

PENGUKURAN KEMIRIPAN SEMANTIK BERBASIS GRAPH PADA *GENE ONTOLOGY*

Dinar Munggaran Akhmad

Program Studi Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas
Pakuan, Jln. Pakuan PO Box 452 Bogor 16143, Indonesia

E-mail: dinar.munggaran@unpak.ac.id

ABSTRAK

Ontologi mendukung suatu sistem Knowledge Management dan membuka kemungkinan untuk berpindah dari pandangan berorientasi dokumen ke arah pengetahuan yang saling terkait dan dapat dimanfaatkan kembali secara lebih fleksibel dan dinamis. Salah satu dokumen ontology yang sangat berperan dalam dunia bioinformatika adalah Gene Ontology. Gene Ontology dibangun berbasis Graph, memuat banyak term/istilah tentang Gen pada makhluk hidup. Gen dapat melakukan mutasi dan hal ini menyebabkan resistensi terhadap segala jenis obat. Gen yang diketahui menyebabkan penyakit malaria resisten terhadap obat antimalarial yaitu gen dhfr dan dhps. Penelusuran kemungkinan gen lain yang resisten dapat dilakukan dengan mengetahui sifat-sifat gen yang berhubungan dengan resistensi tersebut. Salah satu cara adalah dengan menghitung kemiripannya secara semantik melalui pendekatan path length metode Wang. Hasil penelitian menunjukkan bahwa rentang nilai kemiripan kedua gen tersebut adalah 0 – 1. Nilai kemiripan diuji pada salah satu komponen dalam Gene Ontology yaitu pada Molecular Function dengan nilai kemiripan sebesar 0.35 karena terhubung oleh 2 node yang sama yaitu catalytic activity. Dengan demikian penelitian ini diharapkan dapat mendeteksi gen lain yang terindikasi resistensi sebelum penelitian lebih lanjut secara molekuler.

Kata kunci—Graph, Kemiripan, Ontology, Semantik, Wang

ABSTRACT

Ontology supports a Knowledge Management system and opens up the possibility to move from a document-oriented view towards interrelated and reusable knowledge in a more flexible and dynamic manner. One of the ontology documents that plays a very important role in the world of bioinformatics is Gene Ontology. Gene Ontology is built based on Graph, contains many terms / terms about genes in living things. Genes can mutate and this causes resistance to all types of drugs. Genes known to cause malaria resistance to antimalarial drugs are the dhfr and dhps genes. Tracing the possibility of other genes that are resistant can be done by knowing the properties of the genes associated with this resistance. One way is to calculate the semantic similarity through the Wang method's path length approach. The results showed that the similarity value range of the two genes was 0 - 1. The similarity value was tested on one of the components in Gene Ontology, namely the Molecular Function with a similarity value of 0.35 because it is connected by the same 2 nodes, namely catalytic activity. Thus, this study is expected to be able to detect other genes that indicate resistance before further molecular research.

Keywords— Graph, Ontology, Semantic, Similarity, Wang

1. PENDAHULUAN

Ontologi merupakan cara merepresentasikan pengetahuan tentang makna objek, properti dari suatu objek, serta relasi objek tersebut yang mungkin terjadi pada domain pengetahuan [1]. Ontologi mendukung suatu sistem *Knowledge Management* dan membuka kemungkinan untuk berpindah dari pandangan berorientasi dokumen ke arah pengetahuan yang saling terkait, dapat dikombinasikan, serta dapat dimanfaatkan kembali secara lebih fleksibel dan dinamis. Salah satu dokumen ontology yang sangat berperan dalam dunia bioinformatika adalah *Gene Ontology*. *Gene Ontology* (GO) merupakan layanan informasi mengenai gen yang dikembangkan oleh *Gene Ontology Consortium*. Di dalam GO terdiri tiga komponen ontologi yaitu *molecular function*, *biological process*, *cellular component* [2]. GO menjadi lebih penting dalam penelitian biomedis, metode komputasi yang sering dibutuhkan dalam menjelajahi GO yaitu untuk menghitung kesamaan semantik (*semantic similarity*) antara gen maupun protein [3]. Pengukuran *semantic similarity* memiliki peran penting dalam beberapa *task natural language processing*. Pengukuran ini dilatarbelakangi oleh suatu masalah di mana saat ini komputer belum dapat menyamakan persepsi manusia terkait penilaian kesamaan antar kata. Untuk pengukuran kemiripan gen ataupun protein banyak dilakukan terhadap berbagai spesies diantaranya pada bakteri [4] manusia, serta malaria [5].

