imperfect* SSR jauh lebih tinggi dibandingkan dengan *perfect* SSR. Penelitian Ding *et al.*, (2017) menunjukkan keberadaan *imperfect* SSR membuat SSR yang dihasilkan lebih panjang dibanding *perfect* SSR pada serangga. Penelitian tersebut juga menyatakan keberadaan *imperfect* SSR menurunkan tingkat *mismatch* pada lokus-lokus gen Ringkasan *imperfect* SSR disajikan pada Tabel 6.

**Tabel 6.** Ringkasan *imperfect* SSR

Item	Deskripsi	Jumlah
Jumlah total <i>imperfect</i> microsatellite	Jumlah	107,043
Total panjang <i>imperfect</i> microsatellite	Bp	2,810,537
Kelimpahan relatif	Total jumlah iSSR /total panjang valid (loci/Mb)	868.09
Kerapatan relatif	Total panjang iSSR /total panjang valid (bp/Mb)	22792.8

*Imperfect* SSR yang ditemukan terdiri dari tipe yang sama dengan *perfect* SSR. Tipe yang paling banyak ditemukan yaitu trinukleotida (49,640 SSR), mononukleotida (28,863 SSR), dan dinukleotida (17,548 SSR). *Imperfect* SSR terpanjang yaitu 49 nukleotida sebanyak 574 SSR yang terdiri dari 250 trinukleotida, 196 mononukleotida, 59 dinukleotida, 41 tetranukleotida, 15 heksanukleotida, dan 15 pentanukleotida. Distribusi panjang *imperfect* SSR disajikan pada Gambar 4. Adapun motif yang memiliki kelimpahan terbanyak untuk *imperfect* SSR disajikan pada Tabel 7.



**Gambar 4.** Distribusi panjang *imperfect* SSR.

**Tabel 7.** 10 motif *imperfect* SSR dengan kelimpahan tertinggi

Motif	Jumlah	Panjang (bp)	Persentase (%)	Panjang rata-rata (bp)	Kelimpahan relatif (loci/Mb)	Kerapatan relatif (bp/Mb)
A	27,938	830,914	26.1	29.74	226.57	6,738.52
AAG	12,467	298,251	11.65	23.92	101.1	2,418.75
AG	10,067	238,337	9.4	23.68	81.64	1,932.86
ATC	8,586	218,211	8.02	25.41	69.63	1,769.64
AGG	6,004	149,690	5.61	24.93	48.69	1,213.95
AAC	4,622	100,418	4.32	21.73	37.48	814.37
ACG	4,447	106,856	4.15	24.03	36.06	866.58
AAT	4,224	112,828	3.95	26.71	34.26	915.01
AT	4,148	106,170	3.88	25.6	33.64	861.01
ACC	3,438	81,722	3.21	23.77	27.88	662.75

### Gen yang Berpaut dengan situs SSR

Hasil anotasi TSA dengan database dari SwissProt menggunakan BLAST menghasilkan 12,817 transkrip yang homolog dengan database. Hasil penggabungan data SSR dengan data BLAST menunjukkan terdapat 16 SSR yang berada pada situs gen yang berperan dalam sintesis gaharu. Terdapat beberapa SSR yang

mengkode satu gen yang sama yaitu *3-hydroxy-3-methylglutaryl-coenzyme A reductase 1*. Penelitian Luo *et al.*, (2021) menemukan 29 SSR yang berasosiasi dengan gen-gen yang berkaitan dengan aroma bunga pada tanaman peony. Daftar lengkap SSR yang mengkode gen terkait produksi gaharu disajikan pada Tabel 8.

