

Analisis Varian Missense dan Profil Ekspresi Gen Terkait Hipotiroidisme Menggunakan Pendekatan Bioinformatika

Rahmawati ^{1*}, Fauzan Sebastian Ramadhan², Naila Ismatun Hasanah ²

¹ Program Studi Pendidikan Profesi Apoteker, Fakultas Farmasi, Universitas Bakti Tunas Husada, Tasikmalaya, West Java, Indonesia

² Program Studi S1 Farmasi, Fakultas Farmasi, Universitas Bakti Tunas Husada, Tasikmalaya, West Java, Indonesia

*Corresponding author: rahmawati@universitas-bth.ac.id

Abstract

Hypothyroidism is an endocrine disorder influenced by genetic, environmental, and population-related factors. This study aimed to identify and evaluate candidate genetic variants associated with hypothyroidism using an in silico bioinformatics approach. Data were obtained from the NHGRI-EBI GWAS Catalog and filtered using $p < 1 \times 10^{-8}$, followed by removal of duplicate records. Gene expression profiles were evaluated through the GTEx Portal, while genomic locations and variant consequences were examined using Ensembl. Functional effects on proteins were assessed using SNPnexus and PolyPhen-2. The initial search identified 3,280 variant association records. After filtration, 899 SNPs remained, including 560 intron variants, 133 intergenic variants, 89 variants without contextual annotation, 44 non-coding transcript exon variants, 25 candidate missense variants, and 48 variants in other categories. Re-evaluation showed that the 25 candidates could not all be confirmed as missense variants because inconsistencies were found among several rsIDs, mapped gene names, chromosomal locations, and variant consequences. GTEx analysis provided only a descriptive overview of gene expression distribution, while SNPnexus and PolyPhen-2 generated no prediction scores. These findings support bioinformatics as a preliminary screening approach, but standardized annotation and further validation remain necessary.

Keywords: hypothyroidism, SNP, GWAS Catalog, GTEx, Ensembl, SNPnexus

Abstrak

Hipotiroidisme merupakan gangguan endokrin yang dipengaruhi oleh faktor genetik, faktor lingkungan, dan karakteristik populasi. Penelitian ini bertujuan untuk mengidentifikasi serta mengevaluasi kandidat varian genetik terkait hipotiroidisme menggunakan pendekatan bioinformatika in silico. Data diperoleh dari NHGRI-EBI GWAS Catalog dan disaring berdasarkan nilai $p < 1 \times 10^{-8}$ serta penghapusan data duplikat. Profil ekspresi gen dievaluasi melalui GTEx Portal, sedangkan lokasi dan konsekuensi genomik diperiksa menggunakan Ensembl. Dampak fungsional terhadap protein dianalisis menggunakan SNPnexus dan PolyPhen-2. Penelusuran awal menghasilkan 3.280 data asosiasi varian. Setelah filtrasi, diperoleh 899 SNP, terdiri atas 560 varian intron, 133 varian intergenic, 89 varian tanpa konteks, 44 varian exon non-coding transcript, 25 kandidat varian missense, dan 48 varian dalam kategori lainnya. Evaluasi ulang menunjukkan bahwa 25 kandidat belum seluruhnya terkonfirmasi sebagai varian missense karena terdapat ketidaksesuaian antara beberapa rsID, nama gen hasil pemetaan, lokasi kromosom, dan konsekuensi varian. Analisis GTEx hanya memberikan gambaran deskriptif distribusi ekspresi gen, sedangkan SNPnexus dan PolyPhen-2 tidak menghasilkan skor prediksi. Temuan ini mendukung penggunaan bioinformatika sebagai skrining awal, tetapi masih memerlukan anotasi yang terstandar dan validasi lanjutan.

Kata kunci: hipotiroidisme, SNP, GWAS Catalog, GTEx, Ensembl, SNPnexus



PENDAHULUAN

Hipotiroidisme adalah gangguan endokrin yang ditandai oleh berkurangnya produksi hormon tiroid, termasuk tiroksin (T4) dan triiodotironin (T3), yang memengaruhi metabolisme, pertumbuhan, dan fungsi kognitif. Kondisi ini menyebabkan gejala seperti kelelahan, penambahan berat badan, kulit kering, gangguan suasana hati, serta penurunan kemampuan kognitif, sehingga berdampak signifikan pada kualitas hidup pasien (WHO, 2025; Burnstock, 2017). Pengobatan hipotiroidisme masih menghadapi tantangan global karena diagnosis sering terlambat, akses pemeriksaan laboratorium tidak merata, dan respons terapi bervariasi antarpasien (Institute for Health Metrics and Evaluation, 2021).

