



ANALISIS BIOINFORMATIKA BERBASIS WEB PADA SEKUEN GENOM PARSIAL SAGU (*Metroxylon sagu* Rottb.)

WEB-based bioinformatic analysis on partial genome sequence of Sago (*Metroxylon sagu* Rottb.)

Devit Purwoko^{1*}, Imam Civi Cartealy¹, Teuku Tajuddin¹, Diny Dinarti², Sudarsono²

¹Balai Bioteknologi, BPPT. Gedung 630 Kawasan Puspiptek, Tangerang Selatan, Banten 15314

²Lab Biologi Molekuler Tanaman, Departmen Agronomi dan Hortikultura, IPB. Darmaga, Bogor 16680.

*Email: devit.purwoko@bppt.go.id

ABSTRACT

Sago genome sequencing analysis is still very limited. This study is a preliminary study of sago sequence analysis obtained from NGS technology to understand and identify new genetic sequences that have homology to genes in the NCBI database. Sequences were analyzed using Blast2Go to determine the genetic function annotation, putative gene identification was performed on the Arabidopsis database using the BLASTx program with a 10^{-3} e-value limit on The Arabidopsis Information Resource (TAIR) (<http://www.arabidopsis.org/index.jsp>). Gene interactions were analyzed using DAVID and GeneMania programs. Based on sequence analysis with Blast2Go, 33 sequences with Blastx hit consisting of: 29 sequences had a high homology. The sago sequences with a similarity of $\geq 90\%$ are glutamate decarboxylase and HT1-like serine threonine kinase with hit number 10. The distribution of interactions between genes from GeneMania analysis is known to be mostly interconnected in the 65.13% protein domain, predicted 19.83%, genes with 14.47% shared expression and the remaining 0.57% had localization together.

Keywords: bioinformatics, gene annotation, gene ontology, genome sequence, *Metroxylon sagu*

ABSTRAK

Kajian analisis sekuen genom sago hingga saat ini masih amat terbatas. Penelitian ini merupakan riset pendahuluan analisis sekuen sago yang diperoleh dari teknologi NGS untuk mengetahui dan mengidentifikasi sekuen gen baru yang memiliki homologi dengan gen pada database NCBI. Sekuen dianalisis menggunakan perangkat Blast2Go untuk mengetahui anotasi fungsional gen, identifikasi gen putatif dilakukan terhadap database *Arabidopsis* menggunakan program BLASTx dengan batas e-value 10^{-3} pada *The Arabidopsis Information Resource* (TAIR). Interaksi gen dianalisis menggunakan program DAVID dan GeneMania. Berdasarkan analisis sekuen dengan Blast2Go, diperoleh 33 sekuen dengan Blastx hit yang terdiri atas: 29 sekuen memiliki homologi yang tinggi. Gen dengan rataan kemiripan $\geq 90\%$ adalah *glutamate decarboxylase* dan *serine threonine-kinase* HT1-like dengan jumlah hit 10. Persebaran interaksi antar gen hasil analisis GeneMania diketahui sebagian besar saling terkait pada domain protein 65,13%, koneksi yang berhasil diprediksi 19,83%, gen dengan ekspresi bersama 14,47% dan sisanya 0,57% memiliki peranan bersama.

Kata Kunci: anotasi gen, bioinformatika, *Metroxylon sagu*, ontologi gen, sekuen genome

PENDAHULUAN

Sagu (*Metroxylon sagu* Rottb.) merupakan salah satu tanaman Palmae penghasil pati yang mampu mengakumulasi pati pada batangnya 244 sampai 310 kg/pohon (Ehara et al. 2016). Luas areal hutan sagu dunia lebih dari 50% berada di Indonesia dengan potensi 20-40 ton pati sagu/ha/tahun (Bintoro 2011). Dalam 100 gram tepung sagu mengandung nutrisi sebagai berikut: kalori 285, air 77 g, protein 0,2 g, kalsium 30 mg, karbohidrat 71 g, besi 0,7 mg dan serat 0,3 g (Johnson 2010).

Tanaman sagu memiliki potensi ekonomi, energi dan genetik yang tinggi untuk dikembangkan di Indonesia. Secara ekonomi tanaman sagu memberikan keuntungan yang cukup besar dari pati yang dihasilkan persatuan luas dan waktu. Yamamoto (2010) melaporkan bahwa produksi pati kering dari tanaman sagu rata-rata mencapai 337 kg/pohon sedangkan Bintoro (2011) menyatakan potensi tanaman sagu menghasilkan pati 100-900 kg/pohon. Jika jarak tanam 9 m × 9 m akan terdapat 123 pohon/ha maka akan diperoleh 42 ton pati sagu/ha/tahun (Abbas 2006). Potensi produksi pati Indonesia diperkirakan 5 juta ton pati kering sagu/tahun sedangkan konsumsi pati dalam negeri baru mencapai 210 ton (Sumaryono 2007) dan ekspor baru mencapai 9.680,908 ton pada tahun 2015 (Ditjen Perkebunan 2016). Berdasarkan perhitungan tersebut masih terdapat potensi sekitar 4,5 juta ton pati sagu yang belum dimanfaatkan atau setara dengan 607,5 juta US\$ jika diekspor mentah dan bernilai lebih besar lagi jika sudah dalam produk industri.

Potensi sagu yang tidak kalah penting dari sektor energi, jika 4,5 juta ton pati sagu yang belum dimanfaatkan digunakan sebagai bioetanol dengan perhitungan 25 ton/ha/tahun setara dengan 15 kiloliter etanol (Sumaryono 2007) maka akan diperoleh 2,7 juta kiloliter etanol. Tanaman sagu juga memiliki potensi genetik yang belum banyak dikembangkan. Tanaman sagu memiliki mekanisme genetik yang belum banyak dimengerti berkaitan dengan daya adaptasinya pada lingkungan yang tergenang dan pH rendah. Potensi genetik ini dapat dikembangkan untuk mencari gen-gen potensial terhadap cekaman abiotik.