**Tabel 8.** SSR pengkode gen terkait produksi gaharu

No	Sekuens	Standard Motif	Panjang	Gen
1	GHJS01037100.1	A	20	3 AQUCR Delta-guaiene synthase 3 OS=Aquilaria crassna
2	GHJS01015494.1	T	10	MENPI 1-deoxy-D-xylulose 5-phosphate reductoisomerase chloroplastic OS=Mentha piperita
3	GHJS01053266.1	T	11	ARATH 4-hydroxy-3-methylbut-2-enyl diphosphate reductase chloroplastic OS=Arabidopsis thaliana
4	GHJS01036892.1	GA	8	1 GOSHI 3-hydroxy-3-methylglutaryl-coenzyme A reductase 1 OS=Gossypium hirsutum
5	GHJS01083381.1	T	17	1 HEVBR 3-hydroxy-3-methylglutaryl-coenzyme A reductase 1 OS=Hevea brasiliensis
6	GHJS01083380.1	T	20	1 HEVBR 3-hydroxy-3-methylglutaryl-coenzyme A reductase 1 OS=Hevea brasiliensis
7	GHJS01083381.1	T	20	1 HEVBR 3-hydroxy-3-methylglutaryl-coenzyme A reductase 1 OS=Hevea brasiliensis
8	GHJS01083380.1	T	17	1 HEVBR 3-hydroxy-3-methylglutaryl-coenzyme A reductase 1 OS=Hevea brasiliensis
9	GHJS01083380.1	TTGG	4	1 HEVBR 3-hydroxy-3-methylglutaryl-coenzyme A reductase 1 OS=Hevea brasiliensis
10	GHJS01083381.1	TTGG	4	1 HEVBR 3-hydroxy-3-methylglutaryl-coenzyme A reductase 1 OS=Hevea brasiliensis
11	GHJS01015788.1	T	14	ARATH Mevalonate kinase OS=Arabidopsis thaliana
12	GHJS01034941.1	T	22	ARATH Phosphomevalonate kinase peroxisomal OS=Arabidopsis thaliana
13	GHJS01030200.1	A	21	1 LUPAL Farnesyl pyrophosphate synthase 1 OS=Lupinus albus
14	GHJS01037100.1	A	20	3 AQUCR Delta-guaiene synthase 3 OS=Aquilaria crassna
15	GHJS01015718.1	AT	10	ARATH Mitogen-activated protein kinase 3 OS=Arabidopsis thaliana
16	GHJS01084430.1	A	10	CAMSI Phenylalanine ammonia-lyase OS=Camellia sinensis

## Pembahasan

Mikrosatelit terdiri dari 3 macam, *perfect*, *imperfect*, dan *compound*. *Perfect* SSR adalah repetisi nukleotida sepanjang lebih dari 10 basa yang terulang secara sempurna, *imperfect* SSR adalah SSR yang memiliki lebih dari satu interupsi nukleotida di antara repetisinya yang bukan bagian dari repetisi, sedangkan *compound* SSR adalah dua atau lebih SSR yang berada pada satu utas nukleotida yang sama dan dapat diinterupsi oleh maksimum non repetisi dengan jumlah tertentu (Srivastava *et al.*, 2019). Pada hewan primata dan manusia, keberadaan *imperfect* dan *compound* mikrosatelit menurunkan resiko penyakit (Ananda *et al.*, 2014). Interupsi pada motif mikrosatelit merupakan produk dari terjadinya *slipage* saat mutasi dan menunjukkan bahwa region tersebut

mengalami mutasi dengan laju yang lebih lambat dibandingkan *perfect* mikrosatelit (Rossetto, 2001). Hasil analisis dalam penelitian ini menunjukkan kelimpahan dan kerapatan *imperfect* SSR lebih tinggi dibandingkan *perfect* SSR. Jumlah *imperfect* SSR (107,043) yang jauh lebih banyak dibanding *perfect* SSR (31,262) pada penelitian ini menunjukkan bahwa pada gaharu, terdapat lebih banyak region yang memiliki laju mutasi yang lebih lambat.

Studi komparasi gen yang dilakukan oleh (Lawson *et al.*, 2006) situs SSR tersebar tidak secara random pada genom tanaman arabidopsis dan padi, serta banyak terkait dengan region yang mengkode gen tertentu. Sementara itu pada tanaman jagung, SSR terdistribusi tidak secara random dan paling banyak terdapat pada *Untranslated region* (UTR) seperti promotor (Qu and Liu, 2013). Kelimpahan mikrosatelit pada

tanaman tiga kali lebih banyak pada daerah non-koding dibandingkan dengan daerah koding (Hu, Wang and Li, 2011) kecuali trinukleotida dan heksanukleotida karena berkaitan dengan kodon dan asam nukleat (Gao *et al.*, 2013). Trinukleotida yang ditemukan dari penelitian ini pada *perfect* SSR sebanyak 10.1%, sedangkan pada *imperfect* SSR sebanyak 46.37%. Trinukleotida pada SSR kemungkinan adalah kodon triplet yang mengkode basa tertentu (Singh *et al.*, 2012). Trinukleotida yang paling banyak ditemukan pada penelitian ini yaitu AAG yang mengkode asam amino Lysine baik pada *perfect* (875) maupun *imperfect* (12,467 kodon) SSR.