Prevalensi hipotiroidisme pada populasi dewasa dilaporkan bervariasi antarwilayah. Variasi tersebut menunjukkan bahwa kejadian hipotiroidisme dapat dipengaruhi oleh interaksi berbagai faktor, termasuk faktor genetik, faktor lingkungan, dan karakteristik populasi. Oleh karena itu, pemahaman yang lebih mendalam mengenai mekanisme molekuler dan genetik diperlukan untuk mendukung pengembangan strategi diagnosis dan pengobatan yang lebih presisi. Dalam penelitian ini, identifikasi varian genetik terkait hipotiroidisme dilakukan menggunakan NHGRI–EBI GWAS Catalog, yaitu basis data terkurasi yang menghimpun hasil studi asosiasi genom luas pada manusia (Fadista et al., 2016; Cerezo et al., 2025).

Penelitian genetika mengungkap bahwa varian genetik, khususnya SNP missense, dapat mengubah struktur protein atau regulasi ekspresi gen yang terkait dengan jalur hormon tiroid, sehingga memengaruhi risiko hipotiroidisme. Varian ini dapat dikategorikan sebagai benign, possibly damaging, atau probably damaging, sesuai prediksi efeknya terhadap fungsi protein, yang memberikan gambaran penting mengenai mekanisme molekuler penyakit (EMBL-EBI, 2025d; Islam et al., 2025; Adzhubei et al., 2013).

Pendekatan *in silico* berbasis bioinformatika memungkinkan integrasi data genomik dan transkriptomik untuk memetakan variasi genetik dan ekspresi gen secara komprehensif tanpa melibatkan subjek manusia secara langsung. Analisis ini juga menjadi dasar untuk memahami gen-gen kandidat yang berpotensi berkontribusi pada *hypothyroidism* dan memberikan fondasi untuk eksplorasi target terapi lebih lanjut (Rudrapal et al., 2020; Kulkarni et al., 2023; Damian et al., 2023; Ahmad et al., 2021; Ward & Kellis, 2016; Dayem Ullah et al., 2018; Broad Institute, 2023).

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menganalisis varian gen missense dan ekspresi gen pada hipotiroidisme menggunakan pendekatan *in silico*. Diharapkan penelitian ini dapat membuka wawasan baru mengenai mekanisme genetik penyakit serta mendukung pengembangan strategi pengobatan yang lebih presisi dan berbasis bukti genomik.

BAHAN DAN METODE

Desain Penelitian

Penelitian ini menggunakan pendekatan *in silico* dengan desain studi bioinformatika komputasional untuk menganalisis varian genetik (SNP missense) dan ekspresi gen pada hipotiroidisme. Data diperoleh dari berbagai database genomik dan transkriptomik publik, sehingga tidak melibatkan subjek manusia secara langsung. Analisis dilakukan secara bertahap untuk mengidentifikasi gen target dan mengevaluasi variasi genetik yang berpotensi berkontribusi terhadap penyakit (Islam et al., 2025; Dayem Ullah et al., 2018). Penelitian dilaksanakan di Universitas Bakti Tunas Husada pada periode Januari–Mei 2026. Seluruh proses analisis dilakukan menggunakan perangkat keras standar laptop dan basis data daring.

Software

Tabel 1. Website yang Digunakan

No	Software	Fungsi
1	GWAS Catalog (www.ebi.ac.uk/gwas)	Mengidentifikasi SNP terkait hypothyroidism (Hindorff et al., 2009)
2	GTEx Portal (http://gtexportal.org)	Analisis ekspresi gen di berbagai jaringan (Ardlie, 2015)
3	Ensembl (https://www.ensembl.org)	Menentukan lokasi gen pada kromosom dan konteks regulasi genomik (EMBL-EBI, 2025d)
4	SNPnexus (https://www.snp-nexus.org)	Menilai dampak varian genetik pada fungsi protein (Adzhubei et al., 2013)

Prosedur Penelitian

Prosedur penelitian dilakukan melalui lima tahap utama. Tahap pertama memanfaatkan GWAS Catalog untuk mengidentifikasi SNP terkait hypothyroidism, yang disaring berdasarkan p-value $< 1 \times 10^{-8}$ agar hanya data signifikan secara statistik yang dianalisis (Fadista et al., 2016). Tahap kedua menggunakan GTEx Portal untuk menilai ekspresi gen di berbagai jaringan, sehingga peneliti dapat memahami distribusi gen target. Tahap ketiga memanfaatkan Ensembl untuk menentukan posisi SNP pada kromosom secara presisi, termasuk koordinat ekson, intron, serta regulatory region. Tahap keempat menggunakan SNPnexus untuk mengevaluasi efek non-synonymous coding SNP terhadap protein, dengan pengelompokan varian menjadi *benign*, *possibly damaging*, atau *probably damaging*.

Kriteria populasi dalam penelitian ini ditetapkan secara spesifik. Kriteria inklusi mencakup SNP yang terkait dengan hipotiroidisme, berjenis missense, dan memiliki p-value $< 10^{-8}$, sehingga hanya varian signifikan yang dianalisis. Sebaliknya, kriteria eksklusi meliputi SNP non-missense, p-value $\geq 10^{-8}$, atau SNP yang tidak terkait dengan penyakit. Data yang diperoleh kemudian dianalisis untuk menentukan variasi genetik yang signifikan yang memengaruhi fungsi protein, mengevaluasi profil ekspresi gen pada berbagai jaringan, serta memetakan distribusi alel pada populasi yang berbeda. Hasil analisis ini menjadi dasar untuk memahami gen-gen kandidat yang berpotensi berkontribusi pada mekanisme hipotiroidisme dan menyediakan fondasi untuk eksplorasi target terapi pada tahap penelitian lanjutan.

