

Universidad Católica de Santa María

Escuela de Postgrado

Doctorado en Ciencias de la Salud



**“PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL
CÁNCER DIFERENCIADO DE TIROIDES DIAGNÓSTICADO
EN EL HNCASE ESSALUD, AREQUIPA 2019-2020”**

Tesis presentada por el Maestro:

Cárdenas Abarca, Carlos Arturo

Para optar al grado académico de:

Doctor en Ciencias de la Salud

Asesor:

Dr. Bernabé Ortíz, Julio César

Arequipa-Perú

2022

UCSM-ERP

UNIVERSIDAD CATÓLICA DE SANTA MARÍA
ESCUELA DE POSTGRADO
DICTAMEN APROBACIÓN DE BORRADOR DE TESIS

Arequipa, 18 de Mayo del 2022

Dictamen: 004585-C-EPG-2022

Visto el borrador del expediente 004585, presentado por:

2017009141 - CARDENAS ABARCA CARLOS ARTURO

Titulado:

**PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES
DIAGNÓSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020**

Nuestro dictamen es:

APROBADO

**0913 - GUTIERREZ ARANIBAR ROXANA JACQUELINE
DICTAMINADOR**



**1051 - VILLANUEVA SALAS JOSE ANTONIO
DICTAMINADOR**



**1058 - DAVILA DEL CARPIO GONZALO HERMILIO
DICTAMINADOR**



**1484 - MUÑOZ DEL CARPIO TOIA AGUEDA ROSSANGELLA
DICTAMINADOR**



**3308 - CHAVEZ FUMAGALLI MIGUEL ANGEL
DICTAMINADOR**





Epígrafe

No se nos ha pasado por alto que cuánto más exploramos el genoma humano, más nos queda por explorar. "No cesaremos de explorar. Pues al final de toda exploración llegaremos donde empezamos, y conoceremos cuál es nuestro lugar por primera vez".

Thomas Stearns Eliot



Dedicatoria

Dedico esta tesis a mis difuntos y queridos maestros Dr. Percy Simon Márquez Veliz y la Dra. Cecilia Rossana Aguilar Ramírez; es muy complejo explicar con palabras lo valioso que fueron en mi vida, gracias por enseñarme todo lo que sé y, mas que eso, guiarme para ser una mejor persona y profesional, porque me abrieron las puertas a mi carrera a la que me enseñaron amar y a dar el todo de mi por nuestros pacientes con principios, valores y amor.

Agradecimientos

“A Dios por guiar mis pasos y darme fortaleza cada día. A mis queridos maestros y especialmente a mi tutor por su dedicación, paciencia y motivación. Agradecerle también a mi esposa y a mi familia por acompañarme durante este camino. A mis amigos del doctorado que hicieron de estos años un feliz aprender. Así mismo, mencionar que la presente tesis fue financiada por fondos internos para la Investigación de la UCSM, RESOLUCION 27144-R-2020”.

ÍNDICE GENERAL

RESUMEN.....	ix
ABSTRACT	x
INTRODUCCIÓN	1
HIPÓTESIS	3
OBJETIVOS	3
CAPÍTULO I MARCO TEÓRICO	4
1. Marco teórico	5
1.1 Patogénesis molecular del cáncer de tiroides	5
1.1.1 Mutaciones genéticas puntuales	5
1.1.2 Translocaciones de genes.....	7
1.2 Polimorfismos DNA.....	8
1.3 Micro-ARN.....	8
1.4 Gen BRAF V660E.....	9
1.5 Correlación histopatológica y mutaciones.....	14
1.6 Secuenciación de transcriptomas.....	16
1.6.1 Introducción.....	16
1.6.2. Tecnologías transcriptómicas.....	17
1.6.2.1 Microarrays.....	17
1.6.2.2 Perfiles de transcripciones digitales.....	20
1.6.2.3 Secuenciación de ARN.....	22
1.6.3 Protocolos y preparación de la biblioteca de RNA-Seq.....	23
1.6.3.1 Evaluación de la calidad del ARN.....	24
1.6.3.2 Métodos de selección de secuencia de ARN.....	25
1.6.3.3 Fragmentación y Preparación de la biblioteca.....	27

1.6.4 Análisis de datos de la secuencia de ARN.....	29
1.6.5 Montaje del Transcriptoma.....	31
1.6.6 Evaluación de Calidad de la Secuenciación de Transcriptomas.....	33
1.6.6.1 Material de referencia y Control de Calidad.....	35
1.6.7 Expresión Genética y Detección Alternativa de Empalmes.....	38
1.6.8 Detección de genes y variantes de fusión.....	40
1.6.9 Aplicaciones clínicas de RNA-Seq.....	44
1.6.9.1 Oncología de precisión.....	45
1.7 Aplicaciones genómicas en el Cáncer de Tiroides.....	48
1.7.1 Carcinoma Papilar de Tiroides.....	51
1.7.2 Carcinoma Folicular de Tiroides.....	59
1.7.3 Carcinoma de células de Hürthle.....	61
1.7.4 Carcinoma Pobremente Diferenciado.....	62
1.7.5 Carcinoma Anaplásico de Tiroides.....	63
1.7.6 Aplicaciones Clínicas de la Información Genómica en el Cáncer de Tiroides.....	65
1.8 Visualización del Genoma	66
1.8.1 Software IGV	67
1.8.2 Lectura de datos NGS con el software IGV.....	69
1.9 Visualización de fusiones con software Arriba	71
CAPÍTULO II METODOLOGÍA	72
2. Metodología	73
2.1 Lugar y tiempo de realización del estudio.....	73
2.2 Revisión de historias clínicas.....	73
2.3 Población.....	73
2.4 Muestra biológica.....	73

2.4.1 Recolección de la muestra biológica.....	73
2.4.2 Extracción del ARN.....	74
2.4.3 Extracción de ARN total.....	75
2.4.4 Secuenciamiento del Transcriptoma con el Protocolo Illumina.....	76
2.5 Visualización de las alineaciones del formato BAM con el genoma de referencia utilizando el software IGV.....	78
2.6 Identificación de fusiones de genes.....	79
2.7 Evaluación de la calidad de datos secuenciados con NGS.....	79
2.8 Análisis estadístico.....	81
CAPÍTULO III RESULTADOS Y DISCUSIÓN	82
CONCLUSIONES.....	144
RECOMENDACIONES	146
PROPUESTA	147
BIBLIOGRAFÍA.....	148
ANEXOS.....	187
ANEXO N° 1: FICHA DE RECOLECCIÓN DE DATOS	188

ÍNDICE DE TABLAS

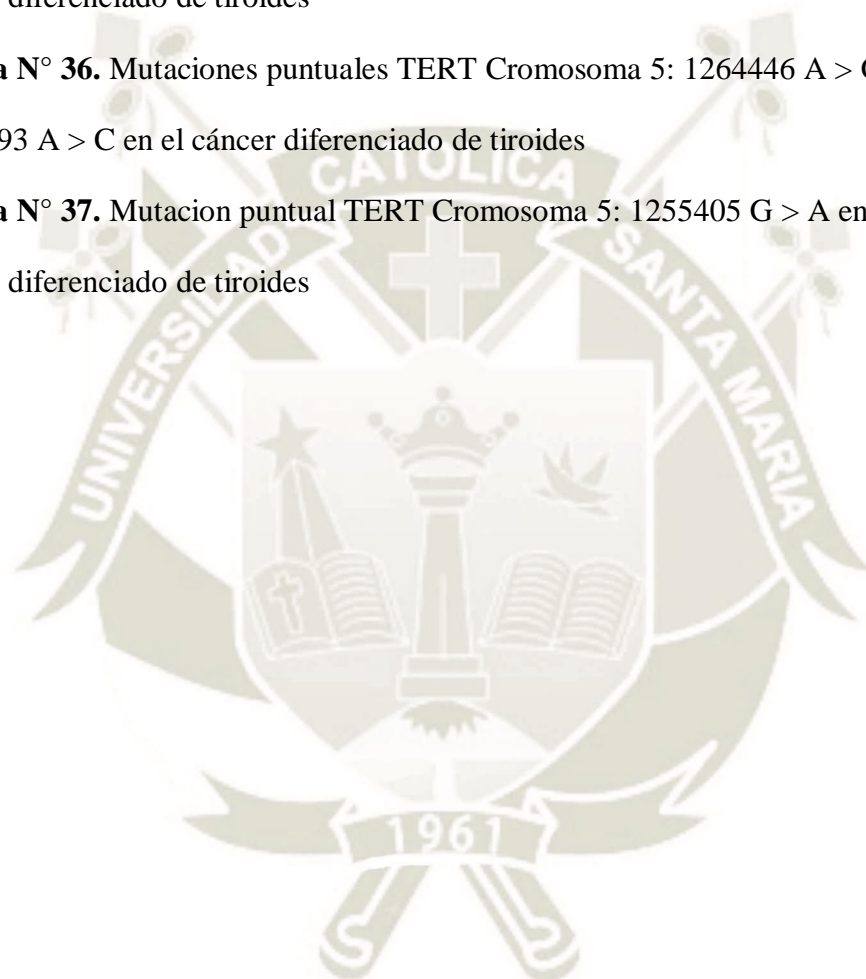
Tabla N°1. Clinical trials dirigidos a inhibidores BRAF en cáncer de tiroides	12
Tabla N°2. Comparación de los métodos de investigación del transcriptoma	17
Tabla N°3. Software bioinformático para análisis de secuenciación del transcriptoma	29
Tabla N°4. Resumen de las características genéticas y genómicas de los cánceres de tiroides	51
Tabla N°5. Visualización de software: una comparación	67
Tabla N°6. Distribución por grupo étnico y sexo de pacientes con el diagnóstico de Cáncer Diferenciado de Tiroides	83
Tabla N°7. Treinta primeros genes con mayor frecuencia de expresión de los casos de cáncer diferenciado de tiroides del 2019 al 2020 (parte 1)	89
Tabla N°7. Treinta primeros genes con mayor frecuencia de expresión de los casos de cáncer diferenciado de tiroides del 2019 al 2020 (parte 2)	90
Tabla N°8. Función de los treinta primeros genes con mayor frecuencia de expresión de los casos de cáncer diferenciado de tiroides del 2019 al 2020 (parte 1)	91
Tabla N°8. Función de los treinta primeros genes con mayor frecuencia de expresión de los casos de cáncer diferenciado de tiroides del 2019 al 2020 (parte 2)	92
Tabla N°9. Mutación P53 y otras asociadas a P53 en el cáncer diferenciado de tiroides	94
Tabla N°8. Mutación P53 y otras asociadas a P53 en el cáncer diferenciado de tiroides	95
Tabla N°9. Traslocaciones tipo fusión en el Cáncer diferenciado de tiroides	96
Tabla N°10. Genes componentes de traslocaciones tipo fusión en el Cáncer diferenciado de tiroides	97
Tabla N°12. Ubicación de las mutaciones puntuales BRAF en el cáncer diferenciado de tiroides	101
Tabla N°13. Mutaciones puntuales TERT en el cáncer diferenciado de tiroides	108
Tabla N°13. Mutaciones puntuales TERT en el cáncer diferenciado de tiroides	109
Tabla N°13. Mutaciones puntuales TERT en el cáncer diferenciado de tiroides	110

ÍNDICE DE FIGURAS

Figura N°1. Desdiferenciación del cáncer de tiroides	10
Figura N° 2. Vía de señalización MAPK, del cáncer de tiroides	10
Figura N° 3. Perfil mutacional de células foliculares tiroideas	15
Figura N° 4. Alteraciones genéticas del cáncer papilar de tiroides según TCGA	15
Figura N° 5. Visión general de la preparación de la Biblioteca de RNA-Seq	23
Figura N° 6. Datos del análisis RNA-Seq	30
Figura N° 7. Montaje de Transcriptoma	32
Figura N° 8. La puntuación BRAF ^{V600E} -RAS (BRS)	56
Figura N° 9. Papel de la diferenciación tiroidea	57
Figura N° 10. Ventana de aplicación de IGV	68
Figura N° 11 Vista de lectura de alineación a 20 kb y resolución de pares de bases	70
Figura N° 12. Análisis de calidad bioinformática de los resultados de NGS	80
Figura N° 13. Distribución por grupo étnico y sexo de pacientes con el diagnóstico de Cáncer Diferenciado de Tiroides	83
Figura N° 14. Características citológicas, histopatológicas y sonográficas de los pacientes con el diagnóstico de Cáncer Diferenciado de Tiroides	85
Figura N° 15. Genes con mayor frecuencia de expresión de los casos de cáncer diferenciado de tiroides del 2019 al 2020 (parte 1)	86
Figura N° 15. Genes con mayor frecuencia de expresión de los casos de cáncer diferenciado de tiroides del 2019 al 2020 (parte 2)	87
Figura N° 15. Genes con mayor frecuencia de expresión de los casos de cáncer diferenciado de tiroides del 2019 al 2020 (parte 3)	88
Figura N° 16. Expresión de TERT, BRAF y P53 en el Cáncer Diferenciado de Tiroides	93
Figura N° 17. Traslocación/fusión TDG-TMEM132B del Cáncer diferenciado de tiroides	98
Figura N° 18. Traslocación/fusión MEIS1-ITSN2 del cáncer diferenciado de tiroides	99
Figura N° 19. Traslocación/fusión KCNE1B-FP671120.4 del cáncer diferenciado de tiroides	100

Figura N° 20. Mutación puntual BRAF Cromosoma 7: 140924572 C > T en el cáncer diferenciado de tiroides	102
Figura N° 21. Mutaciones puntuales BRAF Cromosoma 7: 140850154 T > A; 140850164 T > A en el cáncer diferenciado de tiroides	103
Figura N° 22. Mutaciones puntuales BRAF Cromosoma 7: 140800452 T > A en el cáncer diferenciado de tiroides	104
Figura N° 23. Mutación puntual BRAF Cromosoma 7: 140794401 T > A, C, G en el cáncer diferenciado de tiroides	105
Figura N° 24. Mutación puntual BRAF Cromosoma 7: 140777028 A > T en el cáncer diferenciado de tiroides	106
Figura N° 25. Mutación puntual BRAF Cromosoma 7: 140754195 A > T en el cáncer diferenciado de tiroides	107
Figura N° 26. Mutaciones puntuales TERT Cromosoma 5: 1293763 T > C, 1293765 C > G, 1293766 C > G en el cáncer diferenciado de tiroides	111
Figura N° 27. Mutaciones puntuales TERT Cromosoma: 5:1293910 C > A, 1293913 A > G, 1293914 C > A en el cáncer diferenciado de tiroides	112
Figura N° 28. Mutaciones puntuales TERT Cromosoma 5: 1294136 G > T en el cáncer diferenciado de tiroides	113
Figura N° 29. Mutaciones puntuales TERT Cromosoma 5: 1294238 A > C en el cáncer diferenciado de tiroides	114
Figura N° 30. Mutaciones puntuales TERT Cromosoma 5: 1294313 A > C, 1294315 T > G en el cáncer diferenciado de tiroides	115
Figura N° 31. Mutaciones puntuales TERT Cromosoma 5: 1294430 C > T, 1294479 G > A en el cáncer diferenciado de tiroides	116
Figura N° 32. Mutacion puntual TERT Cromosoma 5: 1280318 T > C en el cáncer diferenciado de tiroides	117

Figura N° 33. Mutacion puntual TERT Cromosoma 5: 1278733 T > A en el cáncer diferenciado de tiroides	118
Figura N° 34. Mutaciones puntuales TERT Cromosoma 5: 1272217 T > C, 1272218 G > T, 1272221 T > C en el cáncer diferenciado de tiroides	119
Figura N° 35. Mutacion puntual TERT Cromosoma 5: 1264568 T > C en el cáncer diferenciado de tiroides	120
Figura N° 36. Mutaciones puntuales TERT Cromosoma 5: 1264446 A > G, 1264493 A > C en el cáncer diferenciado de tiroides	121
Figura N° 37. Mutacion puntual TERT Cromosoma 5: 1255405 G > A en el cáncer diferenciado de tiroides	122



LISTA DE ABREVIATURAS

- ADN:** Ácido desoxirribonucleico
- ADNc:** Ácido desoxirribonucleico complementario
- ARN:** Ácido ribonucleico
- ARNm:** Ácido ribonucleico mensajero
- ARNnc:** Ácido ribonucleico no codificante
- BAM:** Binary Alignment Map (mapa del alineamiento binario)
- NGS:** Next Generation Sequencing (Secuenciamiento de Próxima Generación)
- PE:** Paired End (final pareado)
- RNA-seq:** RNA sequencing
- SAM:** Sequence Alignment Map (Mapa del alineamiento de la secuencia)
- SE:** Single End (final singular)
- TLR:** Receptores de Tipo Toll
- CT:** Cáncer de tiroides
- CDT:** Cáncer diferenciado de tiroides
- CPT:** Cáncer papilar de tiroides
- CFT:** Cáncer folicular de tiroides
- CMT:** Cáncer medular de tiroides
- CAT:** Cáncer anaplásico de tiroides
- TG:** Tiroglobulina
- TPO:** Tiroperoxidasa
- Ab ANTI TG:** Anticuerpos anti tiroglobulina
- TCGA:** The Cancer Genome Atlas
- TSH:** Hormona estimulante de la tiroides
- rTSH:** Receptor de TSH
- I -131:** Iodo radioactivo 131

GRCh38.p7: Genome Reference Consortium Human Build 38 patch release 7

BRAF^{V600E}: B-Raf proto-oncogene valine is replaced by a glutamic acid residue

RAS: Rat sarcoma virus

MEK: Extracellular-signal-regulated kinase

MAPK: Mitogen-activated protein kinases

RET: Ret proto-oncogene

PTC1: Protein patched homolog 1

AKT: Protein kinase B.

PI3K: Phosphatidylinositol-4,5-bisphosphate 3-kinase

PCBD2: Pterin-4 alfa-carbinolamina deshidratasa 2

LRMDA: Asociada a diferenciación de melanocitos ricos en leucina.

WWOX: Dominio WW que contiene oxidorreductasa

RPS29: Proteína ribosomal S29

RBFOX1: RNA de unión al homólogo 1 de Fox

FHIT: Tríada de histidina frágil diadenosina trifosfatasa.

NTM: Neurotrimina

RAD51B: RAD51 paralog B

PRKN: Parkin RBR E3 ubiquitina proteína ligasa

MACROD2: Mono-ADP ribosilhidrolasa 2

ASIC2: Subunidad 2 del canal de iones de detección de ácido

TBC1D5: Miembro 5 de la familia del dominio TBC1

GPC5: Glipicano 5

EXOC4: Componente del complejo de exocistos 4

IMMP2L: Subunidad 2 de la peptidasa de la membrana mitocondrial interna

CAMTA1: Activador de la transcripción de unión a calmodulina 1

CDH13: Cadherina 13

SMYD3: Dominio SET y MYND que contiene 3.

ANKS1B: Repetición de anquirina y dominio de motivo alfa estéril que contiene 1B.

DAB1: Proteína adaptadora DAB 1

DLG2: Discos grandes proteína de andamio MAGUK 2

FRMD4A: Dominio FERM que contiene 4^a

KCNIP4: Proteína que interactúa con el canal dependiente de voltaje de potasio 4

CTNNA2: Catenina alfa 2

AGBL4: Carboxipeptidasa 4 de AGBL

PRKCE: Proteína quinasa C épsilon

SYN3: Sinapsina III

ARHGAP15: Proteína activadora de Rho GTPasa 15

PACRG: Parkin coregulado

C9orf92: Marco de lectura 92 abierto del cromosoma 9

TDG: Timina-ADN glicosilasa

TMEM132B: Proteína transmembrana 132B

MEIS1: Meis homeobox 1

ITSN2: Intersectina 2

KCNE1B: Subfamilia de canales regulados por voltaje de potasio E subunidad reguladora 1B

FP671120.4: RNA ribosomal no codificante

lncRNAs: Long non-coding RNAs

SNV: Single nucleotide variant

TI-RADS: Thyroid Imaging Reporting and Data System

Resumen

Los avances en la tecnología de secuenciación de próxima generación (NGS) RNA-Seq están permitiendo revelar el genotipo del cáncer diferenciado de tiroides (CDT). Poco se conoce acerca de los genes que están activos en esta patología. Por lo que la presente tesis tiene como objetivo describir el perfil de expresión del CDT de la población Arequipeña. Se enroló un total de 11 pacientes con CDT sometidos a tiroidectomía entre 2019-2020 del Hospital Nacional Carlos Alberto Seguí Escobedo EsSalud Arequipa. El ARN total fue aislado de las biopsias obtenidas de los pacientes y estas fueron sometidas a una secuenciación masiva mediante RNA-Seq utilizando el protocolo Illumina. También se revisaron datos clínicos para su caracterización y análisis molecular de dicha patología. El análisis epidemiológico de los pacientes muestra un predominio del grupo etario de 40-49 años de edad (45.45%), a predominio femenino (63.64%). El 100% de casos tuvo escore TI-RADS V, con escores Bethesda V y VI en el 36.36% y 64.4% respectivamente; 72.73% de tumores fueron unifocales. Las variantes histopatológicas más frecuentes fueron papilar (81.82%) y folicular (18.18%), encontrando 2 variantes histológicas agresivas del carcinoma papilar, la columnar y estroma fibroso tipo fascitis-like, ambas en el 9.09%. La tiroiditis crónica coexistió en el 18.18% de casos. El estadio clínico predominante fue el de grado I (72.73%) seguido por el grado II (27.27%). El análisis bioinformático de la secuenciación masiva (RNA Seq), mediante el software IGV y utilizando el genoma GRCh38.p13, mostró una prevalencia de las mutaciones en los genes BRAF, TERT y P53 de todos los pacientes. Los 30 primeros genes con mayor frecuencia de expresión fueron: PCB2, LRMDA, WWOX, RPS29, RBFOX1, FHIT, NTM, RAD51B, PRKN2, MACROD2, ASIC2, TBC1D5, GPC5, EXOC4, IMMP2L, CAMTA1, CDH13, SMYD3, ANKS1B, DAB1, DLG2, FRMDA4A, KCNIP4, CTNNA2, AGLB4, PRKCE, SYN3, ARHGAP15, PACRG, C9orf92. Por otro lado, el análisis bioinformático de las secuencias usando el software Arriba (versión 1.2.0) evidencia la presencia de 3 proteínas de fusión descritas por primera vez en esta patología como son: TDG-TMEM132B, MEIS1-ITSN2, KCNE1B-FP671120.4. La tecnología RNA Seq revela que las muestras analizadas presentan un perfil de expresión estándar, se encontró 5 nuevas mutaciones para el gen BRAF, 18 nuevas mutaciones para el gen TERT y 3 proteínas de fusión que por primera vez son descritas en la presente tesis.

Palabras clave: Cáncer diferenciado de tiroides, perfil de expresión génica, NGS, RNA-Seq.

Abstract

Advances in next-generation sequencing (NGS) RNA-Seq technology are making it possible to reveal the genotype of differentiated thyroid cancer (DTC). Little is known about the genes that are active in this pathology. Therefore, this thesis aims to describe the expression profile of the CDT of the Arequipeña population. A total of 11 patients with DTC who underwent thyroidectomy between 2019-2020 at the Carlos Alberto Seguí Escobedo EsSalud Arequipa National Hospital were enrolled. Total RNA was isolated from the biopsies obtained from the patients and these were subjected to massive sequencing by RNA-Seq using the Illumina protocol. Clinical data was also reviewed for its characterization and molecular analysis of said pathology. The epidemiological analysis of the patients shows a predominance of the age group of 40-49 years of age (45.45%), with a female predominance (63.64%). 100% of cases had a TI-RADS V score, with Bethesda V and VI scores of 36.36% and 64.4%, respectively; 72.73% of tumors were unifocal. The most frequent histopathological variants were papillary (81.82%) and follicular (18.18%), finding 2 aggressive histological variants of papillary carcinoma, columnar and fasciitis-like fibrous stroma, both in 9.09%. Chronic thyroiditis coexisted in 18.18% of cases. The predominant clinical stage was grade I (72.73%) followed by grade II (27.27%). Bioinformatic analysis of massive sequencing (RNA Seq), using the IGV software and using the GRCh38.p13 genome, showed a prevalence of mutations in the BRAF, TERT and P53 genes in all patients. The first 30 genes with the highest expression frequency were: PCB2, LRMDA, WWOX, RPS29, RBFOX1, FHIT, NTM, RAD51B, PRKN2, MACROD2, ASIC2, TBC1D5, GPC5, EXOC4, IMMP2L, CAMTA1, CDH13, SMYD3, ANKS1B, DAB1, DLG2, FRMDA4A, KCNIP4, CTNNA2, AGLB4, PRKCE, SYN3, ARHGAP15, PACRG, C9orf92. On the other hand, the bioinformatic analysis of the sequences using the Arriba software (version 1.2.0) shows the presence of 3 fusion proteins described for the first time in this pathology: TDG-TMEM132B, MEIS1-ITSN2, KCNE1B-FP671120.4. RNA Seq technology reveals that the analyzed samples present a standard expression profile, 5 new mutations were found for the BRAF gene, 18 new mutations for the TERT gene and 3 fusion proteins that are described for the first time in this thesis.

Keywords: Differentiated thyroid cancer, gene expression profile, NGS, RNA-Seq.

INTRODUCCIÓN

El cáncer de Tiroides ocupa actualmente el 9no en frecuencia a nivel mundial, con una incidencia de aproximadamente de 7.4 nuevos casos por cada 100,000 habitantes, realidad que no escapa a nuestro país en el que se encuentra en el 8vo lugar de frecuencia con una incidencia aproximada de 7.5 nuevos casos por 100,000 habitantes, según datos de Globocan (Global Cancer Observatory) para el año 2018 ⁽¹⁾. Es importante resaltar que la incidencia de esta neoplasia ha incrementado a nivel mundial, sin embargo, la mortalidad permanece con cifras muy similares a datas de décadas anteriores, evidencia citada en múltiples revisiones ⁽²⁻³⁾ y textos de consulta actual ⁽⁴⁾. En la literatura nacional se cuenta con poca información relacionada a cifras epidemiológicas de esta enfermedad, sin embargo, hay una revisión actual descrita por Noé Atamary Anahuai, investigador de la Universidad San Ignacio de Loyola, quien en el año 2018 realiza una revisión nacional con data del Ministerio de Salud ⁽⁵⁾, concluyendo que la prevalencia de esta enfermedad ha incrementado en nuestro país, siendo la región de la sierra la más comprometida, pero al igual, con tasas de mortalidad que no se han modificado. Aun así, hay información del incremento de la presentación de variantes agresivas en distintas poblaciones mundiales, donde parece jugar un rol importante la deficiencia de yodo ⁽⁶⁾, muchas veces independientes del tamaño tumoral ⁽⁷⁾.

En lo que respecta a los nuevos avances en diagnóstico y tratamiento del cáncer tiroideo, se han desarrollado novedosas técnicas moleculares ⁽⁸⁾, las cuales son de gran ayuda para la estratificación de agresividad de esta enfermedad ⁽⁹⁾, y búsqueda de blancos terapéuticos ⁽¹⁰⁾, estos de valiosa importancia en situaciones de variantes agresivas de esta neoplasia, en la que se ha visto que expresan mutaciones del Gen BRAF^{V660E} ⁽¹¹⁾.

Técnicas de última generación vienen cumpliendo estos objetivos brindándonos una mejor comprensión de las variables genéticas y un entendimiento de la biología tumoral con

las herramientas de secuenciación de nueva generación NGS ⁽¹²⁾, a través de la aplicación de mRNA-Seq con la tecnología de secuenciación por síntesis, en la que fragmentos de ARN segmentados se copian en el ADNc que servirán para generar ADN bicatenario, lo que nos puede brindar información de la predicción de recurrencia de esta enfermedad ⁽¹³⁾.

Así como se ha visto el incremento de la incidencia de esta enfermedad ⁽¹⁾, también lo ha hecho su agresividad ⁽⁷⁾, en la que los tratamientos habituales, como la cirugía, iodo radioactivo, hormono supresión y quimioterapia, tienen limitaciones, así mismo, en estas situaciones clínicas, en las que en otras realidades sanitarias, sin llegar muy lejos tales como Argentina ⁽¹⁴⁾, se hacen determinaciones moleculares para poder pesquisar así mutaciones que explican esta agresividad, aplicando terapia molecular dirigida a estas alteraciones, mejorando así el tratamiento ⁽³⁾ y calidad de vida de esta enfermedad.

Alumnos del doctorado en Ciencias de la Salud de la Escuela de Posgrado de la Universidad Católica de Santa María (UCSM), Arequipa, lograron ganar un fondo concursable de investigación de la propia universidad, cuyo aporte financiero permitió implementar un laboratorio de Biología Molecular. Con esta fortaleza y oportunidad, y en vista a la problemática descrita, es que nace la iniciativa de poder demostrar la presencia de las diferentes mutaciones descritas en la literatura mundial en pacientes diagnosticados y tratados por cáncer de tiroides en el HNCASE EsSalud Arequipa en el periodo descrito.

HIPÓTESIS

- Dado que en la actualidad contamos con técnicas para estudios del transcriptoma y determinación del perfil de expresión génica de muchos cánceres, como el proyecto del Atlas Genómico del Cáncer (TCGA) para el cáncer de tiroides, es probable que la técnica RNA-Seq permita determinar el perfil de expresión génica mutacional de las células tumorales de pacientes diagnosticados con cáncer diferenciado de tiroides en el HNCASE EsSalud Arequipa 2019-2020.

OBJETIVOS

Objetivo General

- Describir y analizar el perfil de expresión génica de células tumorales de pacientes diagnosticados con cáncer diferenciado de tiroides en el HNCASE EsSalud Arequipa, mediante la tecnología de RNA-Seq.

