



Identifikasi protein signifikan pada interaksi protein-protein penyakit Alzheimer menggunakan algoritme top-k representative skyline query

Identification of significant protein in protein-protein interaction of Alzheimer disease using top-k representative skyline query

Mohammad Romano Diansyah¹⁾, Wisnu Ananta Kusuma^{1,2*)}, Annisa¹⁾

¹⁾Departemen Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Pertanian Bogor
Jl. Meranti Wing 20 Level 5, Kampus IPB Darmaga, Bogor, Indonesia 16680

²⁾Pusat Studi Biofarmaka Tropika, Institut Pertanian Bogor
Jl. Taman Kencana No. 3, Bogor, Indonesia 16128

Cara sitasi: M. R. Diansyah, W. A. Kusuma, and A. Annisa, "Identifikasi protein signifikan pada interaksi protein-protein penyakit Alzheimer menggunakan algoritme top-k representative skyline query," *Jurnal Teknologi dan Sistem Komputer*, vol. 9, no. 3, pp. 126-132, 2021. doi: [10.14710/jtsiskom.2021.13985](https://doi.org/10.14710/jtsiskom.2021.13985), [Online].

Abstract – Alzheimer's disease is the most common neurodegenerative disease. This study aims to analyze protein-protein interaction (PPI) to provide a better understanding of multifactorial neurodegenerative diseases and can be used to find proteins that have a significant role in Alzheimer's disease. PPI data were obtained from experimental and computational predictions and analyzed using centrality measures. The Top-k RSP method was applied to find significant proteins in PPI networks using the dominance rule. The method was applied to the PPI data with the interaction sources from the experimental and experiment+prediction. The results indicate that APP and PSEN1 are significant proteins for Alzheimer's disease. This study also showed that both data sources (experiment+prediction) and the Top-k RSP algorithm proved useful for PPI analysis of Alzheimer's disease.

Keywords – Alzheimer; centrality measures; PPI; skyline query; Top-k RSP

Abstrak – Penyakit Alzheimer merupakan penyakit neurodegeneratif yang paling umum terjadi. Kajian ini bertujuan melakukan analisis protein-protein interaction (PPI) yang dapat memberikan pemahaman lebih baik terhadap penyakit neurodegeneratif dan bisa digunakan untuk menemukan protein yang memiliki peran signifikan pada penyakit Alzheimer. Data PPI diperoleh dari eksperimen dan prediksi komputasional. PPI dapat dianalisis menggunakan centrality measures. Metode Top-k RSP digunakan untuk menemukan protein signifikan dengan menggunakan aturan dominansi dan digunakan pada sumber data interaksi eksperimen dan eksperimen+prediksi. Hasil penelitian ini menunjukkan bahwa APP dan PSEN1 merupakan protein signifikan untuk penyakit Alzheimer. Selain itu,

kedua sumber data (eksperimen+prediksi) dan algoritme Top-k RSP terbukti dapat digunakan untuk analisis PPI dari penyakit Alzheimer.

Kata kunci – Alzheimer; centrality measures; PPI; skyline query; Top-k RSP

I. PENDAHULUAN

Alzheimer merupakan penyakit neurodegeneratif yang paling banyak menyebabkan terjadinya demensia yaitu diperkirakan sebanyak 50 juta penderita Demensia di dunia, 60-70% diakibatkan oleh Alzheimer [1], [2]. Menurut WHO pada 2013, penyakit yang tidak hanya diderita pasien di atas 65 tahun tersebut belum memiliki obat yang dapat menghentikan efek yang terjadi. Penelitian lanjutan masih sangat perlu dilakukan, seperti pencarian biomarker penyakit agar dapat mengukur resiko terjadi, melakukan diagnosis, dan untuk mengukur perkembangan dari penyakit tersebut [2]. Alzheimer ini merupakan penyakit neurodegeneratif yang paling umum terjadi [3]. Penyakit degeneratif dapat terjadi jika terdapat gangguan fungsional pada protein yang berperan dalam proses biologi dan metabolisme di dalam tubuh. Protein-protein, selain berperan dalam regulasi, juga menjadi media penting untuk pertumbuhan [4], [5] dan target dari senyawa obat [6].

Protein dalam menjalankan fungsinya saling berinteraksi satu sama lain membentuk suatu jejaring PPI. PPI merupakan alat yang baik untuk mengungkap fungsi protein, gen penyakit, dan hubungan antar penyakit [7], [8]. Selain itu, PPI dapat merepresentasikan kontak fisik (docking) molekul antarprotein yang terjadi di dalam makhluk hidup [9]. Analisis terhadap PPI dapat digunakan untuk memperoleh biomarker potensial yang dapat menjadi target terapi [6]. Kualitas hasil analisis PPI ditentukan antara lain oleh data PPI yang digunakan, fitur

^{*)} Penulis korespondensi (Wisnu Ananta Kusuma)
Email: ananta@apps.ipb.ac.id

yang digunakan untuk merepresentasikan PPI tersebut, dan algoritme untuk menganalisis PPI tersebut.

Saat ini data interaksi-interaksi antarprotein yang diperlukan untuk membangun PPI telah banyak tersimpan dalam basis data. Salah satu basis data yang menyimpan informasi interaksi protein adalah STRING (string-db.org) yang merupakan basis data dengan jumlah organisme (5.090) dan protein (24,6 juta) terbesar [10]. Data interaksi protein pada STRING didapatkan dari hasil eksperimen dan prediksi. Data interaksi eksperimen diperoleh melalui hasil uji lab sedangkan data interaksi prediksi diperoleh melalui metode *co-expression analysis*, *text-mining*, dan lainnya [11].