HASIL DAN PEMBAHASAN

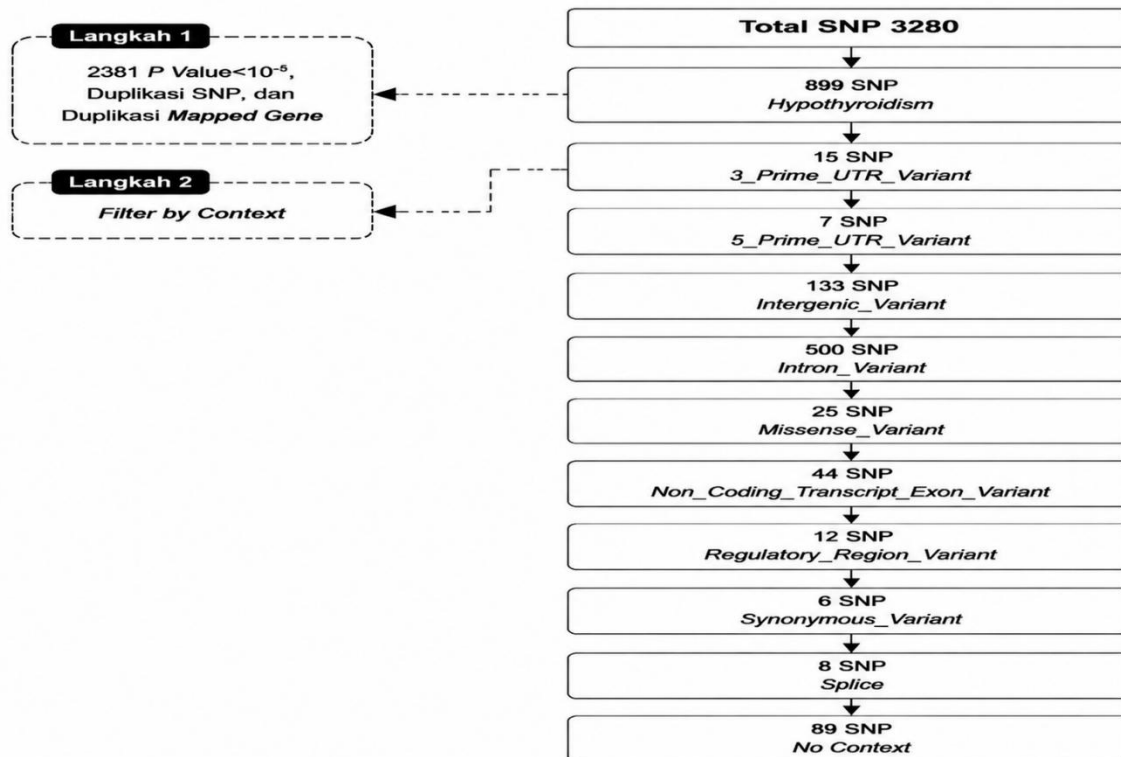
Gambaran Umum Analisis

Penelitian ini menggunakan pendekatan bioinformatika untuk mengidentifikasi kandidat varian genetik yang berasosiasi dengan hipotiroidisme serta mengevaluasi profil ekspresi gen hasil pemetaan varian tersebut pada berbagai jaringan manusia. Analisis dilakukan menggunakan GWAS Catalog, GTEx Portal, Ensembl, dan SNPnexus. GWAS Catalog digunakan untuk memperoleh data asosiasi SNP; GTEx Portal digunakan untuk menggambarkan distribusi ekspresi gen; Ensembl digunakan untuk memeriksa lokasi dan konteks genomik varian; sedangkan SNPnexus digunakan untuk mengevaluasi kemungkinan dampak varian terhadap protein.

Setiap basis data memberikan informasi yang berbeda. Hasil asosiasi pada GWAS menunjukkan hubungan statistik antara lokus dan fenotipe, tetapi tidak secara langsung membuktikan hubungan sebab-akibat. Demikian pula, ekspresi suatu gen pada jaringan tertentu tidak membuktikan bahwa SNP pada lokus tersebut mengubah ekspresi gen. Oleh karena itu, seluruh hasil penelitian ini ditafsirkan sebagai proses identifikasi dan prioritas kandidat untuk analisis lanjutan.

