

BAB 2

LANDASAN TEORI

Bab ini membahas dasar teori yang mendukung analisis jaringan farmakologi menggunakan algoritma *overall centrality* untuk menemukan senyawa-senyawa yang terlibat dalam pengobatan kanker prostat dari tanaman herbal sambiloto, jahe, dan bawang merah. Tinjauan teori disusun secara terstruktur dengan menjelaskan karakteristik serta mekanisme molekuler kanker prostat sebagai penyakit yang kompleks dan melibatkan beberapa gen, protein, serta jalur pensinyalan. Selain itu, dijelaskan pula tentang kandungan bioaktif dan kemampuan antikanker dari tiga tumbuhan herbal yang digunakan. Dalam hal ini, prinsip *network pharmacology* digunakan untuk memetakan hubungan antara senyawa, protein target, dan penyakit. Dijelaskan pula tentang jaringan interaksi protein dan penerapan algoritma *overall centrality* untuk menggabungkan berbagai ukuran sentralitas dalam menentukan protein target yang berperan penting dalam jaringan tersebut. Kerangka teori tersebut berfungsi sebagai dasar konseptual dalam melakukan analisis komputasional pada bab berikutnya.

2.1 Kanker Prostat

Kanker prostat merupakan salah satu jenis kanker yang paling banyak menyerang laki-laki di seluruh dunia dan menjadi penyebab utama kematian akibat kanker pada populasi tersebut. Penyakit ini berkembang melalui interaksi berbagai faktor genetik, molekuler, dan lingkungan yang menyebabkan pertumbuhan sel prostat secara tidak terkendali. Bagian ini membahas pengertian kanker prostat dan faktor-faktor risiko yang berperan dalam perkembangan penyakit sebagai dasar untuk memahami mekanisme molekuler yang dianalisis menggunakan pendekatan *network pharmacology*.

2.1.1 Pengertian Kanker Prostat

Kanker prostat merupakan pertumbuhan sel abnormal dan tidak terkendali yang berkembang pada jaringan kelenjar prostat. Prostat adalah bagian dari sistem reproduksi laki-laki yang terletak di bawah kandung kemih. Kelenjar ini bekerja menghasilkan sebagian cairan semen yang mendukung kelangsungan hidup dan pergerakan sperma. Sebagian besar kanker prostat bermula dari sel-sel kelenjar

pada prostat dan dikenal sebagai *adenokarsinoma*. Jenis yang paling sering ditemukan adalah *adenokarsinoma asinar*, yaitu kanker yang berkembang pada bagian prostat yang menghasilkan cairan [12].

2.1.2 Faktor Risiko Kanker Prostat

Risiko terkena kanker prostat semakin besar jika usia seseorang semakin meningkat. Sebagian besar kasus terjadi pada pria yang sudah tua. Selain faktor usia, keturunan dan riwayat penyakit dalam keluarga juga bisa memengaruhi kemungkinan seseorang terkena penyakit. Laki-laki yang punya keluarga dekat yang pernah mengidap kanker prostat memiliki peluang lebih besar untuk mengalami penyakit yang sama.

Selain faktor genetik, pola makan, obesitas, inflamasi kronis, paparan lingkungan, dan perubahan hormonal diduga berkontribusi terhadap perkembangan penyakit. Akan tetapi, hubungan antara faktor-faktor tersebut dan kanker prostat melibatkan interaksi antara faktor biologis dan lingkungan [12].

2.2 Tumbuhan Herbal Potensial sebagai Antikanker Prostat

Penggunaan tumbuhan herbal sebagai bagian dari upaya pengembangan terapi kanker telah dilakukan sejak lama karena berbagai kandungan senyawa bioaktifnya berpotensi memberikan aktivitas antikanker. Dalam penelitian ini, terdapat tiga tumbuhan herbal yang dianalisis sebagai sumber kandidat senyawa antikanker prostat, yaitu:

1. Jahe (*Zingiber officinale*)
2. Sambiloto (*Andrographis paniculata*)
3. Bawang Merah (*Allium cepa*)

2.3 Interaksi Antar Protein dalam Jaringan PPI

STRINGDB merupakan basis data bioinformatika yang digunakan untuk memperoleh informasi interaksi antar protein atau *protein-protein interaction* (PPI) yang berupa interaksi fisik langsung ataupun asosiasi fungsional antar protein dalam proses biologis tertentu. Untuk mendapatkan hasil interaksi satu protein dengan protein lainnya, pada penelitian ini digunakan rumus

perhitungan *score* yang digunakan oleh STRINGDB, yaitu *combined score* yang diperoleh dari penggabungan beberapa jenis interaksi protein yang aktif, seperti *experiments*, *databases*, *co-expression*, *text mining*, *gene neighborhood*, *gene fusion*, dan *gene co-occurrence*. Nilai *combined score* menunjukkan tingkat kepercayaan terhadap adanya hubungan biologis antar protein. Dalam jaringan PPI, protein direpresentasikan sebagai *node*, sedangkan hubungan antar protein direpresentasikan sebagai *edge*. Nilai *combined score* dari STRINGDB digunakan sebagai bobot (*edge*) untuk menunjukkan tingkat kepercayaan hubungan antar protein, sehingga semakin tinggi nilai *score*, semakin kuat bukti pendukung terhadap asosiasi biologis antar protein tersebut [13, 14]. Perhitungan *combined score* pada STRINGDB dilakukan melalui tiga tahapan, yaitu persamaan untuk menghilangkan *prior probability* (*prior probability removal*), persamaan penggabungan seluruh *evidence channel* (*evidence combination*), dan persamaan pembentukan *combined score* akhir (*final combined score*). Ketiga persamaan tersebut ditunjukkan sebagai berikut.

- Menghilangkan *prior probability* dari setiap *evidence channel*:

$$S'_i = \frac{S_i - p}{1 - p} \quad (2.1)$$

- Menggabungkan seluruh *score evidence channel*:

$$S_{adj} = 1 - \prod (1 - S'_i) \quad (2.2)$$

- Menambahkan kembali *prior probability*:

$$S_{combined} = S_{adj} + p(1 - S_{adj}) \quad (2.3)$$

S_i menunjukkan nilai *score* dari setiap sumber bukti atau *evidence channel*. S'_i menunjukkan nilai *score* yang telah dikoreksi dengan menghilangkan pengaruh *prior probability*. Simbol p menunjukkan nilai *prior probability* yang digunakan oleh STRINGDB yaitu sebesar 0,041. Selanjutnya, S_{adj} menunjukkan nilai gabungan sementara dari seluruh *evidence channel*, sedangkan $S_{combined}$ menunjukkan nilai akhir *combined score*. [14, 15].

2.4 Pendekatan Network Pharmacology

Network pharmacology merupakan pendekatan komputasional yang menggabungkan bidang-bidang seperti biologi sistem, bioinformatika, dan farmakologi untuk mempelajari interaksi yang rumit antara senyawa bioaktif, protein target, serta jalur biologis [16]. Dalam pendekatan *network pharmacology*, analisis diawali dengan mengidentifikasi senyawa bioaktif yang terdapat pada tanaman herbal beserta target protein yang diprediksi berinteraksi dengan senyawa tersebut. Selanjutnya, target protein yang berkaitan dengan penyakit diidentifikasi menggunakan berbagai basis data bioinformatika. Protein yang merupakan irisan antara target senyawa dan target penyakit digunakan untuk membangun jaringan *Protein-Protein Interaction* (PPI). Jaringan tersebut kemudian dianalisis menggunakan metode analisis sentralitas untuk mengidentifikasi protein yang memiliki peran penting dalam jaringan. Pada penelitian ini, hasil analisis sentralitas diintegrasikan menggunakan algoritma *Overall Centrality* sehingga diperoleh protein target yang diprioritaskan untuk dianalisis lebih lanjut melalui *enrichment analysis* dan *molecular docking*. Analisis *network pharmacology* secara umum terdiri dari beberapa tahapan berikut:

1. Membuat jaringan interaksi antara senyawa, target, dan penyakit (*compound-target-disease network*).
2. Menggunakan analisis *centrality* untuk mengidentifikasi protein yang memiliki peran paling penting dalam jaringan tersebut.
3. Menggunakan algoritma *overall centrality* untuk menentukan protein yang terbaik.