Informasi berkenaan dengan karakter

berdiri maupun potensi sagu secara morfologi untuk kandungan pati cukup banyak dikaji (Novero et al. 2012; Kjaer et al. 2004; Karim et al. 2008). Kajian budidaya dan perbanyak bibit baik secara in-vitro (Riyadi et al. 2016) dan ex-vitro (Karyanti et al. 2016) juga telah dilaporkan. Namun Perkembangan kemajuan informasi genomik pada tanaman sagu masih terbatas terutama untuk marka DNA. Kajian molekuler berkenaan dengan keragaman dan kekerabatan sagu di Indonesia menggunakan marka juga belum banyak dilaporkan. Studi penanda DNA tanaman sagu hingga saat ini diantaranya penanda RAPD (Abbas et al. 2009), AFLP (Kjaer et al. 2004) dan DNA kloroplas (cpDNA) (Abbas et al. 2010).

Kajian analisis sekuen genom sagu hingga saat ini masih amat terbatas. Hingga saat ini baru terdeposit 111 sekuen genom dan 413 sekuen EST pada database *National Center for Biotechnology Information* (NCBI, data di akses Maret 2018). Analisis sekuen sagu telah dilaporkan oleh Wee dan Roslan (2012) dengan memanfaatkan EST dari daun muda sagu bahwa terdapat beberapa sekuen sagu yang belum memiliki homologi dengan sekuen pada database NCBI. Analisis sekuen untuk identifikasi gen *ADP-Glucose Phosphorylase* (Budiani et al. 2015) dan gen *GA 20-Oxidase* (Jamel et al. 2011).

Analisis sekuen secara tradisional dengan kloning sekuen kemudian dilanjutkan perunutan dengan metode Sanger membutuhkan tenaga dan biaya. Analisis sekuen memanfaatkan teknologi perunutan DNA terkini yaitu *Next Generation Sequencing* (NGS) berbasis bioinformatika mampu menghasilkan informasi yang efektif dan efisien. Hal ini dapat dilihat pada penambahan data sekuen di NCBI yang dihasilkan melalui teknologi NGS. Hingga tahun 2016 data yang dihasilkan melalui teknologi ini sebesar 3.051,43 TB dengan total basa 5.005,71 Tb (Raza dan Ahmad 2017).

Analisis sekuen dengan bantuan bioinformatika dapat dilakukan dengan dua pendekatan yaitu berbasis *web* dan berbasis algoritma bahasa pemrograman. Bagi pemula pemanfaatan bioinformatika berbasis *web* akan lebih mudah dibandingkan dengan algoritma bahasa pemrograman. Analisis bioinformatika berbasis *web* dapat digunakan dalam mencari anotasi

(penamaan), pemetaan genom, dan analisis sekuen lanjut seperti identifikasi gen homologi serta interaksi antar gen terkait. Analisis bioinformatik berbasis *web* dapat dijalankan secara daring (dalam jaringan) melalui program yang tersedia dengan akses publik di *web*. Keunggulan metode tersebut adalah hemat dan dapat menjadi penelitian pendahuluan sebelum percobaan secara nyata dilakukan (Narita et al. 2012). Namun untuk analisis data sekuen yang besar dibutuhkan algoritma bahasa pemrograman yang ditunjang oleh super komputer agar dapat berjalan secara simultan.

Penelitian ini merupakan riset pendahuluan analisis sekuen sagu yang diperoleh dari teknologi NGS untuk mengetahui dan mengidentifikasi sekuen gen baru yang memiliki homologi dengan gen yang sudah ada pada basis data NCBI. Analisis interaksi gen potensial juga dilakukan untuk memperoleh informasi gen fungsional yang terkait pada jalur biosintesis suatu protein. Analisis bioinformatik sekuen sagu berbasis *web* diharapkan dapat memberi landasan awal untuk diperolehnya gen-gen potensial dari tanaman sagu terkait ketahanan biotik dan abiotik.

BAHAN DAN METODE

Sampel sekuen yang digunakan pada penelitian ini merupakan 85 sekuen sagu yang telah dideposit pada database NCBI. Sekuen tersebut merupakan hasil sekuensing NGS menggunakan metode *paired end* pada mesin Illumina GAIIx dan di *assembly* menggunakan perangkat lunak Ray. Untuk mengidentifikasi gen putatif, sekuen dibandingkan dengan database protein *non-redundant* NCBI menggunakan program BLASTx dengan nilai ambang 10^{-3} . Hasil BLASTx dengan kemiripan rata-rata sama dengan atau kurang dari 60% dikategorikan sebagai homologi rendah sedangkan yang rata-rata lebih dari 60% dikategorikan sebagai homologi tinggi (Wee dan Roslan 2012). Sementara itu, sekuen tanpa kecocokan diklasifikasikan sebagai sekuen tanpa *blast*. Sekuen dengan hasil analisis BLASTx kemudian dipetakan ke *Gene Ontology* (GO) menggunakan program BLAST2GO dan dikategorikan berdasarkan fungsi molekuler, proses biologis dan komponen seluler (Wee dan Roslan 2012).

Analisis sekuen dilakukan untuk mengetahui gen potensial menggunakan program Blastx dengan membandingkan protein *non-redundant* pada database NCBI. Sekuen juga dibandingkan dengan *Arabidopsis* database pada *The Arabidopsis Information Resource* (TAIR) (<http://www.arabidopsis.org/index.jsp>) menggunakan program BLASTx dengan batas e-value 10^{-3} . Hasil *blast* kemudian dipetakan untuk mengetahui anotasi ontology gen yang meliputi kategori fungsi molekuler, proses biologi dan komponen seluler dengan menggunakan identifikasi lokus pada *Bulk Retrieval System of TAIR* (<http://www.arabidopsis.org/tools/bulk/go/index.jsp>).