Identifikasi dan Filtrasi Varian pada GWAS Catalog

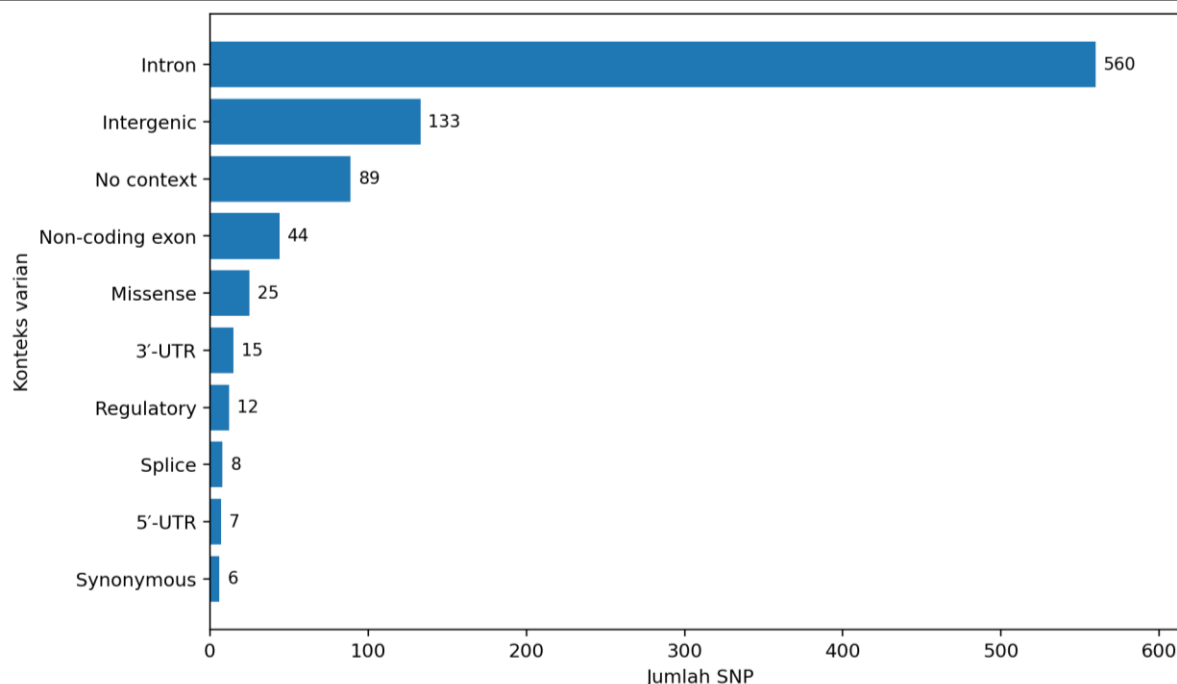
Penelusuran menggunakan kata kunci hypothyroidism pada GWAS Catalog menghasilkan 3.280 data asosiasi SNP. Data kemudian disaring berdasarkan nilai $p < 1 \times 10^{-8}$ serta penghapusan duplikasi SNP dan mapped gene. Setelah proses tersebut, diperoleh 899 SNP yang dikelompokkan berdasarkan konteks genomiknya. Dengan demikian, sebanyak 2.381 data tidak dipertahankan karena tidak memenuhi satu atau lebih kriteria penyaringan.



Gambar 1. Alur identifikasi dan filtrasi varian genetik terkait hipotiroidisme dari GWAS Catalog

Sumber: Pribadi

Sebanyak 899 SNP yang lolos tahap penyaringan terdiri atas 560 intron variants, 133 intergenic variants, 89 varian tanpa konteks, 44 non-coding transcript exon variants, 25 missense variants, 15 3'-UTR variants, 12 regulatory region variants, 8 varian terkait splicing, 7 5'-UTR variants, dan 6 synonymous variants. Jumlah seluruh kategori tersebut adalah 899 SNP, sehingga konsisten dengan jumlah data setelah filtrasi.



Gambar 2. Distribusi 899 SNP berdasarkan konteks genomik.

Varian intronik menjadi kelompok terbesar. Walaupun tidak secara langsung mengubah urutan asam amino, varian intronik dapat berkaitan dengan proses splicing atau elemen regulasi. Varian intergenik juga dapat berada di sekitar enhancer, silencer, atau wilayah pengikatan faktor transkripsi. Varian pada wilayah 3'-UTR dan 5'-UTR dapat berhubungan dengan stabilitas mRNA dan efisiensi translasi, sedangkan varian pada regulatory region dapat memengaruhi regulasi transkripsi.

Penelitian ini memprioritaskan 25 varian missense karena perubahan nukleotida pada wilayah pengode dapat menyebabkan substitusi asam amino. Namun, klasifikasi missense tidak secara otomatis menunjukkan bahwa varian tersebut bersifat patogen. Sebagian varian dapat bersifat netral, sedangkan varian lainnya dapat memengaruhi struktur atau fungsi protein. Oleh karena itu, 25 varian tersebut diperlakukan sebagai kandidat yang memerlukan verifikasi anotasi lebih lanjut.