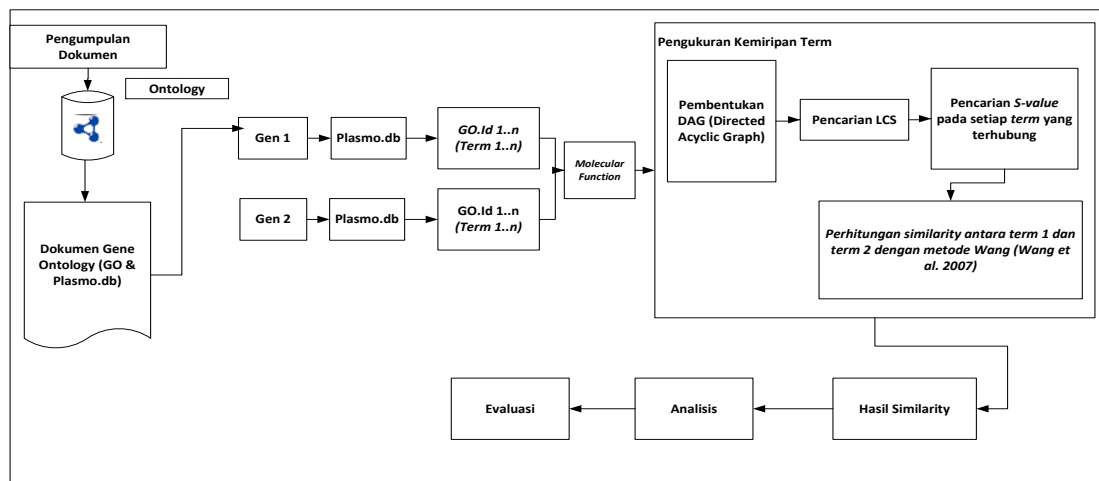
Malaria adalah penyakit yang disebabkan oleh parasit yang dipindahkan dari seorang penderita ke orang lain melalui gigitan nyamuk *Anopheles* yang telah terinfeksi oleh *Plasmodium*. Ada 2 jenis *antimalarial* yang biasa digunakan untuk pengobatan malaria *falciparum* yaitu *chloroquine* dan *sulfadoxine-pyrimethamin* [6]. [7] Telah mengidentifikasi beberapa mutasi yang terkait dengan resistensi terhadap *antimalarial sulfadoxine-pyrimethamin* pada 2 gen *plasmodium falciparum* yaitu *Dihydrofolate reductase (Dhfr)* dan *dihydropteroat synthase (Dhps)*. *Polymorphism* yang menyebabkan 2 gen resisten terhadap obat *antimalaria sulfadoxine-pyrimethamin* memungkinkan adanya kemiripan antar gen salah satunya dalam fungsi molekul.