Gen-gen yang dianotasi dengan *Arabidopsis* dilanjutkan dengan analisis jalur fungsional menggunakan aplikasi Online DAVID (Huang et al. 2009) (DAVID v6.7; <http://david.abcc.ncifcrf.gov/>; Diakses Agustus 2017). Aplikasi ini memiliki kelebihan karena mengandalkan algoritma anotasi fungsional sesungguhnya dan mampu mengelompokkan kumpulan banyak gen ke dalam kelompok fungsional. Penggunaan DAVID dimaksudkan untuk menentukan jalur (*pathway*) mana yang paling banyak dipengaruhi oleh gen yang diprediksi sebelumnya. Prosesnya dengan memasukan daftar gen yang telah diprediksi sebelumnya ke dalam web aplikasi *functional annotation tool* DAVID dan identifikasi gen *Arabidopsis* dipilih sebagai daftar gen. Idealnya, hanya kelompok anotasi dengan nilai *enrichment* 1,3 atau lebih tinggi yang dianggap signifikan namun kelompok anotasi dengan nilai *enrichment* rendah (kurang dari 1,3) masih dapat dipertimbangkan untuk dianalisis lebih lanjut karena ada kemungkinan memiliki gen yang penting (Huang et al. 2009).

Analisis jaringan (*network*) interaksi antar gen dilakukan menggunakan GeneMania (Warde-Farley et al. 2010) (<http://www.genemania.org/>; Diakses Agustus 2017) untuk menentukan apakah ada prediksi *pathway* yang sesuai dalam *network* tertentu. GeneMania merupakan aplikasi berbasis web yang tersedia secara publik untuk mengidentifikasi gen tambahan dari gen prediksi dan hubungan interaksinya dalam bentuk jaringannya. Gen yang diperoleh dari analisis *pathway*

menggunakan DAVID diunggah ke GeneMania dengan memilih database *Arabidopsis*. Semua parameter dibiarkan default untuk analisis. Hasil analisis dengan nilai *false discovery rate* (FDR) < 0,05 saja yang dilaporkan. Alur proses analisis sekuen dapat dilihat pada Gambar 1.

HASIL DAN PEMBAHASAN

Pada penelitian ini, analisis dilakukan pada 85 sekuen sagu yang diperoleh menggunakan teknologi *Next Generation Sequencing* (NGS) dan telah dideposit pada database NCBI dengan nomor identitas sekuen mulai dari MG904300 hingga MG904384. Sekuen dianalisis menggunakan BLASTX menggunakan database NCBI pada perangkat Blast2Go dengan nilai ambang 10^{-3} (Tabel 1). Sebanyak 33 sekuen yang dianalisis diketahui memiliki Blast hit, sedangkan 52 sekuen lainnya tanpa Blast hit. Berdasarkan analisis Blast diperoleh informasi bahwa 17 sekuen memiliki ontologi gen dan berhasil dipetakan. Sekuen yang memiliki kecocokan hanya pada InterProScan yaitu sc-7446. Namun yang memiliki kecocokan dengan database KEGG hanya 2 sekuen dengan 3 enzim yang sudah diketahui yaitu sekuen sc-678470 dan sc-16071 (Lampiran 1). Ketiga enzim tersebut yaitu Adenosinetriphosphatase (EC:3.6.1.3), Nucleoside-triphosphate phosphatase (EC:3.6.1.15) dan Glutamate decarboxylase

(EC:4.1.1.15). Berdasarkan analisis jalur KEGG EC:4.1.1.15 berada dalam jalur metabolisme alanine, aspartate dan glutamate sedangkan EC:3.6.1.3 dan EC:3.6.1.15 berada dalam jalur metabolisme purine.

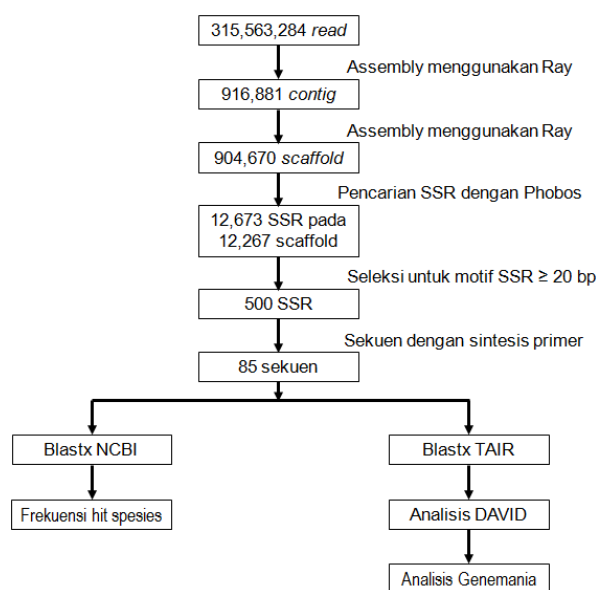
Berdasarkan analisis sekuen dengan Blast2Go dapat dilihat bahwa terdapat 33 sekuen dengan Blastx hit yang terdiri atas: 29 sekuen memiliki homologi yang tinggi dengan rata-rata kemiripan $\geq 60\%$ dan hanya 4 sekuen yang < 60%. Gen yang paling banyak ditemukan pada sekuen sagu dengan rata-rata kemiripan $\geq 90\%$ adalah *glutamate decarboxylase* dan *serine threonine-kinase HT1-like* dengan jumlah hit 10.