Identifikasi dan Filtrasi Varian Genetik Berdasarkan GWAS Catalog

Penelusuran awal pada NHGRI-EBI GWAS Catalog menghasilkan 3.280 data asosiasi varian yang terkait dengan hipotiroidisme. Setelah dilakukan penyaringan berdasarkan nilai $p < 1 \times 10^{-8}$ serta penghapusan data duplikat, diperoleh 899 SNP. Varian kemudian dikelompokkan berdasarkan konteks genomiknya sebagaimana disajikan pada Tabel 2.

Tabel 2. Distribusi kandidat SNP terkait hipotiroidisme berdasarkan konteks genomik setelah filtrasi

No	Konteks genomik	Jumlah SNP, n	Persentase, %
1	<i>Intron variant</i>	560	62,29
2	<i>Intergenic variant</i>	133	14,79
3	<i>No context</i>	89	9,9
4	<i>Non-coding transcript exon variant</i>	44	4,89
5	Kandidat <i>missense variant</i> berdasarkan anotasi awal	25	2,78
6	<i>3'-UTR variant</i>	15	1,67
7	<i>Regulatory region variant</i>	12	1,33

8	<i>Splice-related variant</i>	8	0,89
9	<i>5'-UTR variant</i>	7	0,78
10	<i>Synonymous variant</i>	6	0,67
Total		899	100

Sebagian besar SNP berada pada wilayah intronik, yaitu sebanyak 560 SNP (62,29%), diikuti oleh varian intergenik sebanyak 133 SNP (14,79%). Sebanyak 25 SNP (2,78%) pada anotasi awal dikelompokkan sebagai kandidat *varian missense* dan diprioritaskan untuk pemeriksaan lebih lanjut. Namun, kategori tersebut tidak langsung digunakan untuk menyimpulkan bahwa seluruh kandidat menyebabkan perubahan asam amino karena nama *mapped gene* pada GWAS Catalog belum tentu sama dengan gen pengode yang terdampak oleh varian. GWAS Catalog merupakan basis data asosiasi yang memetakan temuan GWAS berdasarkan lokasi genom, sedangkan konsekuensi varian terhadap transkrip dan protein perlu dikonfirmasi menggunakan alat anotasi seperti Ensembl VEP

Verifikasi Anotasi Kandidat SNP

Sebanyak 25 SNP yang diperoleh dari skrining awal dievaluasi kembali untuk memastikan kesesuaian antara rsID, nilai *p*, nama gen hasil pemetaan, dan konsekuensi varian. Evaluasi ini diperlukan karena nama gen yang tercantum sebagai *mapped gene* belum tentu merupakan gen pengode yang terdampak langsung oleh varian. Ringkasan hasil evaluasi dan keputusan revisi disajikan pada Tabel 3.

Hasil evaluasi menunjukkan bahwa 25 SNP belum dapat dinyatakan sepenuhnya sebagai varian *missense* terverifikasi. Beberapa pasangan SNP–gen tidak sesuai dengan anotasi genomik, sedangkan nilai *p* pada tabel awal tidak dapat dipertahankan secara konsisten. Oleh karena itu, kandidat tersebut dilaporkan sebagai hasil skrining awal. Analisis ekspresi gen dan prediksi dampak protein tidak boleh digunakan untuk menarik kesimpulan biologis sebelum hubungan SNP–gen dan konsekuensi variannya dikonfirmasi.

Tabel 3. Ringkasan evaluasi anotasi dan keputusan analisis terhadap kandidat SNP terkait hipotiroidisme

Aspek yang dievaluasi	Temuan verifikasi	Keputusan revisi
Klasifikasi varian	Sebanyak 25 SNP merupakan kandidat berdasarkan anotasi awal dan belum seluruhnya terkonfirmasi sebagai <i>missense variant</i>	Dilaporkan sebagai 25 kandidat SNP hasil skrining awal
Nilai <i>p</i>	Nilai <i>p</i> setiap asosiasi tidak dapat disajikan seragam sebagai 3×10^{-10}	Nilai seragam dihapus; kriteria seleksi dilaporkan sebagai $p < 1 \times 10^{-8}$
Kesesuaian SNP–gen	Beberapa pasangan SNP–gen tidak sesuai, termasuk rs16903080– <i>TXNDC11</i> dan rs7763540– <i>PSMD14-DT/LINC01806</i>	Pasangan yang tidak terverifikasi tidak digunakan dalam interpretasi biologis
Analisis ekspresi GTEx	Profil ekspresi awal didasarkan pada nama gen yang belum seluruhnya terverifikasi dan belum dilengkapi nilai median TPM	Klaim jaringan dengan ekspresi tertinggi dan kolom fungsi potensial dihapus
Prediksi dampak protein	SNPnexus/PolyPhen-2 menunjukkan <i>no data available</i> untuk seluruh kandidat	Tidak dilakukan klasifikasi <i>benign</i> , <i>possibly damaging</i> , atau <i>probably damaging</i>

Verifikasi Anotasi Genomik Menggunakan Ensembl

Ensembl digunakan untuk memeriksa kembali anotasi genomik kandidat SNP yang diperoleh dari tahap skrining awal. Informasi yang dievaluasi meliputi rsID, kromosom, posisi genom, alel, gen yang bertumpang tindih atau berada di sekitar varian, transkrip yang terdampak, serta konsekuensi varian.