Objetivos Específicos

- Describir las características clínicas y epidemiológicas del cáncer diferenciado de tiroides.
- Determinar el perfil de expresión génica de las biopsias del cáncer diferenciado de tiroides y compararlas con estándares referenciales.
- Analizar y describir las mutaciones encontradas en el gen BRAF.
- Analizar y describir las mutaciones encontradas en el gen TERT.
- Buscar la presencia de las proteínas de fusión en las secuencias obtenidas por RNA Seq mediante softwares bioinformáticos.



1. Marco Teórico

1.1. Patogénesis molecular del cáncer de tiroides

De manera similar a otros cánceres sólidos, el cáncer de tiroides se inicia por alteraciones genéticas y cambios epigenéticos en los oncogenes conductores o genes supresores tumorales ^(15,16). Los recientes avances en el diagnóstico molecular del cáncer de tiroides, proporcionan estrategias de tratamiento individualizados ⁽¹³⁾. Ya es conocida que la mutación genética en la tumorigénesis del cáncer medular de tiroides (CMT) es la activación de la mutación del oncogén RET ⁽¹⁷⁾. La mutación RET no está presente en aquellos tumores derivados de las células foliculares tiroideas, que representan las neoplasias tiroideas más comunes, tales como el adenoma folicular, carcinoma papilar (CPT) y folicular (CFT) bien diferenciados ⁽¹⁷⁾. Los carcinomas pobremente diferenciados y anaplásicos se conocen que se desarrollan a consecuencia de la dediferenciación de un CPT o CFT bien diferenciados ⁽¹⁷⁾. En estos cánceres derivados de las células foliculares, se identifican otras alteraciones moleculares en las vías RAS y PI3K-AKT ^(16,17).

1.1.1. Mutaciones genéticas puntuales

La mutación puntual BRAF (T1799A) en el exón 15 conduce a la expresión de la proteína mutante BRAF^{V600E}, la que da como resultado la activación constitutiva serina/treonina quinasa ^(18,19). La mutación BRAF^{V600E} es una de las alteraciones genéticas más comunes en el cáncer de tiroides, que ocurre en aproximadamente el 45% (30 a 70%) de los CPT esporádicos y el 15% de la variante folicular de CPT ⁽²⁰⁾. En particular, del 80 al 100% de la variante de células altas del CPT, alberga la mutación BRAFV600E ⁽²⁰⁾. La mutación BRAF^{V600E} predice peores resultados clínicos en la evolución del CPT, que incluyen características patológicas agresivas y mayor tasa de recurrencia ^(21,22). Se considera que la mutación BRAF^{V600E} puede causar pérdida de la avidéz del yodo radiactivo por parte

de las células cancerígenas tiroideas y, como resultado, la refractariedad al tratamiento con yodo radiactivo ^(21,22).

La mutación puntual RAS se ubica con frecuencia en el cáncer de tiroides, así como en otros cánceres sólidos ⁽²³⁾. Entre los genes de las tres isoformas RAS (HRAS, KRAS y NRAS), la mutación NRAS predomina en tumores tiroideos ⁽²³⁾. La mutación RAS es relativamente rara (0-20%) en la forma habitual del CPT, mientras que casi la mitad de CFT y la variante folicular del CPT albergan la mutación RAS ^(23, 24). La mutación RAS se observa en aproximadamente el 20% de los adenomas foliculares, lo que sugiere que la mutación RAS ocurre tempranamente en la tumorigenesis ^(23, 24). La mutación RAS amortigua la actividad de la GTPasa, perpetuando constitutivamente su actividad, así mismo la mutación RAS activa la vía PI3K-AKT en la tumorigenesis tiroidea ^(23, 24).

El gen supresor tumoral PTEN es un regulador negativo de la vía de señalización de PI3K-AKT, mediante la función opuesta a la de PI3K. La mutación o la delección de PTEN originan la tumorigenesis de células tiroideas foliculares, entidad bien conocida como la enfermedad (síndrome) de Cowden ⁽²⁵⁾. El síndrome Cowden es una enfermedad hereditaria autosómica causada por mutaciones de línea germinal de PTEN ⁽²⁵⁾. La enfermedad de Cowden es un síndrome cancerígeno estrechamente relacionado con la predisposición a un mayor riesgo de cánceres de tiroides, mama y endometrio, así como de tumores benignos como los hamartomas ⁽²⁵⁾. La alteración PTEN se observa en el 40% de CFT ⁽²⁵⁾. El silenciamiento de PTEN por hipermetilación de su promotor, también se ha encontrado en los CFT y carcinoma de tiroides anaplásico (CAT) ⁽²⁶⁻²⁸⁾.

Otros genes también están mutados en el cáncer de tiroides, tales como CTNNB1, TP53, IDH1 y NDUFA13 (GRIM19) ⁽²⁹⁾. CTNNB1 está involucrado en la vía WNT- β -catenina y a menudo se encuentra mutado en el CAT ⁽³⁰⁾. TP53 codifica el supresor tumoral p53 y está

involucrado en gran variedad de cánceres sólidos. La mutación TP53 se observa con frecuencia en el CAT (70-80%)^(31,32). Las mutaciones CTNNB1 y TP53 se han observado preferentemente en el CAT o el carcinoma pobremente diferenciado de tiroides, lo que sugiere que estas alteraciones genéticas pueden estar asociadas con desdiferenciación o eventos tardíos de la célula folicular en la progresión del cáncer^(31, 32). Aunque el cáncer de tiroides de células de Hürthle no conlleva alteraciones genéticas comunes como BRAFV600E, RAS o RET/PTC⁽³³⁾, un 15% alberga mutaciones de NDUFA13 (GRIM19)⁽³⁴⁾.

1.1.2. Translocaciones de genes

Se encuentran reordenamientos cromosómicos del protooncogén de tirosina quinasa RET, específicamente reordenamientos RET/PTC en el 15-45% de los CPT esporádicos y 80% del CPT inducido por radiación^(35, 36). RET es un protooncogén que codifica el receptor de tirosina. quinasa (RTK). En consecuencia, hay una activación de las vías MAPK y PI3K-AKT^(35, 36). Entre las numerosas translocaciones tipo de RET/PTC, la RET/PTC1 (fusión con CCDC6) y RET/PTC3 (fusión con NCOA4) son las más frecuentes^(37, 38).

En los casos del accidente de Chernobyl, la translocación RET/PTC3 fue comúnmente observada, seguida por la translocación RET/PTC1 tardíamente⁽³⁹⁾. La translocación del gen de la caja emparejada 8 (PAX8) /receptor γ activado por proliferación de peroxisomas (PPARG) es otro oncogén recombinante común en el cáncer de tiroides⁽³⁹⁾. La translocación PAX8/PPARG se observa a menudo en el CFT (25-60%) y en la variante folicular del CPT (~30%)⁽³⁹⁾. La proteína de fusión PAX8/PPARG actúa como un inhibidor dominante negativo de la versión tipo salvaje del supresor tumoral PPAR γ ⁽⁴⁰⁾. Por el contrario, los efectos de la translocación PAX8/PPARG en la función de PAX8 sigue siendo difícil de dilucidar⁽⁴⁰⁾.

1.2. Polimorfismos DNA

Según avances recientes en el análisis de todo el genoma, se ha demostrado que algunas variantes genéticas afectan la susceptibilidad a padecer de cáncer de tiroides diferenciado (CDT), aunque es probable que tenga bajo impacto en la tumorigenesis ⁽⁴¹⁾. El polimorfismo del DNA es una variación de la secuencia de DNA que ocurre comúnmente en parte de la población, donde la frecuencia mínima se toma normalmente como 1% ⁽⁴¹⁾. Se necesitan asociaciones de polimorfismo de DNA con el CDT para ser confirmada ⁽⁴¹⁾. El genotipo combinado GSTN1-nulo/GSTT1-nulo (genes relacionados con la enzima II fase citosólica) y portadores homocigotos del alelo P53 72Pro se reportan que están asociados con un alto riesgo del CDT ⁽⁴²⁻⁴⁴⁾.

En el estudio a gran escala de Islandia, dos variantes comunes localizadas en 9q22.33 y 14q13.3 están asociadas con CDT ⁽⁴⁵⁾. El gen más cercano al locus 9q22.33 es FOXE1 (TTF2), y NKX2-1 (TTF1) se encuentra entre los genes ubicados en el locus 14q13.3 ⁽⁴⁵⁾. Estos genes codifican importantes factores de transcripción asociados con funciones tiroideas. Individuos con las variantes homocigotos para ambos mostraron un riesgo 5,7 veces mayor de CDT que los no portadores ⁽⁴⁵⁾.

1.3. Micro-ARN

El micro-ARN (miR) es un pequeño (19-25 nucleótidos) ARN no codificante que regula negativamente la expresión de genes codificantes ⁽⁴⁶⁾. Se considera que los miR regulan alrededor del 30% del genoma y pueden actuar como genes supresores tumorales u oncogenes ⁽⁴⁷⁾. Un polimorfismo en un miR (miR-146a) y otros numerosos miR implicados en las principales vías de señalización (principalmente PTEN-PI3K-AKT), sugieren que son importantes en la carcinogénesis del CDT ⁽⁴⁸⁾.

1.4. Gen BRAF V660E

El gen BRAF (homólogo B1 del oncogén viral del sarcoma murino v-raf) se ubica en el brazo largo del cromosoma 7 (7q34) y codifica una proteína citoplásmica de 18 exones, una proteína quinasa de serina/treonina (B-Raf) que es reclutada en la membrana celular tras la estimulación por factores de crecimiento ^(49,50). El gen BRAF de tipo salvaje, es un efector corriente abajo dentro de la vía de señalización ERK/MAPK, la cual regula el crecimiento, proliferación, diferenciación y apoptosis de las células humanas ⁽⁵⁰⁾.

La señalización química a través de esta vía es de vital importancia para el desarrollo normal antes del nacimiento ⁽⁵⁰⁾. El gen BRAF proporciona instrucciones para la transmisión de señales químicas desde el exterior de la célula hacia el núcleo, siendo la vía de señalización de MAPK iniciada mediante la activación de un receptor de membrana de tirosina quinasa ⁽⁵¹⁾. Esta señal, a través de la activación de RAS, facilita la homo o heterodimerización de BRAF de tipo salvaje. BRAF activado fosforila a MEK, que, a su vez, fosforila a ERK, lo que resulta en múltiples efectos celulares como la inducción de la proliferación y supervivencia celular ^(51,52). (Figura 1)

Cuando el gen BRAF muta, adquiere propiedades oncogénicas. Los oncogenes tienen el potencial de promover la transformación de células normales a malignas ⁽⁴⁹⁻⁵¹⁾. La activación de mutaciones de BRAF conduce a la dimerización constitutiva de la proteína BRAF, que a su vez activa la cascada de señalización RAF-MEK-ERK, promoviendo así la proliferación celular mientras inhibe la apoptosis ⁽⁵⁰⁾. El resultado final de esta secuencia de eventos es impulsar el desarrollo y crecimiento del cáncer ^(51,52) (Figura 2).

Actualmente, se han identificado más de 40 mutaciones del gen BRAF, la mayoría de las cuales dan como resultado cambios en el dominio de la quinasa y el bucle P de la molécula.

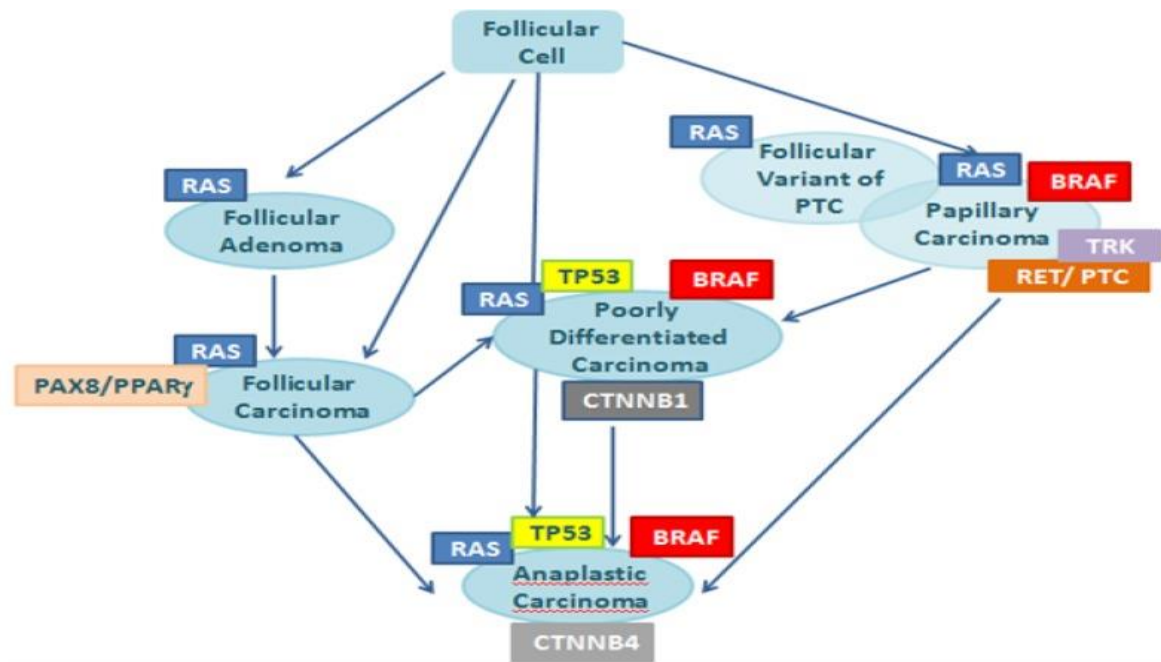


Figura 1. Desdiferenciación escalonada del cáncer de tiroides derivado de células foliculares. Tomadas de Khatami F et al., 2018 ^(51, 52).

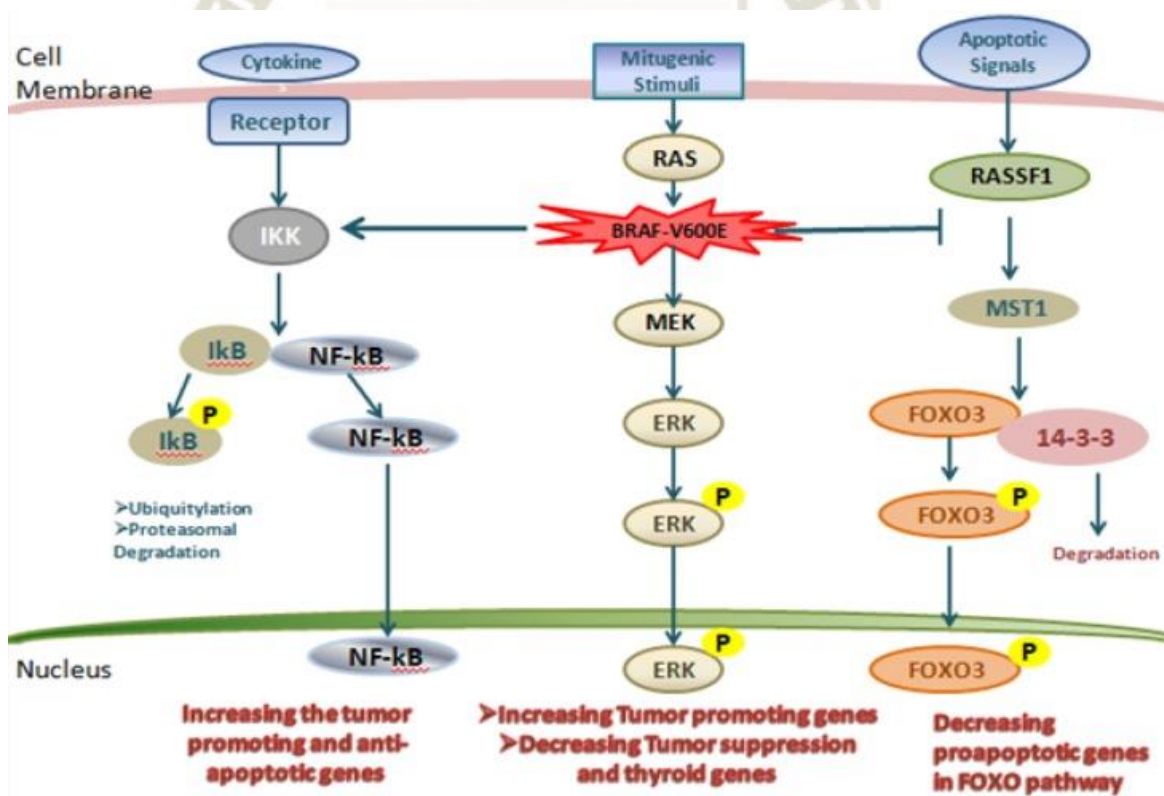


Figura 2. La vía de señalización MAPK y otras relacionadas en la oncogénesis del cáncer de tiroides. Tomadas de Khatami F et al., 2018 ^(51, 52).

Estos productos BRAF mutados fosforilan activamente MEK ⁽⁴⁹⁾. Alrededor del 80% de estas alteraciones genéticas corresponden a la versión trans del hotspot T1799A, que causa la

mutación activadora V600E, que se origina en la sustitución de valina por ácido glutámico en el aminoácido (aa) 600 ⁽⁴⁹⁾. El 20% restante de las alteraciones genéticas de BRAF explican una amplia gama de mutaciones sin sentido; estos residen en las glicinas del bucle G (exón 11) o en el segmento de activación (exón 15) cerca del V600 ⁽⁵³⁾. Una fusión en marco del gen AKAP9 (exones 1 a 8) con el gen BRAF (exones 9 a 18), que se produce a través de una inversión paracéntrica del cromosoma 7, se ha reconocido preferentemente en los CPT inducidos por radiación, en comparación con las mutaciones puntuales de BRAF ⁽⁵³⁾. La mutación V600E confiere actividad transformadora a las células imitando la fosforilación de T599 y/o S602 en el segmento de activación, con la consecuencia de que BRAF permanece constitutivamente activo de manera independiente de RAS ⁽⁵³⁾. Los estudios mutacionales y la cristalografía han demostrado que la mutación BRAFV600E desestabiliza la conformación inactiva de la enzima y produce una quinasa constitutivamente activa con una actividad 500 veces mayor ⁽⁵⁴⁾. Las mutaciones BRAF están vinculadas con una variedad de condiciones clínicas ⁽⁵⁵⁾. El síndrome cardíaco-facies-cutáneo, un trastorno de anomalías congénitas múltiples, se debe a mutaciones de novo en BRAF, MEK, ERK o KRAS ^(55, 56). Es de destacar que las lesiones precancerosas comunes, como los nevos melanocíticos, se correlacionan con una frecuencia sorprendentemente alta de mutaciones BRAF, lo que sugiere que la activación mutacional de la vía RAS/RAF/MAPK es un paso crítico en el inicio de la neoplasia melanocítica, aunque su presencia por sí sola no es suficiente para la tumorigenesis del melanoma ^(55, 56). Además, varias neoplasias como el melanoma, carcinoma de tiroides y, aunque con menor frecuencia, el cáncer colorrectal y el adenocarcinoma de células no pequeñas del carcinoma de pulmón pueden contener la mutación BRAF V600E. ⁽⁵⁷⁻⁶³⁾. Varios estudios indican que los inhibidores de BRAF bloquean el aumento de la proliferación celular inducida por la mutación BRAFV600E ⁽⁵⁷⁻⁶³⁾, según muestra la tabla 1 ⁽⁶⁴⁾.

Tabla 1. Clinical trials dirigidos a inhibidores BRAF en cáncer de tiroides

Fase	Droga	N. Ensayo Clínico	Título	Resultados Publicados	Referencias
1	Vemurafenib + KTN3379	NCT02456701	Mejora de la incorporación de yodo radiactivo en cánceres de tiroides mutantes BRAF con la combinación de vemurafenib y KTN3379	/	/
1	Dabrafenib + Lapatinib	NCT01947023	Dabrafenib y ditosilato de lapatinib en el tratamiento de pacientes con cáncer de tiroides refractario que no se puede extirpar mediante cirugía	Tasa de respuesta parcial del 60%, mediana de supervivencia libre de progresión de 15 meses, con toxicidad aceptable.	Rothenberg et al. 2015
1	Dabrafenib	NCT00880321	Un estudio de fase I para investigar la seguridad, la farmacocinética y la farmacodinámica de GSK2118436 en sujetos con tumores sólidos	Tasa de respuesta parcial del 29%, mediana de supervivencia libre de progresión de 11 meses, con toxicidad aceptable.	Falchook et al. 2015
2	Vemurafenib	NCT01286753	Un estudio de vemurafenib (RO5185426) en participantes con cáncer de tiroides papilar metastásico o irresecable positivo para la mutación BRAF V600	Aumento de la mejor respuesta general, duración de la respuesta y supervivencia libre de progresión tanto en pacientes no tratados previamente como en pacientes tratados con inhibidores de multiquinasa	Brose, et al. 2016
1	Dabrafenib	NCT01534897	Rediferenciación de carcinoma de tiroides papilar mutante BRAF V600E refractario al yodo radiactivo con GSK2118436	Entre 10 pacientes con cáncer de tiroides refractario al yodo radioactivo, 6 pacientes (60 %) demostraron una nueva captación de yodo radioactivo en la exploración de todo el cuerpo después del tratamiento con	Rothenberg et al. 2015

				<p>dabrafenib. Los 6 fueron tratados con 5,5 GBq de yodo-131. Dos pacientes tuvieron respuestas parciales y 4 pacientes tenían enfermedad estable en la reestadificación radiográfica estándar a los 3 meses.</p> <p>La tiroglobulina disminuyó en 4 de 6 pacientes tratados. Un paciente desarrolló carcinoma de células escamosas de la piel. No hubo otros eventos adversos significativos atribuidos a dabrafenib.</p>	
2	Vemurafenib	NCT01709292	Ensayo neoadyuvante de vemurafenib en cáncer de tiroides localmente avanzado	/	/
2	Dabrafenib + Trametinib	NCT01723202	Dabrafenib con o sin trametinib en el tratamiento de pacientes con cáncer de tiroides recurrente	/	/
2	Trametinib + Dabrafenib	NCT03244956	Eficacia de los inhibidores de MEK (trametinib) y BRAFV600E (dabrafenib) con yodo radiactivo (RAI) para el tratamiento del cáncer de tiroides diferenciado metastásico refractario (MERAIODE)	/	/
1	Trametinib + Pazopanib	NCT01438554	Estudio de fase 1 de pazopanib con GSK1120212 en tumores sólidos avanzados, enriquecido con	/	/

			pacientes con cáncer diferenciado de tiroides, sarcoma de partes blandas y colangiocarcinoma		
1	Trametinib + Paclitaxel	NCT03085056	Trametinib en combinación con paclitaxel en el tratamiento del cáncer de tiroides anaplásico	/	/
2	Trametinib	NCT02152995	Trametinib en el aumento de la incorporación de yodo tumoral en pacientes con cáncer de tiroides recurrente o metastásico	/	/

Nota: Esta tabla muestra los estudios clínicos existentes dirigidos al tratamiento del cáncer de tiroides portadores de la mutación BRAF. Tomada de Croce L et al., 2019 ⁽⁶⁴⁾.

1.5. Correlación histopatológica y mutaciones

El carcinoma de tiroides es una neoplasia genéticamente simple con un número bajo de mutaciones ⁽⁶³⁾. Las aberraciones del gen conductor en el cáncer de tiroides bien diferenciado son mutuamente excluyentes, con una media de una mutación por tumor, mientras que los cánceres indiferenciados acumulan alteraciones genéticas adicionales, los llamados eventos tardíos ⁽⁶³⁾. Las mutaciones conductoras y las fusiones de genes se identifican en más del 90% de los cánceres de tiroides, lo que la convierte en una de las neoplasias malignas mejor caracterizadas molecularmente en humanos ⁽⁶⁵⁾. El atlas del genoma del cáncer (TCGA), ha podido correlacionar la presencia de mutaciones y las variantes histopatológicas del cáncer de tiroides ⁽⁶⁶⁾, tal como se muestran en las figuras 3 y 4 ⁽⁶⁷⁾, permitiendo obtener un mejor conocimiento de las vías moleculares en la patogénesis de esta neoplasia, así como elementos para la estratificación de agresividad y blancos terapéuticos ^(65, 66).

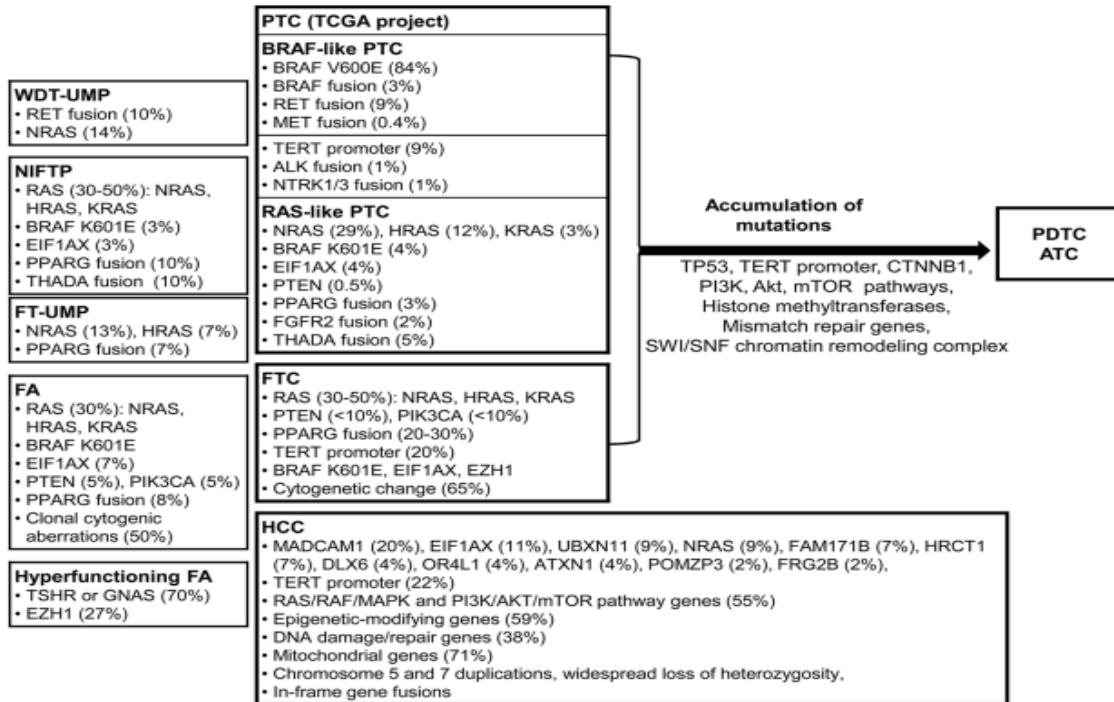


Figura 3: Perfil mutacional de neoplasias derivadas de células foliculares tiroideas. Tomadas de Kakudo K et al., 2018⁽⁶⁷⁾.

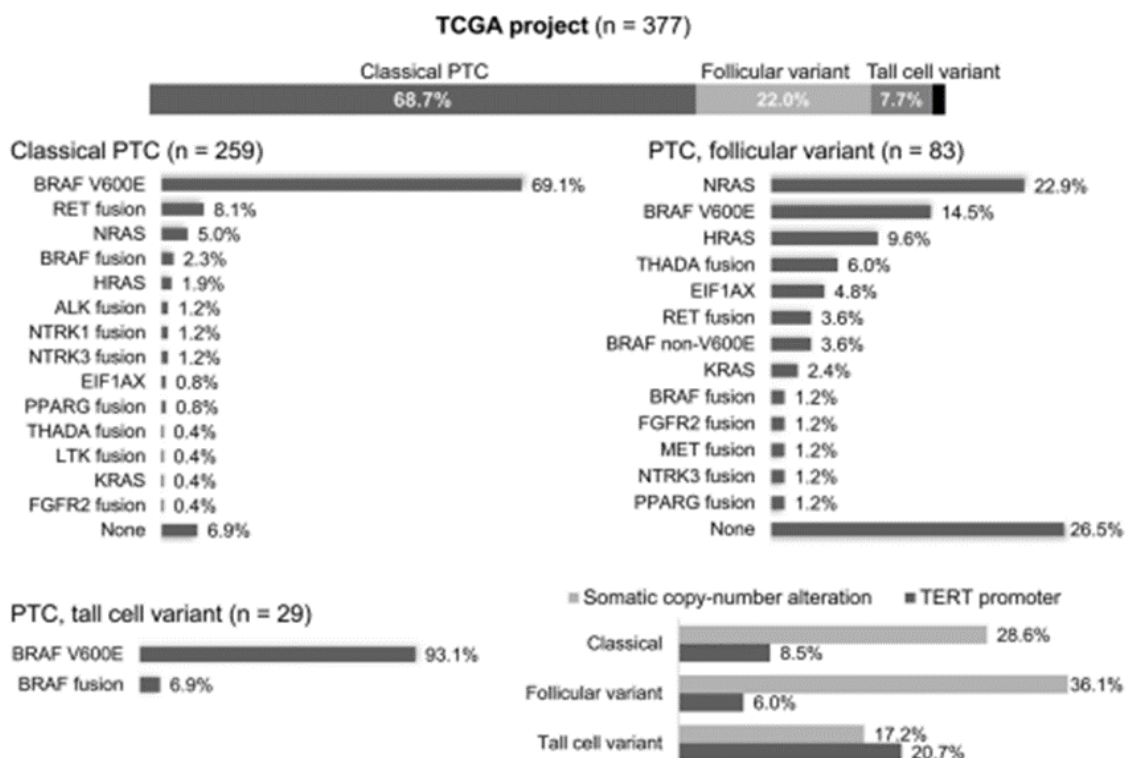


Figura 4: Alteraciones genéticas según las características histológicas del carcinoma papilar de tiroides en el conjunto de datos TCGA. Tomadas de Kakudo K et al., 2018⁽⁶⁷⁾.