EST-SSR yang berada pada *coding region* dalam gen lebih terkonservasi dan tingkat mutasinya lebih rendah, tetapi jika terjadi mutasi dapat menyebabkan penambahan atau pengurangan fungsi gen. Sementara itu, EST-SSR yang berada pada 5' maupun 3' UTR jika terjadi mutasi dapat berpengaruh terhadap splicing serta proses seluler. Informasi SSR yang terkait dengan gen tertentu akan memudahkan dalam membuat penanda molekuler untuk analisis ekspresi gen maupun identifikasi gen. EST-SSR terletak di dalam gen dan lebih terkonservasi dapat menyebabkan marka ini akan menjadi kurang polimorfik jika dibandingkan dengan SSR yang ditemukan pada genom, tetapi jika dibandingkan dengan marka lain EST-SSR masih menunjukkan tingkat variasi yang cukup untuk analisis genetika populasi (Ellis and Burke, 2007).

Penelitian untuk *mining* mikrosatelit pada gaharu *A. malaccensis* telah dilakukan oleh Bordoloi *et al.*, (2022) yang mencari situs mikrosatelit pada *whole genome*, dan (Hishamuddin, 2022) yang mencari situs mikrosatelit pada kloroplas dari 8 spesies *Aquilaria*. Mikrosatelit yang ditemukan dari genom, baik genom nukleus maupun kloroplas memiliki sifat yang lebih polimorfik, cenderung tersebar luas di seluruh genom, tetapi kemampuan transfer antar spesiesnya rendah (Dweikat, 2004). Sementara itu, mikrosatelit dari bagian yang terekspresi (EST) memiliki potensi untuk mendeteksi perubahan gen yang terjadi (Wang *et al.*, 2007), tetapi kurang memiliki banyak informasi karena hanya dihasilkan dari daerah yang ditranskripsikan (Thiel *et al.*, 2003). Mikrosatelit yang ditemukan pada EST

kemungkinan memiliki kaitan dengan gen-gen tertentu yang diekspresikan.

Beberapa gen yang ditemukan berasosiasi dengan situs SSR antara lain 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGR), Mevalonate kinase, Farnesyl pyrophosphate synthase, Delta-guaiene synthase 3, Mitogen-activated protein kinase 3 (MAP Kinase), dan Phenylalanine ammonia-lyase. HMGR berperan dalam biosintesis sesquiterpene sebagai enzim kunci dalam pathway MVP (Xu *et al.*, 2013). Penelitian (Islam and Banu, 2021) menemukan Mevalonate Kinase berperan dalam biosintesis terpenoid backbone pada *A. malaccensis* yang terinfeksi secara alami. Farnesyl pyrophosphate synthase akan diubah menjadi farnesyl pyrophosphate (Shivanand, Arbie and Krishnamoorthy, 2022) yang berperan dalam sintesis sesquiterpene dan phenylethyl chromone (Aminan *et al.*, 2022).

Delta-guaiene synthase merupakan prekursor delta guaiene (Deguerry *et al.*, 2006) yang merupakan salah satu bentuk sesquiterpene (Kumeta and Ito, 2011). MAP Kinase merupakan transcription factor yang berperan dalam signalling setelah terjadi stres akibat wounding sebagai defense response (Xu *et al.*, 2013). Proses perlukaan dan infeksi jamur gaharu menginduksi ekspresi gen Phenylalanine ammonia-lyase (PAL) yang berperan juga dalam sintesis senyawa flavonoid pada *A. sinensis* (Xu *et al.*, 2023). Identifikasi dan karakterisasi EST-SSR yang berpaut dengan gen-gen yang berperan dalam sintesis senyawa penghasil gaharu berguna untuk pemuliaan tanaman gaharu. SSR berbasis gen dapat membantu seleksi tanaman yang mampu menghasilkan gaharu dengan marker-assisted selection (MAS).

## Kesimpulan

Identifikasi EST-SSR dari transkriptom gaharu GHJS00000000.1 memberikan hasil bahwa *imperfect* SSR lebih banyak dibanding *perfect* SSR yang menunjukkan situs SSR pada transkriptom gaharu lebih terkonservasi. Trinukleotida yang ditemukan pada *imperfect* SSR juga lebih tinggi dibandingkan pada *perfect* SSR yang menandakan kodon-kodon penyusun asam amino terkonservasi. Terdapat 16 SSR yang berpaut dengan gen-gen yang berperan dalam sintesis senyawa gaharu yang dapat



dikembangkan menjadi marka molekuler untuk seleksi tanaman gaharu.