Penelitian terkait PPI sering terkendala oleh masalah *interactome coverage*. *Interactome* merupakan suatu set koleksi PPI yang dapat terjadi di dalam sel [12]-[14]. *Interactome coverage* rasio protein yang telah ditemukan dibandingkan dengan *interactome* yang ada dan dinyatakan dalam bentuk persentase (%). Stumpf et al. [15] memperkirakan *interactome* pada manusia terdapat sebanyak 650.000 PPI. *Interactome* jika dibandingkan dengan jumlah interaksi protein pada *Human Protein Reference Database* (HPRD) bulan Desember 2019 hanya terdapat sebanyak 41.327 PPI, sehingga hanya mencakup sebanyak 6,3% prediksi jumlah *interactome*. Dengan demikian, data interaksi antarprotein yang dari eksperimen masih sedikit. Oleh karena itu, pada [16]-[18] digunakan data eksperimen+prediksi untuk mencapai *interactome coverage* yang cukup dan kualitasnya meningkat.

Selanjutnya untuk dapat dilakukan analisis, data PPI harus ditransformasikan ke dalam fitur yang sesuai untuk merepresentasikan graf. Agapito et al. [19] memilih graf sebagai cara utama untuk merepresentasikan PPI dengan *node* sebagai protein dan *edge* sebagai interaksi yang terjadi. Mistry et al. [20] menyatakan *centrality measures* dapat digunakan sebagai fitur yang merepresentasikan pengaruh suatu *node* dalam jejaring interaksi. Transformasi data dapat dilakukan dari graf menjadi objek dengan atribut *centrality measures*.

Saat ini, terdapat banyak nilai perhitungan *centrality measures* dan sudah banyak digunakan untuk menentukan tingkat kepentingan suatu *node*. Banyak penelitian sebelumnya yang menggunakan *centrality measures* untuk mencari PPI yang penting pada jejaring interaksi, seperti pada [21]-[23]. *Centrality measures* yang ada saat ini memiliki karakteristik yang berbeda satu sama lain dalam suatu jejaring interaksi [24].

Faktor ketiga dari kualitas hasil analisis adalah algoritme yang digunakan. Telah banyak algoritme yang digunakan untuk analisis PPI ini, antara lain berbasis topologi dan pembelajaran mesin [12], sedangkan algoritme Skyline Query belum banyak diterapkan pada domain PPI. Borzsonyi et al. [25] menyatakan bahwa Skyline Query dapat memberikan solusi optimal untuk beragam preferensi beragam preferensi peneliti terhadap *centrality measures* berdasarkan dominansi data.

Namun, Skyline Query memiliki kekurangan, yaitu jika diterapkan pada data dengan jumlah atribut banyak

objek *skyline* yang dihasilkan dapat meningkat secara substansial, sehingga tidak relevan dan tidak menarik lagi bagi peneliti [26]. Untuk itu, diperlukan metode pengembangan Skyline Query untuk mencari data yang paling menarik dari hasil Skyline Query. Metode pengembangan yang dapat digunakan untuk mencari objek yang lebih penting dari hasil Skyline Query adalah Top-k *Representative Skyline Query* (RSP). Lin et al. [27] menyatakan jika Top-k RSP mengurutkan objek *skyline* dengan melihat jumlah data yang didominasi suatu objek, semakin banyak mendominasi maka semakin penting.

Penelitian ini bertujuan untuk mengidentifikasi protein signifikan yang terkait dengan penyakit Alzheimer menggunakan Top-k RSP. *Centrality measures* dasar digunakan sebagai fiturnya, seperti [28], yaitu *degree*, *betweenness*, *closeness*, *eigenvector*, dan *eccentricity centrality* dan ditambah dengan *centrality measures* yang diusulkan [29], yaitu *radiality* dan *bridging centrality*. Data yang digunakan merupakan data murni eksperimen dan data gabungan eksperimen+prediksi, sehingga dapat dibandingkan untuk melihat apakah kedua jenis data dapat digunakan untuk mencari protein signifikan pada penyakit degeneratif. Penggunaan data eksperimen yang digabungkan dengan prediksi (prediksi) diharapkan dapat mengatasi masalah *interactome coverage* pada interaksi protein.

III. METODE PENELITIAN

Penelitian ini terdiri atas 4 tahapan utama. Pertama yaitu pengumpulan data dari basis data OMIM dan STRING. Kedua adalah pra-proses data termasuk di dalamnya penghapusan data duplikat dan protein yang tidak terhubung. Ketiga merupakan proses pengolahan data menggunakan Top-k RSP. Tahapan terakhir adalah analisis hasil yang dilakukan dengan studi literatur. Tahapan penelitian dapat dilihat pada Gambar 1. Perangkat keras dan lingkungan operasi yang digunakan adalah komputer dengan spesifikasi Intel i7-4700HQ, NVIDIA GeForce GTX 970, RAM 8 GB, Harddisk penyimpanan 1 TB, dan sistem operasi Windows 8.1.