Glutamate decarboxylase (GAD) merupakan gen dengan sekuen yang terkonservasi dan diketahui memiliki peranan dalam mengkatalisis dekarboksilasi glutamat untuk menghasilkan asam γ -aminobutirat (GABA). Pada tanaman, aktivitasnya dimodulasi oleh pH dan kalsium. Pengaturan aktivitas GAD oleh kalsium adalah melalui domain pengikat calmodulin yang tampaknya spesifik untuk tanaman dan dihasilkan dari proses peristiwa evolusi. GAD pada tanaman biasanya disandikan oleh beberapa gen sehingga menunjukkan pola ekspresi yang berbeda, termasuk spesifisitas jaringan dan respon terhadap stres, yang menunjukkan bahwa ada fenomena spesialisasi dalam kelompok gen ini. Meskipun peran GABA sebagai

Tabel 1. Distribusi sekuen sagu berdasarkan analisis dengan Blast2Go

Kriteria analisis sekuen	Jumlah sekuen
Dianalisis dengan Blast2Go*	85
Memiliki ontologi gen	17
Memiliki Blastx hit	33
Tanpa Blastx hit	52
Berhasil dipetakan	17
Hanya memiliki InterProScan	1
Memiliki hit spesies terbanyak <i>Elaeis guineensis</i>	76
Memiliki anotasi KEGG	2
Anotasi dengan enzim yang sudah diketahui	3

*Nilai ambang $\geq 10^{-3}$.



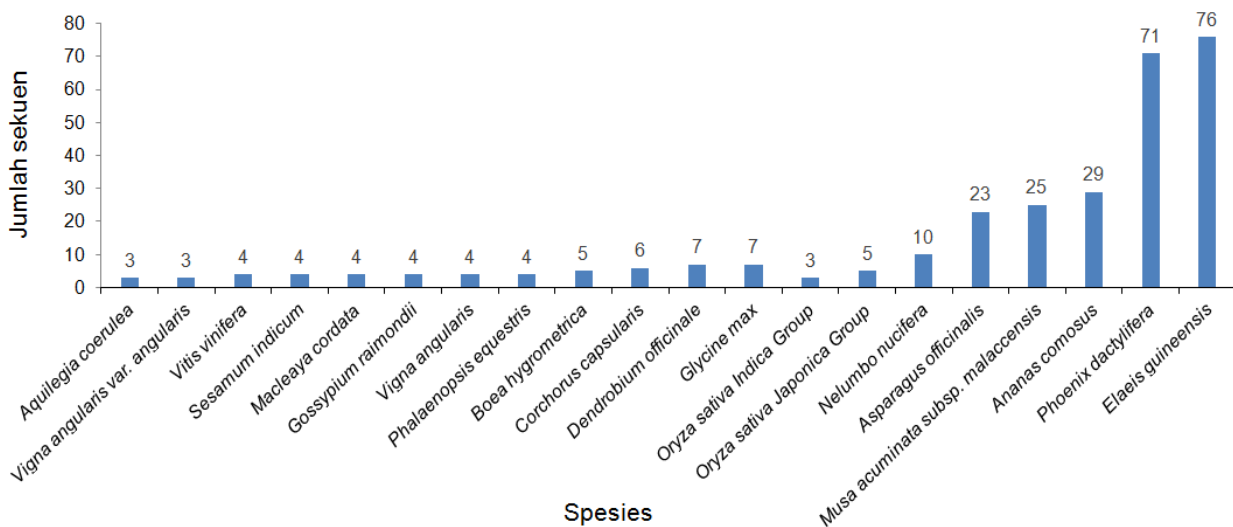
Gambar 1. Bagan alur analisis sekuen

neurotransmitter sangat dikenal pada hewan, beberapa peran untuk GABA dan metabolismenya telah diusulkan pada tanaman selama beberapa dekade terakhir. Peran GABA pada tanaman termasuk berperan sebagai biokimia pH-stat, dalam penyimpanan nitrogen sementara, sebagai osmolyte yang kompatibel, dalam pertahanan melawan stres biotik, dalam mengendalikan spesies oksigen reaktif (*Reactive Oxygen Species / ROS*) dan dalam metabolisme intermediet fotorespirasi. Selain itu, GABA dianggap sebagai sinyal yang mengendalikan ekspresi gen pengatur. Namun, informasi yang tersedia yang mendukung keterlibatan GAD dan / atau GABA dalam banyak proses sangat terbatas, begitu pula kontribusi metabolisme poliamina ke GABA pada tingkat sel (Molina-Rueda et al. 2015). *Serine threonine-kinase HT1-like* diketahui memiliki peranan penting untuk regulasi gerakan stomata dan fungsinya lebih sebagai respon terhadap CO₂ daripada terhadap ABA atau cahaya (Hashimoto et al. 2006). Berdasarkan analisis sekuen dengan Blast2Go, 52 sekuen belum memiliki kecocokan dengan database pada NCBI. Hal ini menunjukkan bahwa sekuen-sekuen tersebut memiliki potensi untuk dianalisis lebih lanjut dan berpeluang besar dalam memperoleh gen-gen baru yang hanya ada pada tanaman sagu.

Analisis BLAST 85 sekuen terhadap spesies pada database NCBI diperoleh informasi kecocokan beberapa spesies tanaman (Gambar 2), yaitu: delapan (40%)

tanaman monokotil (*Oryza sativa Indica Group, Oryza sativa Japonica Group, Nelumbo nucifera, Asparagus officinalis, Musa acuminata subsp. Malaccensis, Ananas comosus, Phoenix dactylifera, Elaeis guineensis*) dan dua belas (60%) tanaman dikotil (*Aquilegia coerulea, Vigna angularis var. Singularis, Vitis vinifera, Sesamum indium, Macleaya cordata, Gossypium raimondii, Vigna singularis, Phalaenopsis equestris, Boea hygrometrica, Corchorus capsularis, Dendrobium officinale, Glycine max*). Sekuen sagu terdistribusi paling banyak pada tanaman dikotil dibandingkan tanaman monokotil. Hal yang berbeda diperoleh pada EST daun muda sagu bahwa sekuen sagu terdistribusi hampir sama antara tanaman monokotil dan dikotil yaitu 41,94% dan 39,24% (Wee dan Roslan 2012). Berdasarkan analisis sekuen hit terbanyak diperoleh pada *Elaeis guineensis* (76) dan *Phoenix dactylifera* (71) yang merupakan tanaman Palma. Hal ini menunjukkan bahwa primer yang disintesis dari genom sagu dapat digunakan pula pada kedua tanaman tersebut.