1.6. Secuenciación de transcriptomas

1.6.1. Introducción:

El transcriptoma es aquel conjunto de los ARN de transcripción de un tipo determinado de célula, que incluye al ARN de codificación de proteínas, como el ARN mensajero (ARNm) y de transcripción no codificante de proteínas como el ARN ribosomal (ARNr), ARN de transferencia (ARNt), microARN (miARN) y otros ARN no codificante (ncRNA) ⁽⁶⁸⁾. A diferencia del genoma, que se comparte por todas las células de un organismo dado, el transcriptoma es específico para un tejido o tipo de célula dada o incluso específico a nivel unicelular ⁽⁶⁹⁾. La secuenciación del transcriptoma (RNA-Seq) es una tecnología de reciente desarrollo que utiliza una secuenciación de alto rendimiento con enfoques para determinar la secuencia de todo el ARN de las transcripciones en un determinado espécimen ⁽⁶⁹⁾. Porque no se basa en la hibridación con sondas existentes, la RNA-Seq se considera una técnica imparcial para poder evaluar la expresión diferencial de genes y así permitir la identificación de transcripciones novedosas como detectar empalmes y/o uso alélico de patrones que están presentes en células o en ciertas situaciones específicas ⁽⁶⁹⁾. Esta tecnología ha expandido rápidamente nuestra comprensión de los perfiles de expresión génica de diversos tejidos y/o células ⁽⁶⁹⁾, incluyendo una mejor comprensión del uso de empalmes alternativos en procesos normales y patológicos, el papel de los elementos funcionales del genoma y el papel de ARN no codificantes, permitiendo así descubrir una gran cantidad de transcripciones de fusión en cáncer ^(70, 71). La investigación y el uso de esta tecnología para transcriptoma, ha permitido nuevos conocimientos en el desarrollo y estudio de enfermedades benignas y malignas, procesos patológicos, migrando al campo clínico con pruebas diagnósticas dirigidas a detectar transcripciones de fusión oncogénicas con un papel en el diagnóstico y tratamiento del cáncer ⁽⁷¹⁻⁷³⁾.

1.6.2. Tecnologías transcriptómicas:

Existen múltiples herramientas tecnológicas para investigar y descifrar el transcriptoma, cada una de ellas con sus propias ventajas y desventajas, según muestra la tabla 2 ⁽⁶⁹⁾. A continuación, se discuten tres de las principales tecnologías, incluyendo las diferencias metodológicas, ventajas, sesgos asociados y limitaciones ⁽⁶⁹⁾.

Tabla 2

Comparación de los métodos de investigación del transcriptoma

Método	Rendimiento	Tiempo requerido	Tipo de información	Ventajas	Desventajas
Microarray	Hasta 20.000 genes	<24 h	Cualitativo	Relativamente más barato, alto rendimiento, cambio rápido, bueno para FFPE	Limitado a la detección de expresión génica, rango dinámico limitado, y se basa en definiciones del transcriptoma existente
NanoString	Hasta 800 genes	<24 h	Cuantitativo	Automatizado, vuelta rápida, capaz de detectar fusiones de genes y miARN, bueno para FFPE	Rango dinámico limitado y se basa en las definiciones existentes del transcriptoma
RNA sequencing	Dirigido a todo el transcriptoma	3-7 días	Cuantitativo	Alto rango dinámico, agnóstico a definiciones del transcriptoma	Caro, soporte computacional necesarios sustancial para el análisis de datos. Uso con FFPE todavía es limitado

Nota. Esta tabla muestra las ventajas y desventajas de RNA sequencing en comparación con dos técnicas para evaluación el transcriptoma. Tomada de Netto G. J. (2019) ⁽⁶⁹⁾.

1.6.2.1. Microarrays:

Gran parte del conocimiento en la actualidad sobre las firmas de expresión génica clínicamente relevantes provienen del uso de microarrays, tecnología que viene ayudando en áreas como la estratificación del riesgo del paciente y en la identificación del origen de aquel

espécimen de tejido tumoral poco diferenciado o indiferenciado ⁽⁷⁴⁾. Los microarrays son capaces de poder determinar niveles de expresión de transcripciones conocidas, mediante hibridación de ARN purificado para oligos complementarios fijados sobre un sustrato sólido ⁽⁶⁹⁾. Las micromatrices se utilizan con extractos de ARN total de tejido congelado, fijado en parafina y embebido en formalina (FFPE), con medidas de transcripción que muestran una consistencia del 80-97% entre los dos tipos de muestra ^(75,76). El ARN total es extraído de los especímenes de interés seguidos por la generación de ADN complementario (ADNc) por medio de transcripción inversa (RT) en el laboratorio ⁽⁶⁹⁾. Luego, las transcripciones inversas se etiquetan con fluorescencia y son hibridadas a la matriz que puede tener sondas generados a partir de ADNc, oligonucleótidos, fragmentos de reacción en cadena de la polimerasa (PCR), fragmentos digeridos con enzimas de restricción, oligómeros o etiquetas de secuencia expresada (EST) entre otras ⁽⁷⁷⁾. Una vez que las transcripciones se hibridan, la matriz se escanea con un láser para medir la intensidad de luz de las transcripciones hibridadas marcadas con fluorescencia que actúan como una medida de transcripción relativa o absoluta de abundancia en la muestra ⁽⁷⁸⁾. Con la mayoría de los microarrays, la expresión génica es una medida acumulativa de todas las transcripciones relacionadas con un gen, a pesar de ello tienden a tener dificultades para diferenciar entre isoformas de transcripción ⁽⁶⁹⁾; sin embargo, se han desarrollado micromatrices especializados para poder capturar estas isoformas ya sea mediante microarrays de mosaico genómico de alta resolución o con sondas que son complementarias a las uniones de exones ⁽⁷⁷⁾. En general, existen dos enfoques para la expresión génica que utilizan microarrays: el primero es comparar directamente una muestra con una muestra control etiquetando cada uno de ellas con una etiqueta fluorescente diferente e hibridando a la misma micromatriz, para posteriormente medir la abundancia relativa de la transcripción en la muestra de prueba en comparación con el control ⁽⁷⁷⁾. El segundo enfoque es hibridar las muestras de prueba y de control por separado, lo que requiere la normalización

de la expresión génica, en todas las matrices que se utilizan en un experimento. En general, los microarrays tienen dificultad para detectar transcripciones de baja abundancia y exhiben capacidades de cuantificación limitadas (rango de dinámicas) ^(77, 79, 80).

Independientemente del enfoque de la expresión génica, con microarrays utilizados, las mediciones de abundancia de transcripción se someterán a un análisis bioinformático mediante modelos clasificadores para crear unas firmas de expresión génica que se pueden utilizar como ayuda en la predicción clínica como subtipificación de tumores o respuesta al tratamiento ⁽⁶⁹⁾. Estos clasificadores basados en microarrays y genes asociados han resultado ser firmas genéticas muy robustas que muestran una consistencia del 80 al 90% en todas las plataformas ⁽⁸¹⁾. La expresión de los ensayos de microarrays de dos genes y las clasificaciones de sus asociaciones ha sido aprobada actualmente por la FDA, incluidos las Pruebas de origen de tejidos y el MammaPrint ^(82, 83). Además de estos dos ensayos, hay una variedad amplia de otros perfiladores de expresión génica basados en microarrays, que aún no han sido aprobados por la FDA, pero están disponibles comercialmente ⁽⁶⁹⁾. El desarrollo de microarrays y sus clasificadores ha ampliado nuestro conocimiento con respecto a la variación intrínseca de expresión entre tipos de tejido, durante la transformación neoplásica, así como ha ayudado en la identificación de los subtipos moleculares de tumores ^(84, 85). Múltiples investigaciones, junto a esfuerzos de la Administración de Medicamentos y Alimentos de los EE. UU (FDA), el proyecto de control de calidad MicroArray (MAQC) y otros grupos han llevado a descripciones detalladas de sesgos y variabilidades asociadas con microarrays y sus análisis de manera detallada ^(86, 87). Sus resultados han llevado a grandes mejoras en el diseño experimental, estrategias de análisis de mejores prácticas, calidad reproducible y precisa de los datos obtenidos de los microarrays ^(77, 86). El preprocesamiento, que incluye la eliminación de áreas con mala señalización, ajuste de imagen para el tamaño/forma de cuadrículas de matriz y normalización para diferencias en la eficiencia del

etiquetado y la calidad del ARN, son vitales para resultados exactos ⁽⁷⁸⁾. Fuentes conocidas de variabilidad incluyen sesgos vinculados a colorantes, eficiencias de hibridación diferencial, etiquetado, amplificación métodos, sesgos relacionados con las fuentes, y preparación de sondas de matriz ^(78, 88). Múltiples variabilidades, sin embargo, pueden ser evitadas mediante el uso de muestras de control adecuadas, réplicas de técnicas biológicas estandarizadas, así como experimentos de intercambio de tintes ^(78, 89). Sin embargo, todavía existen varias limitaciones a pesar de los nuevos métodos que existen para mejorar la calidad de datos ⁽⁶⁹⁾. Estas limitaciones incluyen la dependencia de las definiciones de transcriptomas existentes, los efectos de hibridación cruzada que conduce a niveles de fondo relativamente altos, y un rango dinámico limitado para evaluar la abundancia de la transcripción debido a los niveles de fondo y a la saturación de las señales del informador ⁽⁹⁰⁾. A pesar de estas limitaciones, los microarrays son generalmente precisos, con multiplataforma y reproducibilidad inter-/intra sitio, representando una forma económica y relevante de interrogar clínicamente múltiples perfiles del transcriptoma ⁽⁸⁶⁾.

1.6.2.2. Perfiles de transcripciones digitales

Los avances tecnológicos han llevado al desarrollo del análisis digital de transcripciones, que permite detectar cada molécula de transcripción de ARN individualmente, en lugar de la única detección de una señal combinada de todas las transcripciones hibridadas en el mismo lugar en un microarray ⁽⁶⁹⁾. Este enfoque permite la cuantificación digital dirigida de la expresión de ARN con un alto rendimiento, evaluando la expresión de hasta 800 transcripciones por vez ⁽⁹¹⁾. Esta técnica está disponible comercialmente utilizando el sistema “NanoStringnCounter®” que tiene una sensibilidad comparable a los microarrays y a la PCR en tiempo real ⁽⁹²⁻⁹⁴⁾. Este sistema puede utilizar ARN total extraído como lisados crudos de sangre total y tejido a partir de una gran variedad

de muestras, incluido las obtenidas por FFPE, para la investigación y decodificación de fusiones de genes, expresión de genes, así como la expresión de lncRNA (largas cadenas de ARN no codificantes) y miRNA ⁽⁶⁹⁾. Se ha demostrado que esta plataforma funciona de manera eficiente con tan solo 100 ng de muestra, incluso a partir de muestras con ARN altamente degradado como FFPE ⁽⁹⁵⁾. La detección digital es posible gracias a un sistema de sonda dual que consta de una *sonda informadora* la que etiqueta de forma única cada molécula de transcripción utilizando un código de barras de fluoróforos (sondas fluorescentes a través de tintes o proteínas fluorescentes) específico de la transcripción y una *sonda de captura* que permite la transcripción dirigida a través del código de barras molecular hibridado para ser inmovilizado y detectado mediante una cámara CCD (Dispositivo de Carga Acoplada) y software de análisis de imágenes ⁽⁶⁹⁾.

En general, esta plataforma ha demostrado generar datos altamente reproducibles, con baja variabilidad técnica y biológica intra-muestra ⁽⁹⁵⁾; sin embargo, de manera similar a los microarrays, se ve obstaculizado por un rango dinámico limitado y se basa en definiciones de transcriptoma existentes, limitando su capacidad para identificar transcripciones novedosas, isoformas alternativas de empalme y fusión de genes con novedosas asociaciones ⁽⁹⁶⁾.

Esta plataforma puede ser altamente automatizada y tiene costos más bajos en comparación con otros métodos de elaboración de perfiles transcriptómicos; estas características así como su capacidad para interrogar la expresión génica de muestras de mala calidad en un multiplexado sin necesidad de amplificación y con alta sensibilidad hacen de esta plataforma una opción atractiva para el desarrollo de paneles clínicos específicos como el Ensayo de Cáncer de mama Prosigna® aprobado por la FDA, que proporciona una firma genética pronóstica ^(91, 92, 95, 97).

1.6.2.3. Secuenciación de ARN

Con el desarrollo de la secuenciación masiva paralela y de alto rendimiento de próxima generación (NGS), la transcriptómica basada en secuenciación ahora se vuelve rutinaria y estándar en biología e investigación médica, ganando rápidamente una amplia aceptación en la práctica clínica ⁽⁶⁹⁾. La secuenciación ARN (RNA-Seq) permite la cuantificación digital del transcriptoma completo o dirigido, mediante la expresión de una manera de alto rendimiento a partir de una variedad de tipos de muestras ⁽⁶⁹⁾. De las varias ventajas de RNA-Seq sobre los métodos existentes basados en microarrays, las tres más importantes se relacionan con su naturaleza imparcial sin conocimiento a priori necesario de la secuencia del genoma o de la transcripción, un rango dinámico extremadamente alto que permite la cuantificación de transcripciones de baja y alta expresión, y niveles de ruido relativamente bajos de hibridación cruzada con secuencias del genoma ^(90,98).

Con respecto a las aplicaciones, a diferencia de los microarrays, las metodologías basadas en RNA-Seq no proporcionan solo la cuantificación digital de la expresión de transcripción absoluta y diferencial, sino también datos de secuencia de lectura con resolución de pares de bases que permite la identificación de transcripciones novedosas (incluidas transcripciones alternativas y empalmadas de forma aberrante y ARN no codificante), expresión de variantes somáticas en cáncer, expresión específica de alelos, ARN nuevo que edita nuevos eventos, cambios de secuencia postranscripcional (mutaciones o edición), así como fusiones de genes ^(80,99).

Junto con los costos rápidamente decrecientes de NGS en general y el desarrollo de herramientas altamente evolucionadas para el análisis del transcriptoma, RNA-Seq se ha convertido firmemente en aquella herramienta de elección en la investigación de transcriptomas y el desarrollo de aplicaciones clínicas ⁽⁶⁹⁾.

1.6.3. Protocolos y preparación de la biblioteca de RNA-Seq

"RNA-Seq" es un término técnico genérico que se utiliza para hacer referencia a una variedad de tecnologías de secuenciación de ARN y ADNc, las dos técnicas más comúnmente implementadas son la secuenciación de ARNm en poli-A + fracción de RNA celular total y la secuenciación de ADNc de fracciones de RNA total enriquecidas por captura o amplificación por hibridación, según muestra figura 5 ⁽⁶⁹⁾. Actualmente, la secuenciación de moléculas de ARN no es rutinariamente realizada debido a la alta tasa de error de las tecnologías de secuenciación de una sola molécula; a continuación, se discute la secuenciación de bibliotecas de ADNc típicamente preparadas a partir del ARN total ⁽⁹⁰⁾.

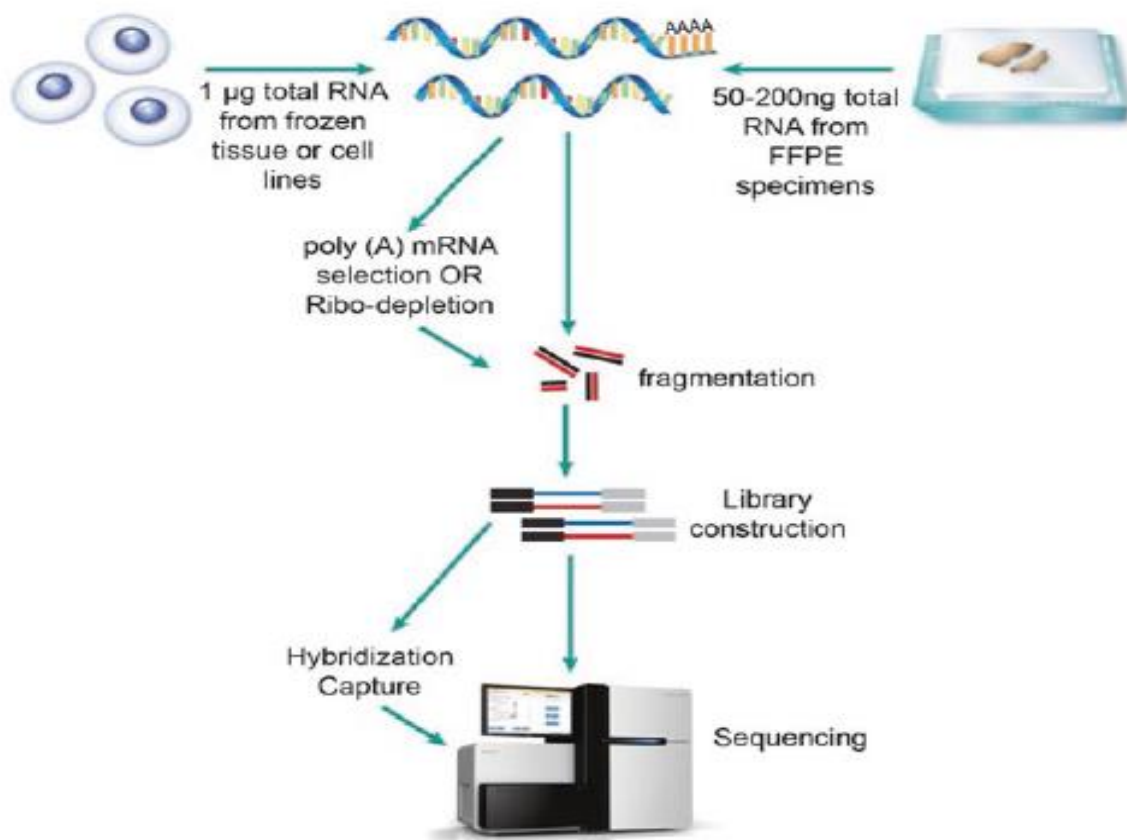


Figura 5: Visión general de la preparación de la Biblioteca de RNA-Seq. Se generan bibliotecas de ARN extraído de líneas celulares, tejido congelado o especímenes de FFPE que se someten a selección y fragmentación de ARN. Seguidamente las bibliotecas de transcriptomas se secuencian usando plataformas NGS. Tomada de Netto G. J. (2019) ⁽⁶⁹⁾.

1.6.3.1. Evaluación de la calidad del ARN

A consecuencia de la inestabilidad inherente del ARN como molécula biológica, la calidad y precisión de los datos de la RNA-Seq dependerán en gran medida de la calidad del RNA de entrada ⁽⁹⁰⁾. La evaluación de la calidad del ARN para los experimentos de RNA-Seq aún no se ha estandarizado, existiendo varias métricas preanalíticas de la integridad del ARN como lo son el número de integridad del ARN (RIN), DV200 (% de fragmentos de ARN > 200 nt) métricas, métodos cuantitativos basados en PCR, así como algoritmos bioinformáticos para la evaluación de la calidad del ARNm de los datos de RNA-Seq en sí, que se han descrito ^(100, 101). Mientras no se demuestre una métrica de estas técnicas que sean altamente precisas en la predicción del éxito de la generación de una biblioteca RNA-Seq de alta complejidad, los ARN con puntuaciones de RIN > 7 y valores de DV200 > 50–60%, por lo general, se considera de "alta" calidad para aplicaciones de RNA-Seq. La elección de un adecuado protocolo RNA-Seq está influenciado por la calidad de ARN; por ejemplo, poli-A + mRNA-Seq no es práctico para ARN de baja calidad y muy fragmentado, como lo es a menudo el caso de FFPE e incluso de muestras congeladas en la práctica clínica ⁽¹⁰⁰⁾.

Los requisitos del ARN de entrada para bibliotecas NGS también varían según el protocolo; basado en la amplificación, las bibliotecas de RNA-Seq dirigidas se pueden preparar a partir de mínimas cantidades como 10-20 ng de ARN total, mientras que el análisis de transcriptoma o transcriptoma codificante requiere cantidad de insumos más altos que van desde 40 ng a 1 ug de ARN total ⁽⁶⁹⁾.

Los niveles de ARN de entrada son típicamente más altos para métodos de selección que emplean depleción de ARNr o para la selección positiva de fragmentos de ARN con colas poli-A intactas, mientras que los métodos de selección basados en captura de hibridación generalmente requieren menos ARN de entrada ⁽⁶⁹⁾.

1.6.3.2. Métodos de selección de secuencia de ARN

El ARN total de muestras biológicas contiene múltiples especies de ARN, incluido el ARN mensajero (ARNm), ARN de transferencia (ARNt), ribosomal (ARNr) y otros ARN no codificantes (ARNc), con ARNr que representa aproximadamente el 85% de todos los ARN de una célula ⁽¹⁰²⁾. Debido a la sobrerepresentación de ARNr, es necesario agotar el ARNr del ARN total extraído, ya que es fundamental para la sensibilidad de la RNA-Seq en la detección de transcripciones de RNA biológica o clínicamente importantes, así como alteraciones dentro de estas transcripciones ⁽⁶⁹⁾.

Los métodos de ribodeplección que se utilizan comúnmente para eliminar eficazmente las transcripciones de ARNr incluyen el enriquecimiento poliadenilado (poli-A) y la depleción del ARN ribosómico (ARNr) ⁽¹⁰⁰⁾. Dependiendo del tipo de aplicación, más estrategias para selectivamente enriquecer ciertas fracciones de ARNm, por ejemplo, fracción de ARN no poliA + (ARN nucleolares pequeños, histonas ARNm, etc.) o el transcriptoma que codifica la proteína, por lo general utilizan la captura de hibridación de ADNc generado por transcripción inversa de ARN o estrategias de enriquecimiento basadas en amplificación (p. ej., PCR Multiplex Anclada) ^(100, 103).

Al considerar el método de selección, se deben considerar cuidadosamente los beneficios e inconvenientes para determinar el mejor enfoque para el tipo de muestra y objetivo experimental ⁽⁶⁹⁾. Como se mencionó arriba, las bibliotecas con poli-A seleccionado y empobrecido de ARNr son los enfoques más comunes para selección de ARNm para ARN-Seq, ambos de manera efectiva excluyendo la mayoría de las transcripciones de ARNr ⁽⁶⁹⁾. La selección de Poli-A utiliza perlas de oligo-dT, enriquecedoras para ARNm adulto con colas 3' poli-A intactas, mientras que los métodos de agotamiento del ARNr utilizan sondas específicas para la captura de hibridación de ARNr para eliminar estas transcripciones por

sustracción basada en cuentas, con ambos tipos de selección que tienen ARNm con sensibilidad muy similar de transcripción ⁽¹⁰³⁻¹⁰⁵⁾.

Mientras la selección de poli-A es muy eficiente para caracterizar transcripciones de ARNm, debido a la dependencia de una cola 3' poli-A intacta, fragmentada y degradada, las moléculas de ARNm con pérdida de colas de poli-A muestran subrepresentación o sesgo 3' en la cobertura de ARNm por ARN-Seq ^(100, 103, 106); este método también no es adecuado para especies de ARN no poliadenilado ^(102, 103).

El agotamiento del ARNr, por otro lado, tiene una cobertura más consistente en todas las transcripciones y se puede utilizar para especies de ARN no poliadenilado. Sin embargo, el agotamiento del ARNr ha sido demostrado ser un poco menos eficiente con respecto a cobertura exónica que las bibliotecas seleccionadas de poli-A ^(102, 104, 106).

Para una selección más específica del transcriptoma, ciertas técnicas de re-secuenciación del ADN dirigidas para capturar regiones seleccionadas para RNA-Seq se han desarrollado ⁽⁶⁹⁾. Dos de los métodos objetivos más comunes de selección incluyen la hibridación basada en la captura de selección y el de secuenciación del transcriptoma basada en amplicones ⁽⁶⁹⁾.

Los rangos de tamaño de los paneles de secuenciación de transcriptomas específicos pueden variar mucho, de unos pocos genes a toda la codificación conocida del transcriptoma ⁽¹⁰⁷⁾. La ARN-Seq dirigida ha demostrado ser muy eficiente para especímenes altamente degradados como FFPE, con hibridación de captura con un mejor rendimiento en muestras degradadas que con poli-A seleccionado y sin bibliotecas de ARNr ⁽¹⁰⁷⁾.

La captura basada en hibridación y la secuenciación del transcriptoma basada en amplicones han demostrado tener una cobertura genómica similar y medidas de expresión

génica normalizadas en comparación con los transcriptomas poli-A seleccionados, sin embargo, todavía a menudo requieren menos niveles de ARN de entrada que la selección poli-A ^(107, 108). Además, estos métodos tienden a tener una mayor sensibilidad con respecto a la técnica con transcriptomas poli-A ⁽¹⁰⁷⁾.

Un método de enriquecimiento específico de RNA-Seq dirigido que ha demostrado ser prometedor en el área clínica del diagnóstico se basa en el método Multiplexado Anclado de PCR (AMP™). La RNA-Seq basada en AMP™ ha demostrado una alta sensibilidad para la detección de fusión de genes independiente del socio de fusión, además a la detección simultánea de variantes de un solo nucleótido, inserciones, eliminaciones y cambios de número de copia ⁽¹⁰⁹⁾.

La secuenciación dirigida permite importantes ahorros en costos, tiempo y potencia informática, resultados incidentales reducidos, y se ha utilizado para el desarrollo de paneles clínicamente útiles ^(110, 111). Sin embargo, la secuenciación dirigida, especialmente la de métodos de enriquecimiento basados en amplificación, puede introducir sesgos que comprometan la complejidad de la biblioteca y pueden resultar en una representación insuficiente de las transcripciones de bajo nivel ⁽⁶⁹⁾.

La utilidad clínica de la secuenciación completa del transcriptoma versus RNA-Seq dirigido es un área actual de investigación que es abordada por varios grupos, incluyendo el programa del Instituto Nacional de Salud de la Secuenciación Clínica de la Investigación Generadora de Evidencia (CSER2) ⁽¹¹²⁾.

1.6.3.3. Fragmentación y Preparación de la biblioteca

La fragmentación del RNA (o ADNc o ADNds) es necesario antes de la reparación del extremo y la ligadura del adaptador para la preparación de la biblioteca de secuenciación

NGS ⁽⁶⁹⁾. Se han desarrollado protocolos específicos para retener información de la cadena de ARNm a través del proceso de preparación de la biblioteca, incluida la incorporación de dUTP (dexosiridina trifosfato) durante la síntesis de la segunda cadena seguida de la digestión de uracil-ADN glicosilasa (UDG) ⁽⁶⁹⁾.

Las moléculas de ADNc trenzado luego de someterse a métodos típicos de preparación de bibliotecas NGS, como la reparación del extremo, la cola A y la ligadura de adaptadores de biblioteca que incluyen muestras de códigos de barras ⁽⁶⁹⁾. Después de la selección de ARN y biblioteca NGS Prep, los fragmentos de ADNc ligados al adaptador son luego amplificados y secuenciados en una plataforma NGS de alto rendimiento ⁽⁹⁰⁾.

Dado que la complejidad y diversidad de los transcriptomas celulares son muy variables y dependientes entre varios factores, incluido el tejido de origen y el estado de enfermedad, generando una biblioteca de ARN complejo con representación adecuada de todas las transcripciones de interés del ARN, incluida la novela baja y transcripciones sin anotaciones, es un gran desafío ⁽⁷⁷⁾.

Los métodos de preparación y selección de bibliotecas, así como la calidad y cantidad del ARN de entrada son los principales contribuyentes a la complejidad de la biblioteca, que puede evaluarse bioinformáticamente utilizando el porcentaje de duplicación como medida de la "unicidad" de las lecturas de secuencia ⁽⁶⁹⁾.

La biblioteca de baja complejidad (altas tasas de duplicación) puede deberse a cantidades limitadas de ARN, ARN muy fragmentado y uso de ciclos de PCR excesivos durante preparación de la biblioteca ⁽⁶⁹⁾. Varios protocolos de preparación de bibliotecas de RNA-Seq incorporan códigos de barras moleculares (o identificadores moleculares únicos) que marcan las moléculas de ARN o ADNc antes de la PCR y puede utilizarse para evaluar con precisión la complejidad de la biblioteca ⁽¹¹³⁾.

1.6.4. Análisis de datos de la secuencia de ARN

Debido a las amplias aplicaciones de los experimentos de RNA-Seq, los pasos del análisis de datos son múltiples, a menudo computacionalmente intensivo, y ningún conjunto de los programas de software ofrecen todas las herramientas analíticas disponibles ⁽⁶⁹⁾, tal como se muestra en la figura 6 y la tabla 3 ⁽⁶⁹⁾.

Tabla 3

Software bioinformático para análisis de secuenciación del transcriptoma

Tipo de análisis		Software aplicado
Control de calidad y preprocesamiento		FastQC
		FASTX-Toolkit
		Trimmomatic
		Picard
		RSeQC
Alineación	Referencia establecida	STAR
		Bowtie2/TopHat
	De nuevo	Trinity
		SOAPdenovo-Trans
		Rnnotator
Transcripción de identificación y cuantificación		Trans-ABYSS
		Cufflinks
Expresión génica diferencial		RSEM
		DESeq2
		edgeR
Empalme alternativo		NOISeq
		CuffDiff2
		rMATS
Descubrimiento de variantes		DESeq2
		GATK
Descubrimiento de fusión de genes		deFuse
		SOAPfuse
		Chimera
		TopHat-Fusion
		PRADA
	Fusion Hunter	

Nota. La presente tabla muestra los componentes de manera secuencial que componen el software bioinformático del RNA-Seq. Tomada de Netto G. J. (2019) ⁽⁶⁹⁾.