A. Pengumpulan data

Data diakuisisi dari basis data OMIM dan STRING pada tanggal 11 Maret 2020. Keduanya merupakan basis data publik. OMIM digunakan untuk mencari protein yang berasosiasi dengan Alzheimer dan STRING digunakan untuk mencari interaksi protein dari protein yang berasosiasi dengan Alzheimer. Tahapan pertama yaitu *query* pada omim digunakan prefix “+” untuk memberikan hasil lebih presisi. Kata kunci yang digunakan adalah “+Alzheimer +Disease”. Selanjutnya, interaksi protein dicari untuk masing-masing protein tersebut pada STRING.

Pengambilan data dilakukan dua kali, yaitu dengan sumber data murni eksperimen dan eksperimen+prediksi. Dengan demikian, untuk setiap protein berasosiasi dengan Alzheimer akan didapatkan dua set koleksi

interaksi protein berdasarkan sumber interaksinya. Seluruh informasi interaksi protein ini akan digabungkan menjadi dua *file* utama hasil pengumpulan data sesuai sumber interaksinya. Pada penelitian ini, proses pengumpulan data dilakukan dengan menggunakan program berbahasa Python 3.7.

B. Praproses data

Tahapan proses ini seluruhnya menggunakan aplikasi Cytoscape. Proses pertama yaitu penghapusan data PPI duplikat dan dilakukan visualisasi untuk menemukan apakah ada jejaring yang tidak terhubung dengan jejaring terbesar. Jejaring yang tidak terhubung dengan jejaring utama (terbesar) dihapuskan. Hal ini karena diasumsikan bahwa protein signifikan ditemukan dari jejaring yang memiliki kerapatan dan diameter terbesar. Selanjutnya, transformasi data dilakukan dari data PPI ke dalam *centrality measures* menggunakan *package* CentiScaPe 2.2 [29]. Hasil dari pra-proses adalah data protein dengan atribut *centrality measures* yang disimpan dalam berkas *comma separated value* (csv).

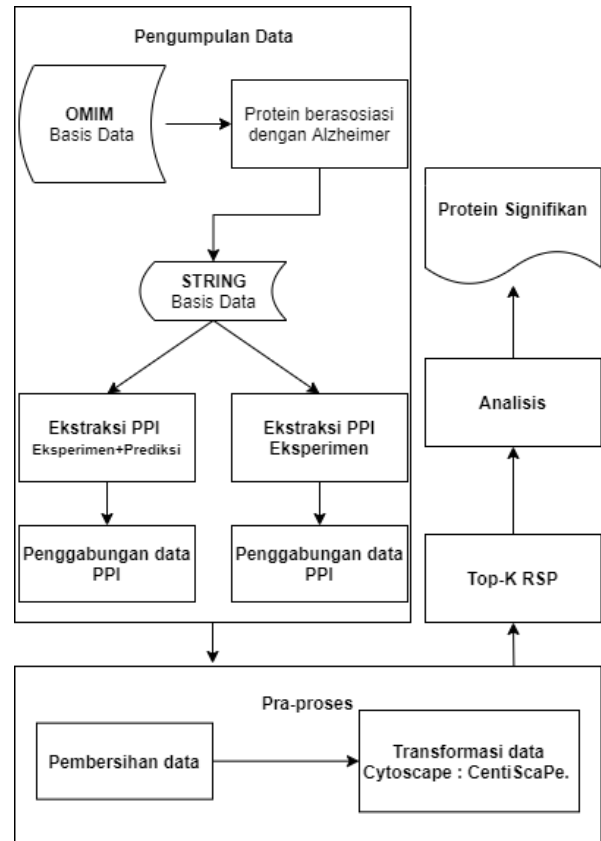
C. Top-k Representative Skyline Query (RSP)

Algoritme Top-k RSP digunakan untuk memaksimalkan objek yang didominasi oleh k objek *skyline* yang diperoleh [27]. Kompleksitas dari algoritme ini adalah $O(kn^2+kn)$ dengan n merupakan jumlah data. Kode semu Top-k RSP dinyatakan dalam Algoritme 1.

Data pada Tabel 1 merupakan contoh objek dengan dua *centrality measures*. Sebagai contoh, objek A yang memiliki *closeness centrality* sebesar 0,0045 memiliki peran yang lebih besar pada jejaring dibandingkan objek B dan C yang bernilai 0,0015 karena pada rumus yang digunakan semakin besar nilai *closeness centrality*, maka semakin dekat posisinya dengan protein lainnya dalam jejaring interaksi. Semakin besar nilai *degree* menandakan semakin banyak protein yang terhubung langsung sehingga meningkatkan kemungkinan untuk menjadi protein pengatur atau juga *hub*.

Penerapan Skyline Query menghasilkan objek A dan C sebagai objek *skyline*. Objek D dan B tidak termasuk ke dalam objek *skyline* karena terdominasi oleh objek lainnya. Objek D terdominasi oleh objek A karena A memiliki skor lebih baik pada semua dimensinya. Objek B terdominasi oleh objek A karena memiliki skor lebih kecil pada *closeness* meskipun bernilai sama pada *degree*. Objek C mendominasi objek B karena memiliki skor sama pada *closeness* tetapi bernilai lebih baik pada *degree*. Objek A dan objek C tidak saling mendominasi (*incomparable*) karena A memiliki skor lebih baik pada *closeness* dan C lebih baik pada *degree*. Tidak ada objek lain yang dapat mendominasi objek A dan C sehingga keduanya merupakan objek *skyline*.