Evaluasi ulang diperlukan karena nama gen yang tercantum sebagai *mapped gene* pada GWAS Catalog tidak selalu menunjukkan gen pengode yang secara langsung mengalami perubahan nukleotida atau asam amino. Hasil pemeriksaan menunjukkan adanya ketidaksesuaian antara beberapa rsID, kode gen, dan lokasi kromosom pada tabel awal. Oleh karena itu, data Ensembl dalam penelitian ini digunakan sebagai tahap pengendalian mutu anotasi sebelum kandidat SNP dilanjutkan ke analisis ekspresi gen dan prediksi dampak protein.

Kandidat hanya dapat diklasifikasikan sebagai *varian missense* apabila anotasi Ensembl menunjukkan adanya substitusi asam amino pada transkrip pengode. Sebaliknya, SNP yang teranotasi sebagai *intron variant*, *intergenic variant*, *regulatory region variant*, atau varian nonpengode tidak dimasukkan dalam analisis prediksi dampak protein. Dengan demikian, hasil pemetaan Ensembl tidak digunakan untuk menyimpulkan bahwa sebagian besar kandidat berada pada ekson nonsinonim atau memiliki dampak fungsional sebelum konsekuensi masing-masing varian terkonfirmasi.

Prediksi Dampak Fungsional Varian Menggunakan SNPnexus

Prediksi dampak fungsional dilakukan menggunakan SNPnexus dengan memilih keluaran PolyPhen-2 pada bagian "*effect of non-synonymous coding SNPs on protein function*". Secara umum, PolyPhen-2 mengelompokkan dampak substitusi asam amino menjadi *benign*, *possibly damaging*, dan *probably damaging* berdasarkan skor prediksi yang dihasilkan.

Hasil analisis menunjukkan bahwa skor dan kategori prediksi PolyPhen-2 tidak tersedia untuk semua 25 kandidat SNP. Dengan demikian, tidak terdapat kandidat yang dapat diklasifikasikan sebagai *benign*, *possibly damaging*, ataupun *probably damaging* berdasarkan hasil analisis ini. Data awal memang menunjukkan keterangan "*no data available*" pada seluruh kandidat yang dimasukkan ke SNPnexus (Adzhubei et al., 2013).

Hasil prediksi ini dapat dilihat pada Tabel 4 yang menyajikan 25 SNP missense beserta gen, kromosom, dan status prediksi. Perlu dicatat bahwa score dan prediction untuk semua varian *No data available* karena beberapa SNP belum tercatat di model prediksi HumDiv/HumVar yang digunakan oleh PolyPhen-2. Hal ini konsisten dengan laporan sebelumnya yang menekankan keterbatasan basis data prediksi *in silico*, sehingga beberapa varian mungkin memerlukan validasi eksperimental untuk memastikan dampak biologisnya (Li et al., 2019; Zhang et al., 2020).

Tabel 4. Identifikasi SNP dan Efek Prediksi Protein

No	SNP	Chromosome	Gene
1	rs16903080	Chr4	TXNDC11
2	rs16903078	Chr4	POP7 - EPO
3	rs1793888	Chr11	YAF2
4	rs1610667	Chr9	SNRPGP8 - CCL20
5	rs16903081	Chr16	RAC1
6	rs17034	Chr7	C12orf42 - LINC02401
7	rs17807628	Chr2	HLA-DMB
8	rs10207044	Chr8	FAP - IFIH1
9	rs11013059	Chr8	FRY

10	rs1364450	Chr9	HLA-DPA1, HLA-DPB1
11	rs140626545	Chr1	SPATA6
12	rs60392867	Chr6	ZNF429
13	rs11675342	Chr8	LEF1
14	rs12118515	Chr6	SLC1A2-AS1
15	rs17106373	Chr5	VEGFC - LINC02509
16	rs2281333	Chr6	CCDC102A
17	rs10511408	Chr14	AFF1
18	rs3059644	Chr19	PGBD1 - SMIM15P2
19	rs2077510	Chr1	CCL20
20	rs6718188	Chr1	PPHLN1
21	rs7763540	Chr16	PSMD14-DT, LINC01806
22	rs2076899	Chr17	FOXK2
23	rs371858405	Chr1	KLF12
24	rs112185105	Chr3	DACT2 - SMOC2
25	rs6062344	Chr3	ZNF652

Sebagaimana disajikan pada Tabel 4, seluruh kandidat SNP menunjukkan hasil "*no data available*". Tidak tersedianya skor dapat disebabkan oleh keterbatasan anotasi pada model PolyPhen-2, belum tersedianya data pendukung pada model HumDiv atau HumVar, atau kandidat tidak teranotasi sebagai varian nonsinonim pada transkrip pengode yang digunakan dalam analisis. Ketidaksesuaian koordinat genom, alel, versi genom, atau transkrip yang dipilih juga perlu dipertimbangkan dalam proses evaluasi ulang.