Untuk mengukur kemiripan antar gen di atas dapat diterapkan dengan menggunakan perhitungan kemiripan antar konsep dalam ontologi. Ada dua jenis perhitungan kemiripan semantik dalam data *Ontology* yaitu berbasis *Information Content* [8,9] dan Berbasis *Path* [10]. Penelitian sebelumnya pernah dilakukan oleh [11] mengukur kemiripan semantic antar 2 gen resistensi yaitu gen *pfmdr1* dan *pfprt* terhadap *antimalarial chloroquine* pada 3 komponen *Gene Ontology*. [12] Memprediksi fungsi-fungsi protein dari spesies manusia dengan memanfaatkan 3 kelompok dari GO yaitu *Molecular function*, *Cellular component* dan *Biological processes*. [13] Melakukan penelitian memprediksi interaksi protein-protein dalam ragi dengan menghitung kemiripan semantik berbasis *path Wu* antar 2 istilah dalam dokumen GO. [14] Melakukan analisis interaksi antar protein dengan menghitung kemiripan semantik antar konsep dalam dokumen GO menggunakan *IC* dan *Path*. Untuk data malaria, [15] melakukan penelitian memprediksi fungsi-fungsi pada gen *Plasmodium falciparum* dengan menghitung nilai kemiripan semantik berbasis *path* untuk meng *cluster* data gen *Plasmodium falciparum* serta melakukan *enrichment analysis*. Sedangkan penelitian mengenai resistensi *antimalarial* pernah dilakukan oleh [7].

Berdasarkan pemaparan di atas pengukuran kemiripan semantik banyak dilakukan untuk memprediksi fungsi dari gen-gen berbagai *species* salah satunya malaria. Maka pada penelitian ini akan melakukan pengukuran *Semantic Similarity* antar gen penyebab resistensi *antimalarial sulfadoxine-pyrimethamin* berdasarkan dokumen GO (*Gene Ontology*) menggunakan metode Wang berbasis *path*. Diharapkan penelitian ini dapat memudahkan untuk mengidentifikasi gen-gen lain yang terindikasi resisten.

2. METODE PENELITIAN

Tahapan-tahapan yang dilakukan pada penelitian ini mengacu pada Gambar 1. Tahapan dimulai dari pengumpulan dokumen, pemasangan *query* gen hingga proses evaluasi.



Gambar 1 Metodologi Penelitian

Data yang dibutuhkan dalam penelitian ini adalah dokumen *gene ontology*. *Gene ontology* mencakup 1 komponen yaitu *molecular function*. *Molecular function* (Fungsi molekul) menggambarkan aktifitas yang terjadi pada tingkat molekuler seperti “aktifitas katalitik” atau “kegiatan mengikat”. *Cellular component* menggambarkan komponen dari sel yang merupakan bagian dari objek yang lebih besar, seperti struktur anatomi (mis kasar retikulum endoplasma atau inti) atau kelompok produk gen (misalnya ribosom, proteasome atau dimer protein)[2].

Data selanjutnya yang dibutuhkan yaitu *database* gen spesies *plasmodium falciparum*. Pasangan gen 1 dan gen 2 yang dimaksud yaitu *dhfr* dan *dhps*, dengan melihat pada *database* diperoleh GO id yang merepresentasikan masing-masing gen dengan masing-masing komponen dalam *gene ontology*. Kedua gen merujuk pada penelitian sebelumnya yang dilakukan oleh [7].

Setelah didapatkan GO id pada masing-masing komponen *gene ontology* maka langkah selanjutnya adalah pembentukan *Directed Acyclic Graph* (DAG) sesuai dengan metode dari [10]. Jika DAG sudah terbentuk maka tahap selanjutnya adalah menghitung kemiripan antar 2 gen secara semantik.

Gene Ontology disusun sebagai grafik acyclic yang diarahkan (DAG) di mana setiap istilah yang didefinisikan memiliki hubungan dengan satu atau lebih hal lain dalam domain yang sama dan kadang-kadang ke domain yang lain dengan relasi “*is-a*” dan “*part-of*” [2]. Garis titik-titik menunjukkan hubungan “*part-of*” dan garis tebal menunjukkan hubungan “*is-a*”. Untuk mengukur nilai kesamaan semantik (*semantic similarity*) pada istilah GO (GO *term/node*), pertama kita harus merubah semantiknya terlebih dahulu ke dalam format numerik, Untuk itu [10] mendefinisikan kontribusi tersebut sebagai S-Value dari istilah GO *t* yang terhubung dengan istilah A. Untuk banyak istilah *t* dalam DAG_A (A, T_A, E_A), S-Value yang terhubung dengan istilah A, SA(*t*) didefinisikan sebagai:

$$\begin{cases} S_A(A) = 1 \\ S_A(t) = \max\{W_e + S_A(t') \mid t' \in \text{childrenof}(t)\} \text{ if } t \neq A \end{cases} \quad (1)$$