Tabel 1 saat diproses menggunakan Top-k RSP memberikan hasil objek A sebagai hasil Top-1 RSP, sedangkan untuk Top-2 RSP menghasilkan objek A diikuti oleh objek C. Untuk kedua *centrality measures*, objek A mendominasi sebanyak dua data, yaitu objek D dan B,



Gambar 1. Tahapan penelitian

Algoritme 1. Pseudocode Top-k RSP

```

N = Input data
top_k = {}
1: from 1 to k
2: find skyline with the highest domination score from N
3: sky = skyline with the highest domination score
4: find data dominated by sky
5: dom = data dominated by sky
6: top_k = append{top_k, sky}
7: remove sky and dom from N
end

```

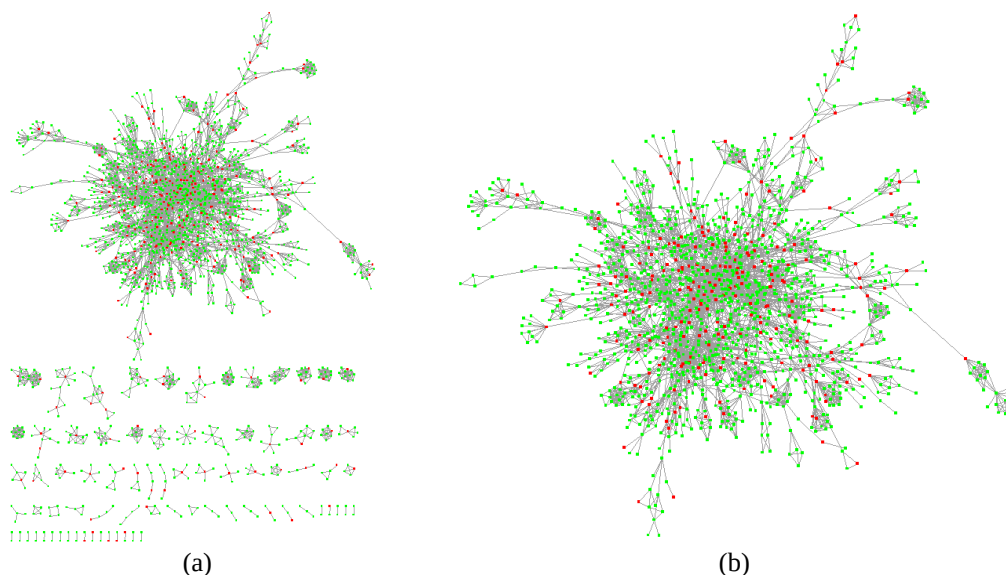
Tabel 1. Contoh data menggunakan dua *centrality measures*

Objek	Degree	Closeness
A	30	0,0045
B	30	0,0015
C	50	0,0015
D	20	0,0035

sedangkan objek C hanya mendominasi 1 data yaitu objek B. Dengan demikian, rank tertinggi adalah objek A.

D. Analisis hasil

Analisis ini bertujuan untuk melihat apakah hasil ranking tertinggi dari Top-k RSP sesuai dengan literatur. Hasil yang paling banyak mendominasi pada Top-k RSP adalah hasil dari Top-1 RSP. Analisis



Gambar 2. Visualisasi data: (a) sebelum dan (b) setelah pembersihan data

dilakukan terhadap hasil data eksperimen murni dengan eksperimen+prediksi untuk menentukan apakah sumber data tersebut dapat digunakan untuk analisis PPI pada penyakit Alzheimer.

III. HASIL DAN PEMBAHASAN

Basis data OMIM yang diperoleh adalah sebanyak 353 protein yang berasosiasi dengan Alzheimer, namun hanya ada 326 protein yang memiliki data interaksi pada basis data STRING. Protein-protein yang tidak memiliki data interaksi pada STRING tidak dilanjutkan dalam penelitian ini. Hasil dari penggabungan data interaksi protein dapat dilihat pada [Tabel 2](#). Protein yang berasosiasi dengan Alzheimer untuk sumber data interaksi eksperimen adalah sejumlah 1.889 protein dan untuk sumber data eksperimen+prediksi sejumlah 2.374 protein. Jumlah interaksinya adalah 5.224 untuk data eksperimen dan 10.546 untuk data eksperimen+prediksi.

Tahap selanjutnya adalah proses visualisasi untuk menghapus data duplikat dan jejaring yang tidak terhubung dengan jejaring utama. Contoh visualisasi pada sumber data eksperimen pada [Gambar 2\(a\)](#) menunjukkan bahwa masih terdapat jejaring yang tidak terhubung dengan jejaring utama. Setelah dilakukan pembersihan data, hasil visualisasinya ditunjukkan pada [Gambar 2\(b\)](#). Penghapusan data duplikat dan jejaring yang tidak terhubung dengan jejaring utama mengakibatkan tereduksinya jumlah protein dan interaksi protein.

Jumlah protein dan interaksi setelah dilakukan pembersihan data dapat dilihat pada [Tabel 3](#). Pembersihan data dilakukan dengan menghapus data yang tidak terhubung dengan jejaring utama. Transformasi data selanjutnya dilakukan dari bentuk jejaring interaksi ke dalam *centrality measures*. Hasil dari transformasi data yang berupa berkas csv digunakan sebagai masukan pada program Top-k RSP yang telah diimplementasikan.