Ontologi gen adalah klasifikasi gen sesuai kriteria yang telah ditentukan dengan menghubungkan semua gen yang tersedia ke dalam gen tertentu. Klasifikasi dilakukan dengan menghimpun serangkaian gen dalam tiga kelompok kategori: *Celluler Componen (CC), Biological Process (BP)* dan *Molecular Function (MF)*. Program BLASTx pada TAIR digunakan untuk mencari anotasi ontologi gen. Hasil anotasi ontologi gen sekuen produk primer tanaman



Gambar 2. Frekuensi 20 spesies tanaman dengan hit terbanyak

sagu dengan nilai homologi yang tinggi digambarkan pada peta kromosom *Arabidopsis thaliana* (Gambar 3).

Sekuen tanaman sagu yang memiliki anotasi dengan gen *Arabidopsis thaliana* tersebar pada lima kromosom dengan penyebaran terbanyak pada kromosom pertama. AT1G01510 merupakan GO ID yang terbanyak untuk kategori BP, gen ini berperan dalam morfogenesis daun, bunga dan buah. AT5G58300 merupakan GO ID terbanyak untuk kategori MF, gen ini berperan dalam mengikat ATP dan aktivitas protein kinase. AT5G50850 merupakan GO ID yang terbanyak untuk kategori CC, gen ini berperan dalam komponen sel mitokondria.

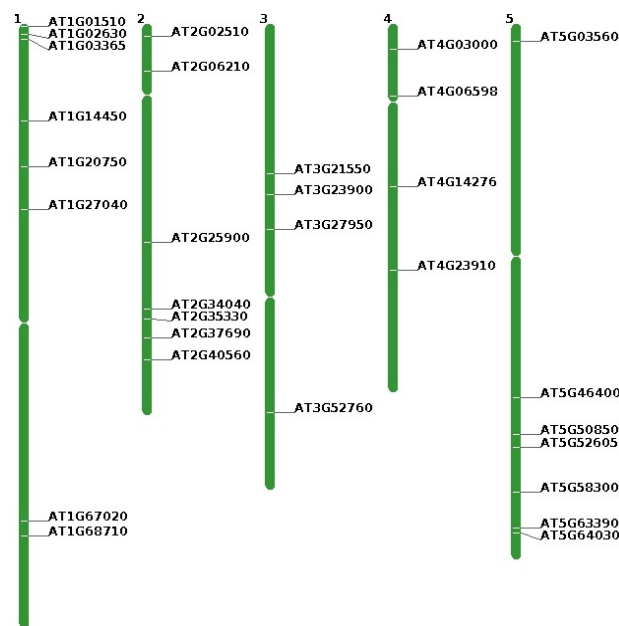
Hasil anotasi ontologi gen dan kategori fungsional berdasarkan identifikasi lokus dapat dilihat pada Gambar 4. Berdasarkan hasil anotasi sekuen, total 41 ontologi gen dapat ditentukan dan terdistribusi kepada tiga kategori: fungsi molekuler (60), proses biologi (93) dan komponen seluler (103). Jumlah total ketiga kategori tersebut lebih banyak dari jumlah total ontologi gennya (41) karena beberapa ontologi gen yang dipetakan berjumlah satu atau lebih dari masing-masing kategori.

Berdasarkan klasifikasi fungsional ontologi gen diperoleh bahwa kategori komponen seluler (Gambar 4A) membentuk 14 kelompok fungsional yang didominasi oleh komponen intraseluler (18,4%) dibandingkan ekstraseluler (3,9%). Beberapa gen juga diketahui berada pada bagian kloroplas dan mitokondria sekitar 6-7%. Sebagian kecil gen berada pada lokasi plastid sitosol dan retikulum endoplasma berkisar antara 1-2% saja.