Setelah menentukan S-Values dari semua istilah dalam DAG_A , kita menghitung nilai semantik dari istilah A pada GO, $SV(A)$ sebagai:

$$SV(A) = \sum_{t \in T_A} S_A(t) \quad (2)$$

Kita asumsikan bahwa faktor kontribusi semantik untuk relasi “*is-a*” dan “*part-of*” adalah 0.8 dan 0.6 [10] kemudian kita gunakan persamaan (1) dan (2) untuk menghitung S-value pada istilah GO pada setiap DAG.

Diberikan $DAG_A = (A, T_A, E_A)$ dan $DAG_B = (B, T_B, E_B)$ untuk masing-masing istilah A dan B pada GO, sehingga *semantic similarity* antara 2 istilah GO, $S_{GO}(A, B)$, didefinisikan pada persamaan 3.

$$S_{GO}(A, B) = \frac{\sum_{t \in T_A \cap T_B} (S_A(t) + S_B(t))}{SV(A) + SV(B)} \quad (3)$$

Di mana $S_A(t)$ adalah nilai dari S-value dari istilah t GO yang terhubung ke istilah A dan $S_B(t)$ adalah nilai S-value dari istilah t yang terhubung dengan istilah B.

Biasanya satu gen dijelaskan oleh banyak istilah GO. Kemiripan semantik antara istilah GO dan satu set istilah GO yaitu $GO = \{go_1, go_2, \dots, go_k\}$, $Sim(go, GO)$, didefinisikan sebagai kemiripan semantik maksimum antara istilah go dan banyak istilah dalam set/kumpulan istilah GO dengan persamaan 4.

$$Sim(go, GO) = \max_{1 \leq i \leq k} (S_{GO}(go, go_i)) \quad (4)$$

Oleh karena itu, mengingat dua gen G_1 dan G_2 dijelaskan dengan kumpulan istilah GO, $GO_1 = \{go_{11}, go_{12}, \dots, go_{1m}\}$ dan $GO_2 = \{go_{21}, go_{22}, \dots, go_{2n}\}$ maka [10] mendefinisikan kemiripan fungsional ke dalam persamaan 5.

$$Sim(G_1, G_2) = \frac{\sum_{1 \leq i \leq m} Sim(go_{1i}, GO_2) + \sum_{1 \leq j \leq n} Sim(go_{2j}, GO_1)}{m + n} \quad (5)$$

Hasil kemiripan yang diperoleh akan dianalisis serta dievaluasi dengan melihat studi literatur terkait.

3. HASIL DAN PEMBAHASAN

Dalam penelitian ini, ada 2 gen dari species penyakit malaria yang menyebabkan resisten terhadap obat antimalarial *sulfadoxine-pyrimethamin* yaitu *Dihydrofolate reductase (Dhfr)* dan *dihydropteroat synthase (Dhps)* yang merupakan bagian dari penelitian [7]. Dua gen tersebut akan dihitung kemiripannya secara semantik berbasis ontologi dengan metode *graph*. Berdasarkan pemaparan sebelumnya pada metodologi penelitian, satu gen biasanya direpresentasikan oleh banyak *term* / istilah GO dalam *molecular function*. Tabel 1 berisi informasi GO *terms* yang menjelaskan gen *Dhfr* dan *Dhps* pada komponen GO *molecular function*.