Tabel 2. Hasil penggabungan data interaksi protein

Sumber Interaksi	Jumlah Protein	Jumlah Interaksi
Eksperimen	1.889	5.224
Eksperimen+Prediksi	2.374	10.546

Tabel 3. Jumlah protein dan interaksi setelah pembersihan data

Sumber Interaksi	Jumlah Protein	Jumlah Interaksi
Eksperimen	1.445 76,50%	4.400 84,23%
Eksperimen+Prediksi	2.125 89,51%	9.679 91,78%

Proses pengolahan dengan Top-k RSP dilakukan untuk kedua jenis sumber data interaksi. Langkah pertama dilakukan terhadap data interaksi dengan sumber eksperimen yang terdapat pada [Tabel 3](#), yaitu dengan jumlah protein 1445 dan jumlah interaksi 4400. Dengan menerapkan algoritme Top-k RSP pada data interaksi protein-protein tersebut, diperoleh nilai k maksimum sebesar 29 karena maksimum objek *skyline* yang bisa dihasilkan dari jejaring interaksi protein-protein tersebut adalah 29. Hal ini menunjukkan bahwa ada 29 protein yang berpotensi sebagai protein penting (signifikan) yang berperan dalam regulasi terkait penyakit Alzheimer. Pengurutan dengan Top-k RSP dilakukan untuk mencari protein paling signifikan dari 29 objek *skyline* yang ada seperti ditunjukkan dalam [Gambar 3](#). Top-1 RSP memberikan hasil bahwa protein APP berhasil mendominasi sebanyak 1339 protein lainnya dari 1445 protein yang ada, diikuti ESR1 dan PTPRC yang mendominasi sebanyak masing-masing 61 dan 6 protein. Dengan demikian, protein signifikan untuk Alzheimer berdasarkan metode Top-k RSP terhadap sumber data interaksi eksperimen adalah APP.

Proses selanjutnya adalah memproses data eksperimen+prediksi menggunakan Top-k RSP dengan jumlah protein sebelumnya adalah 2125 dengan interaksi sebanyak 9679 (Tabel 3). Dengan menerapkan algoritme Top-k RSP, diperoleh nilai k maksimum sebesar 14. Hal ini menunjukkan bahwa maksimum hanya 14 objek *skyline* yang dapat dihasilkan dari data protein dan interaksi tersebut. Dominansi keempat belas protein tersebut dapat dilihat pada Gambar 4. Dengan menggunakan Top-1 RSP, protein PSEN1 merupakan sebagai protein paling signifikan karena protein PSEN1 berhasil mendominasi sebanyak 2037 protein lainnya dari 2125 protein yang ada.

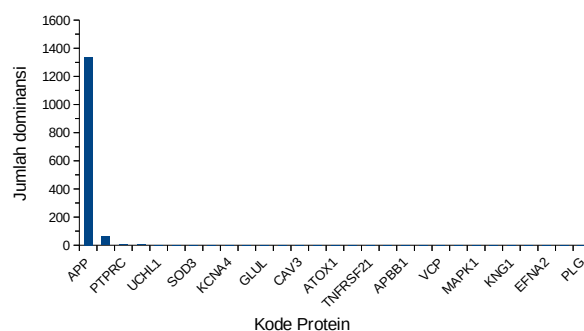
APP merupakan protein signifikan dari sumber data interaksi eksperimen, sedangkan PSEN1 merupakan protein signifikan dari sumber data interaksi eksperimen+prediksi. Mutasi pada APP dan PSEN1 menyebabkan terjadinya *autosomal dominant forms of early-onset Alzheimer disease* (AD-EOAD) [30]. APP dan PSEN1 sudah terverifikasi sebagai 2 dari 3 protein penyebab terjadinya EOAD [31].

Alzheimer sendiri merupakan penyakit demensia yang terjadi akibat adanya produksi berlebih dan penumpukan β -amyloid peptide ($A\beta$) pada otak [32], [33]. $A\beta$ muncul sebagai hasil proses *proteolytic* dari APP (β -amyloid precursor protein) [34]. Goate et al. [35] menunjukkan jika mutasi pada protein APP dapat memicu terjadinya Alzheimer. Di sisi lain, Kelleher dan Shen [36] menyatakan bahwa mutasi pada PSEN1 merupakan penyebab umum terjadinya *familial Alzheimer's Disease* (FAD). Kerusakan pada PSEN1 menyebabkan penyakit Alzheimer berat dengan umur mula terjadinya penyakit adalah pada umur 30 tahun [37].