### Ucapan Terima Kasih

Peneliti ucapkan terima kasih kepada semua pihak yang terlibat dalam penelitian ini baik secara moral maupun materil.

### Referensi

- Abdul Kadir, F.A., Azizan, K.A. and Othman, R. (2021) 'Transcriptome of *Aquilaria malaccensis* containing agarwood formed naturally and induced artificially', *BMC Research Notes*, 14(1), pp. 1–4. Available at: <https://doi.org/10.1186/s13104-021-05532-9>.
- Altschul, S.F. (1990) 'Basic local alignment search tool', *Journal of Molecular Biology*, 215(3), pp. 403–410. Available at: [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
- Aminan, A. W., Hamid, A. A. A., Tajuddin, S. N., & Ramli, A. N. M. (2022). Active site and structural analysis of sesquiterpene synthase towards synthesis of aromatic compounds related to agarwood formation in *Aquilaria malaccensis*.1–25. <https://doi.org/10.21203/rs.3.rs-1810009/v1>.
- Ananda, G., Hile, S. E., Breski, A., Wang, Y., Kelkar, Y., Makova, K. D., & Eckert, K. A. (2014). Microsatellite interruptions stabilize primate genomes and exist as population-specific single nucleotide polymorphisms within individual human genomes. *PLoS genetics*, 10(7), e1004498. <https://doi.org/10.1371/journal.pgen.1004498>.
- Anwar, M. S., Wardati, W., & Ardian, A. (2017). *Efek pemberian pupuk kascing dan urea terhadap pertumbuhan bibit gaharu (Aquilaria malaccensis Lamk.)* (Doctoral dissertation, Riau University).
- Aqmarina, A., Siregar, U. J., & Turjaman, M. (2018). *Kandungan Kimia Gaharu dan Ekspresi Gen Sesquiterpene synthase 1 (SesTPS1) pada Aquilaria malaccensis Lamk dan Gyrinops versteegii Domke* (Doctoral dissertation, Bogor Agricultural University (IPB)).
- Bordoloi, S., Ravi, N., Modi, M. K., & Jayaraj, R. S. C. (2022). Genome-wide mining of simple sequence repeats and development of polymorphic SSR markers in *Aquilaria malaccensis*. *Indian Journal of Genetics and Plant Breeding*, 82(02), 236-239. <https://doi.org/10.31742/IJGPB.82.2.15>.
- Borisova, O. F., Shchyolkina, A. K., Chernov, B. K., & Tchurikov, N. A. (1993). Relative stability of AT and GC pairs in parallel DNA duplex formed by a natural sequence. *FEBS letters*, 322(3), 304-306. [https://doi.org/10.1016/0014-5793\(93\)81591-M](https://doi.org/10.1016/0014-5793(93)81591-M).
- CITES (2018) *CITES Trade Database [2018]. Compiled by UNEP-WCMC for the CITES Secretariat*. Available at: [trade.cites.org](http://trade.cites.org).
- Deguerry, F., Pastore, L., Wu, S., Clark, A., Chappell, J., & Schalk, M. (2006). The diverse sesquiterpene profile of patchouli, *Pogostemon cablin*, is correlated with a limited number of sesquiterpene synthases. *Archives of biochemistry and biophysics*, 454(2), 123-136. <https://doi.org/10.1016/j.abb.2006.08.006>.
- Ding, S., Wang, S., He, K., Jiang, M., & Li, F. (2017). Large-scale analysis reveals that the genome features of simple sequence repeats are generally conserved at the family level in insects. *BMC genomics*, 18, 1-10. <https://doi.org/10.1186/s12864-017-4234-0>.
- Du, L., Zhang, C., Liu, Q., Zhang, X., & Yue, B. (2018). Krait: an ultrafast tool for genome-wide survey of microsatellites and primer design. *Bioinformatics*, 34(4), 681-683. <https://doi.org/10.1093/bioinformatics/btx665>.
- Kuleung, C., Baenziger, P. S., & Dweikat, I. (2004). Transferability of SSR markers among wheat, rye, and triticale. *theoretical and Applied Genetics*, 108, 1147-1150. <https://doi.org/10.1007/s00122-003-1532-5>.
- Ellis, J. R., & Burke, J. M. (2007). EST-SSRs as a resource for population genetic analyses. *Heredity*, 99(2), 125-132. <https://doi.org/10.1038/sj.hdy.6801001>.
- Feng, S. P., Li, W. G., Huang, H. S., Wang, J. Y., & Wu, Y. T. (2009). Development, characterization and cross-species/genera transferability of EST-SSR markers for