Tabel 1 Informasi untuk definisi *node* / *term* yang menganotasikan gen *Dhfr* dan *Dhps* pada komponen *Molecular Function*

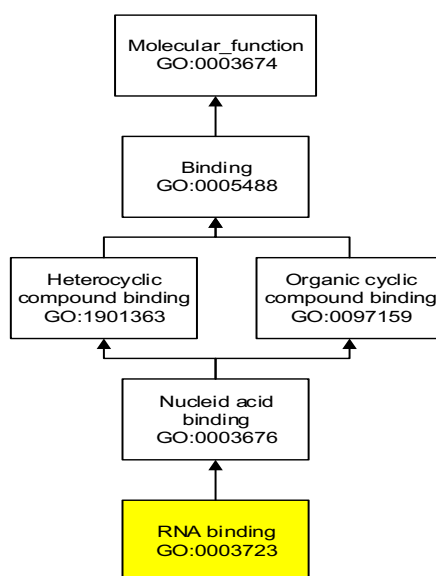
GO ID	Definisi
<i>Dhfr</i>	
GO:0003723 (RNA binding)	<i>Interacting selectively and non-covalently with an RNA molecule or a portion thereof.</i>
GO:0004146	<i>Catalysis of the reaction: 5,6,7,8-tetrahydrofolate + NADP+ = 7,8-</i>

(<i>Dihydrofolate reductase activity</i>)	<i>dihydrofolate + NADPH + H⁺.</i>
GO:0004799 (<i>Thymidylate synthase activity</i>)	<i>Catalysis of the reaction: 5,10-methylenetetrahydrofolate + dUMP = 7,8-dihydrofolate + thymidylate.</i>
<i>Dhps</i>	
GO:004156 (<i>Dihydropteroate synthase activity</i>)	<i>Catalysis of the reaction: 2-amino-4-hydroxy-6-hydroxymethyl-7,8-dihydropteridine diphosphate + 4-aminobenzoate = diphosphate + dihydropteroate.</i>

Pembentukan DAG

Pembentukan DAG GO:0003723

Berdasarkan metode [10] setelah mengetahui pasangan gen yang akan dihitung kemiripannya maka tahap selanjutnya adalah membentuk *Directed Acyclic Graph* (DAG). Mengacu pada Tabel 1 yaitu akan ada 4 DAG yang terbentuk dari gen *Dhfr* dan *Dhps* yaitu untuk GO id GO:0003723 yang terlihat pada Gambar 2, GO:0004146 pada Gambar 3, GO:0004799 pada Gambar 4 dan GO:004156 pada Gambar 5.



Gambar 2 DAG untuk node / term GO RNA binding:0003723

Istilah/*term* yang merepresentasikan 2 gen malaria yang resisten terhadap obat antimalarial yaitu *DHPS* dan *DHFR* tidak banyak sehingga bentuk DAG pun lebih sederhana.

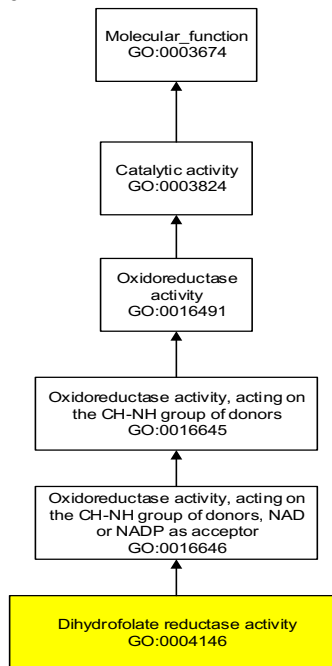
Menghitung S-values Setiap Node GO

Setelah DAG terbentuk maka tahap selanjutnya menghitung *S-values* pada masing-masing *node / term* dalam DAG dengan menggunakan Persamaan 5 dan 6. Tabel 2 memuat nilai *S-values* untuk GO:0003723 .