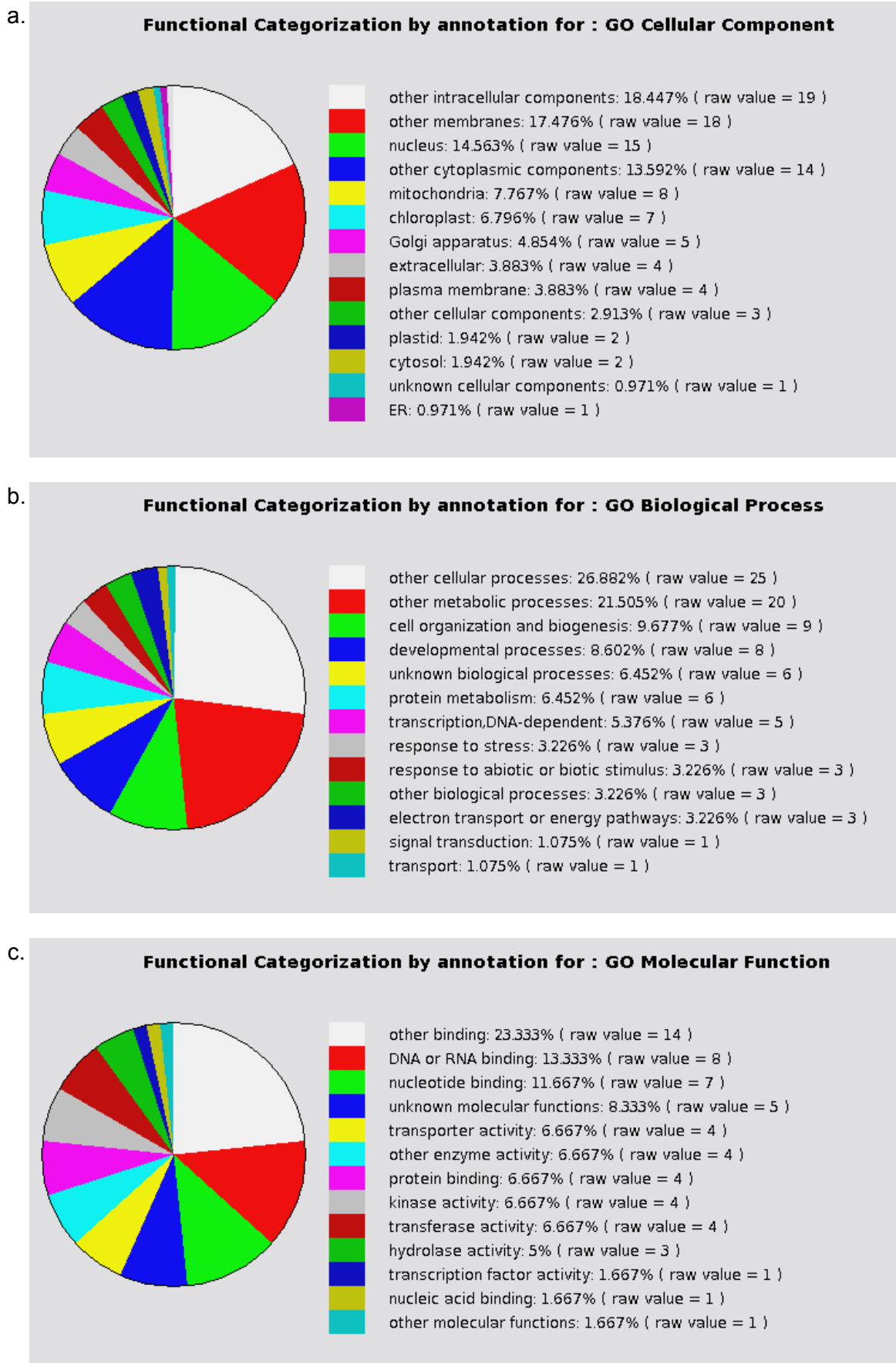
Kategori proses biologi (Gambar 4B) membentuk 13 kelompok fungsional dengan proses seluler (26,9%) mendominasi, diikuti proses metabolisme (21,5%). Gen yang berkaitan dengan proses perkembangan dan organisasi sel juga relatif banyak dijumpai. Hal ini dimungkinkan karena sekuen diperoleh dari daun muda yang masih mengalami pertumbuhan dan perkembangan. Gen-gen terkait respon terhadap cekaman biotik dan abiotik juga ditemukan, sebagaimana diketahui bahwa tanaman sagu memiliki ketahanan terhadap cekaman lingkungan seperti lingkungan dengan pH rendah (Harsanto 1986).

Pada kategori fungsi molekuler (Gambar 4C) membentuk 13 kelompok fungsional dimana gen-gen terkait dengan aktivitas faktor transkripsi dan pengikat asam nukleat relatif sedikit dijumpai. Sedangkan gen terkait dengan pengikat RNA dan DNA banyak dijumpai, lebih dari 10%. Analisis sekuen pada ketiga kategori menunjukkan adanya sekuen yang tidak ada kecocokan sama sekali dengan sekuen pada database NCBI (*unknown cellular components, unknown biological processes, unknown molecular functions*). Hal ini dapat diduga bahwa sekuen yang tidak diketahui atau belum terklasifikasikan tersebut merupakan sekuen dengan gen-gen yang hanya terdapat pada tanaman sagu saja.

Sudah diketahui umum bahwa gen tidak terekspresi sendiri tetapi ada gen-gen lain yang saling berinteraksi. Maka itu analisis interaksi antar gen juga dilakukan menggunakan GeneMANIA (www.genemania.org) berdasarkan bobot *co-expression* untuk mengetahui seberapa besar interaksi antar gen yang dianalisis. Metode pembobotan jaringan secara otomatis dipilih untuk menampilkan 20 gen terkait dan paling banyak 10 atribut terkait. Sebelumnya sebanyak 30 gen hasil anotasi ontologi gen pada TAIR dianalisis lebih lanjut menggunakan DAVID untuk mengetahui gen fungsional terkait.



Gambar 3. Penyebaran GO ID hasil ontologi gen pada kromosom *Arabidopsis thaliana*