- rubber tree (Hevea brasiliensis). *Molecular breeding*, 23, 85–97. <https://doi.org/10.1007/s11032-008-9216-0>.
- Gao, T., Ma, X., & Zhu, X. (2013). Use of the psbA-trnH region to authenticate medicinal species of Fabaceae. *Biological and Pharmaceutical Bulletin*, 36(12), 1975–1979. <https://doi.org/10.1248/bpb.b13-00611>.
- Gao, X., Su, Q., Yao, B., Yang, W., Ma, W., Yang, B., & Liu, C. (2022). Development of EST-SSR markers related to polyphyllin biosynthesis reveals genetic diversity and population structure in Paris polyphylla. *Diversity*, 14(8), 589. <https://doi.org/10.3390/d14080589>.
- Hidayat, H., Siburian, R. and Indah Yuliana, C. (2020) ‘Gaharu Alam, Jaringan Perdagangan, dan Gaharu Budidaya: Studi Kasus Kalimantan Timur’, *Jurnal Biologi Indonesia*, 16(1), pp. 99–110. Available at: <https://doi.org/10.47349/jbi/16012020/99>.
- Hishamuddin, M.S. (2022) ‘Mining and analysis of chloroplast simple sequence repeats ( SSRs ) from eight species of Aquilaria Mining and analysis of chloroplast simple sequence repeats ( SSRs ) from eight species of’, 46(1). Available at: <https://doi.org/10.3906/bot-2108-61>.
- Hu, J., Wang, L. and Li, J. (2011) ‘Comparison of genomic SSR and EST-SSR markers for estimating genetic diversity in cucumber’, *Biologia Plantarum*, 55(3), pp. 577–580. Available at: <https://doi.org/10.1007/s10535-011-0129-0>.
- Islam, R. and Banu, S. (2021) ‘Transcript profiling leads to biomarker identification for agarwood resin - loaded Aquilaria malaccensis’, *Trees* [Preprint], (0123456789). Available at: <https://doi.org/10.1007/s00468-021-02180-1>.
- Kamaruddin, K., Titawael, R. and Gawariah, G. (2022) ‘Contribution of Gaharu (Aquilaria sp) to Community Income in Fatmite Village, Namrole District, South Buru Regency’, *Jurnal Agrohut*, 12(1), pp. 23–30.
- Kotwal, S. *et al.* (2016) ‘De novo transcriptome analysis of medicinally important plantago ovata using RNA-seq’, *PLoS ONE*, 11(3). Available at: <https://doi.org/10.1371/journal.pone.0150273>.
- Kumeta, Y. and Ito, M. (2011) ‘Genomic organization of d -guaiene synthase genes in Aquilaria crassna and its possible use for the identification of Aquilaria species’, pp. 508–513. Available at: <https://doi.org/10.1007/s11418-011-0529-7>.
- Lawson, M.J., Zhang, L. and Tech, V. (2006) ‘Distinct patterns of SSR distribution in the Arabidopsis thaliana and rice genomes’, 7, pp. 1–11. Available at: <https://doi.org/10.1186/gb-2006-7-2-r14>.
- Li, Z., Yun, L., Gao, Z., Wang, T., Ren, X., & Zhao, Y. (2022). EST-SSR primer development and genetic structure analysis of Psathyrostachys juncea Nevski. *Frontiers in plant science*, 13, 837787. <https://doi.org/10.3389/fpls.2022.837787>.
- Liao, B., Lee, S. Y., Meng, K., Yin, Q., Huang, C., Fan, Q., ... & Chen, S. (2019). Characterization and novel Est-SSR marker development of an important Chinese medicinal plant, Morinda officinalis How (Rubiaceae). *Biotechnology & Biotechnological Equipment*, 33(1), 1311–1318. <https://doi.org/10.1080/13102818.2019.1664322>.
- Liu, Y. Y., Wei, J. H., Gao, Z. H., Zhang, Z., & Lyu, J. C. (2017). A review of quality assessment and grading for agarwood. *Chinese Herbal Medicines*, 9(1), 22–30. [https://doi.org/10.1016/S1674-6384\(17\)60072-8](https://doi.org/10.1016/S1674-6384(17)60072-8).
- López-Sampson, A. and Page, T. (2018) ‘History of Use and Trade of Agarwood’, *Economic Botany*, 72(1), pp. 107–129. Available at: <https://doi.org/10.1007/s12231-018-9408-4>.
- Luo, L., Yang, Y., Zhao, H., Leng, P., Hu, Z., Wu, J., & Zhang, K. (2021). Development of EST-SSR markers and association analysis of floral scent in tree peony. *Scientia Horticulturae*, 289, 110409.