Tabel 1 S-values untuk node / term GO dalam DAG RNA Binding:0003723

GO terms	GO:0003723	GO:0003676	GO:1901363	GO:0097159
S-values	1.0	0.8	0.64	0.64
GO terms	GO:0005488	GO:0003674		
S-values	0.512	0.41		

Pembentukan DAG GO:0004146



Gambar 3. DAG untuk node / term GO Dihydrofolate reductase activity:0004146

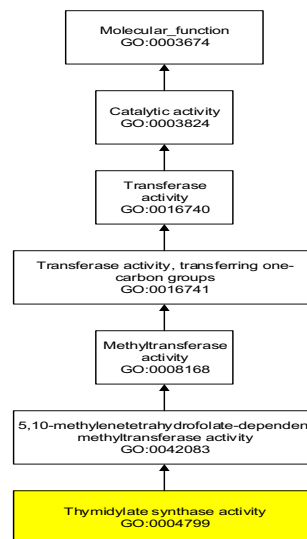
Menghitung S-values Setiap Node GO

Setelah DAG terbentuk maka tahap selanjutnya menghitung *S-values* pada masing-masing node / term dalam DAG dengan menggunakan Persamaan 5 dan 6. Tabel 3 memuat nilai *S-values* untuk DAG GO:0004146 .

Tabel 3 *S-values* untuk node / term GO dalam DAG Dihydrofolate reductase activity:0004146

GO terms	GO:0004146	GO:0016646	GO:0016645	GO:0016491
<i>S-values</i>	1.0	0.8	0.64	0.512
GO terms	GO:0003824	GO:0003674		
<i>S-values</i>	0.409	0.327		

Pembentukan DAG GO:0004799



Gambar 4 DAG untuk node GO Thymidilate synthase activity:0004799

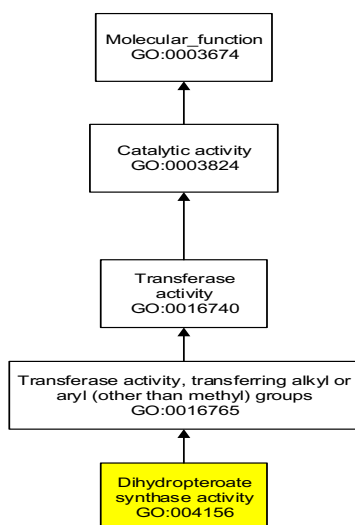
Menghitung S-values Setiap Node GO

Setelah DAG terbentuk maka tahap selanjutnya menghitung S-values pada masing-masing node / term dalam DAG dengan menggunakan Persamaan 5 dan 6. Tabel 19 memuat nilai S-values untuk DAG GO:0004799 .

Tabel 4 S-values untuk node / term GO dalam DAG Thymidilate synthase activity:0004799

GO terms	GO:0004799	GO:0042083	GO:0008168	GO:0016741
S-values	1.0	0.8	0.64	0.512
GO terms	GO:0016740	GO:0003824	GO:0003674	
S-values	0.409	0.327	0.261	

Pembentukan DAG GO:004156



Gambar 5 DAG untuk node / term GO Dihydropteroate synthase activity:004156

Menghitung S-values Setiap Node GO

Setelah DAG terbentuk maka tahap selanjutnya menghitung S-values pada masing-masing node / term dalam DAG dengan menggunakan Persamaan 5 dan 6. Tabel 20 memuat nilai S-values untuk GO:004156.

Tabel 5 S-values untuk node / term GO dalam DAG Dihydropteroate synthase activity:004156

GO terms	GO:004156	GO:0016765	GO:0016740	GO:0003824
S-values	1.0	0.8	0.64	0.512
GO terms	GO:0003674			
S-values	0.41			

Perhitungan Kemiripan Node pada DAG / Kemiripan Gen

Sama seperti pasangan gen yang pertama, setelah pembentukan DAG dan perhitungan S-values, maka tahap selanjutnya adalah menghitung nilai semantic similarity antar node / term dengan menggunakan Persamaan 7 dibantu dengan fungsi yang telah disediakan dalam Bahasa Pemrograman R. Kemiripan dihitung pada satu komponen GO (molecular function). Hasil kemiripan antara term gen Dhfr dan Dhps dengan menggunakan Persamaan 3 pada komponen GO Molecular Function sebagaimana terlampir pada Tabel 6.