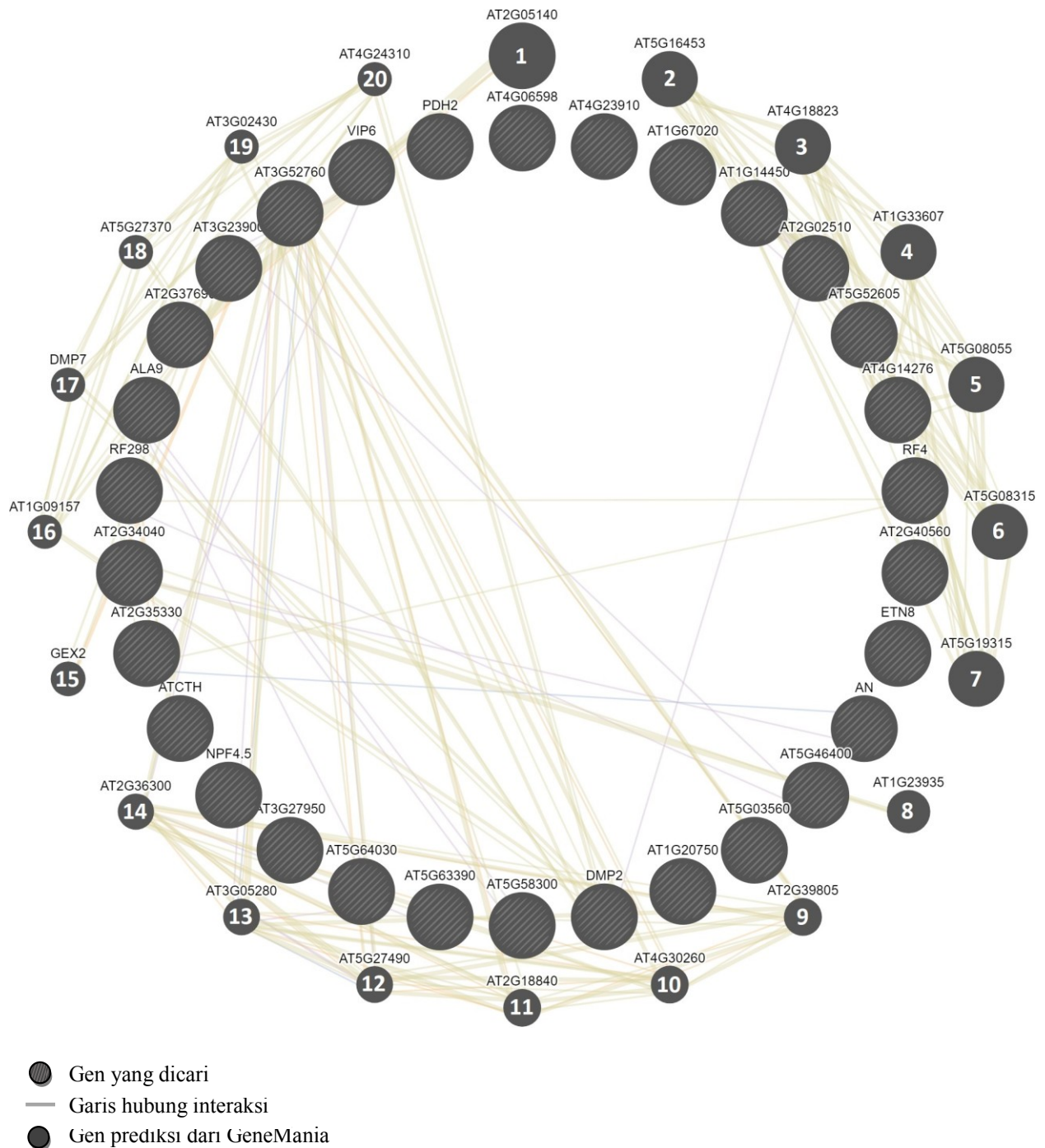


Gambar 4 Pengelompokan fungsional genom sagu pada 85 sekuen A: Komponen seluler, B: Proses biologi, C: Fungsi molekuler

aplikasi DAVID dipergunakan sebagai kumpulan gen untuk analisis lebih lanjut. Keempat kluster ini memiliki nilai *Enrichment* antara 0,32 hingga 1,56 (Lampiran 2). Analisis gen fungsional dari DAVID yang berjumlah 30 gen kemudian dianalisis pada GeneMania sehingga diperoleh gambaran hubungan interaksi berdasarkan *co-expression* masing-masing gen (Gambar 5). Dari 30 gen yang disubmit terdapat 20 gen

yang saling berinteraksi dari total 50 gen dengan jumlah 158 koneksi antar gen yang diprediksi oleh GeneMania. Persebaran interaksi antar gen hasil analisis GeneMania diketahui sebagian besar saling terkait pada domain protein 65,13%, koneksi yang berhasil diprediksi 19,83%, yang memiliki ekspresi bersama 14,47% dan sisanya 0,57% memiliki lokalisasi bersama.

Peranan interaksi gen dapat dilihat dari



Gambar 5. Hubungan interaksi antar gen terkait hasil analisis ontologi sekuen sagu dengan gen pada *Arabidopsis* melalui GeneMania. Ukuran node menunjukkan besarnya peranan suatu gen dalam hubungan interaksi antar gen dari sagu.

ukuran node, makin besar node menandakan besarnya peranan gen pada hubungan interaksi tersebut. Berdasarkan besar peranan 20 gen yang saling berinteraksi, gen phosphoribosyl-aminoimidazole carboxylase berada pada peringkat pertama dengan bentuk node yang paling besar. Gen ini diketahui berperan dalam biosintesis purin (<https://www.ncbi.nlm.nih.gov/gene/10606>). Peringkat ke-2 hingga ke-7 diketahui merupakan kelompok defensin-like protein, peringkat ke-9 hingga ke-14 diketahui merupakan kelompok Yip1 protein, peringkat ke-16 hingga ke-20 merupakan kelompok DUF679: protein yang belum diketahui fungsinya. Sedangkan peringkat ke-8 merupakan gen Apoptosis Inhibitory 5 yang berkaitan dengan kematian sel dan berperan sebagai penghambat endogen Caspase-2 dalam mengatur autophagy, stabilitas genomik dan penuaan sel (Imre et al. 2017). GEX2 pada peringkat 15 berperan pada faktor perkembangan reproduksi (Mori et al. 2014).

KESIMPULAN

Analisis bioinformatika sekuen sagu berbasis web mampu memberikan informasi homologi dan anotasi gen potensial. Sebanyak 33 sekuen sagu memiliki homologi pada database NCBI. Gen yang paling banyak ditemukan pada sekuen sagu dengan rataan kemiripan $\geq 90\%$ adalah *glutamate decarboxylase* dan *serine threonine-kinase HT1-like* dengan jumlah *hit* 10.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada BPPT atas beasiswa pada program pascasarjana IPB 2014.