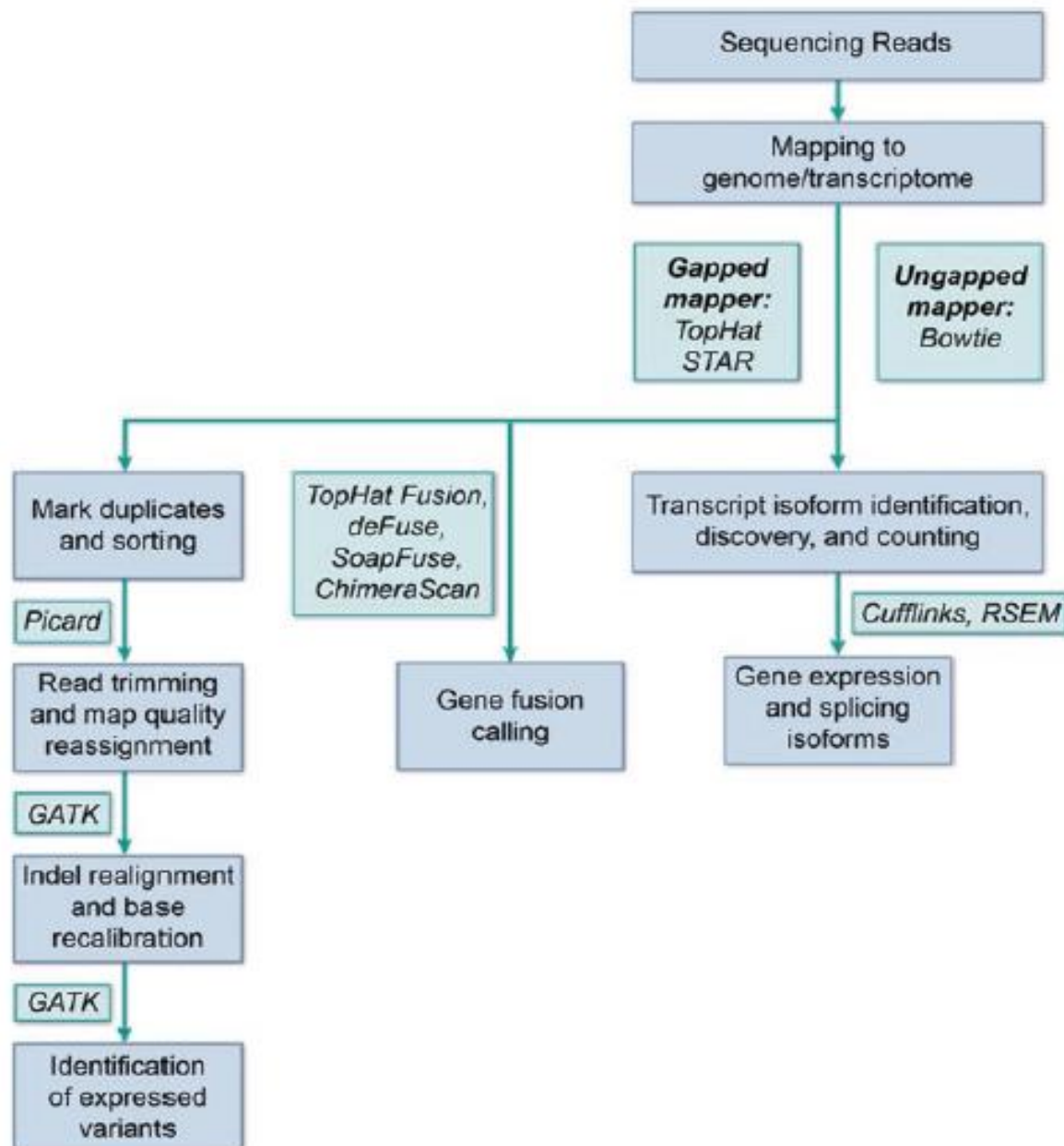


Figura 6: Datos del análisis RNA-Seq. Las vías de secuenciación informática del transcriptoma son altamente dependientes de la naturaleza del experimento. Después los datos de la secuenciación del transcriptoma son alineados con el genoma humano o la secuencia de referencia del transcriptoma; programas específicos son integrados en la tubería para realizar análisis especializado para perfiles de expresión y detección de mutaciones somáticas, fusiones de genes, y empalme alternativo de eventos. Tomada de Netto G. J. (2019) ⁽⁶⁹⁾.

Los procesos clave comunes a todas las vías son el control de calidad y la lectura de alineación ⁽⁶⁹⁾. Dependiendo de la naturaleza del experimento, los programas específicos se incorporan a la vía para realizar el perfil de expresión y expresión diferencial, detección de

mutaciones somáticas y fusiones, y alternativas de empalme, por nombrar algunos ⁽⁶⁹⁾. Una evaluación completa del análisis de datos de la RNA-Seq y las mejores prácticas de las directrices se pueden encontrar en otros lugares ⁽¹¹⁴⁾; a continuación, se analizan algunos pasos y aplicaciones claves.

1.6.5. Montaje del Transcriptoma

Las lecturas de secuencia obtenidas de las plataformas NGS utilizadas suelen ser cortas y, por lo tanto, deben reconstruirse en transcripciones completas, con la excepción de clase de ARN de secuencia corta, como miRNA ⁽¹¹⁵⁾.

Antes del montaje, las lecturas de secuenciación cruda se preprocesan para eliminar lecturas de baja calidad, duplicados de PCR, secuencias adaptadoras, errores de secuenciación y otros artefactos mediante herramientas como como FastQC, FASTX-Toolkit y Trimmomatic ⁽¹¹⁵⁻¹¹⁸⁾.

En particular, los errores de secuenciación se eliminan o corrigen mediante el uso de la puntuación de calidad para cada lectura, una función de probabilidad de que una base específica en la secuencia es correcta, y/o la frecuencia k-mer, que es el número de veces que aparece un oligonucleótido corto de longitud k en un conjunto de secuencias de DNA ⁽⁶⁹⁾.

Los k-mers de muy baja frecuencia suelen originarse en errores de secuenciación, y las lecturas que contienen estos errores se pueden quitar. Sin embargo, esto podría eliminar, pero en muy raras ocasiones, transcripciones genuinas ⁽¹¹⁵⁾.

Después del preprocesamiento, el ensamblaje del transcriptoma se puede lograr mediante un ensamblaje basado en el genoma de referencia o un ensamblaje "de novo" según muestra la figura 7 ⁽¹¹⁴⁾.

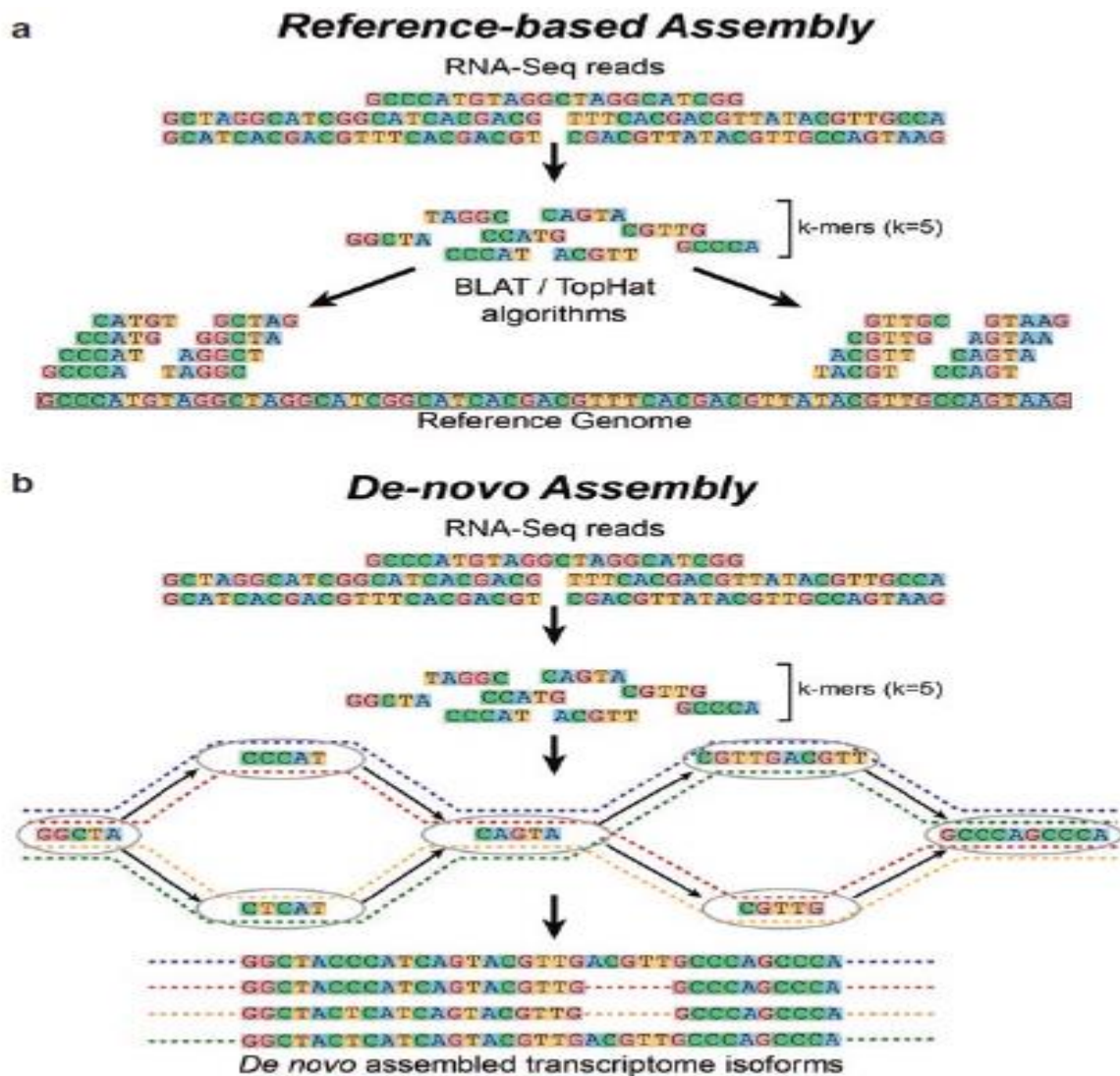


Figura 7: Montaje de Transcriptoma. Las lecturas de secuenciación de ARN pueden ser ensambladas por un sistema de referencias (a) o un montaje de novo (b). Métodos de ensamblaje basados en referencias dividen las lecturas de RNA-Seq en subcadenas (k-mers), mapeando estos fragmentos más pequeños a una referencia de genoma para crear un transcriptoma ensamblado. Métodos de montaje de novo, similar al ensamblaje basado en referencias, dividen la RNA-Seq en k-mers; sin embargo, la siguiente generación de los k-mers, en forma de subcadenas se organizan en base a la presencia de al menos uno base superpuesta en k-mers adyacentes ensamblándose en un Gráfico de Bruijn para determinar todas las posibles secuencias de combinaciones de isoformas. Tomada de Netto G. J. (2019) ⁽⁶⁹⁾.

El método basado en referencias comprende dos partes. Primero, las lecturas están alineadas con una referencia del genoma o transcriptoma con un "empalme-consciente alineador" como TopHat o STAR ⁽¹¹⁹⁻¹²¹⁾. En segundo lugar, las lecturas superpuestas del mismo lugar se agrupan en un gráfico para llegar a todas las posibles isoformas, seguidas de análisis mediante

programas como Gemelos o RSEM para la resolución, descubrimiento y cuantificación de isoformas de transcripción ^(122, 123). Los métodos basados en referencias requieren menos poder de computación, elimina algunos artefactos y errores como la no alineación con el genoma de referencia, y son muy sensibles a las transcripciones raras ⁽⁶⁹⁾. Es importante tener en cuenta que los errores causados por los alineadores de lectura corta pueden trasladarse al ensamblaje y que las lecturas empalmadas que abarcan intrones más largos pueden ser omitidas. "Lecturas múltiples", donde una secuencia se alinea igualmente bien para varios loci en el genoma, pueden ser excluidas, sin embargo, esto dejará lagunas en el conjunto de la secuencia final ⁽¹¹⁵⁾. El montaje "de novo" no utiliza una referencia del genoma (Fig. 7). En cambio, aprovecha la redundancia en las lecturas cortas y utiliza las superposiciones para ensamblar el transcriptoma con herramientas de software como Rnnotator, SOAPdenovo-Trans, TransABySS y Trinity ⁽¹²⁴⁻¹²⁷⁾. Trinity fue desarrollado específicamente para datos de RNA-Seq y previene genes superpuestos en la misma hebra que se interpretan erróneamente como transcripciones de fusión ⁽⁶⁹⁾. También agrupa secuencias lineales relacionadas que representan isoformas alternativas o familias de genes parálogos en estructuras no lineales que contienen "burbujas" y fines alternativos. En general, las estrategias "de novo" exigen mucha potencia de cálculo ^(127, 128). La nube informática es una alternativa y se han desarrollado ensambladores de genomas basados en la nube ⁽¹¹⁵⁾.

1.6.6. Evaluación de Calidad de la Secuenciación de Transcriptomas

Después del montaje, es fundamental evaluar la calidad y alineación de las secuencias resultantes ⁽¹¹⁴⁾. Existen múltiples programas para evaluar la calidad de la secuenciación incluyendo Picard (Broad Institute), RNASeQC (Broad Institute), FastQC (Babraham Institute) y RSeQC (Baylor College of Medicina) ^(116, 129-131). Secuenciación de métricas de control de calidad debe usarse para evaluar regularmente la calidad de secuenciación para

cada ejecución, así como para determinar tasas de precisión para cada plataforma de secuenciación y kit de preparación de la biblioteca para ayudar en la detección de problemas de instrumentación y variabilidad de lote a lote en kits de preparación y secuenciación de bibliotecas ⁽⁶⁹⁾. El consorcio Europeo de Variación Genética de la Enfermedad (GEUVADIS) ha recomendado varios parámetros que se evalúan como control de calidad, incluida la distribución de las puntuaciones de calidad base y contenido de GC, tasa de mapeo, desviación estándar del tamaño de inserto y distribución de cobertura a lo largo de la transcripción ⁽¹³²⁾. De estos, uno de los parámetros más importantes es el porcentaje de lecturas mapeadas, que tiende a ser un indicador general de la precisión general de la secuenciación ⁽⁶⁹⁾. Bajo tasas de mapeo, por debajo del 70% al mapear el transcriptoma humano, puede ser una indicación de Contaminación por ADN ⁽¹²⁰⁾.

Otro importante parámetro, fuera de los recomendados por GEUVADIS, es el porcentaje de unicidad con niveles bajos que posiblemente indiquen una biblioteca baja complejidad que puede ser el resultado de una mala Calidad del ARN ⁽⁶⁹⁾. La sensibilidad de RNA-Seq, como todas las técnicas de secuenciación de próxima generación, depende altamente de la complejidad de la biblioteca y la profundidad de la secuenciación ⁽⁶⁹⁾. La complejidad de la biblioteca, como se indicó anteriormente, puede cuantificarse mediante el software de control de calidad (QC), como RNA-SeQC, que detalla la tasa única de transcripciones mapeadas en función de sitios únicos de inicio de lectura ⁽¹³⁰⁾. En nuestra experiencia, una tasa de ARN única de aproximadamente 40% o más tiende a indicar bibliotecas NGS de alta calidad ⁽⁶⁹⁾. Desafortunadamente, determinar la profundidad óptima de la secuenciación necesaria para RNA-Seq no está bien definida aún ⁽⁶⁹⁾.

Como la expresión de ARN es variable y puede abarcan varios órdenes de magnitud, la profundidad de secuenciación necesaria para detectar las transcripciones de interés se

correlaciona inversamente con los niveles naturales de la expresión de la transcripción en la muestra ^(123, 133). Las transcripciones altamente expresadas serán más fácilmente detectadas que aquellas transcripciones con un nivel de expresión de moderado a bajo y, por lo tanto, requieren menos profundidad de secuenciación ⁽⁶⁹⁾. Transcripciones expresadas de bajas a moderadas representan una proporción menor de la población total de transcripciones, la detección de las transcripciones expresadas requerirá mayores profundidades de secuenciación ^(127, 134). Una profundidad de secuenciación insuficiente puede conducir a una expresión génica errónea y a variantes de métricas de frecuencia de alelos, así como una incapacidad para detectar transcripciones bajas expresadas, fusiones de genes y variantes ⁽⁶⁹⁾. Actualmente, no existen estándares o pautas con respecto a la profundidad de secuenciación requerida para una alta confianza en la detección de la representación completa del transcriptoma o para aplicaciones específicas ⁽⁶⁹⁾.

El uso de estudios de mezcla de células y picos sintéticos utilizando variantes verdaderas positivas y fusiones de genes con una amplia gama de tasas de expresión ayudará a informar sobre la cobertura óptima necesaria para cada ensayo ⁽⁶⁹⁾.

1.6.6.1. Material de referencia y Control de Calidad

Dado que la RNA-Seq es una tecnología en evolución con mejoras en constante cambio en cuanto a avances en instrumentación, química de secuenciación, nuevas metodologías de construcción de bibliotecas, y análisis computacional, para que la RNA-Seq sea clínicamente viable, existe una inminente necesidad del desarrollo de estándares, "la mejor práctica" y materiales de referencia para garantizar el mantenimiento de la calidad entre las ejecuciones de secuenciación ⁽⁶⁹⁾.

Mientras que los estándares de referencia para el ADN se han vuelto más ampliamente establecidos, los estándares para el ARN se han quedado atrás debido a la

complejidad y diversidad del transcriptoma, así como debido a la mayor variación de la calidad en la muestra de ARN, diferentes métodos de preparación de bibliotecas y selección de ARN, y el requisito de un mayor análisis bioinformático complejo ⁽¹³⁵⁾.

Uno de los primeros materiales de referencia de ARN desarrollados para microarrays y secuenciación del transcriptoma fue las muestras de ARN de referencia humana universal compuesto de ARN derivado de una mezcla equimolar de múltiples líneas celulares ⁽⁶⁹⁾. Las muestras de referencias de ARN han sido desarrolladas y analizadas por muchos consorcios, incluidos SeQC y la Asociación de Instalaciones de Recursos Biomoleculares (ABRF), con los datos resultantes disponibles para descargar de Gene Expression Omnibus (GEO) y el repositorio de Sequence Read Archive (SRA) para utilizar esta información como control de calidad ^(99, 105, 136, 137). Sin embargo, una de las principales limitaciones del uso de material biológico y genético humano como la referencia universal para las muestras de ARN es la incapacidad de estos controles para combinarse directamente con muestras de pacientes sin contaminar los análisis posteriores ⁽¹³⁵⁾.

Poder combinar un estándar de referencia directamente con muestras de pacientes proporciona el beneficio agregado de que el control se someta al mismo tiempo a todos los pasos de preparación y secuenciación de la biblioteca como espécimen clínico, actuando, así como un control interno que refleja con precisión todos los procesos y variabilidades técnicas asociadas a la muestra de cada paciente ⁽⁶⁹⁾. Los controles de aumento exógeno sirven como una buena alternativa a las muestras de ARN de referencia humano, ya que estos controles a menudo comprenden secuencias no humanas o artificiales que permiten distinguir las lecturas derivadas de la muestra del paciente, lo que permite añadir a estos controles a la muestra de ARN de cada paciente ⁽¹³⁵⁾. Además, el uso de un estándar de ARN exógeno en lugar de transcripciones endógenas como "genes de limpieza" proporciona un método

estándar más confiable debido a la naturaleza idéntica y constante expresión entre muestras ⁽¹³⁸⁾. El Consorcio de Controles de ARN Externo (ERCC) se formó para desarrollar controles de ARN externo para su uso en la evaluación del rendimiento técnico de los ensayos de expresión génica, incluidos microarrays y RNA-Seq ⁽⁶⁹⁾. Estos controles externos se desarrollaron para ayudar a determinar la precisión, reproducibilidad y límites de detección de los ensayos de perfiles transcriptómicos, así como para servir como medida estándar para la secuenciación de tasas de error, sesgos de cobertura y cuantificación de la transcripción ⁽¹³⁸⁾. El esfuerzo de ERCC ha llevado al desarrollo de estándares de ARN externos conocidos como los picos de ERCC que consisten en dos conjuntos preformulados de 92 transcripciones poliadeniladas de la biblioteca ERCC de referencia de plásmidos generada a partir de los genomas de *Bacillus subtilis* y *Methanocaldococcus jannaschii*, así como de ADN sintético ^(138, 139). Grandes estudios que incluyen la secuenciación de próxima generación SeQC y proyectos ABRF han utilizado los picos de ERCC para evaluar la sensibilidad y precisión de RNA-Seq, y estos han sido recomendados en última instancia por la ABRF como métrica de calidad útil basada en muestras ^(99, 105, 138).

Fuera de estos múltiples esfuerzos, los laboratorios han generado controles de ARN spike-ins (transcripción de ARN de secuencia conocida y cantidad utilizada para calibrar mediciones en ensayos de hibridación de ARN) sintético personalizados para evaluar la sensibilidad de RNA-Seq con respecto a la detección de isoformas de transcripción, clases pequeñas de ARN y alteraciones clínicamente relevantes que incluyen fusiones de genes ^(138, 140-142). Estos picos sintéticos generados para de la variación genética humana puede ser importante para evaluar la sensibilidad de RNA-Seq en regiones que son a menudo difíciles de caracterizar por las tecnologías NGS ⁽⁶⁹⁾, como en áreas ricas en GC (guanina-citosina). Muchos tipos de alteraciones genómicas también se pueden multiplexar dentro de un solo conjunto de picos de ARN que permiten evaluar una variedad de tipos de alteración a la vez

⁽⁶⁹⁾. A pesar de todos los beneficios relacionados con los controles de picos externos, hay un desafío constante para estos controles de referencias para mantener la conmutabilidad entre diferentes métodos de preparación de bibliotecas y plataformas de secuenciación, ya que las construcciones sintéticas no lo hacen siempre y funcionan de manera similar a las transcripciones de ARN nativo, como se demostró en el estudio ABRF en cuáles ERCC spike-ins se desempeñaron mejor en bibliotecas con ribodepleción que en bibliotecas poli-A ^(99, 105).

Es importante señalar que el desarrollo de los estándares de referencia RNA-Seq es un proceso, como lo demuestra el recientemente lanzado Proyecto ERCC 2.0, para iniciar el desarrollo de controles de aumento de ARN actualizados que incluyen imitadores de ARNm mejorados, así como nuevos genes de fusión del cáncer y controles de ARN pequeño ⁽¹⁴³⁾.

1.6.7. Expresión Genética y Detección Alternativa de Empalmes

Después de la evaluación de calidad de la secuenciación, el análisis para la cuantificación de la transcripción, identificación de isoformas de empalme, variantes y genes de fusión se pueden iniciar ⁽⁶⁹⁾. Una de las más comunes aplicaciones de la secuenciación del transcriptoma está relacionada con la estimación de la expresión de los niveles de transcripción ⁽⁶⁹⁾.

La expresión de la transcripción de RNA-Seq se cuantifica agregando el número total de lecturas que se alinean con cada transcripción seguida de la normalización según el tamaño y la función de la biblioteca ⁽⁶⁹⁾, con expresión de transcripción expresada en términos de "fragmentos esperados por kilobase de transcripción por millón de fragmentos mapeados" (FPKM) o "Lecturas por kilobase por millón de lecturas asignadas" (RPKM) ⁽⁷³⁾. Programas especializados como "La suite de gemelos" y RSEM han desarrollado algoritmos sofisticados que ayudan en la asignación precisa de lecturas asignadas a más de una transcripción por el

uso de un enfoque de maximización de expectativas que explica los sesgos particulares asociados con la distribución de lectura no uniforme a lo largo de la longitud del gen ^(122, 123). Una vez que las transcripciones han sido cuantificadas y normalizadas, los datos pueden luego ser utilizados para evaluar la expresión diferencial de firmas de genes entre diferentes muestras, que son normalmente calculadas utilizando distribuciones de probabilidad discretas como Poisson o binomio negativo. Métodos populares de expresión diferencial como el análisis de los datos de RNA-Seq incluye edgeR, DESeq2 y NOISeq ⁽¹⁴⁴⁻¹⁴⁶⁾. El análisis diferencial del nivel de transcripción también puede detectar eventos alternativos comparando la expresión de diferentes isoformas de transcripción.

Existen programas desarrollados para la detección de eventos de empalme alternativo los que incluyen CuffDiff2, DSGSeq y rMATS, todos ellos utilizan diferentes algoritmos para identificar isoformas alternativas ⁽¹⁴⁷⁻¹⁴⁹⁾. *CuffDiff2*, parte de la suite de gemelos, integra la estimación de abundancia de la transcripción y expresión diferencial del análisis de isoformas, con métodos incorporados para el sesgo de corrección de secuencia y uso de una distribución binomio beta-negativo para asignar fragmentos a cada isoforma ⁽¹⁴⁷⁾. *DSGSeq* compara los recuentos de lectura en los exones y sus uniones para la detección de diferencias significativas en la abundancia de isoformas, mientras que *rMATS* identifica las uniones de exones presentes en las lecturas para detectar eventos de empalme diferencial de exones ^(148, 149).

Los métodos basados en exones o uniones tienden a tener mayor precisión en la identificación de alternativas específicas de eventos de empalme; sin embargo, todos estos enfoques tienden a verse obstaculizados en general por los factores intrínsecos como limitaciones de la secuenciación de lectura corta ⁽¹⁵⁰⁾. Múltiples estudios han demostrado una fuerte concordancia de la expresión de la transcripción normalizada entre RNA-Seq y

microarrays, con algunos estudios que atribuyen RNA-Seq con mayor sensibilidad ^(77, 151). Sin embargo, cabe señalar que esfuerzos de la FDA y su proyecto SeQC, así como por la ABRF han demostrado que las mediciones absolutas de la expresión génica de RNA-Seq son poco fiables, mientras que las medidas de la expresión génica relativa son precisas y reproducibles en sitios y varias plataformas ^(99, 105). Más, RNA-Seq es susceptible a sesgos comunes de secuenciación NGS, como contenido de GC, sesgos posicionales, efectos por lotes, ruido de fondo y otros sesgos de secuencia, además de los sesgos asociados con métodos de selección de transcripciones ^(93, 152).

Muchos de los filtros de preprocesamiento y métodos analíticos han sido desarrollados para reducir el ruido de fondo, aunque es importante tener en cuenta que a menudo mejoran su exactitud a expensas de la precisión ⁽⁹⁹⁾. Se han desarrollado paquetes como NOISEq R para ayudar en la identificación de sesgos específicos de conjuntos de datos individuales y ayudar en la normalización de los datos ⁽¹⁴⁶⁾.

Se recomienda encarecidamente el uso de múltiples réplicas biológicas en todos los experimentos de detección de isoformas y expresión génica, ya que a menudo los datos de pequeños conjuntos de muestras pueden ser ruidosos. Además, los estudios han demostrado que la elección del método e incluso la versión del software pueden afectar notablemente los resultados del análisis y ninguna metodología puede servir como la mejor opción para todos los conjuntos de datos ⁽¹¹⁴⁾. Por lo tanto, es importante documentar el software, su versión y configuración utilizado para el uso de análisis futuro ⁽⁶⁹⁾.

1.6.8. Detección de genes y variantes de fusión

Utilizando los avances de la secuenciación de próxima generación, RNA-Seq permite la resolución y cuantificación de pequeñas a grandes niveles de nucleótidos en la detección y expresión enfermedades asociadas con alteraciones tales como variantes de un solo

nucleótido, inserciones, deleciones y eventos de fusión de genes, determinando la expresión específica de alelos ⁽¹¹⁴⁾. En particular, RNA-Seq también puede capturar variantes que a veces son difíciles de detectar en el genoma, como translocaciones o eventos de empalme ⁽⁶⁹⁾.

La detección de la presencia y/o expresión de variantes de un solo nucleótido clínicamente relevantes y fusiones de genes, representa una de las más fáciles vías traducibles para secuenciación del transcriptoma clínico, especialmente como estrategias de tratamiento, las cuales se vuelven cada vez más dependiente de los datos moleculares y la presencia de alteraciones genéticas específicas en el tejido enfermo ⁽⁶⁹⁾. De hecho, el ARN dirigido a los paneles NGS diseñados específicamente para la detección de variantes y fusiones de genes ya comenzó a integrarse en la práctica clínica actual ⁽⁶⁹⁾. Debido a la complejidad del análisis RNA-Seq, diferentes flujos de trabajo de análisis de "mejores prácticas" se han desarrollado para cada tipo de variante ⁽¹⁵³⁾.

La detección de variantes de un solo nucleótido (SNV) y pequeñas deleciones en los datos de RNA-Seq tienen similares flujos de trabajo a los utilizados en la secuenciación del ADN, caracterizados por la alineación con el genoma humano/transcriptoma y pasos de preprocesamiento como el recorte de lectura, seguido de las variantes de llamadas que se utilizan normalmente para Secuencias de ADN ⁽⁶⁹⁾. Es de destacar que algunas variantes de llamadas, como el HaplotypeCaller de GATK, tienen modos de desarrolló específicos de RNA-Seq con el fin de aumentar la especificidad, combatiendo así las llamadas variantes erróneas asociadas con la dificultad para resolver sitios de unión de empalme ⁽¹⁵⁴⁾.