Tabel 2 Kemiripan antara node / term yang menganotasikan gen *Dhfr* dan *Dhps* pada komponen *Molecular Function*

<i>dhps</i>		
	GO_ID (MF)	
<i>dhfr</i>	GO:0003723	GO:004156
	GO:0004146	0.111
	GO:0004799	0.235
		0.35

Berdasarkan Tabel 21 dengan menggunakan Persamaan 9 maka diperoleh hasil kemiripan antar gen *Dhfr* dan *Dhps* dalam komponen *Molecular Function* adalah sebesar 0.35

Analisis

Kemiripan dihitung berdasarkan kedekatan antar *node* dalam *Graph* dalam hal ini *Gene Ontology* dengan melihat garis yang menghubungkan node tersebut (*is-a* dan *part-of*), maka berdasarkan Tabel 6 nilai kemiripan terbesar diperoleh antara GO:0004799 dan GO:004156 sebesar 0.35. Hasil itu diperoleh karena terhubung oleh 2 *node term* yang sama yaitu merupakan bagian dari atau *part-of node Transferase activity* dan *Catalytic Activity*.

EVALUASI

Mekanisme resistensi antimikroba secara garis besar dibagi dua yaitu mekanisme *intrinsic/innate* dan perolehan/*acquire* [16]. Mekanisme resistensi Perolehan/*acquire* salah satunya dapat melalui mutasi.

Berdasarkan hasil penelitian yang diperoleh maka evaluasi ini mengacu pada penuturan literatur di atas mengenai mekanisme resistensi yang dituangkan dalam bentuk tabel. Evaluasi dilakukan dengan menilai kesesuaian antara informasi mekanisme resistensi literatur dengan GO id yang diperoleh terhadap setiap *gen*. Tabel 7 memuat hasil evaluasi penelitian yang telah dilakukan.

Tabel 7 Evaluasi Hasil Penelitian GO dan Studi Literatur

Mekanisme Resistensi	Informasi <i>Gene Ontology</i>		
	Kesesuaian	GO id	Komponen
Mutasi	✓	GO:0004799 (<i>Thymidylate synthase activity</i>)	<i>Catalysis of the reaction: 5,10-methylenetetrahydrofolate + dUMP = 7,8-dihydrofolate + thymidylate.</i> Mutasi pun mempengaruhi perolehan <i>Thymidylate</i> . * <i>Molecular Function</i> *

Berdasarkan Tabel 7 dapat terlihat kesesuaian antara informasi mengenai mekanisme resistensi dalam literatur dengan informasi GO id dari *Gene Ontology* pada ketiga komponennya. Maka dengan ini menunjukkan bahwa *Gene Ontology* bisa digunakan untuk memperoleh informasi awal terkait resistensi untuk *gen-gen* yang ada pada penyakit malaria khususnya.

4. KESIMPULAN

Penelitian dengan teknik molekuler memiliki banyak kendala diantaranya pengadaan alat dan bahan memerlukan biaya yang sangat tinggi dan waktu pengerjaan memerlukan waktu yang relatif lama. Hadirnya teknologi *knowledge* dalam bentuk ontology dapat membantu peneliti ataupun orang awam untuk memperoleh informasi awal mengenai suatu gen melalui *Gene Ontology*.

Berdasarkan hasil evaluasi, *Gene Ontology* dapat digunakan untuk mendapatkan informasi awal terkait resistensi pada gen-gen penyakit malaria khususnya. Sehingga dapat digunakan pula untuk mendeteksi gen lain yang terindikasi resistensi sebelum dilakukan penelitian secara molekuler. Penelitian diharapkan dapat dilakukan terhadap berbagai gen yang mempengaruhi resistansi penyakit lain. Penelitian dapat dilakukan dengan melakukan penggabungan metode antara *graph* dan *Information Content* untuk mencapai hasil pengukuran yang lebih akurat.