- <https://doi.org/https://doi.org/10.1016/j.scienta.2021.110409>.
- Lv, F., Yang, Y., Sun, P., Zhang, Y., Liu, P., Fan, X., ... & Wei, J. (2022). Comparative transcriptome analysis reveals different defence responses during the early stage of wounding stress in chi-nan germplasm and ordinary *Aquilaria sinensis*. *BMC Plant Biology*, 22(1), 464. <https://doi.org/10.1186/s12870-022-03821-4>.
- Qu, J. and Liu, J. (2013) 'A genome-wide analysis of simple sequence repeats in maize and the development of polymorphism markers from next-generation sequence data', *BMC Research Notes*, 6(1), pp. 1–10. Available at: <https://doi.org/10.1186/1756-0500-6-403>.
- R Core Team (2021) 'R: A language and environment for statistical computing.', *R Foundation for Statistical Computing* [Preprint]. Available at: <https://www.r-project.org/>.
- Ranade, S. S., Lin, Y. C., Zuccolo, A., Van de Peer, Y., & García-Gil, M. D. R. (2014). Comparative in silico analysis of EST-SSRs in angiosperm and gymnosperm tree genera. *BMC Plant Biology*, 14, 1-10.
- Rossetto, M. (2001) 'Sourcing of SSR markers from related plant species.', in J.C. Henry (ed.) *Plant genotyping: the DNA fingerprinting of plants*. Oxford: CABI Publishing, pp. 211–224. Available at: <https://doi.org/10.1079/9780851995151.0211>.
- Shen, S., Chai, X., Zhou, Q., Luo, D., Wang, Y., & Liu, Z. (2019). Development of polymorphic EST-SSR markers and characterization of the autotetraploid genome of sainfoin (*Onobrychis viciifolia*). *PeerJ*, 7, e6542. <https://doi.org/10.7717/peerj.6542>.
- Shivanand, P., Arbie, N.F. and Krishnamoorthy, S. (2022) 'Agarwood — The Fragrant Molecules of a Wounded Tree', pp. 1–23.
- Singh, S., Gupta, S., Mani, A., & Chaturvedi, A. (2012). Mining and gene ontology based annotation of SSR markers from expressed sequence tags of *Humulus lupulus*. *Bioinformation*, 8(3), 114.
- Srivastava, D., Ahmad, M. M., Shamim, M., Maurya, R., Srivastava, N., Pandey, P., ... & Siddiqui, M. H. (2019). Modulation of Gene Expression by Microsatellites in Microbes. In *New and Future Developments in Microbial Biotechnology and Bioengineering* (pp. 209-218). Elsevier. <https://doi.org/10.1016/B978-0-444-63503-7.00012-7>.
- Thiel, T., Michalek, W., Varshney, R., & Graner, A. (2003). Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and applied genetics*, 106, 411-422. <https://doi.org/10.1007/s00122-002-1031-0>.
- Vranová, E., Coman, D. and Grissem, W. (2013) 'Network analysis of the MVA and MEP pathways for isoprenoid synthesis', *Annual Review of Plant Biology*, 64(March 2013), pp. 665–700. Available at: <https://doi.org/10.1146/annurev-arplant-050312-120116>.
- Wang, H. Y., Wei, Y. M., Yan, Z. H., & Zheng, Y. L. (2007). EST-SSR DNA polymorphism in durum wheat (*Triticum durum* L.) collections. *Journal of Applied Genetics*, 48, 35-42.
- Xu, J., Du, R., Wang, Y., & Chen, J. (2023). Wound-Induced Temporal Reprogramming of Gene Expression during Agarwood Formation in *Aquilaria sinensis*. *Plants*, 12(16), 2901.
- Xu, Y., Zhang, Z., Wang, M., Wei, J., Chen, H., Gao, Z., ... & Li, W. (2013). Identification of genes related to agarwood formation: transcriptome analysis of healthy and wounded tissues of *Aquilaria sinensis*. *BMC genomics*, 14, 1-16. <https://doi.org/10.1186/1471-2164-14-227>.
- Zhang, J., Kuo, C.C.J. and Chen, L. (2011) 'GC content around splice sites affects splicing through pre-mRNA secondary structures', *BMC Genomics*, 12. Available at: <https://doi.org/10.1186/1471-2164-12-90>