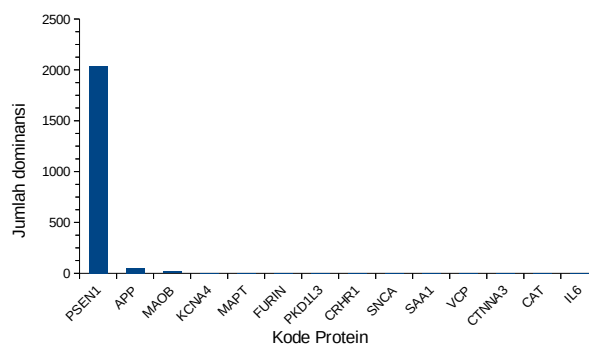
Secara umum, mutasi yang terjadi pada PSEN1 dan APP berhubungan dengan meningkatnya produksi $A\beta$ -42 peptide yang berasosiasi dengan perubahan aktivitas sekresi gamma [38]. Saat ini, $A\beta$ telah ditandai sebagai biomarker untuk *drug target* pada terapi Alzheimer [39]. Pada APP, $A\beta$ merupakan hasil dari pemecahan APP oleh tiga enzim yaitu α -, β -, dan γ -secretase. APP dipecah oleh β -secretase kemudian oleh γ -secretase akan menghasilkan $A\beta$ yang bila dipecah oleh α -secretase akan menghasilkan peptide *nontoxic* [40]. Pada kasus mutasi ganda di APP menyebabkan peningkatan produksi $A\beta$ yang meningkat akibat pemecahan APP oleh β -secretase yang juga meningkat, sehingga terjadi penumpukan berlebih pada $A\beta$ [40].

Berdasarkan studi literatur tersebut terbukti jika protein APP dan PSEN1 memiliki pengaruh signifikan akan terjadinya penyakit Alzheimer. Kedua sumber data memberikan hasil protein yang terbukti berpengaruh signifikan terhadap penyakit Alzheimer meskipun hasilnya berbeda.

Kedua jenis data, yaitu data interaksi yang diperoleh dari eksperimen saja dan eksperimen+prediksi, memberikan hasil protein yang memang berperan dalam regulasi penyakit Alzheimer. Dengan demikian, kedua jenis data tersebut dapat digunakan pada penelitian PPI terkait dengan penyakit Alzheimer. Hal ini menunjukkan pula bahwa *interactome coverage* pada



Gambar 3. Dominansi protein dalam objek *skyline* dari sumber data eksperimen



Gambar 4. Dominansi protein dalam objek *skyline* dari sumber data eksperimen+prediksi

penyakit Alzheimer sudah cukup baik, namun masih memerlukan pembuktian dengan melakukan penelitian lebih lanjut. Oleh karena itu, meskipun data eksperimen dapat digunakan untuk mencari protein signifikan dari PPI penyakit Alzheimer, tetapi penggunaan sumber data eksperimen+prediksi untuk penelitian terkait PPI tetap lebih dianjurkan untuk meningkatkan *interactome coverage*.

IV. KESIMPULAN

Top-k RSP dapat digunakan untuk menemukan protein yang memiliki peran penting pada penyakit Alzheimer. Protein signifikan yang ditemukan adalah APP dan PSEN1 yang keduanya telah terbukti memiliki pengaruh signifikan akan terjadinya penyakit Alzheimer. Sumber data interaksi eksperimen dan eksperimen+prediksi penyakit Alzheimer juga dapat digunakan untuk analisis PPI menggunakan metode pada penelitian ini. Namun, penggunaan data eksperimen+prediksi lebih dianjurkan dibanding data eksperimen.

UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada Kementerian Riset, Teknologi, dan Pendidikan Tinggi yang membiayai penelitian ini melalui skema Hibah Penelitian Tesis Magister tahun 2020.

MATERI PENDUKUNG

Naskah ini mempunyai berkas pendukung berupa hasil pengumpulan data gen/protein yang terkait penyakit Alzheimer dari basis data OMIM, informasi interaksi antarprotein dari basis data STRING, dan hasil analisis topologi jejaring interaksi protein-protein yang diperoleh dengan menggunakan aplikasi Cytoscape.

