

Menjelajahi Gen Inti Melalui Analisis Transkriptomik Komparatif untuk Diagnosis Dini, Prognosis, dan Terapi Kanker Kolorektal

Latar Belakang

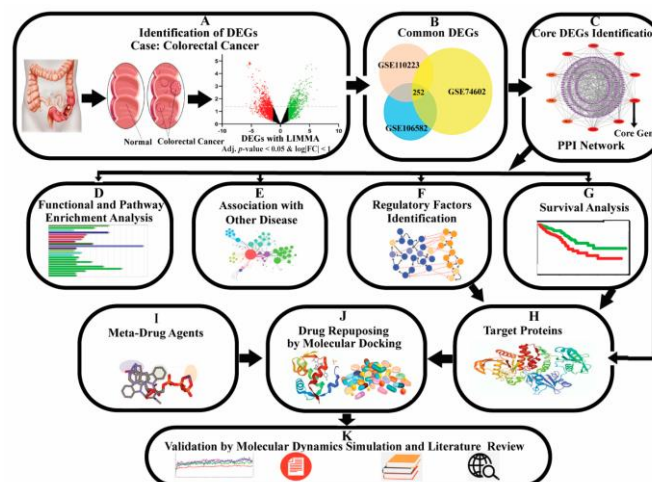
Kanker kolorektal (CRC) merupakan salah satu jenis kanker paling umum dan mematikan di dunia dengan angka kejadian dan kematian yang terus meningkat. Tingginya mortalitas CRC terutama disebabkan oleh keterlambatan diagnosis dan keterbatasan biomarker yang mampu mendeteksi penyakit pada tahap awal. Deteksi dini terbukti dapat meningkatkan tingkat kelangsungan hidup secara signifikan serta menurunkan morbiditas dan biaya pengobatan.

Konteks Penelitian

Perkembangan kanker kolorektal berkaitan erat dengan perubahan genetik dan epigenetik yang menyebabkan perbedaan ekspresi gen antara jaringan kanker dan jaringan normal. Gen yang diekspresikan secara berbeda (Differentially Expressed Genes/DEGs) berperan sebagai onkogen atau gen penekan tumor dan terlibat dalam berbagai proses patogenetik kanker. Meskipun banyak penelitian transkriptomik telah mengidentifikasi DEGs pada CRC, sebagian besar belum difokuskan pada penerapannya untuk diagnosis dini, prognosis, dan terapi, serta sering menghasilkan terlalu banyak gen sehingga sulit untuk divalidasi secara eksperimental.

Tujuan Penelitian

Penelitian ini bertujuan untuk mengidentifikasi sejumlah kecil gen inti yang diekspresikan secara berbeda (core-DEGs) yang berperan sebagai penyebab atau pemicu kanker kolorektal, serta mengkaji peran biologisnya. Hasil penelitian diharapkan dapat memberikan dasar yang lebih efektif dan efisien untuk pengembangan diagnosis dini, penentuan prognosis, dan strategi terapi kanker kolorektal.



Gambar 1. Alur kerja penelitian ini.

Metode Penelitian

Penelitian ini menggunakan pendekatan bioinformatika terintegrasi untuk mengidentifikasi gen inti penyebab kanker kolorektal (CRC) serta mengeksplorasi potensi obat yang dapat digunakan kembali. Tiga dataset microarray CRC (GSE106582, GSE110223, dan GSE74602) diperoleh dari basis data NCBI Gene Expression Omnibus (GEO), yang mencakup sampel jaringan kanker dan jaringan normal di sekitarnya dari berbagai platform microarray. Identifikasi gen yang diekspresikan secara berbeda (Differentially Expressed Genes/DEGs) dilakukan menggunakan alat GEO2R berbasis metode LIMMA, dengan koreksi nilai p menggunakan pendekatan Benjamini–Hochberg. DEGs yang signifikan ditentukan berdasarkan nilai p tersesuaikan $< 0,05$ dan ambang perubahan ekspresi \log_2 fold change ± 1 . DEGs umum dari ketiga dataset selanjutnya dianalisis untuk memperoleh gen inti (core genes).

Jaringan interaksi protein–protein (PPI) dibangun menggunakan basis data STRING dan divisualisasikan dengan Cytoscape. Gen inti dipilih berdasarkan analisis topologi jaringan menggunakan metode Maximum Clique Centrality (MCC). Hubungan gen inti dengan tahapan perkembangan CRC dianalisis menggunakan UALCAN berbasis data TCGA, sedangkan kekuatan prognosinya dievaluasi melalui analisis kelangsungan hidup Kaplan–Meier menggunakan SurvExpress. Analisis pengayaan dilakukan untuk mengidentifikasi keterkaitan gen inti dengan ontologi gen (GO), jalur KEGG, penyakit, serta regulator molekuler menggunakan alat Enrichr. Analisis pan-kanker dilakukan dengan TIMER 2.0 untuk mengevaluasi peran gen inti pada berbagai jenis kanker. Selain itu, jaringan regulasi gen yang melibatkan faktor transkripsi (TF) dan mikroRNA (miRNA) dibangun menggunakan NetworkAnalyst berbasis database JASPAR dan TarBase.