2.5 Analisis Centrality

Analisis *centrality* dilakukan menggunakan Python. Dalam penelitian ini digunakan empat jenis ukuran sentralitas, yaitu *degree*, *betweenness*, *closeness*, dan *eigenvector centrality*. Beberapa metrik sentralitas digunakan karena pengertian "*important nodes*" dalam sebuah jaringan bisa berbeda-beda, tergantung pada jenis metrik yang digunakan dan juga bentuk struktur jaringan tersebut. Ini memungkinkan pemilihan "*important nodes*" yang lebih lengkap dan akurat [11].

A. Degree Centrality

Degree centrality mengukur tingkat kepentingan suatu simpul berdasarkan jumlah hubungan langsung yang dimilikinya dengan simpul lain dalam jaringan. Semakin banyak hubungan langsung yang dimiliki, semakin tinggi nilai *degree centrality* simpul tersebut. Dalam jaringan interaksi protein, nilai yang tinggi menunjukkan bahwa suatu protein memiliki banyak interaksi langsung dan berpotensi berperan sebagai *hub protein*. Oleh karena itu, *degree centrality* dapat digunakan untuk menggambarkan tingkat aktivitas atau keterhubungan langsung suatu simpul dalam jaringan [17]. Rumus *Degree Centrality* untuk sebuah simpul v adalah:

$$DC(v) = \frac{d_v}{N - 1} \quad (2.4)$$

Perhitungan *degree centrality* dilakukan dengan menghitung jumlah hubungan langsung (*degree*) yang dimiliki setiap simpul terhadap simpul lainnya dalam jaringan. Nilai tersebut kemudian dibagi dengan jumlah maksimum hubungan yang mungkin dimiliki suatu simpul, yaitu $N - 1$, sehingga diperoleh nilai yang telah dinormalisasi. Semakin banyak hubungan langsung yang dimiliki suatu protein, semakin tinggi nilai *degree centrality* dan semakin besar perannya sebagai *hub protein* dalam jaringan.

B. Betweenness Centrality

Betweenness centrality mengukur seberapa sering suatu simpul berada pada jalur terpendek yang menghubungkan dua simpul lainnya. Dalam jaringan interaksi protein, protein dengan nilai tinggi dapat menjadi penghubung antara beberapa kelompok protein yang berbeda dan berpotensi memiliki peran penting dalam mengatur hubungan antar kelompok tersebut [17].

Rumus *Betweenness Centrality* untuk sebuah simpul v adalah:

$$BC(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad (2.5)$$

Nilai *betweenness centrality* dihitung dengan menentukan seluruh jalur terpendek yang menghubungkan setiap pasangan simpul dalam jaringan.

Selanjutnya dihitung berapa banyak jalur terpendek tersebut yang melewati simpul yang sedang dianalisis. Semakin sering suatu simpul dilalui oleh jalur terpendek, semakin tinggi nilai *betweenness centrality*, yang menunjukkan bahwa simpul tersebut berperan sebagai penghubung antarbagian jaringan.

C. Closeness Centrality

Closeness centrality mengukur tingkat kedekatan suatu simpul terhadap seluruh simpul lain berdasarkan jarak terpendek dalam jaringan. Dalam jaringan interaksi protein, nilai yang tinggi menunjukkan bahwa suatu protein berada pada posisi yang relatif dekat dengan banyak protein lain dan berpotensi menyebarkan atau menerima informasi biologis secara lebih cepat [17]. Rumus *Closeness Centrality* untuk sebuah simpul v adalah:

$$CC(v) = \frac{N - 1}{\sum_{u \neq v} d(v, u)} \quad (2.6)$$

Perhitungan *closeness centrality* dilakukan dengan menghitung jarak terpendek dari suatu simpul ke seluruh simpul lain di dalam jaringan. Jumlah seluruh jarak tersebut kemudian digunakan sebagai penyebut dalam persamaan sehingga menghasilkan ukuran kedekatan suatu simpul terhadap simpul lainnya. Semakin kecil total jarak yang dimiliki suatu simpul, semakin tinggi nilai *closeness centrality*, yang menunjukkan bahwa simpul tersebut dapat menjangkau simpul lain secara lebih efisien.