DAFTAR PUSTAKA

- [1] Chandrasekaran B, Josephon JR, Benjamins VR. 1999. What are ontologies, and why do we need them?. *IEEE Intelligent Systems*. 9:20-26.doi:1094-7167/99/\$10.00.
- [2] Consortium GO. 2015. Gene ontology consortium: going forward. *Nucleic acids research*. 43:D1049-D1056.
- [3] Guzzi, P. H., Mina, M., Guerra, C., & Cannataro, M. 2012. Semantic similarity analysis of protein data: assessment with biological features and issues. *Briefings in bioinformatics*, 13(5), 569-585.
- [4] Wu X, Zhu L, Guo J, Zhang DY, Lin K. 2006. Prediction of yeast protein-protein interaction network: insights from the gene ontology and annotations. *Nucleic Acids Research*. 34:2137-2150.Doi:10.1093/nar/gkl219.
- [5] Tedder PMR, Bradford JR, Needham CJ, McConkey GA, Bulpitt AJ, Westhead DR. 2010. Gene function prediction using semantic similarity clustering and enrichment analysis in the malaria parasite plasmodium falciparum. *Bioinformatics*. 26(19):2431-2437.doi:10.1093/bioinformatics/btq450.
- [6] Sidhu ABS, Verdier-Pinard D , Fidock DA. 2002. Chloroquine resistance in Plasmodium falciparum malaria parasites conferred by pfcrt mutations. *Science*. 298:210-213.
- [7] Syafruddin D, Asih PB, Casey GJ, Maguire J, Baird JK, Nagesha HS, Cowman AF , Reeder JC. 2005. Molecular epidemiology of Plasmodium falciparum resistance to antimalarial drugs in Indonesia. *The American journal of tropical medicine and hygiene*, p.2.
- [8] Lin D. 1998. *An information-theoretic definition of similarity*. In fifteenth International Conference on Machine Learning;98: 296-304.
- [9] Resnik P. 1999. *Semantic similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language*. *J Artif Intell Res(JAIR)*; 11: 95-130.
- [10] Wang JZ, Du Z, Payattakool R, Yu PS, Chen C-F. 2007. *A new method to measure the semantic similarity of GO terms*. *Bioinformatics*; 23(10): 1274-81.
- [11] Akhmad, D. M., Herdiyeni, Y., & Sudarnika, E. 2018. *Similarities of Antimalarial Resistance Genes in Plasmodium Falciparum Based on Ontology*. *Telkomnika*, 16(1), 415-423.
- [12] Jensen LJ, Gupta R, Staerfeldt H-H, Brunak S. 2003. *Prediction of human protein function according to Gene Ontology categories*. *Bioinformatics*; 19(5): 635-42.

- [13] Wu X, Zhu L, Guo J, Zhang D-Y, Lin K. 2006. *Prediction of yeast protein-protein interaction network: insights from the Gene Ontology and annotations*. *Nucleic acids research*; 34(7): 2137-50.
- [14] Guzzi PH, Mina M, Guerra C, Cannataro M. 2011. *Semantic similarity analysis of protein data: assessment with biological features and issues*. *Briefings in bioinformatics*.; 13(5): 569-85.
- [15] Tedder PM, Bradford JR, Needham CJ, McConkey GA, Bulpitt AJ, Westhead DR. 2010. *Gene function prediction using semantic similarity clustering and enrichment analysis in the malaria parasite Plasmodium falciparum*. *Bioinformatics*; 26(19): 2431-7.
- [16] Cavaleri J, Rankin D, Harbeck J, Sautter L, McCarter S, Sharp S, et al. 2005. *Manual of antimicrobial susceptibility testing*. American Society for Microbiology, Seattle, Washington. 12: 53-42.