Hasil tersebut menunjukkan bahwa dampak kandidat SNP terhadap struktur dan fungsi protein belum dapat ditentukan melalui analisis PolyPhen-2. Oleh karena itu, penelitian ini tidak menyatakan adanya kandidat yang termasuk dalam kategori *probably damaging*, *possibly damaging*, atau *benign*. Klaim sebelumnya mengenai kemungkinan efek merusak pada varian yang dikaitkan dengan *RAC1*, *C12orf42-LINC02401*, dan *FAP-IFIH1* harus dihapus karena tidak didukung oleh skor prediksi yang tersedia.

Tidak tersedianya hasil PolyPhen-2 juga perlu dipertimbangkan bersama dengan hasil verifikasi anotasi genomik. PolyPhen-2 terutama digunakan untuk menilai dampak substitusi asam amino pada protein. Kandidat yang teranotasi sebagai varian intronik, intergenik, regulatori, atau transkrip nonpengode tidak sesuai untuk dinilai sebagai substitusi asam amino menggunakan PolyPhen-2. Oleh karena itu, analisis prediksi fungsional sebaiknya diulang hanya pada kandidat yang telah terkonfirmasi sebagai *missense variant* pada transkrip pengode (Rand et al., 2025).

Secara keseluruhan, hasil SNPnexus dalam penelitian ini menunjukkan keterbatasan ketersediaan informasi prediksi fungsional, bukan bukti mengenai ada atau tidaknya dampak biologis. Penelitian lanjutan perlu melakukan verifikasi rsID, koordinat genom, alel referensi dan alternatif, serta transkrip yang terdampak sebelum mengulangi prediksi menggunakan PolyPhen-2 atau metode komputasional lainnya. Validasi eksperimental tetap diperlukan untuk memastikan pengaruh suatu varian terhadap struktur, aktivitas, atau stabilitas protein.

KESIMPULAN

Penelusuran menggunakan GWAS Catalog menghasilkan 3.280 data asosiasi varian terkait hipotiroidisme. Setelah penyaringan berdasarkan nilai $p < 1 \times 10^{-8}$ dan penghapusan data duplikat,

diperoleh 899 SNP yang dikelompokkan berdasarkan konteks genomiknya. Sebanyak 25 SNP diprioritaskan sebagai kandidat berdasarkan anotasi awal. Evaluasi ulang menunjukkan bahwa 25 kandidat tersebut belum dapat seluruhnya dinyatakan sebagai varian *missense* karena ditemukan ketidaksesuaian antara beberapa rsID, nama gen hasil pemetaan, lokasi kromosom, dan konsekuensi varian. Oleh karena itu, kandidat tersebut diposisikan sebagai hasil skrining awal dan tidak digunakan secara langsung untuk menetapkan gen penyebab atau varian patogenik pada hipotiroidisme. Analisis GTEx memberikan gambaran deskriptif mengenai distribusi ekspresi gen pada berbagai jaringan manusia. Namun, hasil tersebut tidak dapat digunakan untuk menyimpulkan bahwa suatu SNP menyebabkan perubahan ekspresi gen atau berperan langsung dalam hipotiroidisme karena hubungan SNP–gen belum seluruhnya terverifikasi dan data GTEx bukan merupakan perbandingan antara pasien hipotiroidisme dan kelompok kontrol. Analisis SNPnexus dan PolyPhen-2 tidak menghasilkan skor maupun kategori prediksi untuk seluruh kandidat. Dengan demikian, penelitian ini belum dapat mengklasifikasikan kandidat sebagai *benign*, *possibly damaging*, atau *probably damaging*, serta belum dapat menyimpulkan adanya kerusakan fungsi protein. Secara keseluruhan, penelitian ini memberikan kerangka skrining awal varian genetik terkait hipotiroidisme melalui integrasi GWAS Catalog, GTEx Portal, Ensembl, dan SNPnexus. Hasil penelitian bersifat eksploratif dan menegaskan pentingnya verifikasi rsID, koordinat genom, alel, gen pengode, serta transkrip sebelum dilakukan interpretasi biologis. Penelitian lanjutan perlu melibatkan anotasi genomik yang lebih terstandar, analisis eQTL, metode prediksi fungsional tambahan, dan validasi eksperimental untuk memastikan relevansi kandidat varian terhadap hipotiroidisme.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada mahasiswa Universitas Bakti Tunas Husada yang telah berkontribusi aktif dalam penyusunan artikel dan pelaksanaan penelitian ini. Dedikasi dan partisipasi mahasiswa menjadi bagian penting dalam keberhasilan penelitian ini.