Untuk mengeksplorasi potensi terapi, dilakukan studi penambatan molekuler antara protein target gen inti dan molekul obat yang dikumpulkan dari basis data DSigDB dan PubChem. Protein target diperoleh dari Protein Data Bank dan SWISS-MODEL. Interaksi protein–ligan dianalisis menggunakan AutoDock dan Discovery Studio Visualizer. Kandidat obat potensial dipilih berdasarkan skor afinitas pengikatan rata-rata tertinggi. Selanjutnya, kompleks protein–ligan teratas disimulasikan menggunakan simulasi dinamika molekuler (MD) selama 100 ns dengan perangkat lunak YASARA berbasis medan gaya AMBER14. Stabilitas kompleks dan energi bebas pengikatan dianalisis menggunakan metode MM-PBSA untuk memvalidasi kekuatan dan kestabilan interaksi molekuler kandidat obat.

Hasil Penelitian

Analisis ekspresi gen dari tiga dataset microarray kanker kolorektal (GSE106582, GSE110223, dan GSE74602) berhasil mengidentifikasi sejumlah besar gen yang diekspresikan secara berbeda (DEGs) antara jaringan kanker dan jaringan normal. Integrasi ketiga dataset menghasilkan 252 gen yang diekspresikan secara berbeda secara konsisten (common DEGs/cDEGs), yang menunjukkan keterlibatan kuat gen-gen tersebut dalam patogenesis kanker kolorektal. Berdasarkan analisis jaringan interaksi protein–protein (PPI) dan pendekatan topologi Maximum Clique Centrality (MCC), sepuluh gen inti (core genes/CG) diidentifikasi sebagai pengendali utama kanker kolorektal, yaitu **AURKA**, **TOP2A**, **CDK1**, **PTTG1**, **CDKN3**, **CDC20**, **MAD2L1**, **CKS2**, **MELK**, dan **TPX2**. Analisis ekspresi menunjukkan bahwa seluruh CG ini secara signifikan terdisregulasi pada semua tahap perkembangan kanker kolorektal, dari tahap awal hingga tahap lanjut.

Analisis prognosis menggunakan kurva Kaplan–Meier mengungkapkan bahwa ekspresi CG mampu membedakan secara signifikan kelompok pasien dengan risiko tinggi dan rendah, menunjukkan nilai prognostik yang kuat. Selain itu, analisis asosiasi penyakit dan pan-kanker menunjukkan bahwa CG tidak hanya terkait dengan kanker kolorektal, tetapi juga dengan berbagai jenis kanker lainnya, menandakan peran pan-kanker dari gen-gen inti tersebut. Analisis pengayaan fungsional menunjukkan bahwa CG terutama terlibat dalam regulasi siklus sel, pembelahan mitosis, organisasi spindel, dan aktivitas protein kinase, yang merupakan proses biologis kunci dalam perkembangan kanker. Jalur KEGG yang diperkaya, seperti siklus sel dan jalur kanker spesifik, semakin menegaskan peran sentral CG dalam proliferasi sel kanker. Analisis jaringan regulasi mengidentifikasi faktor transkripsi utama (**NFIC**, **FOXC1**, **YY1**, dan **GATA2**) serta beberapa miRNA kunci (**hsa-miR-34a-5p**, **hsa-miR-124-3p**, **hsa-miR-16-5p**, dan lainnya) sebagai regulator ekspresi CG pada tingkat transkripsional dan pasca-transkripsional. Dalam konteks terapi, studi penambatan molekuler mengidentifikasi tujuh kandidat obat potensial dengan afinitas pengikatan tinggi terhadap protein target CG, dengan **Manzamine A** menunjukkan afinitas pengikatan tertinggi secara konsisten. Simulasi dinamika molekuler (MD) selama 100 ns mengonfirmasi stabilitas kompleks protein–ligan teratas, khususnya kompleks **CDC20–Cardidigin** dan **MELK–Staurosporine**, yang menunjukkan stabilitas struktural dan energi pengikatan yang kuat.