DAFTAR PUSTAKA

Abbas B (2006) Keragaman genetik tanaman sagu di Indonesia berdasarkan marka molekuler genom kloroplas dan genom inti [disertasi]. Bogor (ID): Institut Pertanian Bogor

Abbas B, Bintoro MH, Sudarsono, Surahman M, Ehara H (2009) Genetic relationship of sago palm (*Metroxylon sagu* Rottb.)

in Indonesia based on RAPD markers. Biodiversitas. 10: 168–174. doi: 10.13057/biodiv/d100402

Abbas B, Yanuarius R, Bintoro MH, Sudarsono, Surahman M, Ehara H (2010) Genetic diversity of sago palm in Indonesia based on chloroplast DNA (cpDNA) Markers. Biodiversitas. 11: 112–117. doi: 10.13057/biodiv/d110302

Bintoro MH (2011) Progress of sago research in Indonesia. The 10th International Sago Symposium. Institut Pertanian Bogor Press, Bogor

Budiani A, Putranto RA, Minarsi H, Riyadi I, Sumaryono, Abbas B (2015) Expression and cloning of gene encoding ADP-Glucose Phosphorylase from sago palm (*Metroxylon sagu* Rottb). Menara Perkebunan 83: 76-85

Dirjenbun (2016) Statistik Perkebunan Indonesia 2015-2017: Sagu. Direktorat Jenderal Perkebunan, Kementerian Pertanian, Jakarta

Ehara H, Naito H, Mishima T, Toyoda Y, Mizota C, Susanto S, Bintoro MH, Pasolon YB, Abbas B, Suwignyo RA, Munandar (2016) Comparison of growth parameters and yield components of sago palms grown in the islands in Southeast Asia and Melanesia. Seminar Ilmiah dan Lokakarya Nasional Sagu 9-10 November 2016, Bogor, Indonesia

Harsanto PB (1986) Budidaya dan Pengelolaan Sagu. Kanisius, Jakarta

Hashimoto M, Negi J, Young J, Israelsson M, Schroeder JI, Iba K (2006) Arabidopsis HT1 kinase controls stomatal movements in response to CO₂. Nat Cell Biol 8: 391-397. doi:10.1038/ncb1387

Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4: 44-57. doi: 10.1038/nprot.2008.211

Imre G, Berthelet J, Heering J, Kehrlöesser S, Melzer IM, Lee BI, Thiede B, Dötsch V, Rajalingam (2017) Apoptosis inhibitor 5 is an endogenous inhibitor of caspase-2. EMBO Rep 18: 733-744. doi:10.15252/embr.201643744

- Jamel B, Hussain MH, Salleh MA, Busri N (2011) Isolation and characterization of the GA 20-oxidase cDNA from sago palm (*Metroxylon sago* Rottb.). *AsPac J Mol Biol Biotechnol* 19: 83-93
- Johnson DV (2010) Tropical Palms. Food and Agriculture Organization (FAO) of The United Nations, Rome
- Karim AA, Tie APL, Manan DMA, Zaidul ISM (2008) Starch from the sago (*Metroxylon sago*) palm tree—properties, prospects, and challenges as a new industrial source for food and other uses. *Comprehensive Reviews in Food Science and Food Safety* 7: 215-228. doi: 10.1111/j.1541-4337.2008.00042.x
- Karyanti, Sigit Y, Tajuddin T, Erwinda, Minaldi, Haska N (2016) Penanganan anakan muda pada kultur ex vitro untuk menghasilkan bibit sago (*Metroxylon sago* Rottb.) siap tanam. *J Bioteknol Biosains Indones* 3: 13-19. doi: 10.29122/jbbi.v3i1.20
- Kjær A, Barfod AS, Asmussen CB, Seberg O (2004) Investigation of genetic and morphological variation in the sago palm (*Metroxylon sago*; *Arecaceae*) in Papua New Guinea. *Ann Bot* 94: 109–117. doi: 10.1093/aob/mch112
- Molina-Rueda JJ, Garrido-Aranda A, Gallardo F (2015) Glutamate decarboxylase. *In: Amino Acids in Higher Plants*. Chapter 7: 129. Spain. doi: 10.1079/9781780642635.0129
- Mori T, Iqawa T, Tamiya G, Miyaqishima SY, Berger F (2014) Gamete attachment requires GEX2 for successful fertilization in *Arabidopsis*. *Curr Biol* 24: 170-175. doi: 10.1016/j.cub.2013.11.030
- Narita V, Arum AL, Isnaeni S, Fawzya NY (2012) Analisis bioinformatika berbasis web untuk eksplorasi enzim kitosanase berdasarkan kemiripan sekuens. *J Al-Azhar Indones* 1: 197-203
- Novero AU, Mabras MB, Esteban HJ (2012) Epigenetic inheritance of spine formation in sago palm (*Metroxylon sago* Rottb.). *Plant Omics J* 5: 559-566
- Raza K, Ahmad S (2017) Recent advancement in next generation sequencing techniques and its computational analysis. Tersedia pada: <https://arxiv.org/ftp/arxiv/papers/1606/1606.05254.pdf>
- Riyadi I, Efendi D, Purwoko BS, Santoso D (2016) Embriogenesis somatik tidak langsung pada tanaman sago (*Metroxylon sago* Rottb.) menggunakan sistem kultur suspensi, perendaman sesaat, dan media padat. *J AgroBiogen* 12: 37–44
- Sumaryono (2007) Tanaman sago sebagai sumber energi alternatif. *Warta Penelitian dan Pengembangan Pertanian* 29: 4
- Warde-Farley D, Donaldson SL, Gomes O, Zuberi K, Badrawi R, Chao P, Franz M, Grouios C, Kazi F, Lopes CT, Maitland A, Mostafavi S, Montojo J, Shao Q, Wright G, Bader GD, Morris Q (2010) The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res* 38: W214 - W220. doi: 10.1093/nar/gkq537
- Wee CC, Roslan HA (2012) Expressed sequence tags (ESTs) from young leaves of *Metroxylon sago*. *3 Biotech*. 2: 211–218. doi: 10.1007/s13205-012-0048-6
- Yamamoto Y (2010) Starch Productivity In Sago Palm. *In: Sago Palm As The Resource Crop In The 21st Century*. Pp 218-232. The society of sago palm studies ed. Kyoto University press, Kyoto