DAFTAR PUSTAKA

- Adzhubei, I., Jordan, D. M., & Sunyaev, S. R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. *Current Protocols in Human Genetics*, 2(SUPPL.76). <https://doi.org/10.1002/0471142905.hg0720s76>
- Ahmad, S., Qazi, S., & Raza, K. (2021). Translational bioinformatics methods for drug discovery and drug repurposing. In *Translational Bioinformatics in Healthcare and Medicine* (Vol. 13, pp. 127–139). <https://doi.org/10.1016/B978-0-323-89824-9.00010-0>
- Ardlie, K. G., Deluca, D. S., Segre, A. V., Sullivan, T. J., Young, T. R., Gelfand, E. T., ... GTEx Consortium. (2015). The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science*, 348(6235), 648–660. <https://doi.org/10.1126/science.1262110>
- Broad Institute. (2023). HaploReg v4.2. Pubs.Broad Institute. <https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php>
- Buniello, A., MacArthur, J. A. L., Cerezo, M., Harris, L. W., Hayhurst, J., Malangone, C., ... Parkinson, H. (2019). The NHGRI-EBI GWAS Catalog. *Nucleic Acids Research*, 47(D1), D1005–D1012. <https://doi.org/10.1093/nar/gky1120>
- Burnstock, G. (2017). Purinergic signalling: Therapeutic developments. *Frontiers in Pharmacology*, 8, 1–55. <https://doi.org/10.3389/fphar.2017.00755>
- Damian, J. W., Thorp, J. G., et al. (2023). Identification of drug repurposing candidates for the treatment of anxiety: a genetic approach. *Psychiatry Research*, 326, 115343. <https://doi.org/10.1016/j.psychres.2023.115343>
- Dayem Ullah, A. Z., Oscanoa, J., Wang, J., et al. (2018). SNPnexus: Assessing the functional relevance of genetic variation to facilitate the promise of precision medicine. *Nucleic Acids Research*, 46(W1), W109–W113. <https://doi.org/10.1093/nar/gky399>



- EMBL-EBI. (2025a). Ensembl Genome Browser. Asia.Ensembl.
https://asia.ensembl.org/Homo_sapiens
- EMBL-EBI. (2025b). Ensembl Glossary. Plants.Ensembl.
<https://plants.ensembl.org/info/website/glossary.html>
- Fadista, J., Manning, A. K., Florez, J. C., & Groop, L. (2016). The (in)famous GWAS P-value threshold revisited and updated for low-frequency variants. *European Journal of Human Genetics*, 24(8), 1202–1205. <https://doi.org/10.1038/ejhg.2015.269>
- GWAS. (2025). GWAS Catalog. EMBL-EBI. <https://www.ebi.ac.uk/gwas/>
- Haddad, N., Gamaethige, S. M., Wehida, N., & Elbediwy, A. (2024). Drug repurposing: Exploring potential anti-cancer strategies by targeting cancer signalling pathways. *Biology*, 13(6), 1–15. <https://doi.org/10.3390/biology13060386>
- Institute for Health Metrics and Evaluation. (2021). Global Burden of Disease (GBD) [Online database]. Vizhub. <https://vizhub.healthdata.org/gbd-results/>
- Islam, S. M., Hasan, M. M., Alam, J., Dey, A., & Molineaux, D. (2025). In silico screening, molecular dynamics simulation and binding free energy identify single-point mutations that destabilize p53 and reduce binding to DNA. *Proteins: Structure, Function, and Bioinformatics*, 93(2), 498–514. <https://doi.org/10.1002/prot.26747>
- Kulkarni, V. S., Alagarsamy, V., Solomon, V. R., et al. (2023). Drug repurposing: An effective tool in modern drug discovery. *Russian Journal of Bioorganic Chemistry*, 49(2), 157–166. <https://doi.org/10.1134/S1068162023020139>
- Li, H., Chen, W., & Zhao, Q. (2019). Exonic variants affecting protein function and their role in endocrine disorders. *Frontiers in Genetics*, 10, 1124. <https://doi.org/10.3389/fgene.2019.01124>
- Rand, S. A., Ahlberg, G., Tragante, V., et al. (2025). Genome-wide association study and polygenic risk prediction of hypothyroidism. *Nature Genetics*, 57, 3007–3015. doi:10.1038/s41588-025-02410-Z.
- Rudrapal, M., Khairnar, J., & Jadhav, A. (2020). Drug repurposing (DR): An emerging approach in drug discovery. IntechOpen. <https://doi.org/10.5772/intechopen.89698>
- Ward, L. D., & Kellis, M. (2016). HaploReg v4: Systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Research*, 44(D1), D877–D881. <https://doi.org/10.1093/nar/gkv1340>
- WHO. (2025). Hypothyroidism fact sheet. World Health Organization. <https://www.who.int/news-room/fact-sheets/detail/hypothyroidism>
- Zhang, X., Wang, Y., Li, Z., et al. (2020). Functional impacts of missense variants in human exome across tissues. *Human Molecular Genetics*, 29(3), 402–414. <https://doi.org/10.1093/hmg/ddz248>