La detección de grandes deleciones en los datos de RNA-Seq es actualmente mucho más desafiante debido a la dificultad en la alineación local en los datos de RNA-Seq, así como las complicaciones debidas a eventos de empalme de exones que pueden conducir a un

gran grado de llamadas falsas positivas ⁽⁶⁹⁾. A pesar de algunos de los desafíos bioinformáticos asociados con la variante, invocando datos de RNA-Seq, las tuberías de RNA pueden exhibir una alta sensibilidad en la detección de variantes, pero a menudo requieren un filtrado de variantes para aumentar la especificidad ^(153, 155).

Otra clase de variante que puede ser detectada por RNA-Seq son *eventos de fusión de genes*. Las fusiones de genes son conductores conocidos en cáncer y, a menudo, pueden servir como biomarcadores diagnósticos o predictivos de la enfermedad ⁽¹⁵⁶⁾. Se ha demostrado que RNA-Seq es altamente sensible para la detección de eventos de fusión ⁽⁶⁹⁾.

El descubrimiento de fusiones de genes por RNA-Seq, mientras se da el descubrimiento de nuevas isoformas análogas, se complica por el hecho de que los segmentos de transcripción a menudo necesitan mapear a múltiples cromosomas o están separados por grandes distancias en el mismo cromosoma ⁽⁶⁹⁾. Como tal, se han desarrollado programas específicos para la detección de fusiones de genes en datos de RNA-Seq, incluidos deFuse, TopHat-Fusion, Fusion Hunter, PRADA, Chimera y SOAPfuse ⁽¹⁵⁷⁻¹⁶²⁾.

Cada algoritmo tiene sus propios beneficios, con algunas herramientas superando a otros en cuanto a sensibilidad, valor de predicción positiva, consumo de tiempo y uso de memoria ⁽¹⁶³⁾. Además, el desempeño de cada programa también puede depender de la muestra calidad, longitud de lectura, calidad de las lecturas y número total de lecturas de secuenciación logradas ^(163, 164).

Los artefactos también son muy comunes, como resultado de la desalineación de lecturas debido a homología, errores de secuenciación y polimorfismos ⁽¹¹⁴⁾. Por lo tanto, similar a las canalizaciones de las llamadas “RNA-Seq SNV”, metodologías de posprocesamiento que utilizan filtros heurísticos son necesarios para maximizar la especificidad ⁽⁶⁹⁾. Uno de los principales beneficios de RNA-Seq para la detección de

fusiones de genes en comparación con los ensayos casos clínicos más específicos como la reacción en cadena polimerasa de transcripción inversa (RT-PCR) es el beneficio adicional de la detección agnóstica de alteraciones de una manera de alto rendimiento ⁽⁶⁹⁾.

Los ensayos de RT-PCR son altamente sensibles pero específico para una sola alteración con múltiples reacciones individuales requeridas cuando se prueban varios socios de fusión o varios posibles puntos de interrupción, que pueden ser un gran inconveniente para genes que se sabe que tienen un conjunto muy diverso de genes secundarios asociados que generan múltiples diferentes transcripciones de fusión, como ALK, KIT, y ROS1 en el cáncer de pulmón ⁽¹⁶⁵⁾.

Además, la hibridación fluorescente in situ (FISH), un método alternativo a RT-PCR para la detección de translocaciones, es capaz de detectar reordenamientos que presentan el gen objetivo independientemente del gen secundario, aún es incapaz de identificar el gen secundario asociado específico.

RNA-Seq representa un avance sobre estas dos metodologías a través de su naturaleza agnóstica y de alto rendimiento, manteniendo al mismo tiempo un grado similar de sensibilidad ⁽¹⁶⁵⁾. Actualmente, el análisis de RNA-Seq está asociado con análisis bioinformáticos complejos, que a menudo limitan el grado de adopción de la tecnología ⁽⁶⁹⁾. A fin de reducir parte de la carga bioinformática asociada con RNA-Seq, se han desarrollado programas de análisis fáciles de usar "conecta y reproduce" que puede hacerlo más accesible para los típicos laboratorios clínicos ⁽⁶⁹⁾.

Por ejemplo, Illumina e Ion Torrent ha integrado una amplia gama de complementos en sus aplicaciones BaseSpace y Torrent Browser, respectivamente, para la detección de variantes, fusiones de genes y expresión génica diferencial en una manera fácil de usar que requiere muy pocas habilidades informáticas ⁽⁶⁹⁾. Otras empresas, como ArcherDx y

Asuragen, han desarrollado conjuntos de análisis bioinformáticos fáciles de usar que utilizan interfaces de gráficos de usuario como compañeros de su ARN objetivo, paneles NGS para realizar el análisis informático más accesible para los laboratorios ⁽⁶⁹⁾.

Esta tendencia se espera que continúe a medida que más y más empresas comiencen a desarrollar paneles NGS para aplicaciones clínicas ⁽⁶⁹⁾. Sin embargo, para aplicaciones clínicas, es importante comprender los conceptos básicos y suposiciones de cada una de estas herramientas, ya que pueden tener un impacto significativo en la sensibilidad y especificidad de cada prueba clínica ⁽⁶⁹⁾.

Laboratorios clínicos que implementan estas herramientas listas para usar, así como aquellos que desarrollan sus propias tuberías, deben adherirse a las pautas de la Asociación de Patología Molecular (AMP) y del Colegio de Patólogos Americanos (CAP) en el desarrollo y validación bioinformática de canales para NGS ⁽¹⁶⁶⁾.

1.6.9. Aplicaciones clínicas de RNA-Seq

Se desarrolló el perfil de expresión génica clínica en la última década, con aplicaciones para el diagnóstico del cáncer, pronóstico del cáncer y detección del rechazo de trasplantes, algunos de los cuales han obtenido la Autorización de la Administración de Drogas y Alimentos de los Estados Unidos (FDA) ⁽¹⁶⁷⁻¹⁷¹⁾.

Sin embargo, a pesar de la gran promesa, el uso generalizado de perfiles de transcriptomas para las enfermedades crónicas inflamatorias, neurológicas e infecciosas, por ejemplo, no se han convertido en realidad ⁽¹⁷²⁻¹⁷⁴⁾. Ayudado por la rápida disminución del costo de la generación de datos, y la técnica continua de avances de NGS, RNA-Seq tiene el potencial de convertirse en una poderosa herramienta en la gestión y tratamiento de

enfermedades humanas, aunque esta tecnología aún no se ha generalizado en el uso clínico (175, 176).

1.6.9.1. Oncología de precisión

El impacto de RNA-Seq ha sido posiblemente el más importante en el campo de la biología del cáncer y de la oncología clínica ⁽⁶⁹⁾. Del descubrimiento de varias fusiones de genes novedosos como impulsores oncogénicos en diferentes tipos de tumores para la identificación de firmas de expresión predictivas de la respuesta a la inmunoterapia, los estudios basados en RNA-Seq han promovido una influencia transformadora en nuestra comprensión del cáncer ⁽¹⁷⁷⁻¹⁷⁸⁾.

Cada vez más se ha llevado a intentos de aprovechar el poder de RNA-Seq para su uso clínico en oncología de precisión ⁽⁶⁹⁾. Mientras que el diagnóstico molecular en el cáncer sigue siendo dependiente en gran medida de la secuenciación del ADN, RNA-Seq ofrece varias ventajas complementarias en lo que ahora llamamos medicina personalizada, por la capacidad de detectar variaciones estructurales que conducen a fusiones de genes dirigibles que en su mayoría no son detectadas por los enfoques estándar basados en secuenciación de ADN dirigida (incluida la secuenciación completa del exoma) y la capacidad para evaluar el microambiente del tumor y predecir así una respuesta a la inmunoterapia evaluando los niveles de expresión de mutaciones oncogénicas ⁽⁶⁹⁾.

En virtud de proporcionar tanto una instantánea estructural de aberraciones genómicas y una imagen dinámica de procesos celulares a través del análisis de expresión, RNA-Seq ofrece una funcionalidad incomparable por la descripción general del genoma del cáncer ⁽⁶⁹⁾.

Se han realizado múltiples grandes ensayos de secuenciación genómica en el cáncer, lo que permitió incorporar a RNA-Seq en los diagnósticos basados en ADN con resultados prometedores como el iCat y ensayos MI-ONCOSEQ, y varios otros que están en curso, por ejemplo, NCI-MATCH y Texas KidsCanSeq ⁽¹⁷⁹⁻¹⁸³⁾. De los varios usos de RNA-Seq en el diagnóstico del cáncer, ninguno es más evidente que la capacidad para detectar genes de fusión de manera no sesgada.

Porque RNA-Seq proporciona una transcripción completa de secuencias y es capaz de ensamblar transcripciones sin depender de referencias de secuencias preexistentes, se ha utilizado para la detección de transcripciones de fusión, tanto para las que anteriormente se conocen que están asociadas con tumores específicos y para fusiones no identificadas previamente ⁽¹⁷⁵⁾.

Aunque las fusiones de genes han sido reiteradamente bien documentadas en neoplasias hematológicas y sarcomas durante varias décadas, el descubrimiento de fusiones de genes recurrentes en tumores epiteliales sólidos es relativamente reciente ⁽⁶⁹⁾. En 2005, Tomlins et al. informó el descubrimiento de transcripciones de fusión entre los TMPRSS2 y los genes del factor de transcripción ETS en el cáncer de próstata ⁽¹⁸⁴⁾.

Este descubrimiento transformó nuestra comprensión de los tumores sólidos y abrió la puerta a una avalancha de estudios que informaron fusiones recurrentes en una variedad de tipos de tumores ⁽¹⁸⁵⁻¹⁹¹⁾. Algunos de estos descubrimientos, como la identificación de los reordenamientos de ALK en el cáncer de pulmón, ya se han incorporado a los algoritmos de diagnóstico y las estrategias de gestión y manejo de pacientes con cáncer ^(192, 193).

Basado en los ensayos de RNA-Seq, tanto del transcriptoma completo como del dirigido con paneles RNA-Seq, se están utilizando para detectar transcripciones de fusión en tumores sólidos y hematológicos, neoplasias malignas por igual ⁽⁶⁹⁾.

Es probable que el número de las transcripciones de fusión informadas en cánceres humanos continúan aumentando a medida que el uso de RNA-Seq se vuelve más prevalente ⁽⁶⁹⁾. Como en el caso del cáncer de pulmón y tiroides, algunas de estas transcripciones de fusión podrían ser de utilidad para el diagnóstico, pronóstico o selección de terapias dirigidas ⁽⁶⁹⁾. Las detecciones de transcripciones de fusión se han convertido rápidamente en una rutina en la práctica de la patología clínica ^(165, 194).

Otro uso del perfil del transcriptoma es la identificación del *origen tisular en cánceres de sitio primario desconocido*. Varias plataformas comerciales están disponibles clínicamente para este propósito ^(167, 195, 196). Aunque esta aplicación aún no ha migrado a plataformas de RNA-Seq, es previsible que esto ocurra en un futuro próximo, porque esta tecnología permite no solo la identificación del sitio de origen basado en el patrón de expresión sino también la detección de mutaciones expresadas y eventos de empalme alternativo que podría ser de utilidad para la selección de la terapia ⁽¹⁹⁷⁾.

El cáncer de mama, cáncer de pulmón, cáncer de tiroides, melanoma y el tratamiento del cáncer de próstata son otras áreas donde el perfil del transcriptoma se ha incorporado en el manejo clínico de rutina, con el uso de perfiles de expresión génica que determinan la probabilidad de un diagnóstico de cáncer, con valor pronósticos de recurrencia tumoral o que son utilizados para el manejo terapéutico en la etapa temprana de estos pacientes ^(168, 194, 198-205).

Mientras que la mayor parte de estas pruebas utilizan PCR cuantitativa con transcriptasa inversa (RT-PCR) o microarrays, esfuerzos para traducir algunos de estos paneles en RNA-Seq, en plataformas ya están en marcha ⁽²⁰⁶⁾. Esto tiene el potencial para ampliar el contenido de estos paneles e incorporar biomarcadores terapéuticos en estas pruebas de pronóstico ^(169, 206).

La utilidad de RNA-Seq se aprecia mejor cuando se integra con el perfil genómico multimodal. Cuando se realiza con toda la secuenciación del exoma o la secuenciación dirigida, RNA-Seq ofrece la capacidad de evaluar el efecto funcional de variantes de empalme, omisión de exón y del empalme aberrante ⁽⁶⁹⁾. Cuando se combina con el número de copias evaluadas, RNA-Seq permite la correlación entre alteraciones de ganancia de función como amplificaciones de alto nivel, regulación ascendente de expresión y descubrimiento de alteraciones de pérdida de función como la expresión monoalélica en el contexto del silenciamiento del promotor ⁽⁶⁹⁾.

En resumen, la investigación básica y clínica con RNA-Seq en oncología nos proporciona un tesoro de información que debería permitirnos comprender mejor el inicio y la progresión del tumor, y resistencia terapéutica ⁽⁶⁹⁾.

La complejidad del cáncer ha sido evidenciada por investigaciones recientes y ahora está claro que la comprensión biología del cáncer y nuestra capacidad para personalizar tratamiento y de impactar en los resultados requerirá el uso de todas las tecnologías "ómicas" disponibles porque no todas las alteraciones moleculares que impulsan el comportamiento tumoral se detectan mediante un único enfoque ⁽²⁰⁷⁾.

Ejemplos de esto son los tumores pediátricos, en qué las mutaciones son menos frecuentes que en los adultos y donde los análisis de transcriptomas han identificado genes desregulados que podrían descubrir nuevos enfoques terapéuticos dirigidos ⁽²⁰⁸⁾.

1.7. Aplicaciones genómicas en el Cáncer de Tiroides

En la patogenia del cáncer participan algunos procesos que implican la adquisición continua de variación genética, principalmente a través de mutaciones aleatorias en el ADN y la selección natural que actúa sobre la consiguiente diversidad en el fenotipo ⁽⁶⁹⁾.

A pesar de las diversas defensas del cuerpo contra la transformación células cancerígenas, se acumulan células raras con suficientes mutaciones ventajosas que permitan su supervivencia y aumento de la proliferación celular, resultando clínicamente en un tumor ⁽⁶⁹⁾. Estas y otras propiedades otorgan a estas células la capacidad de invadir tejidos adyacentes y hacer metástasis a distancia, las que son características patológicas del cáncer ^(209, 210). En consecuencia, mucho esfuerzo en la investigación del cáncer sobre las últimas décadas se ha dirigido a identificar estas alteraciones genéticas y comprender sus consecuencias sobre la función celular ⁽⁶⁹⁾.

Estos esfuerzos se han acelerado enormemente debido a los fenomenales avances en tecnologías genómicas como la secuenciación de ADN de próxima generación ⁽²¹¹⁾. Como consecuencia y a través de los esfuerzos coordinados de redes de investigadores como el Atlas del genoma del cáncer (TCGA) y el Consorcio Internacional del Genoma del Cáncer (ICGC), la última década ha sido testigo de una rápida caracterización de las alteraciones de todo el genoma que ocurren en los tipos de cánceres más comunes ⁽²¹²⁾, incluidos los cánceres derivados de las células foliculares tiroideas.

La glándula tiroides contiene dos tipos de células endocrinas, y los cánceres de tiroides se clasifican de acorde a esta división ⁽⁶⁹⁾. Tumores derivados de las células foliculares que representan la gran mayoría de las neoplasias tiroideas, que pueden ser benignas (adenomas foliculares) o poseer potencial maligno, es decir carcinomas ⁽⁶⁹⁾.

Por el contrario, los tumores derivados de las células neuroendocrinas de la tiroides, denominadas células parafoliculares o C, son tumores neuroendocrinos denominado carcinoma medular ⁽⁶⁹⁾. Dada la naturaleza relativamente rara del carcinoma medular, casi todos los esfuerzos de caracterización del genoma se han dirigido a los cánceres de tiroides de células foliculares ⁽²¹³⁾.

Las neoplasias foliculares de tiroides se clasifican según un esquema taxonómico simple basado sobre los parámetros histopatológicos tradicionales ⁽⁶⁹⁾. Tumores compuestos por células foliculares que carecen potencial maligno se diagnostican como adenoma folicular de tiroides (AFT). Dado su benigno comportamiento, los AFTs todavía no han sido genómicamente caracterizado en cualquier grado significativo ⁽⁶⁹⁾. El carcinoma papilar de tiroides (CPT) representa el cáncer de tiroides más común y, en consecuencia, fue el tipo de tumor seleccionado por TCGA para su proyecto sobre el cáncer de tiroides ⁽⁶⁹⁾. Más allá del CPT, otros cánceres de tiroides altamente diferenciados incluyen el carcinoma folicular de tiroides (CFT) y el carcinoma de células de Hurthle (CCH) por mencionar ⁽⁶⁹⁾.

Alguna información genética está disponible para estos tumores, pero se están realizando esfuerzos de caracterización genómica multidimensional del CFT y CCH. Colectivamente, CPT, CFT y CCH se han agrupado como cánceres de tiroides diferenciados porque retienen importantes diferenciaciones celulares ⁽⁶⁹⁾.

Más allá de estos cánceres diferenciados de tiroides, el cáncer de tiroides anaplásico (CTA) representa una forma clínicamente agresiva de cáncer de tiroides en el que las células tumorales han perdido su diferenciación tiroidea. Respectivamente, los CTA se denominan como sinónimo cáncer de tiroides indiferenciado ⁽⁶⁹⁾.

El carcinoma de tiroides pobremente diferenciado (CAPD) se refiere a tumores intermedio con reducida diferenciación tiroidea, que clínica y patológicamente caen entre carcinomas diferenciados e indiferenciados ⁽⁶⁹⁾.

Dada su naturaleza agresiva, el CAPD y el CTA, han sido objeto de importantes pruebas de investigación genómica, aunque se están realizando estudios adicionales. Información esencial sobre la genética y genómica de estos cánceres de tiroides se presentan en detalle y se resumen en la Tabla 4 ⁽⁶⁹⁾.

Tabla 4

Resumen de las características genéticas y genómicas de los cánceres de tiroides

Tipo de carcinoma tiroideo	Alteraciones somáticas			BRAF ^{V600E} - como firma de expresión	Carga mutacional tumoral
	SSNVs	Fusión de genes	SCNAs		
Carcinoma papilar	<i>BRAF^{V600E}</i> , <i>RAS</i>	<i>RET</i> , <i>NTRK1/3</i> , <i>BRAF</i> , <i>ALK</i>	Pocas; pérdida de 22q, ganancia de 1q	Presente	Baja
Carcinoma folicular	<i>RAS</i> , <i>PTEN</i> , promotor <i>TERT</i> , <i>RBI</i>	<i>PAX8-PPARG</i>	Muchas; pérdida de 22q, ganancia de 1q	NA	Baja
Carcinoma de células de Hürthle	<i>RAS</i> , <i>TP53</i> , promotor <i>TERT</i> , DNA mitocondrial	Desconocida	Muchas	NA	Baja
Carcinoma pobremente diferenciado	<i>BRAF^{V600E}</i> , <i>RAS</i> , promotor <i>TERT</i> , <i>TP53</i> , <i>EIF1AX</i>	Desconocida	Desconocida	Presente	Moderada
Carcinoma anaplásico	<i>BRAF^{V600E}</i> , <i>RAS</i> , <i>TP53</i> , promotor <i>TERT</i> , <i>EIF1AX</i> , mutaciones heterogéneas	<i>RET</i>	Muchas (aneuploidía)	Ausente	Alta

Nota. La presente tabla muestra las diferentes mutaciones que fueron encontradas en los distintos tipos de cáncer de tiroides. Tomada de Netto G. J. (2019) ⁽⁶⁹⁾.

1.7.1. Carcinoma Papilar de Tiroides

El carcinoma papilar de tiroides (PTC) fue caracterizado genómicamente por la Red de Investigación TCGA como 1 de sus 23 proyectos comunes de cáncer, además de sus 10 proyectos de cáncer poco común ⁽²¹⁴⁾. La cohorte de estudio constaba de 496 tumores primarios y 8 metastásicos que afectaban a los ganglios linfáticos regionales ⁽²¹⁴⁾. Se obtuvieron datos clínico-patológicos asociados para cada tumor, y los CPT se clasificaron

adicionalmente como uno de los tres subtipos patológicos predominantes, es decir, CPT clásico o habitual ($n = 324$), CPT variante de células altas ($n = 35$) o variante folicular ($n = 99$) del CPT ⁽²¹⁴⁾. Los tumores se evaluaron mediante las plataformas moleculares TCGA estándar, que incluían secuenciación de ADN de todo el exoma, secuenciación de ARNm, secuenciación de miARN, perfil de número de copias, y perfiles de metilación del ADN ⁽²¹⁴⁾. El análisis de los datos moleculares resultantes fueron ubicados en tres líneas. *Primero*, mutaciones somáticas, que incluían variantes de un solo nucleótido (SSNV), pequeñas inserciones y deleciones (indels), fusiones de genes y número de alteraciones de copias. (SCNA), fueron identificadas para caracterizar el paisaje genómico del CPT, con un especial énfasis en aquellos casos sin mutaciones previas controladoras conocidas ⁽²¹⁴⁾. En *segundo* lugar, se exploraron las consecuencias biológicas de las mutaciones utilizando datos moleculares multidimensionales ⁽²¹⁴⁾. Finalmente, se utilizaron datos moleculares para obtener conocimientos novedosos en la clasificación de CPT incorporando datos sobre genotipo, señalización celular, diferenciación, y riesgo clínico de comportamiento agresivo ⁽²¹⁴⁾. Usando datos de secuenciación de ADN del exoma completo, el estudio TCGA demostró un nivel relativamente bajo de carga mutacional tumoral (0,41 mutaciones no sinónimas por Mb) en comparación con otros cánceres, especialmente aquellos asociados con carcinógenos como el humo del cigarrillo (pulmón) y la luz solar (melanoma) ⁽²¹⁵⁾. La carga de mutaciones tumorales aumenta con la edad del paciente ⁽²¹⁵⁾.

El estudio TCGA confirmó el papel dominante de mutaciones BRAF en el CPT. BRAF^{V600E} fue de lejos la mutación más común, aunque el estudio ilustró las muchas formas en que BRAF puede transformarse en un oncogén ⁽⁶⁹⁾. Por ejemplo, indels (una mutación nombrada con la combinación de inserción y deleción) BRAF, así como diversas fusiones de genes, resultaron ser mutuamente excluyente con las alteraciones más comunes, lo que sugiere que estas raras mutaciones son oncogénicas ⁽⁶⁹⁾. Se identificó un solo caso de una variante

folicular de CPT con BRAF^{K601E}. Colectivamente, alguna forma de mutación BRAF estuvo presente en > 60% de los tumores ⁽⁶⁹⁾.

La mayoría de los otros SSNV ocurrió más allá de BRAF^{V600E} dentro de uno de los genes RAS, con la siguiente frecuencia relativa (NRAS, 8,5%; HRAS, 3,5%; KRAS, 1,2%). Como se esperaba, las mutaciones RAS ocurrieron como mutaciones hotspot en los codones 12 y 61 ⁽⁶⁹⁾. Las mutaciones somáticas de EIF1AX fueron identificadas por primera vez en el cáncer de tiroides por el estudio TCGA, con 1,5% de los casos de CPT, estas portan mutaciones puntuales de este factor de iniciación de la traducción ligado al cromosoma X ⁽⁶⁹⁾. Desde el informe TCGA, las mutaciones de EIF1AX han sido confirmadas en el CPT y encontradas en otros tipos del cáncer de tiroides ⁽²¹⁶⁻²¹⁹⁾. El mecanismo oncogénico de EIF1AX mutado puede involucrar expresiones génicas aberrantes a través de la estabilización de la traducción de sitios de preiniciación ⁽²²⁰⁾. Otros genes mutados significativamente en consecuencia al algoritmo MutSig ⁽²²¹⁾ incluyen a PPM1D y CHEK2. Los genes de cáncer conocidos con mutaciones incluyen TP53, ARIDB1, MLL, PTEN, ATM, RB1, EZH1, MEN1, MLL3, APC y NF1 ⁽⁶⁹⁾. Además, se identificaron mutaciones del promotor TERT en el 9% de los casos y se asociaron con características agresivas. El estudio TCGA reafirmó y amplió el papel de los reordenamientos y translocaciones de genes en el CPT ⁽⁶⁹⁾.

Las fusiones de genes se encontraron en el 15,3% de tumores, y ningún tumor contenía más de una fusión. Además de ser mutuamente excluyentes entre sí, las fusiones de genes no coexistieron con mutaciones BRAF, RAS o EIF1AX ⁽⁶⁹⁾. Esta relación de mutaciones impulsoras mutuamente excluyentes representa uno de los ejemplos más sorprendentes en todos los tipos de cáncer e ilustra el papel dominante que juegan estas mutaciones en la patogénesis del CPT. Además de identificar nuevos socios de fusiones de genes comunes

(por ejemplo, FKBP15-RET), se descubrieron fusiones completamente nuevas, por ejemplo, UCAC-LTK ⁽⁶⁹⁾.

La gran mayoría de estos tumores demostrada preservación y sobreexpresión de dominios quinasa en la fusión resultante, proporcionando evidencia de su función como quinasas activadas constitutivamente. Además de las mutaciones puntuales y fusiones de genes, una minoría significativa (27,2%) de los CPT también contenía SCNA, que estaban preferentemente presente en casos sin el controlador (driver) común, mutaciones descritas anteriormente ⁽⁶⁹⁾.

Esta relación sugiere un papel para los SCNA como eventos impulsores en el CPT. Los tumores con SCNA se enriquecieron para la variante folicular del CPT, de acuerdo con las observaciones anteriores y proporcionando apoyo para una cierta relación entre la variante folicular del CPT y las verdaderas neoplasias foliculares ⁽²²²⁾. Estudios comparativos recientes de pan-cáncer en muchos tipos de cáncer destacó las tasas excepcionalmente bajas de aneuploidia observadas en el CPT ⁽²²³⁾.

Las vías alteradas somáticamente se categorizaron utilizando sofisticados métodos computacionales enfoques que combinan datos sobre SSNV, fusión de genes y SCNA. El algoritmo HotNet2 ⁽²²⁴⁾ identificó numerosas subredes relevantes, incluyendo la vía MAPK, las interacciones ECM-receptor y el complejo proteico asociado a FANCA. Con este enfoque, RAPIGAP fue identificado como un raro gen impulsor potencial en la vía MAPK y la base de datos COMSIC confirmó que este gen rara vez está mutado en los cánceres de tiroides ⁽²²⁵⁾.

Combinando variantes de un solo nucleótido, indels, fusiones de genes y alteraciones en el número de copias, se encontraron mutaciones impulsoras (drivers) en la gran mayoría del

CPT (98,8%). Este resultado colectivo representó un gran avance con importantes implicaciones clínicas para los ensayos de genotipificación de nódulos tiroideos ⁽²²⁶⁻²²⁸⁾.

El estudio de TCGA también arrojó información sobre la consecuencia de los defectos genómicos multidimensionales y de la exclusión mutua de las mutaciones BRAF^{V600E} y RAS ⁽⁶⁹⁾. Una puntuación basada en la expresión génica fue desarrollada, realizando mediciones para ver si el perfil de expresión génica de un tumor dado se asemejaba a perfiles mutaciones de BRAF^{V600E} o RAS ⁽⁶⁹⁾. Usando una gran cohorte de tumores (n = 391) que tenían tanto al exoma y la secuenciación de ARN, la firma de expresión génica de 71 genes se derivó comparando tumores con mutación BRAF^{V600E} y RAS ⁽⁶⁹⁾.

Los datos de estos 71 genes se utilizaron luego para derivar una puntuación continua, denominada puntuación BRAF^{V600E}-RAS (BRS), siendo las puntuaciones negativas llamada similar a BRAF^{V600E} y similar a RAS positivo ⁽⁶⁹⁾. La BRS se utilizó como escala de referencia para interrogar las consecuencias de la señalización de las mutaciones observadas en la cohorte, como se muestra en la figura 8 ⁽⁶⁹⁾.

La vista resultante arrojó varias perspectivas interesantes. Todas las mutaciones de BRAF^{V600E} fueron similar a BRAF^{V600E}, y todas las mutaciones de RAS eran similares a RAS, con una excepción de triple mutante ⁽⁶⁹⁾. Las mutaciones BRAF (indels y BRAF^{K601E}) fueron similar a RAS, un resultado que ilustra la complejidad de la señalización MAPK con diferencias en la retroalimentación de ERK ^(229, 230). Por el contrario, las fusiones BRAF y RET fueron débilmente similares a BRAF^{V600E}, y otras las fusiones fueron neutrales en el esquema BRS ⁽⁶⁹⁾. Estos resultados destacan las diferencias en la funcionalidad de estas mutaciones impulsoras (driver) y son reforzadas por correlaciones entre genotipo-fenotipo, incluyendo la observación de que los cánceres con BRAF^{K601E} mutado están altamente enriquecidos para la variante folicular del CPT ⁽²³¹⁾.