Pembahasan

Hasil penelitian ini menunjukkan bahwa pendekatan integratif berbasis multi-dataset microarray dan analisis jaringan mampu mengidentifikasi gen inti yang berperan penting dalam patogenesis kanker kolorektal. Sepuluh CG yang diidentifikasi sebagian besar berhubungan dengan regulasi siklus sel dan mitosis, proses fundamental yang sering mengalami disregulasi pada sel kanker. Konsistensi ekspresi CG pada semua tahap perkembangan CRC menegaskan potensi gen-gen ini sebagai biomarker diagnosis dini. Kemampuan CG dalam memprediksi prognosis pasien menunjukkan bahwa gen-gen ini tidak hanya berperan dalam inisiasi kanker, tetapi juga dalam progresi dan agresivitas tumor. Temuan bahwa beberapa CG menunjukkan peran pan-kanker mengindikasikan adanya mekanisme molekuler bersama antar berbagai jenis kanker, yang membuka peluang pengembangan terapi lintas kanker.

Identifikasi TF dan miRNA utama memberikan wawasan tambahan mengenai mekanisme regulasi ekspresi gen CRC, yang penting untuk memahami kompleksitas kontrol molekuler kanker. Selain itu, pendekatan reposisi obat berbasis penambatan molekuler dan simulasi MD menunjukkan potensi penggunaan kembali obat-obatan yang ada sebagai terapi CRC, sehingga dapat menghemat waktu dan biaya pengembangan obat baru. Secara keseluruhan, penelitian ini menyediakan kerangka kerja komputasional yang komprehensif untuk mengidentifikasi biomarker molekuler dan kandidat obat potensial kanker kolorektal. Namun, validasi eksperimental lebih lanjut secara *in vitro* dan *in vivo* masih diperlukan untuk mengonfirmasi fungsi biologis gen inti dan efektivitas kandidat obat yang diusulkan.

Kesimpulan

Penelitian ini berhasil mengidentifikasi sepuluh gen inti yang diekspresikan secara berbeda, yaitu **AURKA**, **TOP2A**, **CDK1**, **PTTG1**, **CDKN3**, **CDC20**, **MAD2L1**, **CKS2**, **MELK**, dan **TPX2**, yang menunjukkan potensi prognostik yang kuat, khususnya pada tahap awal kanker kolorektal (CRC).

Analisis asosiasi penyakit dan pan-kanker menunjukkan bahwa gen-gen inti tersebut tidak hanya berperan dalam CRC, tetapi juga terlibat dalam berbagai jenis kanker lainnya, sehingga mengindikasikan peran pan-kanker yang signifikan. Analisis jaringan regulasi mengungkapkan keterlibatan beberapa faktor transkripsi utama (**NFIC, FOXC1, YY1, dan GATA2**) serta miRNA kunci (**hsa-miR-147a, hsa-miR-129-2-3p, hsa-miR-124-3p, hsa-miR-34a-5p, hsa-miR-23b-3p, dan hsa-miR-16-5p**) dalam pengaturan ekspresi gen inti pada tingkat transkripsional dan pasca-transkripsional. Selain itu, analisis pengayaan fungsional menunjukkan bahwa gen-gen inti tersebut terutama terlibat dalam proses biologis penting seperti regulasi siklus sel dan pembentukan spindle mitosis, yang berperan krusial dalam perkembangan kanker kolorektal.

Lebih lanjut, analisis penambatan molekuler mengidentifikasi tujuh kandidat molekul obat potensial, yaitu **Manzamine A, Cardidigin, Staurosporine, Sitosterol, Benzo[a]pyrene, Nocardiosis sp., dan Riccardin D**, yang menunjukkan afinitas pengikatan tinggi terhadap protein target gen inti. Secara keseluruhan, temuan penelitian ini memberikan kontribusi penting dalam pengembangan biomarker molekuler dan kandidat terapi potensial untuk diagnosis dini, penentuan prognosis, serta pengobatan kanker kolorektal. Namun demikian, validasi eksperimental lebih lanjut masih diperlukan untuk mengonfirmasi efektivitas dan keamanan temuan ini.

Daftar Pustaka

- Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell*, 144(5), 646–674. <https://doi.org/10.1016/j.cell.2011.02.013>
- Gray, J. W., & Collins, C. (2000). Genome changes and gene expression in human solid tumors. *Carcinogenesis*, 21(3), 443–452. <https://doi.org/10.1093/carcin/21.3.443>
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 68(6), 394–424. <https://doi.org/10.3322/caac.21492>
- Arnold, M., Sierra, M. S., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2017). Global patterns and trends in colorectal cancer incidence and mortality. *Gut*, 66(4), 683–691. <https://doi.org/10.1136/gutjnl-2015-310912>
- Siegel, R. L., Miller, K. D., Goding Sauer, A., Fedewa, S. A., Butterly, L. F., Anderson, J. C., Cercek, A., Smith, R. A., & Jemal, A. (2020). Colorectal cancer statistics, 2020. *CA: A Cancer Journal for Clinicians*, 70(3), 145–164. <https://doi.org/10.3322/caac.21601>
- Joranger, P., Nesbakken, A., Sorbye, H., Hoff, G., Oshaug, A., & Aas, E. (2020). Survival and costs of colorectal cancer treatment and effects of changing treatment strategies: A model approach. *European Journal of Health Economics*, 21(3), 321–334. <https://doi.org/10.1007/s10198-019-01125-7>
- Mo, S., Dai, W., Wang, H., Lan, X., Ma, C., Su, Z., Xiang, W., Han, L., Luo, W., Zhang, L., et al. (2023). Early detection and prognosis prediction of colorectal cancer by circulating tumor DNA methylation haplotypes: A multicenter cohort study. *eClinicalMedicine*, 55, 101717. <https://doi.org/10.1016/j.eclinm.2022.101717>