DAFTAR PUSTAKA

- [1] J. Weller and A. Budson, "Current understanding of Alzheimer's disease diagnosis and treatment," *F1000Research*, Vol 7, 1161, 2018. doi: [10.12688/f1000research.14506.1](https://doi.org/10.12688/f1000research.14506.1)
- [2] -, "Alzheimer disease and other dementia," World Health Organization, 2013. [Online]. Available: https://www.who.int/medicines/areas/priority_medicines/BP6_11Alzheimer.pdf. [Accessed: Mar. 14, 2021]
- [3] M. T. Hayes, "Parkinson's disease and parkinsonism," *The American Journal of Medicine*, vol. 132, pp. 802-807, 2019. doi: [10.1016/j.amjmed.2019.03.001](https://doi.org/10.1016/j.amjmed.2019.03.001)
- [4] G. Glaever et al., "Functional profiling of the *Saccharomyces cerevisiae* genome," *Nature*, vol. 418, no. 6896, pp. 387-391, 2002. doi: [10.1038/nature00935](https://doi.org/10.1038/nature00935)
- [5] M. L. Acencio and N. Lemke, "Towards the prediction of essential genes by integration of network topology, cellular localization and biological process information," *BMC Bioinformatics*, vol. 10, 290, 2009. doi: [10.1186/1471-2105-10-290](https://doi.org/10.1186/1471-2105-10-290)
- [6] C. Qin, Y. Sun, and Y. Dong, "A new method for identifying essential proteins based on network topology properties and protein complexes," *PLoS One*, vol. 11, no. 8, e0161042, 2016. doi: [10.1371/journal.pone.0161042](https://doi.org/10.1371/journal.pone.0161042)
- [7] J. W. Chang, Y. Q. Zhou, M. T. Ul Qamar, L. L. Chen, and Y. D. Ding, "Prediction of protein-protein interactions by evidence combining methods," *International Journal of Molecular Sciences*, vol. 17, no. 11, 1946, 2016. doi: [10.3390/ijms17111946](https://doi.org/10.3390/ijms17111946)
- [8] W. Liu, A. Wu, M. Pellegrini, and X. Wang, "Integrative of human protein, function, and disease networks," *Scientific Report*, vol. 5, 14344, 2015. doi: [10.1038/srep.14344](https://doi.org/10.1038/srep.14344)
- [9] J. D. L. Rivas and C. Fontanillo, "Protein-protein interactions essentials: key concepts to building and analyzing interactome networks," *PLoS Computational Biology*, vol. 6, no. 6, e1000807, 2010. doi: [10.1371/journal.pcbi.1000807](https://doi.org/10.1371/journal.pcbi.1000807)
- [10] D. Szklarczyk et al., "STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets," *Nucleic Acids Research*, vol 47, pp. 607-613, 2019. doi: [10.1093/nar/gky/1131](https://doi.org/10.1093/nar/gky/1131)
- [11] D. Szklarczyk et al., "The STRING database in 2017: quality controlled protein-protein association networks, made broadly accessible," *Nucleic Acids Research*, vol. 45, pp. D362-D368, 2017. doi: [10.1093/nar/gkw937](https://doi.org/10.1093/nar/gkw937)
- [12] J. Yu and F. Fotouhi, "Computational approaches for predicting protein-protein interactions: a survey," *Journal of Medical Systems*, vol. 30, pp. 39-44, 2006. doi: [10.1007/s10916-006-7402-3](https://doi.org/10.1007/s10916-006-7402-3)
- [13] W. Ali and C. M. Deane, "Evolutionary analysis reveals low coverage as the major challenge for protein interaction network alignment," *Molecular BioSystems*, vol. 6, no. 11, pp. 2296-2304, 2010. doi: [10.1039/c004430j](https://doi.org/10.1039/c004430j)
- [14] M. E. Cusick, N. Klitgord, M. Vidal, and D. E. Hill, "Interactome: gateway into systems biology," *Human Molecular Genetics*, vol. 14, no. 2, pp. R171-R181, 2005. doi: [10.1093/hmg/ddi335](https://doi.org/10.1093/hmg/ddi335)
- [15] M. P. H. Stumpf et al., "Estimating the size of human interactome," *PNAS*, vol. 105, no. 19, pp. 6959-6964, 2008. doi: [10.1073/pnas.0708078105](https://doi.org/10.1073/pnas.0708078105)
- [16] R. Jansen, N. Ian, J. Qian, and M. Gerstein, "Integration of genomic datasets to predict protein complexes in yeast," *Journal of Structural and Functional Genomics*, vol. 2, pp. 71-81, 2002. doi: [10.1023/A:1020495201615](https://doi.org/10.1023/A:1020495201615)
- [17] L. J. Lu, Y. Xia, A. Paccanaro, H. Yu, and M. Gerstein, "Assessing the limits of genomic data integration for predicting protein networks," *Genome Research*, vol. 15, pp. 945-953, 2005. doi: [10.1101/gr.3610305](https://doi.org/10.1101/gr.3610305)
- [18] Y. Qi, Z. Bar-Joseph, and J. Klein-Seetharaman, "Evaluation of different biological and computational classification methods for use in protein interaction prediction," *Proteins*, vol. 62, no. 3, pp. 490-500, 2006. doi: [10.1002/prot.20865](https://doi.org/10.1002/prot.20865)
- [19] G. Agapito, P. H. Guzzi, and M. Cannataro, "Visualization of protein interaction networks: problems and solutions," *BMC Bioinformatics*, vol. 14, no. 1, S1, 2013. doi: [10.1186/1471-2105-14-S1-S1](https://doi.org/10.1186/1471-2105-14-S1-S1)
- [20] D. Mistry, R. P. Wise, and J. A. Dickerson, "DiffSLC: a graph centrality method to detect essential proteins of a protein-protein interaction network," *PLoS ONE*, vol. 12, no. 11, e0187091, 2017. doi: [10.1371/journal.pone.0187091](https://doi.org/10.1371/journal.pone.0187091)
- [21] M. W. Hahn and A. D. Kern, "Comparative genomics of centrality and essentiality in three eukaryotic protein-interaction networks," *Molecular Biology and Evolution*, vol. 22, no. 4, pp. 803-806, 2005. doi: [10.1093/molbev/msi072](https://doi.org/10.1093/molbev/msi072)
- [22] A. Ozgur, T. Vu, G. Erkan, and D. R. Radev, "Identifying gene-disease associations using centrality on a literature mined gene-interaction network," *Bioinformatics*, vol. 24, pp. i277-i285, 2008. doi: [10.1093/bioinformatics/btn182](https://doi.org/10.1093/bioinformatics/btn182)
- [23] M. R. Diansyah, W. A. Kusuma, and A. Annisa, "Analysis of protein-protein interaction using skyline query on parkinson disease," in *International Conference on Advanced Computer*