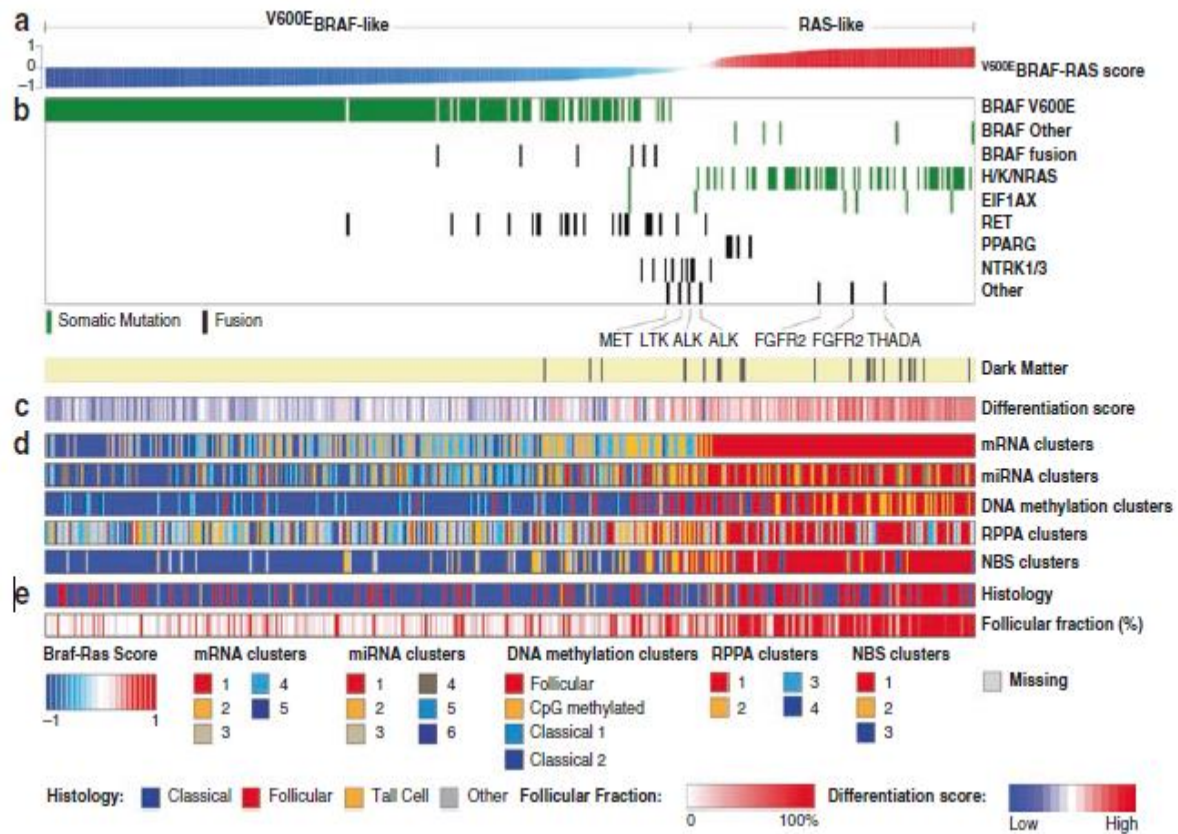


Figura 8: La puntuación $BRAF^{V600E}$ -RAS (BRS). (a) 391 CPTS se clasificaron de acuerdo al puntaje $BRAF^{V600E}$ -RAS, siendo negativo similar al $BRAF^{V600E}$ y siendo positivo similar a RAS. El BRS se asoció fuertemente con (b) estado mutacional, (c) puntuación de diferenciación tiroidea, (d) agrupaciones genómicas, (e) tipo histológico y la fracción folicular fracción. (De Cancer Genome Atlas Research ⁽⁶⁶⁾ con permiso de Elsevier). Tomada de Netto G. J. (2019) ⁽⁶⁹⁾.

Dado el papel esencial de la diferenciación de células foliculares en el cáncer de tiroides, el estudio TCGA exploró el impacto de la diferenciación en la cohorte extrayendo los datos de expresión génica para 16 genes de función y metabolismo de la tiroides ⁽⁶⁹⁾. La puntuación resultante, denominada diferenciación tiroidea Score (TDS), se utilizó nuevamente como escala de referencia a fin de interrogar el papel de la diferenciación ⁽⁶⁹⁾, a través de los varios grupos de mutaciones (figura 9).

Este análisis confirmó una mayor diferenciación de células foliculares en tumores con mutación RAS en comparación con tumores con fusiones RET o BRAF y $BRAF^{V600E}$ ⁽⁶⁹⁾. Sin embargo, la diferenciación fue muy variable dentro de la Cohorte $BRAF^{V600E}$, un hallazgo

inesperado que ha dado implicaciones del poder pronóstico de BRAF^{V600E} y la respuesta a la terapia con RAI ⁽⁶⁹⁾.

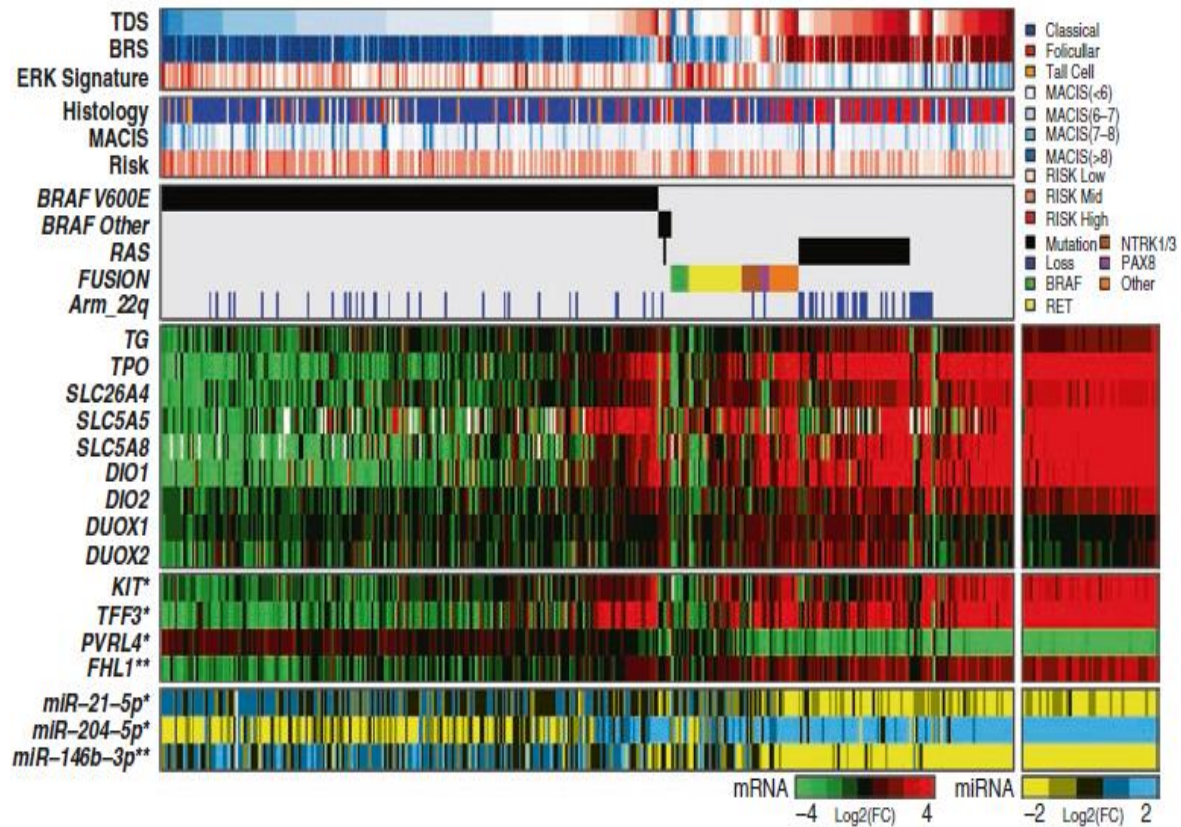


Figura 9: Papel de la diferenciación tiroidea. La puntuación de diferenciación tiroidea (TDS) en la cohorte TCGA con tumores ordenados por estado mutacional. Debajo del TDS está el BRS, firma ERK. Ver ⁽⁶⁾, histología, puntuación MACIS, riesgo de recurrencia, mutaciones y datos de expresión génica para genes relacionados con la tiroides y otros. (Del genoma del cáncer Atlas Research ⁽⁶⁶⁾ con permiso de Elsevier). Tomada de Netto G. J. (2019) ⁽⁶⁹⁾.

Los datos genómicos también se utilizaron para derivar varias clasificaciones moleculares del CPT. El BRS y puntuaciones de TDS se incorporaron en los análisis para ilustrar las diferencias biológicas entre los clusters moleculares resultantes y para informar cualquier relación entre grupo tumoral, histología tumoral, genotipo, señalización y diferenciación ⁽⁶⁹⁾.

Un ejemplo basado en la expresión de miARN logro identificar diferentes clusters con diferenciales en la expresión oncogénica de miR (p. ej., miR-146b) y medidas bajas de BRS

y TDS, lo que sugiere que la expresión de miR oncogénicos puede modular el fenotipo del CPT ⁽⁶⁹⁾.

Al comienzo del proyecto TCGA, hubo preocupación de que los resultados puedan ser peatonales debido a la naturaleza indolente de CPT ⁽⁶⁹⁾. De lo contrario, el genoma silencioso general de CPT con tumor de baja carga mutacional y genoma estable con pocas alteraciones en el número de copias, permitió lograr conocimientos críticos sobre muchos aspectos de la biología del CPT, incluidas las consecuencias de las mutaciones conductoras o driver ⁽⁶⁹⁾. Tales conocimientos son difíciles de extraer en tipos de tumores más complicados con genomas inestables, grados más altos de carga mutacional y múltiples vías activadas ⁽⁶⁹⁾. Por lo tanto, al final el estudio TCGA tuvo éxito identificando mutaciones novedosas, revelando conocimientos en la patogenia del CPT, proporcionando así una base para la comparación de estudios adicionales de otros tipos de cáncer de tiroides ⁽⁶⁹⁾. Perfiles genómicos tumorales de pacientes con cáncer avanzado se han convertido rápidamente en el estándar de atención oncológica ⁽⁶⁹⁾. En consecuencia, los datos de estos programas de secuenciación clínica, tanto de laboratorios académicos como comerciales, han estado acumulando información, tanto así que ahora el análisis de grandes cohortes es factible e informativo ⁽⁶⁹⁾. Dicho informe ⁽²³²⁾ sobre las alteraciones genéticas presentes en una gran cohorte de CPTS en estadio avanzado utilizando dos ensayos genómicos similares (MSK-IMPACT ⁽²³³⁾ y FoundationOne ⁽²³⁴⁾) reforzó los hallazgos de TCGA, pero también destacó las diferencias en la enfermedad en estadio avanzado en comparación con la cohorte TCGA ⁽⁶⁹⁾. Similar a los hallazgos del TCGA, una baja carga tumoral mutacional global, también se observó. La carga mutacional de nuevo aumenta con la edad del paciente, con CPT pediátricos exhibiendo el menor número de mutaciones ⁽⁶⁹⁾. Las mutaciones que involucran a BRAF dominaron el paisaje mutacional somático, con 74% de los casos que posee una forma de esta mutación. Mutaciones RAS y fusiones RET se encontraron en 9% y 7% de casos, respectivamente ⁽⁶⁹⁾. Similar a los

resultados de TCGA, mutaciones conductoras que involucran BRAF, RAS y RET eran fuertemente excluyentes mutuamente ⁽⁶⁹⁾. Colectivamente, 84% y 18% de los casos tenían alteraciones de MAPK y genes de la vía PI3K/AKT mTOR, respectivamente ⁽⁶⁹⁾. De acuerdo con el modelo emergente de las mutaciones del promotor TERT asociadas con tumores clínicamente agresivos, estas mutaciones estaban presentes en el 61% de los casos y estaban enriquecidos en pacientes mayores ⁽⁶⁹⁾. Del mismo modo, mutaciones de TP53 fueron mucho más frecuente (10%) en comparación con la cohorte TCGA (0,7%). Mutaciones de CDKN2A y RBM10, no observadas en TCGA, estuvieron presentes en 8% y 7% de los casos, respectivamente ⁽⁶⁹⁾. Colectivamente, los genes mutados observados en gran parte superpuestos con los de TCGA, estaban mutados a tasas notablemente más altas, un hallazgo que casi ciertamente refleja las verdaderas diferencias biológicas presentes en la enfermedad en estadio avanzado ⁽⁶⁹⁾.

Los CPT pediátricos poseen un paisaje genético algo distinto, caracterizado por una menor carga mutacional tumoral; fusiones de genes quinasa más frecuente observados que implican a RET, ALK y NTRK; y mutaciones BRAF menos frecuentes ^(232, 235-237). Información adicional sobre la patogénesis de la PTC se derivará de los estudios recientemente publicados del proyecto TCGA Pan-Cancer Atlas, donde encuentran diferentes aspectos de la genómica del cáncer ⁽²³⁸⁻²⁴¹⁾.

(<https://www.cell.com/pbassets/consortium/pancanceratlas/pancani3/index.html?code=cell-site>).

1.7.2. Carcinoma Folicular de Tiroides

El carcinoma folicular de tiroides (CFT) es mucho menos común que el CPT y, en consecuencia, sus características genómicas se han estudiado y definido menos a fondo ⁽⁶⁹⁾. El CFT se caracteriza por mutaciones puntuales de RAS y reordenamientos de PPARG ⁽²⁴²⁾,

así como otras alteraciones somáticas que involucran la vía PI3K/AKT mTOR ⁽⁶⁹⁾. Los pacientes con síndrome de Cowden tienen línea germinal de inactivación de PTEN y pueden desarrollar tanto AFT como CFT con una distribución multifocal, apoyando un modelo en el que el CFT surge de AFT preexistentes ⁽²⁴³⁾. En una secuenciación de ADN dirigida a todo el genoma de los 48 pacientes con CFT, revelaron una heterogeneidad de alteraciones somáticas de conducción, muchas que eran novedosas, incluidas mutaciones de oncogenes (MDM2, FLI1), reguladores transcripcionales (MITF, FLI1, ZNF331), enzimas epigenéticas (KMT2A, NSD1, NCOA1, NCOA2) y quinasas (JAK3, CHEK2, ALK) ⁽²⁴⁴⁾.

Curiosamente, la población polaca tenía una tasa muy baja (4%) de mutación RAS ⁽⁶⁹⁾. Un estudio genómico reciente de adenomas foliculares de tiroides (AFT) y CFT se realizó para caracterizar sus genomas y explorar la relación etiológica entre estas dos patologías ⁽²⁴⁵⁾. Utilizando secuenciación del exoma completo, no se observaron diferencias significativas en la frecuencia de mutaciones no sinónimas en el AFT y el CFT ^(69, 245). Además, la frecuencia de mutaciones impulsoras conocidas fue similar entre los dos tipos de tumores, y no se observaron correlaciones entre las características genómicas y clínico-patológicas ⁽⁶⁹⁾.

Sin embargo, hubo diferencias significativas en SCNA con CFT que poseen seis veces más tasa de SCNA, apoyando un papel para su participación en la progresión de AFT al CFT ^(246, 247). Los SCNA recurrentes incluyeron una ganancia de 1q y pérdida de 22q, las cuales se observaron en el estudio TCGA ⁽⁶⁹⁾. Un trabajo elegante ha demostrado que la pérdida de NF2, ubicada en 22q, promueve la oncogénesis a través de un aumento de la señalización de las MAPK específicamente en cánceres de tiroides con mutación de RAS ⁽²⁴⁸⁾. Las mutaciones del promotor TERT también se han reportado en una minoría significativa de CFT y se han asociado con una enfermedad agresiva ^(249, 250). Datos de la secuenciación clínica de 65 casos de CFT en estadio avanzado brindó la oportunidad de

examinar su genética ⁽²³²⁾. Las mutaciones RAS fueron las más frecuentes (66%) y las mutaciones BRAF (7,6%) menos frecuentes ⁽⁶⁹⁾. Mutaciones raras de BRAF conocidas de estar asociadas con tumores de patrón folicular (por ejemplo, BRAF^{K601E} en la variante folicular CPT) fueron sobrerrepresentados, y solo un caso tenía BRAF^{V600E} ⁽⁶⁹⁾. En esta cohorte avanzada, las mutaciones del promotor TERT fueron notablemente más frecuentes (71%), lo que proporciona más apoyo a su papel en la agresión de la enfermedad ^(251, 252). Mutaciones en PTEN (9%) y RB1 (9%) fueron más frecuentes en comparación con el CPT ⁽⁶⁹⁾. Enfoques genómicos de ADN libre de células circulantes recién comienzan a aplicarse al CFT, con resultados prometedores ⁽²⁵³⁾.

1.7.3. Carcinoma de células de Hürthle

El carcinoma de células de Hurthle (CCH) es tipo de cáncer de tiroides con una característica distintiva que se compone de grandes células con abundante citoplasma granular y nucléolos prominentes, denominados oncocitos ⁽²⁵⁴⁾. El fenotipo oncocítico está relacionado con la acumulación de mitocondrias ⁽²⁵⁵⁾.

Los datos colectivos sugieren que el CCH tiene un genoma distintivo caracterizado por una mayor frecuencia de mutaciones mitocondriales ^(256, 257), diferente impulsor de perfiles mutacionales ⁽²⁵⁸⁾ y altas tasas de aneuploidia ^(247, 258, 259). Los estudios de transcriptomas han revelado tres clases de tumores que corresponden aproximadamente al adenoma, CCH mínimamente invasivo y el CCH ampliamente invasivo ⁽²⁵⁸⁾. Datos de secuenciación clínica de 35 pacientes con cáncer de CCH avanzado, afirman la opinión de que el CCH es un tipo distinto de cáncer de tiroides ⁽⁶⁹⁾.

Las mutaciones RAS fueron menos frecuentes (15%) y no se identificaron mutaciones en BRAF o RET. TP53 estuvo presente en el 20% de los tumores ⁽⁶⁹⁾. Las mutaciones del promotor TERT fueron frecuentes (59%), lo que nuevamente sugiere un papel

universal para el promotor TERT en el desarrollo de cánceres de tiroides agresivos de todo tipo ⁽⁶⁹⁾. Como se señaló anteriormente, en la genómica multidimensional, se están realizando estudios de CCH y se espera que proporcionen información adicional sobre el genoma de este enigmático tipo de cáncer de tiroides ⁽⁶⁹⁾.

1.7.4. Carcinoma Pobrementemente Diferenciado

El CAPD es un cáncer de tiroides relativamente raro que representa un tipo intermedio que taxonómicamente se encuentra entre los cánceres de tiroides diferenciados y el carcinoma anaplásico ⁽⁶⁹⁾. Estudios genómicos recientes del CAPD confirma este punto de vista informado por patología ⁽⁶⁹⁾.

En un estudio de una gran cohorte de CAPD (n = 84) el uso de MSK-IMAPCT sugirió que el CAPD contenía un mayor número de alteraciones de conducción (driver) que se concentra en unas pocas vías o grupos funcionales de genes, incluidos PI3K/AKT/mTOR, SWI/SNF complejo remodelador de cromatina, histonas metiltransferasas y de reparación de desajustes de ADN ⁽²¹⁸⁾. Además, las mutaciones TP53 (8%) y el promotor TERT (40%) se encontraron con mayor frecuencia en comparación con el CPT ⁽⁶⁹⁾.

Así mismo, las mutaciones del promotor TERT fueron clonales, en contraste con el CPT en el que eran subclonales según datos del TCGA ⁽⁶⁹⁾. Similar al CPT, los datos genómicos apoyan la vista que CAPD se puede dividir en similares BRAF^{V600E} (BRAF^{V600E}-like) y subgrupos similares a RAS (RAS-like) ⁽²⁶⁰⁾. Esta distinción exhibe una alta correlación con la clasificación histopatológica ⁽⁶⁹⁾.

La forma original del CAPD, denominada carcinoma insular, ahora se sabe que representa una forma de CAPD similar a RAS con una alta frecuencia de mutaciones RAS ^(218, 261, 262).

Por el contrario, ahora se reconoce una forma no insular del CAPD basada en la sobre tasa

mitótica, necrosis y alto grado nuclear ⁽²⁶³⁾. A diferencia de los CAPD similares a RAS, dichos CAPD con frecuencia contienen mutaciones BRAF^{V600E} y tienen un fenotipo similar a BRAF^{V600E} ⁽²¹⁸⁾. CAPD también contiene mutaciones frecuentes de EIF1AX (11%) en comparación con 1,5% en el CPT ⁽⁶⁹⁾.

En el CPT, las mutaciones EIF1AX fueron mutuamente excluyentes con la mutación conductora común; sin embargo, en el CAPD, eran coexistentes con mutaciones RAS ⁽⁶⁹⁾. Además, los tumores con mutaciones EIF1AX estuvieron asociados con tumores más grandes y una supervivencia más corta ⁽²¹⁸⁾.

1.7.5 Carcinoma Anaplásico de Tiroides

CAT es la forma más agresiva y letal del cáncer de tiroides y, por definición, se compone de células foliculares indiferenciadas ⁽⁶⁹⁾. Dado su estado indiferenciado, el CAT representa la última etapa en el esquema de progresión del cáncer de tiroides, que afortunadamente solo ocurre en un pequeño número de pacientes con una incidencia de 1 a 2 casos por millón ⁽⁶⁹⁾. Sin embargo, ATC es responsable para una proporción significativa de cáncer de tiroides mortalidad ⁽⁶⁹⁾.

En consecuencia, el CAT ha sido el tema de investigación genética y genómica significativa en un intento por comprender mejor su patogénia y desarrollar nuevos enfoques terapéuticos ⁽⁶⁹⁾. Se cree que el CAT surge de cánceres de tiroides preexistentes, un punto de vista respaldado por sus impulsores genéticos con cánceres de tiroides diferenciados (por ejemplo, BRAF^{V600E} y RAS) y varias observaciones patológicas ⁽⁶⁹⁾.

Muchos CAT contienen componentes diferenciados, más comúnmente del CPT de alto riesgo, como variantes de células altas, tal es así que pequeños CAT que no coexisten con cánceres diferenciados son esencialmente inexistentes ⁽⁶⁹⁾. El genoma del cáncer de ATC

muestra un grado notablemente alto de heterogeneidad genética según numerosos estudios, un hallazgo que en gran medida explica los desafíos en el tratamiento eficaz del CAT ⁽⁶⁹⁾. Entre todos los cánceres de tiroides, el CAT posee la frecuencia más alta de mutación de TP53 (54–73%) ^(218, 232, 264). TP53 es uno de los genes mutados con mayor frecuencia en todos los cánceres (50%) y ha sido referido como el "guardián del genoma" debido a su papel en la activación de un programa transcripcional antiproliferativo en respuesta a estímulos de estrés ⁽²⁶⁵⁾. En consecuencia, los tumores con la pérdida de la función de p53, incluida la mayoría de los CAT, acumulan una amplia variedad de cambios en el genoma en comparación con los tumores sin pérdida de p53 ^(266, 267).

Un estudio inicial de NGS de 22 pacientes con CAT definió el perfil mutacional, que incluía los genes impulsores (driver) comunes como BRAF^{V600E} y RAS, así como genes del cáncer no asociados previamente con el cáncer de tiroides, como genes de reparación de desajustes ⁽⁶⁹⁾. Los tumores mostraban un fenotipo hipermutador con una alta carga de mutaciones tumorales ⁽⁶⁹⁾.

Una investigación genética similar de una gran cohorte francesa (n = 144 de diez centros) revelaron una heterogeneidad genética significativa con el siguiente perfil mutacional: TP53 (54%), RAS (43%), BRAF (13,8%) y 17% de la Vía PI3K/AKT/mTOR ⁽²⁶⁴⁾. Usando el ensayo genómico MSK-IMPACT, Landa y col. lograron conocimientos importantes sobre el CAT utilizando una cohorte de 33 tumores ⁽²¹⁸⁾. Se observaron frecuencias altas de mutaciones de 73% en el TP53 y promotor TERT ⁽⁶⁹⁾.

Similar a otros estudios, las mutaciones BRAF^{V600E} y RAS fueron mutuamente excluyentes ⁽⁶⁹⁾. Usando datos del transcriptoma, la distinción entre la mutación similar a BRAF^{V600E} (BRAF^{V600E}-like) y similar a RAS (RAS-like) que se encontraba conservada en el CAPD se perdió en CAT, presumiblemente por la mayor carga mutacional del tumor y la

aneuploidia en el CAT ^(268, 269). La frecuencia mutacional de EIF1AX (9%) fue similar a la observada en el CAPD ⁽⁶⁹⁾. Las mismas cuatro vías y genes de los grupos identificados en el CAPD también estaban mutados en el ATC, pero con frecuencias mutacionales más altas ⁽⁶⁹⁾.

Se proporcionó un gran informe de 196 CAT avanzados más conocimientos sobre su genética y clasificación de tumores ⁽¹³²⁾. Usando una agrupación jerárquica y un análisis de aprendizaje automático de los datos mutacionales, se lograron identificar cuatro grupos moleculares. *Grupo 1* se caracterizó por mutaciones BRAF y se cree que representan los CAT que evolucionaron de CPT preexistentes ⁽⁶⁹⁾.

Asimismo, el *grupo 3* contenía mutaciones RAS y se cree que son derivados del CFT con mutación de RAS. CAT del *grupo 4* tienen una diversidad de mutaciones (RAS, PTEN, NF1, RB1) con mayores cargas mutacionales relacionadas con defectos de genes de reparación de desajustes ⁽⁶⁹⁾. En este grupo se especuló que los CAT se originaban del HCC, pero esta asociación requiere validación ya que el genoma del CCH no está completamente definido. Finalmente, los CAT del *grupo 2* se caracterizan por la pérdida de mutaciones funcionales de los genes reguladores del ciclo celular, CDKN2A y CDKN2B, y no tienen bien definido una lesión precursora ⁽⁶⁹⁾.

1.7.6 Aplicaciones Clínicas de la Información Genómica en el Cáncer de Tiroides

La disponibilidad de datos moleculares de todo el genoma para el cáncer de tiroides ha catalizado la aplicación de información molecular seleccionada para ayudar en la decisión de someterse a una extirpación quirúrgica de un nódulo tiroideo ⁽⁶⁹⁾.

La evaluación clínica de los nódulos tiroideos implica un examen clínico, imágenes de ultrasonido y citología por aspiración con aguja fina (FNA) ^(270, 271). A pesar de la eficacia

general de este enfoque, un número significativo de pacientes tiene resultados de citología por FNA indeterminados y, posteriormente, son sometidos a una cirugía por lo que resulta en el examen final de patología definir si corresponde ser una enfermedad benigna, ya sea hiperplasia nodular o adenomas, o maligna ⁽⁶⁹⁾.

Ya se ha habido previsto hace una década que la información molecular de todo el genoma produciría avances significativos en la clasificación de los tumores tiroideos ⁽¹¹⁰⁾ y se han desarrollado pruebas moleculares, específicamente para abordar estos problemas ⁽⁶⁹⁾. Porque existe una fuerte correlación genotipo-fenotipo en los cánceres de tiroides diferenciados, una variedad de los enfoques de prueba es factibles y se han desarrollado ^(272, 273). Éstas incluyen identificación directa de mutaciones impulsoras comunes (es decir, *genotipificación*) ^(227, 274-276) y/o identificación de *perfiles de expresión génica* asociados con *nódulos* benignos o malignos ⁽²⁷⁷⁻²⁷⁹⁾.

El impacto clínico y económico de las pruebas moleculares es un tema de investigación activa ^(275, 280-283), especialmente a medida que estos ensayos continúan sometiéndose a un desarrollo continuo con refinamientos ^(276, 277). En líneas similares, los perfiles de *metilación del ADN* de todo el genoma también se ha propuesto que tiene importancia *pronostica* en cánceres tiroideos diferenciados ⁽²⁸⁴⁻²⁸⁶⁾.

1.8. Visualización del Genoma

Con respecto a la elección del software de visualización del genoma, se muestran las diferentes plataformas que se utilizan para el análisis genómico, características que se resumen en la tabla 5 ⁽²⁸⁷⁾. Si bien, como se muestra en la tabla 5, los softwares analizados tienen características muy similares entre sí, IGV (visor integrador del genoma), tiene ventajas por su capacidad de integración, especialmente para el soporte que ofrece para control remoto ⁽²⁸⁷⁾. Esta característica proponderante, en un área donde los grandes

conjuntos de datos producidos por NGS son complejos de administrar ⁽²⁸⁷⁾. Esto significa que, cuando se inicia IGV, ejecuta un dominio de servidor web que acepta solicitudes remotas. Esta función se puede utilizar para enviar comandos a IGV y recibir resultados. Esta característica con la idea de desarrollo futuro a implementar una plataforma de análisis NGS como servicio web ⁽²⁸⁷⁾.

Tabla 5
Visualización de software: una comparación

Características	IGV	Savant	Artemis
Fuente abierta	Si	Si	Si
Idioma	Java	Java	Java
Plataforma	Todos	Todos	Todos
Tecnología	Illumina, 454, Sanger, ChIP-Seq, RNA-Seq	Illumina, 454, Sanger	Illumina, 454, Sanger
Formato de archivo	BAM, SAM, GOBY, BED, GFF, GTF, PSL, CN, GCT, FASTA, . . . , . . .	SAM, BAM, FASTA, WIG, GFF, BED, tab-delimited	BAM, VCF, BCF, FASTA, tab-delimited
Control remoto	HTTP, HTTPS, FTP	No	No
Soporte	Si	Si	Si

Nota. La presente tabla muestra las diferentes características de los softwares que se utilizan para visualización del genoma. Tomada de A. Agliata et al. (2014) ⁽²⁸⁷⁾.

1.8.1. Software IGV

IGV es un visor de alto rendimiento para datos genómicos, capaz de manejar grandes conjuntos de datos heterogéneos, al tiempo que proporciona una interfaz simple e intuitiva para navegación en el genoma ⁽²⁸⁷⁾. Está escrito en Java, por lo que es multiplataforma. IGV admite varios formatos de archivo (hasta 30 diferentes tipos) de alineación de lectura, incluidos SAM, BAM y Goby ⁽²⁸⁸⁾. Es compatible para manejar conjuntos de datos desde el almacenamiento local y remoto. Además, está disponible para su uso libre con la licencia GNU LGPL, y los recursos de hardware requeridos permiten su instalación incluso en

computadoras de escritorio ⁽²⁸⁷⁾. IGV también permite la interacción con los datos a través de diferentes niveles de detalle gestionados por zoom, que va desde todo el genoma a la única base/nucleótido, utilizando un enfoque similar al utilizado por mapas de Google ^(287, 288). IGV utiliza imágenes preprocesadas de datos genómicos, que representan diferentes resoluciones, dejando la pantalla en tiempo de ejecución con una mejor resolución solo para las partes requeridas ^(287, 288). Este enfoque se llama mosaico de datos ⁽²⁸⁷⁾, y se muestra en la figura 10.