D. Eigenvector centrality

Eigenvector centrality merupakan ukuran sentralitas yang menentukan tingkat kepentingan suatu simpul berdasarkan keterhubungannya dengan simpul-simpul lain yang juga memiliki pengaruh tinggi. Oleh karena itu, suatu simpul tetap dapat memperoleh nilai *eigenvector centrality* yang tinggi meskipun jumlah hubungannya tidak banyak, selama simpul tersebut terhubung dengan simpul-simpul yang memiliki posisi penting dalam jaringan. Dalam jaringan interaksi protein, nilai yang tinggi menunjukkan bahwa suatu protein berada di sekitar protein-protein berpengaruh sehingga berpotensi memiliki peran penting dalam struktur dan penyebaran informasi biologis pada jaringan [18]. Rumus

Eigenvector Centrality adalah sebagai berikut:

$$EC(v) = \frac{1}{\lambda} \sum_{u=1}^N A_{vu} x_u \quad (2.7)$$

Perhitungan *eigenvector centrality* mempertimbangkan tidak hanya jumlah hubungan yang dimiliki suatu simpul, tetapi juga tingkat kepentingan simpul-simpul yang terhubung dengannya. Nilai sentralitas dihitung secara iteratif menggunakan *adjacency matrix* hingga diperoleh nilai yang konvergen. Oleh karena itu, suatu simpul dapat memiliki nilai *eigenvector centrality* yang tinggi apabila terhubung dengan simpul-simpul lain yang juga memiliki pengaruh besar dalam jaringan.

2.6 Algoritma Overall Centrality

Overall centrality merupakan ukuran sentralitas gabungan yang digunakan untuk menilai tingkat kepentingan suatu simpul secara menyeluruh berdasarkan beberapa ukuran sentralitas tradisional. Gan dan Djauhari mengembangkan ukuran tersebut karena setiap metrik sentralitas memiliki definisi dan karakteristik yang berbeda dalam menggambarkan posisi suatu simpul di dalam jaringan [19]. Oleh karena itu, penggunaan satu ukuran sentralitas saja belum tentu mampu memberikan gambaran yang lengkap mengenai peran suatu simpul. *Overall centrality* dibentuk sebagai kombinasi linear optimal dari *degree centrality*, *betweenness centrality*, *closeness centrality*, dan *eigenvector centrality* [19]. Karena keempat metrik sentralitas tersebut memiliki berbagai rentang nilai dan skala, maka normalisasi data (*Min-Max Normalization*) harus dilakukan terlebih dahulu sebelum menghitung *overall centrality*. Proses normalisasi dilakukan menggunakan persamaan *Min-Max Normalization* sebagaimana ditunjukkan pada Persamaan 2.8.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (2.8)$$

X_{norm} adalah nilai metrik setelah diubah menjadi bentuk normal, X adalah nilai metrik sebenarnya dari node tersebut, X_{min} adalah nilai terkecil dari metrik di seluruh jaringan, dan X_{max} adalah nilai tertinggi dari metrik di seluruh jaringan. Nilai *Overall Centrality* kemudian dihitung menggunakan persamaan *Overall*

Centrality sebagaimana ditunjukkan pada Persamaan 2.9.

$$OC = DC_{\text{norm}} + BC_{\text{norm}} + CC_{\text{norm}} + EC_{\text{norm}} \quad (2.9)$$

Nilai *overall centrality* yang lebih tinggi menunjukkan bahwa suatu simpul memiliki tingkat kepentingan yang lebih besar berdasarkan gabungan karakteristik dari keempat ukuran sentralitas. Dalam jaringan interaksi protein, protein dengan nilai *overall centrality* yang tinggi diprioritaskan sebagai protein target utama karena memiliki nilai yang tinggi, baik berdasarkan jumlah hubungan langsung, peran sebagai penghubung, kedekatan dengan simpul lain, dan keterhubungannya dengan protein yang berpengaruh [19].