- Science and Information Systems*, Bali, Indonesia, Oct. 2019, pp. 175-180. doi: [10.1109/ICACISIS47736.2019.8979892](https://doi.org/10.1109/ICACISIS47736.2019.8979892)
- [24] G. S. Lee and M. A. Djauhari, "An overall centrality measures: the case of U.S stock market," *International Journal of Basic & Applied Sciences*, vol. 12, no. 6, pp. 99-103, 2012.
- [25] S. Borzsonyi, D. Kossman, and K. Stocker, "The skyline operator," in *17th International Conference on Data Engineering*, Heidelberg, Germany, Apr. 2001, pp. 421-430.
- [26] M. Kontaki, A. N. Papadopoulos, and Y. Manolopoulos, "Continuous k-dominant skyline computation on multidimensional data streams," in *ACM Symposium on Applied Computing (SAC)*. Ceara, Brazil, Jun. 2008, pp. 956-960. doi: [10.1145/1363686.1363908](https://doi.org/10.1145/1363686.1363908)
- [27] X. Lin, Y. Yuan, Q. Zhang, and Y. Zhang, "Selecting stars: the k most representative skyline operator," in *23rd International Conference on Data Engineering*, Istanbul, Turkey, Apr. 2007, pp. 86-95. doi: [10.1109/ICDE.2007.367.854](https://doi.org/10.1109/ICDE.2007.367.854)
- [28] P. Sharma, D. K. Bhattacharyya, and J. K. Kalita, "Centrality analysis in PPI networks," in *2016 International Conference on Accessibility to Digital World*, Guwahati, India, Dec. 2016, pp. 135-140. doi: [10.1109/ICADW.2016.7942528](https://doi.org/10.1109/ICADW.2016.7942528)
- [29] G. Scardoni, F. Fabbri, C. Laudanna, and G. Tossadori, "CentiScaPe: Network centralities for Cytoscape," Univ of Verona, 2009.
- [30] H. M. Lanoiselee *et al.*, "APP, PSEN1, and PSEN2 mutations in early-onset Alzheimer disease: A genetic screening study of familial and sporadic cases," *PLoS Medicine*, vol. 14, no. 3, e1002270, 2017. doi: [10.10371/journal.pmed.1002270](https://doi.org/10.10371/journal.pmed.1002270)
- [31] V. V. Giau, J. M. Pyun, J. Suh, E. Bagyinszky, S. S. A. An, and S. Y. Kim, "A pathogenic PSEN1 Trp165Cys mutation associated with early-onset Alzheimer's disease," *BMC Neurology*, vol. 19, 188, 2019. doi: [10.1186/s12883-019-1419-y](https://doi.org/10.1186/s12883-019-1419-y)
- [32] Y. Zhang, R. Thompson, H. Zhang, and H. Xu, "APP processing in Alzheimer's disease," *Molecular Brain*, vol. 4, 3, 2011. doi: [10.1186/1756-6606-4-3](https://doi.org/10.1186/1756-6606-4-3)
- [33] R. J. Brien and P. C. Wong, "Amyloid precursor protein processing and Alzheimer's disease," *Annual Review of Neuroscience*, vol. 34, pp. 185-204, 2011. doi: [10.1146/annurev-neuro-061010-113613](https://doi.org/10.1146/annurev-neuro-061010-113613)
- [34] C. Haass, E. H. Koo, A. Mellon, A. Y. Hyung, and D. J. Selkoe, "Targeting of cell-surface β -amyloid precursor protein to lysosomes: alternative processing into amyloid-bearing fragments," *Nature*, vol. 357, pp. 500-503, 1992. doi: [10.1038/357500a0](https://doi.org/10.1038/357500a0)
- [35] A. Goate *et al.*, "Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease," *Nature*, vol. 349, no. 6311, pp. 704-706, 1992. doi: [10.1038/349704a0](https://doi.org/10.1038/349704a0)
- [36] R. J. Kelleher and J. Shen, "Presenilin-1 mutations and Alzheimer's disease," *PNAS*, vol. 114, no. 4, pp. 629-631, 2017. doi: [10.1073/pnas.1619574114](https://doi.org/10.1073/pnas.1619574114)
- [37] L. M. Bekris, C. E. Yu, T. D. Bird, and D. W. Tsuang, "Genetics of Alzheimer disease," *Journal of Geriatric Psychiatry and Neurology*, vol. 23, no. 4, pp. 213-227, 2010. doi: [10.1177/0891988710383571](https://doi.org/10.1177/0891988710383571)
- [38] V. M. Giau, E. Bagyinszky, Y. C. Youn, S. S. A. An, and S. Kim, "APP, PSEN1, and PSEN2 mutations in Asian patients with early-onset Alzheimer disease," *International Journal of Molecular Sciences*, vol. 20, no. 19, 4757, 2019. doi: [10.3390/ijms20194757](https://doi.org/10.3390/ijms20194757)
- [39] E. Tamagno, M. Guglielmotto, D. Monteleone, G. Manassero, V. Vasciaveo, and M. Tabaton, "The unexpected role of $A\beta_{1-42}$ monomers in the pathogenesis of Alzheimer's disease," *Journal of Alzheimer's Disease*, vol. 62, pp. 1241-1245, 2017. doi: [10.3233/JAD-170581](https://doi.org/10.3233/JAD-170581)
- [40] C. L. Masters and K. Beyreuther, "Alzheimer's centennial legacy: prospects for rational therapeutic intervention targeting the $A\beta$ amyloid pathway," *Brain*, vol. 129, no. 11, pp. 2823-2839, 2006. doi: [10.1093/brain/awl251](https://doi.org/10.1093/brain/awl251)



©2021. This is an open access article distributed under the terms and conditions of the [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).