- Porcu, E., Sadler, M. C., Lepik, K., Auwerx, C., Wood, A. R., Weihs, A., Sleiman, M. S. B., Ribeiro, D. M., Bandinelli, S., Tanaka, T., et al. (2021). Differentially expressed genes reflect disease-induced rather than disease-causing changes in the transcriptome. *Nature Communications*, 12, 5647. <https://doi.org/10.1038/s41467-021-25895-5>
- Bogaert, J., & Prenen, H. (2014). Molecular genetics of colorectal cancer. *Annals of Gastroenterology*, 27(1), 9–14.
- Lu, A. G., Feng, H., Wang, P. X., Han, D. P., Chen, X. H., & Zheng, M. H. (2012). Emerging roles of ribonucleotide reductase M2 in colorectal cancer and ultraviolet-induced DNA damage repair. *World Journal of Gastroenterology*, 18(34), 4704–4713. <https://doi.org/10.3748/wjg.v18.i34.4704>
- Liu, X., Zhang, H., Lai, L., Wang, X., Loera, S., Xue, L., He, H., Zhang, K., Hu, S., Huang, Y., et al. (2013). Ribonucleotide reductase small subunit M2 serves as a prognostic biomarker and predicts poor survival in colorectal cancer. *Clinical Science*, 124(9), 567–579. <https://doi.org/10.1042/CS20120340>
- Gan, Y., Li, Y., Li, T., Shu, G., & Yin, G. (2018). CCNA2 acts as a novel biomarker in regulating the growth and apoptosis of colorectal cancer. *Cancer Management and Research*, 10, 5113–5124. <https://doi.org/10.2147/CMAR.S176833>
- Branchi, V., García, S. A., Radhakrishnan, P., Györfy, B., Hissa, B., Schneider, M., Reißfelder, C., & Schölch, S. (2019). Prognostic value of DLGAP5 in colorectal cancer. *International Journal of Colorectal Disease*, 34(8), 1455–1465. <https://doi.org/10.1007/s00384-019-03334-3>
- Hozhabri, H., Lashkari, A., Razavi, S. M., & Mohammadian, A. (2021). Integration of gene expression data identifies key genes and pathways in colorectal cancer. *Medical Oncology*, 38, 1–14. <https://doi.org/10.1007/s12032-021-01522-9>
- Wei, F. Z., Mei, S. W., Wang, Z. J., Chen, J. N., Shen, H. Y., Zhao, F. Q., Li, J., Liu, Z., & Liu, Q. (2020). Differential expression analysis reveals CLCA1 as a prognostic and diagnostic biomarker for colorectal cancer. *Frontiers in Oncology*, 10, 573295. <https://doi.org/10.3389/fonc.2020.573295>
- Xu, H., Ma, Y., Zhang, J., Gu, J., Jing, X., Lu, S., Fu, S., & Huo, J. (2020). Identification and validation of core genes in colorectal cancer. *BioMed Research International*, 2020, 8082697. <https://doi.org/10.1155/2020/8082697>
- Rahman, M. R., Islam, T., Gov, E., Turanli, B., Gulfidan, G., Shahjaman, M., Banu, N. A., Mollah, M. N. H., Arga, K. Y., & Moni, M. A. (2019). Identification of prognostic biomarker signatures and candidate drugs in colorectal cancer: Insights from systems biology analysis. *Medicina*, 55(1), 20. <https://doi.org/10.3390/medicina55010020>
- Patil, A. R., Leung, M. Y., & Roy, S. (2021). Identification of hub genes at different stages of colorectal cancer using an integrated bioinformatics approach. *International Journal of Environmental Research and Public Health*, 18(11), 5564. <https://doi.org/10.3390/ijerph18115564>
- Smyth, G. K. (2004). Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, 3(1), Article 3. <https://doi.org/10.2202/1544-6115.1027>

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57(1), 289–300.

Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguéz, P., Doerks, T., Stark, M., Müller, J., & Bork, P. (2011). The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Research*, 39(Suppl. 1), D561–D568. <https://doi.org/10.1093/nar/gkq973>

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, 13(11), 2498–2504. <https://doi.org/10.1101/gr.1239303>