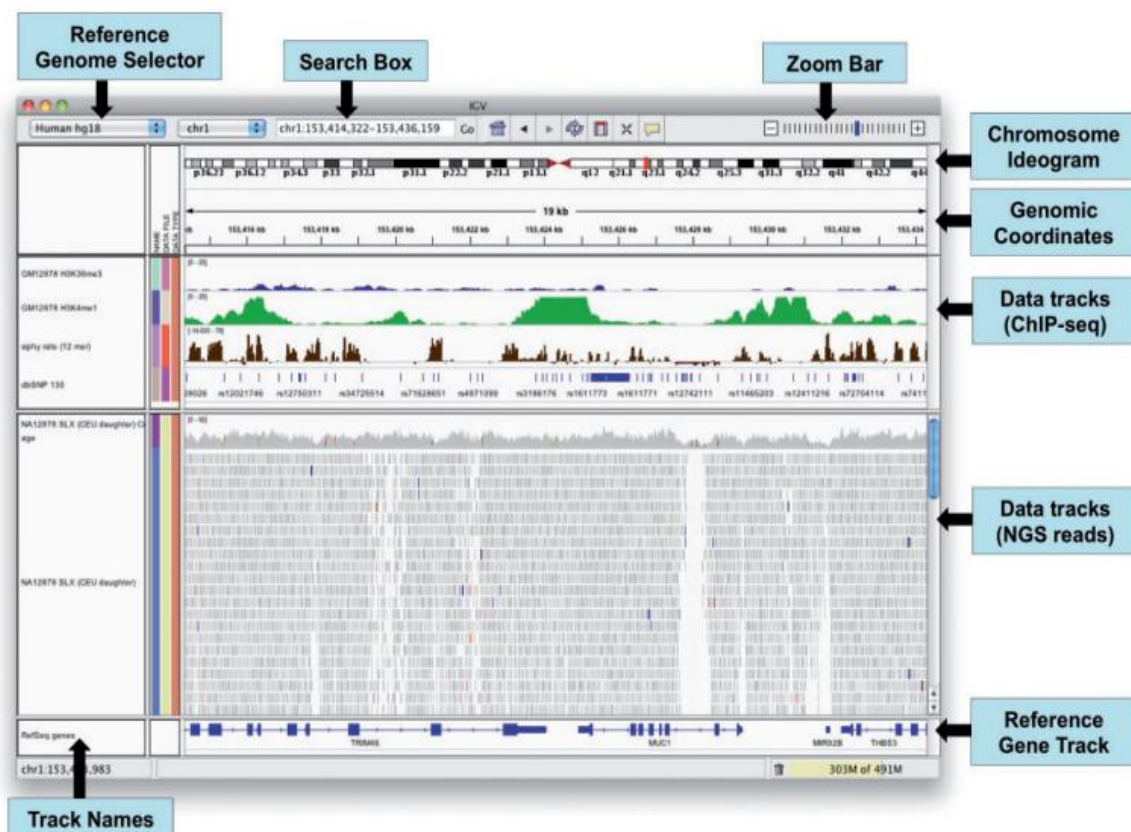


Figura 10: Ventana de aplicación de IGV. Tomada de A. Agliata et al. (2014) ⁽²⁸⁷⁾.

Como para todos los visores de genomas, IGV se usa principalmente para verificar la alineación de los datos analizados al genoma de referencia ^(287, 288). Ya que en el proceso de alineación los artefactos son introducidos por la plataforma de secuenciación, los SNP (polimorfismos de nucleótido único) deben detectarse a partir de errores ⁽²⁸⁷⁾. La desalineación, en la práctica, se destaca adecuadamente, en ciertos niveles de acercamiento (zoom) ^(287, 288). IGV también ofrece la capacidad de mostrar diferentes características de las

lecturas alineadas, como la calidad del mapeo, la frecuencia de alelos y muchas otras características utilizadas por los biólogos para validar mutaciones ⁽²⁸⁷⁾.

1.8.2. Lectura de datos NGS con el software IGV

IGV incluye una gran cantidad de características especializadas para explorar las alineaciones de lectura de NGS, incluidas características adaptadas para la visualización y validación de variantes genómicas, empalme de transcritos de ARN y metilación de secuenciación de bisulfito ⁽²⁸⁸⁾. Cuando se carga más de un archivo en IGV, puede mostrar lecturas de cada archivo en paneles separados o fusionarlos como si vinieran del mismo archivo ⁽²⁸⁸⁾. Debido a la magnitud de archivos de datos almacenados en NGS que deben ser alineados, IGV muestra diferentes niveles de detalles dependiendo del nivel de zoom ⁽²⁸⁷⁾. Al ver un cromosoma completo, no es útil mostrar todas las lecturas, ya que las lecturas individuales no se distinguen en la vista a este nivel de zoom ^(287, 288). Por lo tanto, cuando se amplía toda la salida, se visualiza en un solo gráfico de barras la cobertura de lectura individual. IGV proporciona herramientas para precalcular esta cobertura. Cuando se amplía más allá de un umbral de visibilidad configurable por el usuario (por defecto, 30 kb), las alineaciones de lectura aparecen y se muestran como barras horizontales ⁽²⁸⁸⁾. En este nivel de zoom, IGV calcula dinámicamente la cobertura de lectura en una región visualizada. Por lo tanto, zoom revela aún más la lectura individual de bases ^(287, 288). IGV usa color y transparencia para resaltar eventos interesantes en los datos de alineación y para restar énfasis visualmente a otros, aplicada esta propiedad en el gráfico de cobertura, en el nivel de lectura y para bases individuales ⁽²⁸⁸⁾. Se pueden usar varias propiedades para cambiar el esquema de color de lectura, para agrupar, ordenar y filtrar interactivamente las lecturas ⁽²⁸⁷⁾. Estas características se pueden usar solas o en combinación, usando una o más propiedades. Las propiedades incluyen identificador de muestra, hebra, lectura de grupo, calidad de mapeo,

determinación de una base en una posición particular, distancia y orientación de pares, etiquetas personalizadas y más ⁽²⁸⁷⁾. El acercamiento (zooming) más allá del umbral de visibilidad de la alineación, también agregará barras de color a la pista de cobertura gris en ubicaciones donde una gran cantidad de bases de lectura no coinciden con la referencia, útil para identificar supuestos SNP ⁽²⁸⁸⁾. El tamaño relativo y el color de estas barras indican la frecuencia alélica de cada base en esa ubicación ^(287, 288), esto se puede mostrar en la figura 11A. En Figura 11B, la vista se ha ampliado aún más para mostrar las lecturas y bases individuales en la región de uno de los SNP putativos identificados en la cobertura pista.

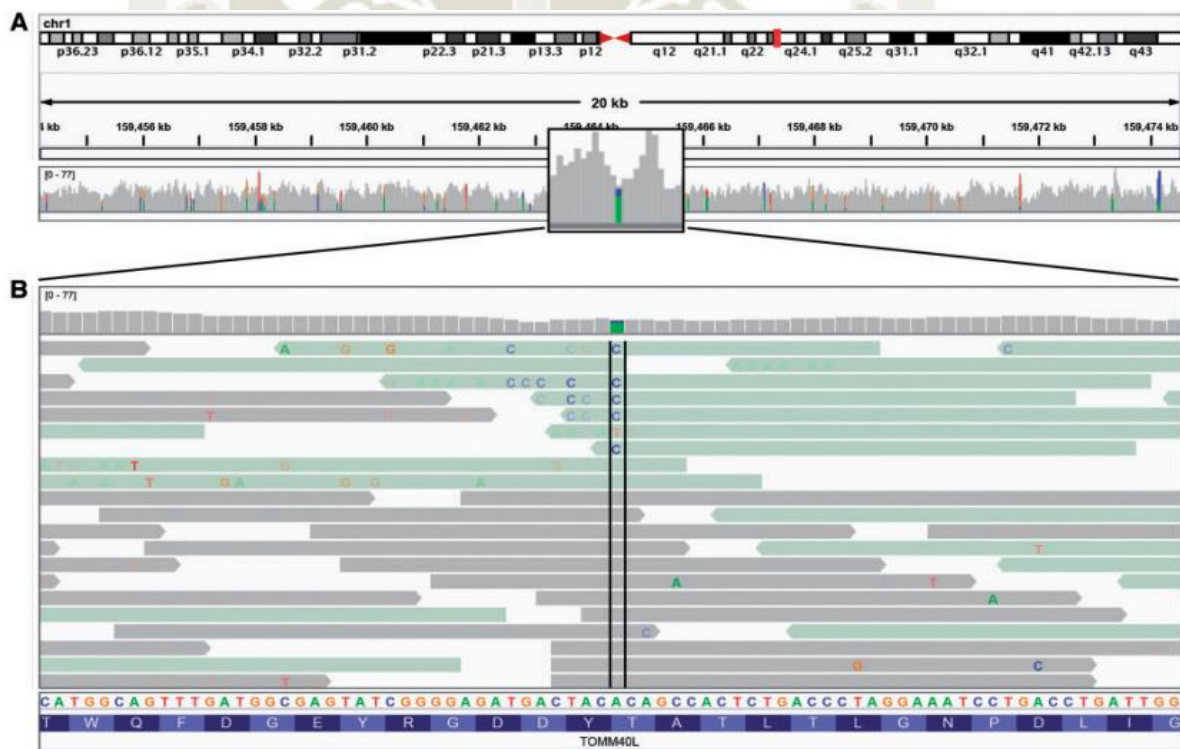


Figura 11: Vista de lectura de alineación a 20 kb y resolución de pares de bases. (A) Las lecturas se resumen como un gráfico de cobertura. Las posiciones con un número importante de desajustes con respecto a la referencia se han resaltado con barras de color que indican tanto la presencia de desajustes como la frecuencia del alelo. (B) Las discrepancias individuales de base se muestran con transparencia alfa proporcional a la calidad. Tomada de Thorvaldsdóttir et al. (2012) ⁽²⁸⁸⁾.

Las bases de lectura individuales que coinciden con el genoma de referencia, se muestran en el mismo color que la lectura, mientras que los desajustes están codificados por colores configurados para una determinada base nitrogenada, y se les asigna un valor de

transparencia proporcional a la puntuación base, esto en relación a una escala que evalúa los desajustes de calidad (nivel de calidad de Phred) ⁽²⁸⁸⁾. Eventos intracromosómicos, como inserciones, deleciones, inversiones y duplicaciones, pueden afectar tanto la distancia genómica entre las alineaciones de pareja, así como su orientación; ara resaltar estos eventos, IGV toma muestras del archivo de alineación para determinar dinámicamente la distancia y las orientaciones esperadas, y luego usa el color para marcar pares aberrantes ⁽²⁸⁸⁾.

1.9. Visualización de fusiones con software Arriba

Arriba, es un algoritmo de detección de fusión diseñado específicamente para cumplir con los exigentes requisitos de la oncología de precisión asistida por secuenciación de alto rendimiento (HTS) ⁽²⁸⁹⁾. Debido a un flujo de trabajo altamente optimizado, puede procesar muestras contemporáneas de RNA-seq en menos de una hora ⁽²⁸⁹⁾. Los filtros sofisticados detectan fusiones incluso en condiciones desfavorables, como muestras de baja pureza ⁽²⁸⁹⁾. Además, Arriba es capaz de detectar transcripciones aberrantes que no son encontradas por la mayoría de los métodos de detección de fusión, pero que pueden ser clínicamente relevante ⁽²⁸⁹⁾. Esto incluye genes supresores de tumores que ocasionalmente se inactivan por reordenamientos dentro del gen o por translocaciones a intrones o regiones intergénicas ⁽²⁸⁹⁾. Aunque el mismo enfoque técnico puede aplicarse para detectar dichas transcripciones, la mayoría de las herramientas de detección de fusión disponibles no las reportan; en consecuencia, las aberraciones clínicamente relevantes pueden pasarse por alto ⁽²⁸⁹⁾. Arriba mejora la detección sobre los métodos existentes, en los que puede encontrar inversiones/duplicaciones intragénicas y translocaciones a intrones/regiones intergénicas ⁽²⁸⁹⁾.



2. Metodología

2.1 Lugar y tiempo de realización del estudio

Este es un estudio de tipo correlacional y transversal que inició con la recolección de muestras de tejido tumoral y data clínica de todos los casos diagnosticados como cáncer diferenciado de tiroides, en el Servicio de Anatomía Patología del Hospital Nacional Carlos Alberto Segúin Escobedo EsSalud Arequipa (HNCASE), entre julio de 2019 y febrero del 2020.

El protocolo de investigación fue aprobado por el comité de investigación del HNCASE, posterior a lo cual se obtuvieron muestras del banco de tejidos del servicio de Anatomía patológica de este nosocomio, de un total de 11 pacientes con diagnóstico histopatológico de cáncer diferenciado de tiroides (CDT) intervenidos quirúrgicamente en esta institución.

2.2 Revisión de historias clínicas

La revisión de la historia clínica, permitió acceder a información epidemiológica y clínica como edad, sexo, antecedente de tiroiditis crónica, score sonográfico TI-RADS, citología BETHESDA e informe del diagnóstico histopatológico según la 8va edición de la AJCC (The American Joint Committee on Cancer).

2.3. Población

11 pacientes operados con diagnóstico histopatológico de CDT en el periodo de estudio.

2.4 Muestra biológica

Tejido tumoral de cada uno de los 11 pacientes operados con el diagnóstico histopatológico de CDT en el periodo de estudio.

2.4.1 Recolección de la muestra biológica

Para la recolección de la muestra se contó con el apoyo del Servicio de Anatomía patológica, quienes recibieron las piezas operatorias en el banco de tejidos, inmediatamente culminada

la cirugía, corroborando la presencia de cáncer diferenciado de tiroides, posterior a la cual se recolectó un fragmento de tejido fresco de 30 mg, el que fue colocado en un tubo estéril eppendorf con hielo seco, almacenado en cadena de frío; inmediatamente en los primeros 30 minutos la muestra fue trasladada al laboratorio de biología molecular del campus de la Universidad Católica de Santa María (UCSM).

2.4.2 Extracción del ARN

La extracción de ARN se realizó en el laboratorio de Biología Molecular de la UCSM Arequipa. Para la extracción del ARN, se procedió a trozar el tejido en el tubo eppendorf, se añadió 500 ul de Buffer Fosfato Salino (PBS), se centrifugó a 12500 rpm por 10 minutos, se aspiró el sobrenadante, el cual se colocó en un nuevo tubo eppendorf de 1.5, se eliminó el nuevo sobrenadante obteniendo pellet (pequeñísimas partículas de tejido triturado), a este se le añadió 700 ul de buffer de lisis de células (10mM Tris – HCL, 10mM EDTA, 50 mM NaCl, 5% SDS, pH 7.5) y 10 μ L de proteinasa K, se vortexió durante 6 minutos y se llevó a incubación por 30 minutos a 60° C.

Transcurrido el tiempo de incubación se agregó 500 μ L de fenol cloroformo: alcohol isoamílico, se centrifugó a 12000 rpm por 10 minutos, se transfirió la fase acuosa a un nuevo eppendorf, sin tocar la interfase, añadiendo isopropanol en una cantidad de 1 a 2.5 volúmenes de la fase acuosa recuperada, esta mezcla se centrifugó a 12000 rpm por 10 minutos, se eliminó el isopropanol por inversión en un papel toalla una sola vez, quedando solamente pellet, al cual se le agregó 500 μ L de etanol al 70-75%, se centrifugó a 12000 rpm por 10 minutos eliminando el sobrenadante por inversión en papel toalla, para quedarse únicamente con el nuevo pellet, a este se agregó 15 μ L de RNAsa, se le incubó en termoblock a 37° por unos 10 minutos, se centrifugó a 12000 rpm por 10 minutos, para finalmente ser sometido a

pruebas de integridad y cantidad de RNA por electroforesis y espectrofotometría en gel de agarosa al 1.5%.

El pellet final de RNA que pasó control de calidad, fue rotulado y almacenado a una temperatura de -20°C . Los 11 pellets preparados fueron enviados, suspendidos en 1 cc de agua libre de RNasa, con sistema de transporte hielo seco al laboratorio de NOVOGEN y Admera Health Biopharma Servicesen de USA. Paralelamente, para tener un apoyo o muestra de reserva en caso que el ARN total no cumpliera con las condiciones requeridas para el secuenciamiento, se enviaron tres tubos con tejido tumoral (entre 100 – 300 mg) en tubos eppendorf de 1.5 suspendidos en 1 cc de reactivo RNA later, para que el ARN del tejido se conservase, a temperatura de refrigeración (4°C) ⁽²⁹⁰⁾.

2.4.3 Extracción de ARN total

Las muestras de ARN y tejido tumoral adicional llegaron al laboratorio NOVOGEN y Admera Health Biopharma Servicesen de USA, donde se realizaron controles de calidad, utilizando el Kit de inicio de cuantificación invitrogen fluorómetro Qubit 4, que nos permite medir con precisión la cantidad de ARN, y ahora también la integridad y calidad del ARN. Seguidamente, se verificó la extracción efectiva del ARN, para ello se corrió lo extraído en gel de agarosa teñido con el reactivo fluorescente no mutagénico (CSL-RUNSAFE), leído al transiluminador LED de luz azul ⁽²⁹⁰⁾. Posteriormente, se determinó la cantidad de ARN en la plataforma Picogreen usando el método fluométrico, mientras que el número de integridad del RNA (RIN) fue obtenido mediante el Bioanalizador Agilent Technologies 2100 ⁽²⁹⁰⁾. Para efectos de obtener un ARN sin contaminantes, el ADN es digerido por las DNasas que contiene el reactivo utilizado en el protocolo y, además, previa elaboración de la librería, algún remanente de este es removido durante la purificación del ARN mensajero ⁽²⁹⁰⁾.

Los valores de RIN encontrados se encontraron por debajo de 7, en las 11 muestras, por lo que, según los estándares de calidad, se optó por el método de la deplección de RNA ribosomal de Illumina, el cual necesita RIN: ≥ 2 , para la construcción de la librería de RNAseq.

2.4.4 Secuenciamiento del Transcriptoma con el Protocolo Illumina

Los ARNs extraídos que cumplieron con los requisitos necesarios por Admera health y Novogen (número de integridad de ARN con un RIN ≥ 2 y una concentración de ARN de 200 ng – 1 ug) fueron sometidos a la preparación de la librería (RNA-seq).

Para poder realizar este proceso, el ARN total obtenido tuvo que ser convertido a ADN complementario, pasando previamente por el protocolo indicado por la guía de preparación de la muestra: “TruSeq Stranded Total RNA”, para así poder preparar la librería ⁽²⁹⁰⁾. La plataforma Next Generation Sequencing utilizada para el secuenciamiento fue la “NovaSeq 6000” ⁽²⁹⁰⁾.

El sistema NovaSeq 6000 representa la plataforma de secuenciación de la tecnología de Illumina de alto rendimiento, el instrumento, se basa en la secuenciación por síntesis (*sequencing by synthesis*) ⁽²⁹⁰⁾, en la que los fragmentos de ARN segmentados se copian en el ADNc de la primera cadena mediante cebadores aleatorios y transcriptasa inversa y, seguidamente, se produce la síntesis de ADNc de la segunda cadena mediante ADN polimerasa I y ARNasa H (se incorpora dUTP en lugar de dTTP para generar ADNc bicatenario, obteniendo ADNc con extremos romos) ⁽²⁹⁰⁾. Seguidamente, se añade una base única “A” en estos fragmentos de ADNc y se produce la posterior ligadura del adaptador ⁽²⁹⁰⁾.

Los productos se purifican y se enriquecen con PCR para crear la biblioteca definitiva de ADNc ⁽²⁹⁰⁾. El método Truseq stranded de ARN total con el kit Ribo-Zero fue empleado para la preparación de la librería tipo *Paired-end* (PE) ⁽²⁹⁰⁾. Se usó el kit Ribo-Zero, porque mejora

los resultados de RNA-seq al eliminar el mtrRNA (RNA mitocondrial) y el rRNA (RNA ribosomal) citoplásmico, e incluso de muestras de ARN degradadas o archivadas, lo que permite un mejor mapeo de las lecturas exómicas, maximizando la cantidad de lecturas útiles de ARN-seq del ARN total derivado de células de mamíferos.

El procedimiento para el secuenciamiento por la tecnología NGS de Illumina incluye 4 pasos ⁽²⁹⁰⁾:

- **Preparación de la muestra:** Para la construcción de la librería, el ARN es extraído de la muestra ⁽²⁹⁰⁾. Después de aplicar el control de calidad (quality control, QC), las muestras que califican, proceden para la construcción de la librería ⁽²⁹⁰⁾.
- **Construcción de la librería:** La librería es construida a partir de la fragmentación aleatoria de la muestra del ADN complementario, seguida de la ligación 5' y 3' ⁽²⁹⁰⁾. Alternativamente, la “tagmentación” fusiona las reacciones de fragmentación y ligación en un solo paso, que incrementa enormemente la eficiencia del proceso de preparación de la librería ⁽²⁹⁰⁾. Posteriormente, los fragmentos ligados al adaptador se amplifican por PCR y se **purifican en gel** ⁽²⁹⁰⁾. La electroforesis en gel, es el procedimiento de laboratorio estándar para separar el ADN por tamaño (p. ej., longitud en pares de bases) para su visualización y purificación ⁽²⁹⁰⁾. La electroforesis utiliza un campo eléctrico para mover el ADN con carga negativa a través de una matriz de gel de agarosa hacia un electrodo positivo ⁽²⁹⁰⁾. Los fragmentos de ADN más cortos migran a través del gel más rápidamente que los más largos ⁽²⁹⁰⁾. Por lo tanto, puede determinar la longitud aproximada de un fragmento de ADN ejecutándolo en un gel de agarosa junto con una escalera de ADN (una colección de fragmentos

de ADN de longitudes conocidas), lo que permite la recuperación de ADN del gel que es directamente adecuado para PCR y el secuenciamiento ⁽²⁹⁰⁾.

- **Secuenciamiento:** Para la obtención de agrupamientos, la biblioteca se carga en una celda de flujo donde los fragmentos se capturan en un campo de oligos unidos a la superficie y que son complementarios a los adaptadores de la librería ⁽²⁹⁰⁾. A continuación, cada fragmento se amplifica por grupos. Cuando se completa la generación del clúster, las plantillas están listas y preparadas para la secuenciación propiamente dicha en el equipo ⁽²⁹⁰⁾. El número de lecturas fue de 40 a 60 millones por muestra durante el secuenciamiento.
- **Datos sin procesar (*raw data*):** La data resultado del secuenciamiento previamente construida, es procesada por el software y se genera un archivo de tipo FastQ ⁽²⁹⁰⁾.

2.5 Visualización de las alineaciones del formato BAM con el genoma de referencia utilizando el software IGV

Una vez obtenida la base de datos del secuenciamiento y pasadas las pruebas de control de calidad respectivas, se guardarán los datos obtenidos en formato BAM (formato binario para el almacenamiento de datos de secuenciación) ⁽²⁹⁰⁾. La extensión de archivo (BAM) contiene información sobre lecturas de secuencias después de haber sido estas alineadas contra un genoma de referencia ⁽²⁹⁰⁾.

Para proceder a la visualización de las alineaciones encontradas se utilizó el software IGV (integrative genomics viewer), comparando contra la base genómica GRCh38.p13 (versión 13 del consorcio 38 del genoma de referencia humano), para poder detectar mutaciones.

2.6 Identificación de fusiones de genes

Se utilizó el software Arriba versión 1.2.0, una herramienta de línea de comandos para la detección de fusiones de genes a partir de datos de RNA-Seq, desarrollada para su uso en un entorno de la investigación clínica, con alta sensibilidad y un tiempo de ejecución corto.

2.7 Evaluación de la calidad de datos secuenciados con NGS

Después de realizar el mapeo de las lecturas del genoma, se examinó la calidad de los datos de RNA-seq obtenidos, utilizando los softwares FastQC y MultiQC v1.8, tanto en el laboratorio de Novogene USA como del Instituto Pasteur en Francia.

Se encontraron los siguientes resultados:

- Se encontró un *porcentaje de lecturas mapeadas* mayor al 80% por muestra (considerandose de calidad un porcentaje mayor de 80-90%), para aquellas con porcentajes menores al 80%, se hizo una duplicación de lecturas hasta llegar al 100%.
- La *tasa de secuencias duplicadas* fue de 37-81% (considerandose de calidad tasas entre 30-90%).
- El número de *lecturas con puntajes de calidad* promedio en *cada posición base*, presentaron un Phread score de 36 (considerandose de calidad socores entre 2-40).
- El *contenido promedio de GC* de las lecturas estudiadas (figura 13c) se encontró en 44% (los rangos de normalidad para el genoma humano varían entre el 35-60% con un promedio del 41%).
- Los *porcentajes de alineaciones* se encontraron mayores al 80% (considerandose óptimos mayores a 70% e idealmente mayores o iguales a

90%); la distribución del porcentaje de bases primarias que se *alinean con regiones en el genoma* de referencia, muestra mayores porcentajes de alineaciones de lectura con las áreas intrónicas e intergénicas, resultados que son esperados según el tipo de muestra procesada, en el caso de nuestro estudio con el método de la depleción de rRNA.

- La *cobertura génica normalizada* (número de veces que se lee un nucleótido durante la secuenciación), muestra que en el mayor porcentaje de genes se encontró entre 125 a 150 millones de lecturas (considerandose óptimo > 100 millones).

En la figura 12, se muestra una tabla resumen de los parámetros antes descritos, encontrando que los indicadores de calidad bioinformática mas relevantes mostraron valores adecuados, por lo que el procesamiento pasó el control de calidad, validando los resultados obtenidos.

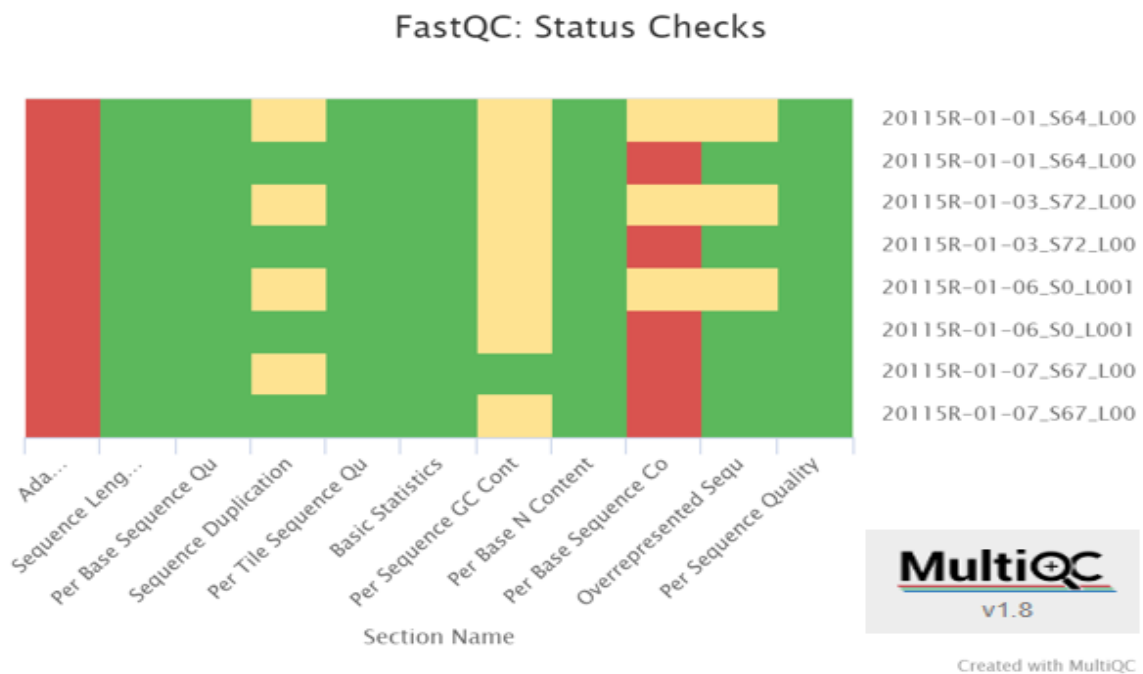


Figura 12. Análisis de calidad bioinformática de los resultados con NGS. Estado de cada sección del FastQC que muestra si los resultados se muestran completamente normales (verde), ligeramente anormales (naranja) o muy inusuales (rojo). Fuente Reporte MultiQC v1.8 Admera health.

2.8 Análisis estadístico

El análisis y procesamiento de datos, se realizó utilizando el paquete SPSS (ver 20.0 para Windows, SPSS Inc., Chicago, IL, EE. UU.). Se aplicó estadística descriptiva, utilizando tablas de distribución de frecuencias en la que los datos categóricos se presentaron como frecuencia y porcentaje.





CAPÍTULO III
RESULTADOS Y DISCUSIÓN

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

Tabla 6

Distribución por grupo etáreo y sexo de pacientes con el diagnóstico de Cáncer Diferenciado de Tiroides

EDAD	MASCULINO		FEMENINO		TOTAL	
	Nº	%	Nº	%	Nº	%
< 40 AÑOS	0	0.00	1	09.09	1	9.09
40-49 AÑOS	1	9.09	4	36.36	5	45.45
50-59 AÑOS	0	0.00	1	9.09	1	9.09
60-69 AÑOS	1	9.09	1	9.09	2	18.18
70-79 AÑOS	2	18.18	0	0.00	2	18.18
TOTAL	4	36.36	7	63.64	11	100.00

En la tabla 6 se muestra la distribución de la población estudiada con el diagnóstico de cáncer diferenciado de tiroides (CDT) por grupo etáreo y por sexo. El mayor porcentaje de la población estudiada correspondió al sexo femenino (63.64%) en comparación al sexo masculino (36.36%). El grupo etáreo que predominó fue el de 40 a 49 años de edad (45.45%).