2.7 Enrichment Analysis

Enrichment analysis merupakan metode bioinformatika yang digunakan untuk mengidentifikasi fungsi biologis, aktivitas molekuler, lokasi seluler, dan jalur pensinyalan yang secara statistik lebih banyak direpresentasikan dalam sekumpulan gen atau protein. Dalam analisis jaringan farmakologi, metode ini membantu menjelaskan peran biologis protein target yang telah diprioritaskan melalui analisis jaringan [20]. Analisis pengayaan umumnya mencakup *Gene Ontology* (GO), yang terdiri atas tiga kategori utama, yaitu *biological process*, *molecular function*, dan *cellular component*. *Biological process* menjelaskan proses biologis yang melibatkan produk gen, *molecular function* menggambarkan aktivitas protein pada tingkat molekuler, sedangkan *cellular component* menunjukkan lokasi protein di dalam sel [21].

Dalam penelitian ini, *enrichment analysis* dilakukan menggunakan aplikasi berbasis *web* Enrichr untuk menjelaskan fungsi dari protein target utama yang telah dipilih melalui algoritma *overall centrality*. Enrichr merupakan platform analisis pengayaan set gen berbasis *web* yang dikembangkan oleh Ma'ayan Laboratory [22]. Protein dimasukkan ke dalam Enrichr, kemudian dianalisis menggunakan pustaka *Gene Ontology* untuk mengidentifikasi proses biologis, fungsi molekuler, dan komponen seluler yang berkaitan dengan protein target [21]. Selain itu, KEGG digunakan untuk mengetahui jalur molekuler yang berhubungan dengan perkembangan kanker prostat [23]. Hasil analisis Enrichr ditampilkan dalam bentuk daftar fungsi atau jalur beserta gen yang terlibat yaitu, *p-value*, *adjusted p-value*, *odds ratio*, dan *combined score* [24]. Dengan demikian, penggunaan Enrichr

membantu menjelaskan peran protein target dan jalur molekuler yang dipengaruhi oleh senyawa dari sambiloto, jahe, dan bawang merah.

2.8 Molecular Docking dalam Interaksi Senyawa-Target

Molecular docking adalah metode komputasi yang digunakan untuk memprediksi cara molekul kecil atau ligan berinteraksi dengan makromolekul target, khususnya protein reseptor. Metode ini bisa memberikan informasi tentang letak tempat ikatan, arah ligan, bentuk struktur kompleks, serta estimasi daya ikat antara ligan dan protein. Dalam membuat obat dari bahan alami, teknik *molecular docking* digunakan untuk membantu menemukan protein yang mungkin menjadi sasaran molekuler dari senyawa aktif biologis [25].

2.9 Integrasi Analisis Jaringan Farmakologi

Pendekatan ini terintegrasi untuk menggabungkan penelitian pada tingkat sistem menggunakan *network pharmacology* dengan analisis interaksi pada tingkat molekuler melalui *molecular docking*. Integrasi ini memberikan pemahaman yang lebih luas mengenai kemampuan senyawa bioaktif dari sambiloto, jahe, dan bawang merah dalam memengaruhi berbagai jenis protein serta jalur biologis yang terkait dengan kanker prostat. Analisis *overall centrality* digunakan untuk menentukan prioritas protein berdasarkan gabungan sifat struktur jaringan, sedangkan *molecular docking* digunakan untuk memperkirakan kekuatan ikatan dan cara interaksi antara senyawa bioaktif dengan protein utama yang dituju. Menggabungkan kedua metode tersebut membantu proses pengenalan dan pengecekan komputasional terhadap kandidat senyawa lebih terorganisir, sehingga senyawa yang memiliki peluang terbesar dapat dipilih untuk uji lanjutan. Sebab itu, menggabungkan analisis *network pharmacology* dengan teknik *molecular docking* bisa menjadi dasar yang rasional dalam menciptakan terapi berbasis tanaman herbal untuk mengobati kanker prostat.