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

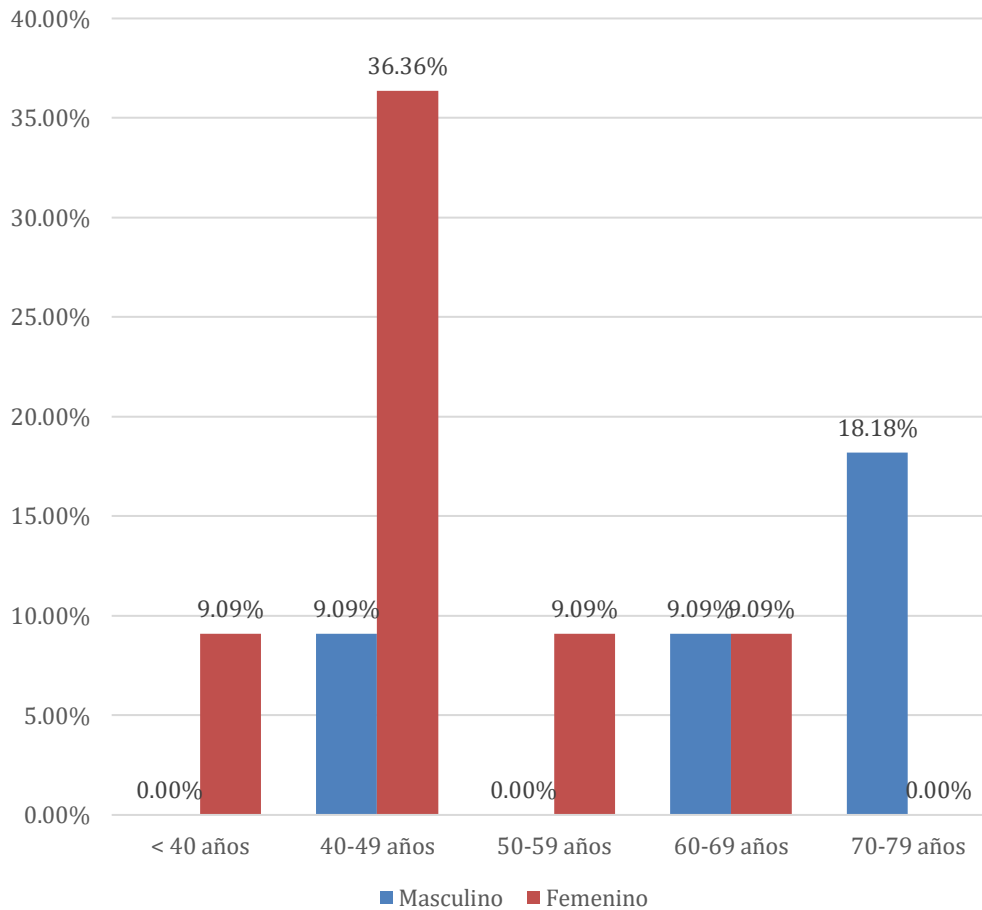


Figura 13. Distribución por grupo etáreo y sexo de pacientes con el diagnóstico de Cáncer Diferenciado de Tiroides. El mayor porcentaje de mujeres con CDT se encontró entre los 40 a 49 años de edad, a diferencia del CDT en varones que se presentó con mayor frecuencia entre los 70 a 79 años de edad. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

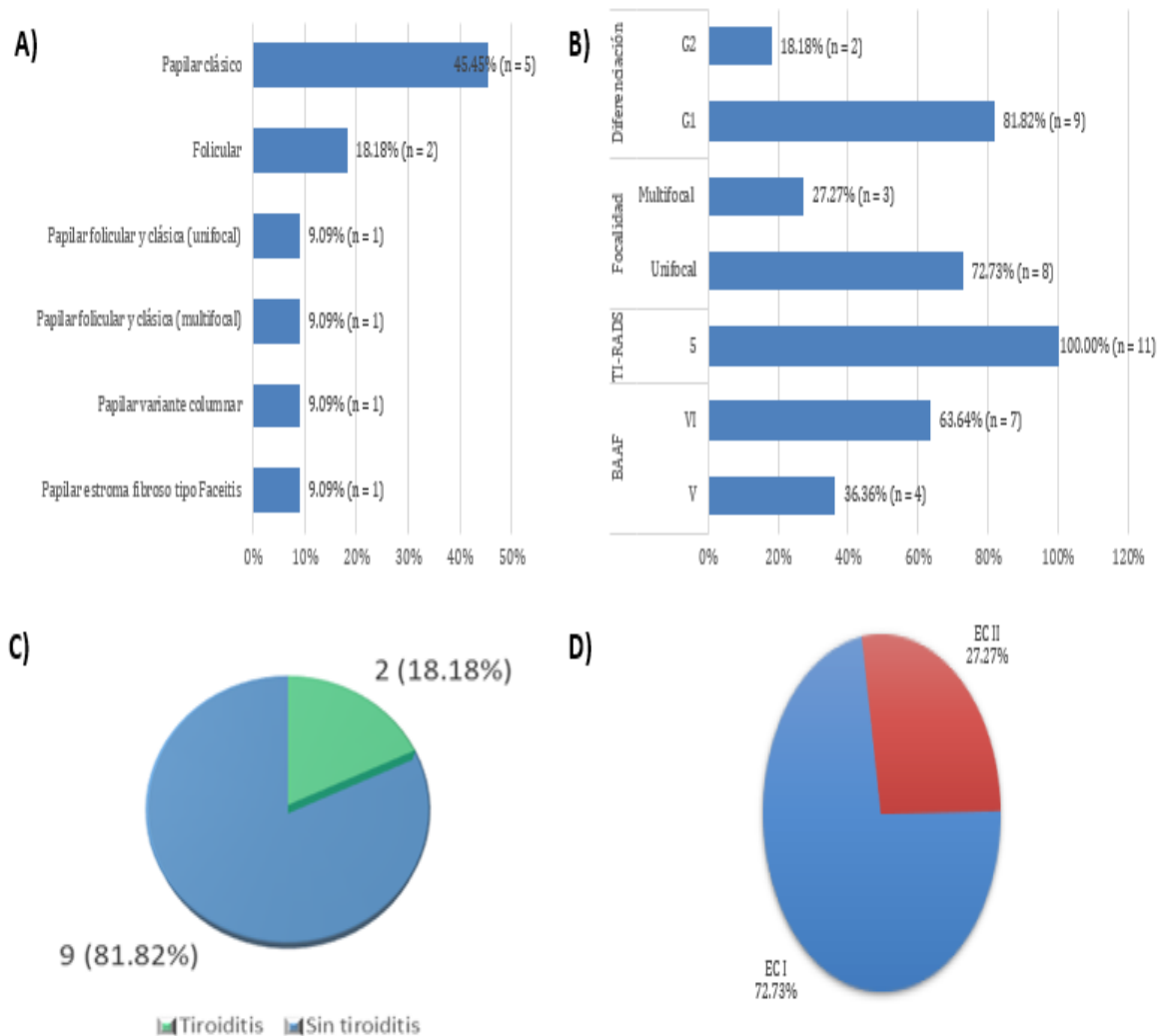


Figura 14. Características citológicas, histopatológicas y sonográficas de los pacientes con el diagnóstico de Cáncer Diferenciado de Tiroides. Distribución de la variante histopatológica (A). Se encontró una predominancia de la variante papilar (81.82%), frente a la variante folicular (18.18%). Características ecográficas, citológicas y tumorales (B). El 81.82% de tumores fueron bien diferenciados (G1), unifocales (72.73%); TI-RADS 5 (100%); con predominio de la categoría VI (63.64%) y V (36.36%) del sistema Bethesda. Presencia de tiroiditis crónica asociada (C). En el 18.18% de pacientes con diagnóstico de CDT coexistió el diagnóstico de tiroiditis crónica. Estadiaje clínico (D). Los estadios clínicos que se presentaron fueron el I y II, predominando el EC I en el 72.73% versus un 27.27% para el EC II. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

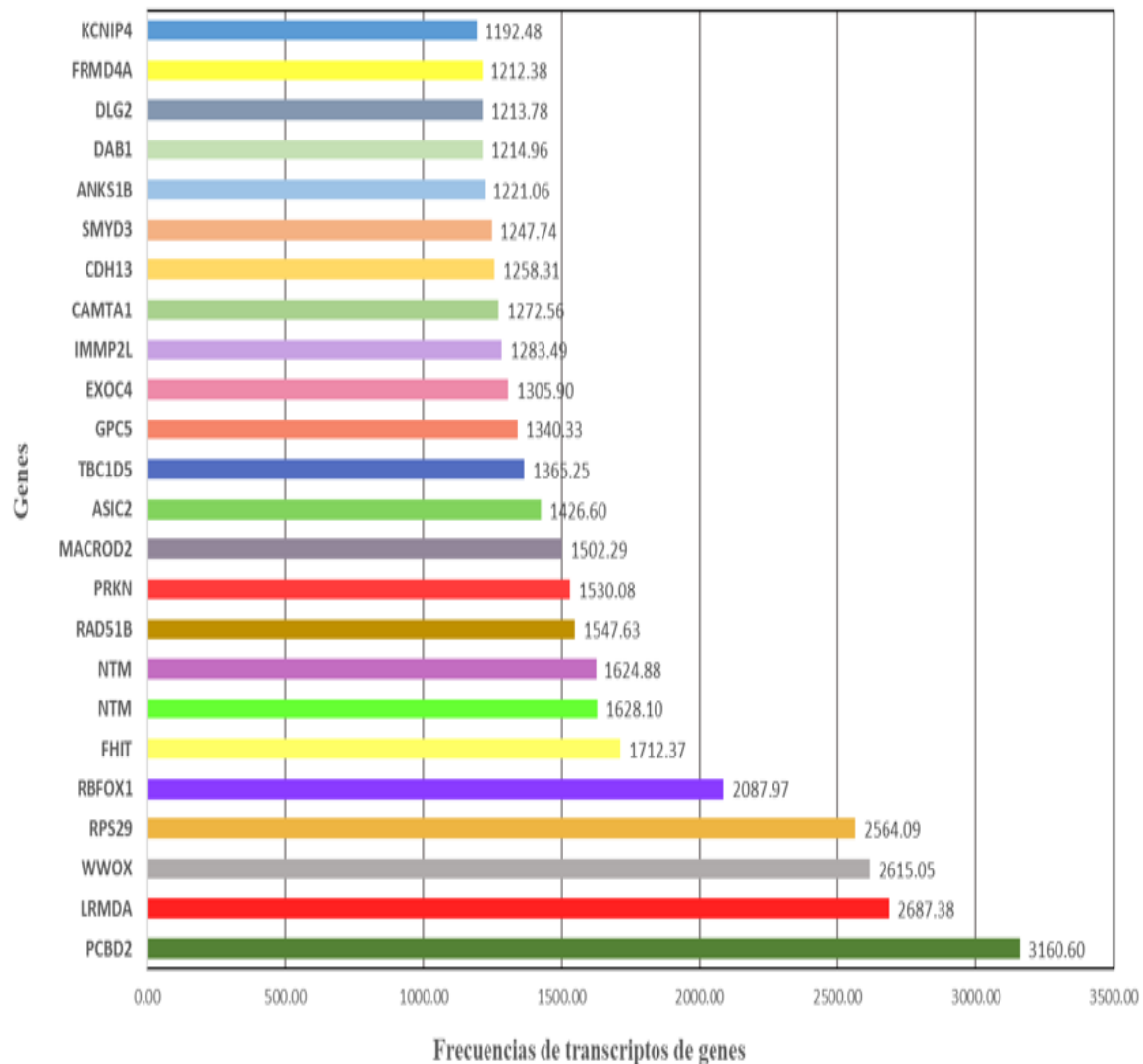


Figura 15 (parte 1 de 3). Genes con mayor frecuencia de expresión comunes de los 11 casos estudiados de Cáncer Diferenciado de Tiroides del 2019 al 2020. En la parte 1 de la figura 20 se muestra las frecuencias de transcripción de los primeros 24 genes expresados comunes en los 11 casos de pacientes con CDT estudiados. El gen PCB2 es el gen con mayor número de transcripciones de los 11 pacientes portadores de cáncer diferenciado de tiroides, encontrando al gen KCNIP4 en el puesto 24. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

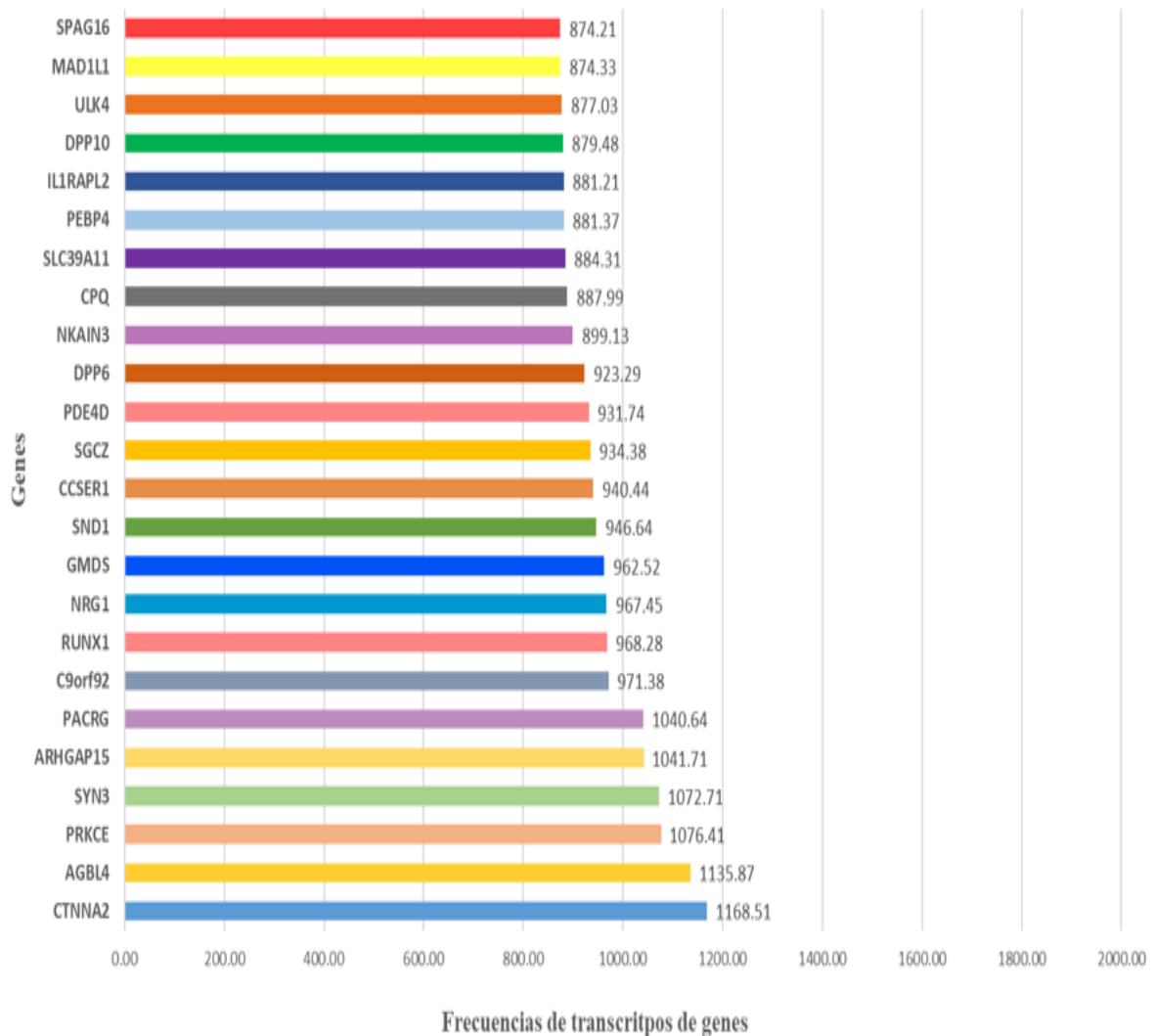


Figura 15 (parte 2 de 3). Genes con mayor frecuencia de expresión comunes de los 11 casos de Cáncer Diferenciado de Tiroides del 2019 al 2020. En la parte 2 de la figura 20 se muestra las frecuencias de transcripción de los genes del puesto 25 al puesto 48 expresados comunes en los 11 casos de pacientes con CDT estudiados. El gen CTNNA2 ocupa el puesto 25, mientras que el gen SPAG16 ocupa el puesto 48, en los 11 casos de la población estudiada. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

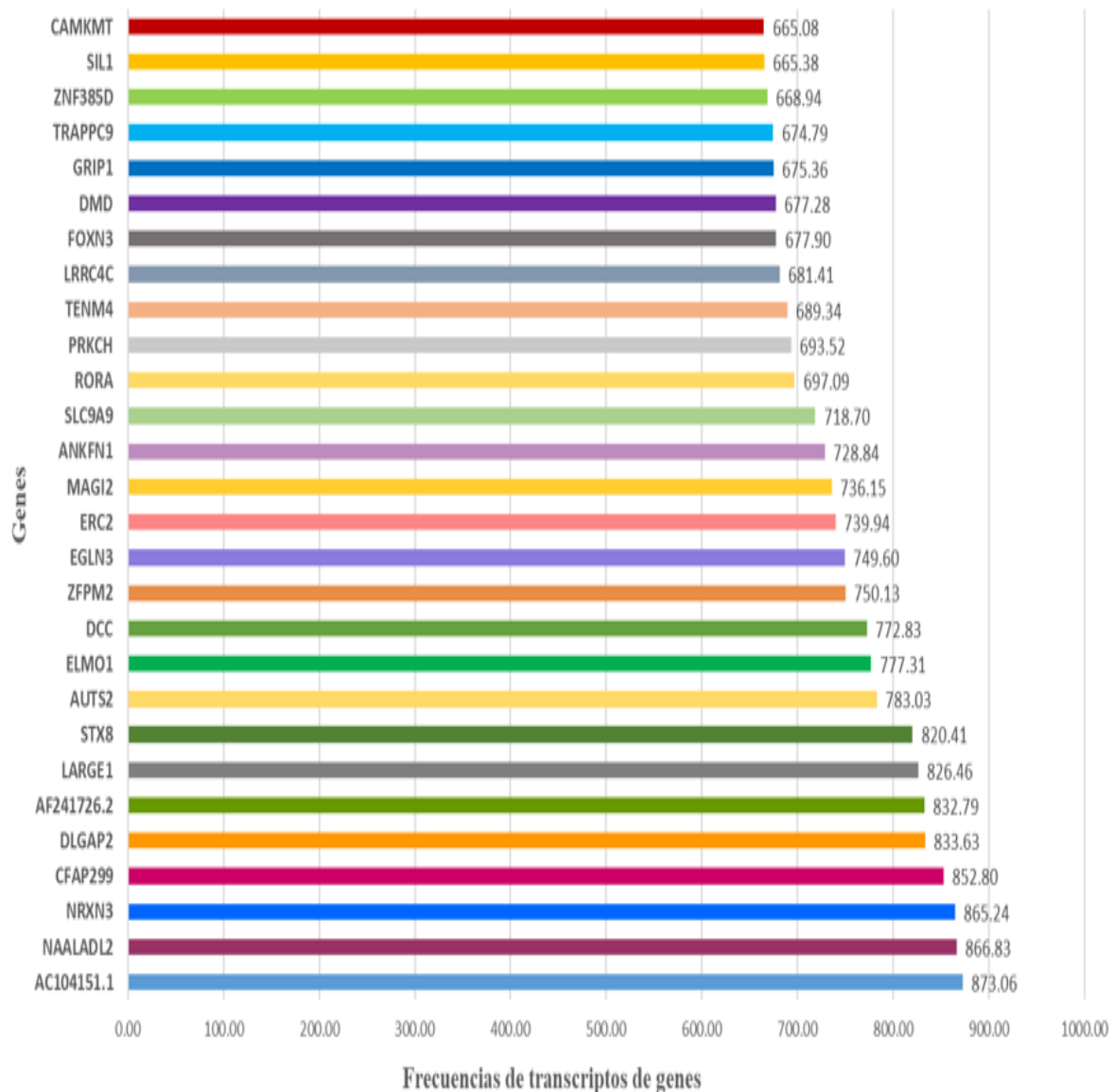


Figura 15 (parte 3 de 3). Genes con mayor frecuencia de expresión comunes de los 11 casos de Cáncer Diferenciado de Tiroides del 2019 al 2020. En la parte 3 de la figura 20 se muestra las frecuencias de transcripción de los genes del puesto 49 al puesto 76 expresados comunes en los 11 casos de pacientes con CDT estudiados. El gen AC104151.1 ocupa el puesto 49, mientras que el gen CAMKMT ocupa el puesto 76, en los 11 casos de la población estudiada. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

Tabla 7 (parte 1 de 2)

Treinta primeros genes con mayor frecuencia de expresión comunes de los 11 casos de Cáncer Diferenciado de Tiroides del 2019 al 2020

Gen	Nombre	Sinónimos	Localización	Referencia
PCBD2	Pterin-4 alfa-carbinolamina deshidratasa 2	PHS2; DCOH2; DCOHM.	5q31.1	(315)
LRMDA	Asociada a diferenciación de melanocitos ricos en leucina.	CDA017; C10orf11.	10q22.2-q22.3	(315)
WVOX	Dominio WW que contiene oxidorreductasa	FOR; WOX1; DEE28; EIEE28; FRA16D; SCAR12; HHCMA56; PRO0128; SDR41C1; D16S432E.	16q23.1-q23.2	(315)
RPS29	Proteína ribosomal S29	S29; uS14; DBA13.	14q21.3	(315)
RBFOX1	RNA de unión al homólogo 1 de Fox	2BP1; FOX1; A2BP1; FOX-1; HRNBPI.	16p13.3	(315)
FHIT	tríada de histidina frágil diadenosina trifosfatasa.	FRA3B; AP3Aase.	3p14.2	(315)
NTM	Neurotrimina	HNT; NTRI; CEPU-1; IGLON2	11q25	(315)
RAD51B	RAD51 paralog B	REC2; R51H2; RAD51L1.	14q24.1	(315)
PRKN	Parkin RBR E3 ubiquitina proteína ligasa	PDJ; AR-JP; LPRS2; PARK2.	6q26	(315)
MACROD2	Mono-ADP ribosilhidrolasa 2	C2orf133; C20orf133.	20p12.1	(315)
ASIC2	Subunidad 2 del canal de iones de detección de ácido	ACCN; BNC1; MDEG; ACCN1; BNaC1; ASIC2a; hBNaC1.	17q11.2-q12	(315)
TBC1D5	Miembro 5 de la familia del dominio TBC1	KIAA0210.	3p24.3	(315)
GPC5	Glipicano 5	GLYPICAN 5	13q31.3	(315)
EXOC4	Componente del complejo de exocitos 4	SEC8; Sec8p; SEC8L1.	7q33	(315)
IMMP2L	Subunidad 2 de la peptidasa de la membrana mitocondrial interna	IMP2; IMP2-LIKE; IMMP2L-IT1.	7q31.1	(315)

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

Tabla 7 (parte 2 de 2)

Treinta primeros genes con mayor frecuencia de expresión comunes de los 11 casos de Cáncer Diferenciado de Tiroides del 2019 al 2020

Gen	Nombre	Sinónimos	Localización	Referencia
CAMTA1	Activador de la transcripción de unión a calmodulina 1	CANPMR.	1p36.31-p36.23	(315)
CDH13	Cadherina 13	CDHH; P105.	16q23.3	(315)
SMYD3	Dominio SET y MYND que contiene 3.	KMT3E; ZMYND1; ZNFN3A1; bA74P14.1.	1q44	(315)
ANKS1B	Repetición de anquirina y dominio de motivo alfa estéril que contiene 1B.	EB1; AIDA; EB-1; ANKS2; AIDA-1; cajalin-2.	12q23.1	(315)
DAB1	Proteína adaptadora DAB 1	SCA37.	1p32.2-p32.1	(315)
DLG2	Discos grandes proteína de andamio MAGUK 2	PSD93; PSD-93; PPP1R58; chapsyn-110.	11q14.1	(315)
FRMD4A	Dominio FERM que contiene 4A	FRMD4; CCAFCA; bA295P9.4.	10p13	(315)
KCNIP4	Proteína que interactúa con el canal dependiente de voltaje de potasio 4	CALP; KCHIP4.	4p15.31-p15.2	(315)
CTNNA2	catenina alfa 2	CAPR; CTNR; CAP-R; CT114; CDCBM9.	2p12	(315)
AGBL4	Carboxipeptidasa 4 de AGLB	CCP6.	1p33	(315)
PRKCE	Proteína quinasa C épsilon	PKCE; nPKC-epsilon.	2p21	(315)
SYN3	Sinapsina III	Synapsin – 3, CN28H9.2.	22q12.3	(315)
ARHGAP15	Proteína activadora de Rho GTPasa 15	BM046.	2q22.2-q22.3	(315)
PACRG	Parkin coregulado	GLUP; PACRG2.1; PARK2CRG; HAK005771.	6q26	(315)
C9orf92	Marco de lectura 92 abierto del cromosoma 9	Em:AL513424.1.	9p22.3	(315)

En la tabla 7 (parte 1 y 2) se muestran los 30 primeros genes con mayor frecuencia de expresión comunes de los 11 casos de CDT. Se describen sus nombres, sinónimos de nomenclatura, ubicación cromosómica y referencias bibliográficas.

**PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE
TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020**

Tabla 8 (parte 1 de 2)

**Función de los treinta primeros genes con mayor frecuencia de expresión comunes de
los 11 casos de Cáncer Diferenciado de Tiroides del 2019 al 2020**

Gen	Función y patologías asociadas	Referencia
PCBD2	Cofactor de dimerización del factor nuclear de hepatocitos-1. Precursor del cofactor de fenilalanina hidroxilasa BH4 (tetrahidrobiopterina). Mutación provoca hiperfenilalaninemia, vitíligo. Sobre expresada en cáncer de colon y melanoma maligno.	(315)
LRMDA	Codifica una proteína repetida rica en leucina. Papel en la diferenciación de los melanocitos. Mutación se relaciona al albinismo oculocutáneo autosómico recesivo 7 (OCA7).	(315)
WWOX	Gen supresor de tumores al inducir apoptosis. Mutaciones asociadas con múltiples tipos de cáncer y con la ataxia espinocerebelosa autosómica recesiva 12.	(315)
RPS29	Proteína ribosómica que puede mejorar la actividad supresora de tumores de la proteína 1A relacionada con Ras (KREV1). Mutación se relaciona con cánceres colorrectales.	(315)
RBFOX1	Proteínas de unión al ARN. Mutación asociada a la ataxia espinocerebelosa tipo 2 (SCA2) y enfermedades neurodegenerativas familiares.	(315)
FHIT	Gen supresor de tumores. Proteína del metabolismo de las purinas. Daño inducido por carcinógenos puede conducir a translocaciones y transcripciones aberrantes. Mutación asociada a carcinomas de esófago, estómago y colon.	(315)
NTM	Moléculas de adhesión celular. Promueve el crecimiento y la adhesión de neuritas, proteína de unión a opioides/similar a una molécula de adhesión celular (OPCML).	(315)
RAD51B	Reparación del ADN mediante recombinación homóloga. La sobreexpresión de este gen provoca un retraso del ciclo celular G1 y apoptosis celular, lo que sugiere un papel de esta proteína en la detección de daños en el ADN. Mutación asociada a leiomiomas.	(315)
PRKN	Direccionamiento de proteínas sustrato para la degradación proteasomal. Mutaciones causan la enfermedad de Parkinson y la enfermedad de Parkinson juvenil autosómica recesiva.	(315)
MACROD2	Deacetilasa involucrada en la eliminación de ADP-ribosa de proteínas mono-ADP-ribosiladas. La proteína codificada se traslada del núcleo al citoplasma tras el daño del DNA.	(315)
ASIC2	Asociada a canales de sodio sensibles a amilorida y neurotransmisión.	(315)
TBC1D5	Recuperación de proteínas de membrana desde el endosoma al Golgi. Tubula la membrana endosomal.	(315)
GPC5	Proteoglicanos de heparán sulfato de la superficie celular. Papel en el control de la división celular y la regulación del crecimiento. Sobre expresado en linfomas y otros tumores.	(315)
EXOC4	Componente del complejo de exocistos. Remodelación citoesquelética de actina y la maquinaria de transporte de vesículas. Biogénesis de la polaridad de la superficie de las células epiteliales.	(315)
IMMP2L	Procesamiento de las secuencias de péptidos señal que se utilizan para dirigir las proteínas mitocondriales a las mitocondrias. Actividad catalítica del complejo de peptidasa de la membrana interna mitocondrial (IMP).	(315)
CAMTA1	Factor de transcripción y supresor de tumores. Translocación entre este gen y el gen WWTR1, con la oncoproteína WWTR1-CAMTA1 resultante que conduce a un hemangioendotelio epitelioide, un cáncer vascular maligno.	(315)

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

Tabla 8 (parte 2 de 2)

Función de los treinta primeros genes con mayor frecuencia de expresión comunes de los casos 11 casos de Cáncer Diferenciado de Tiroides del 2019 al 2020

Gen	Función y patologías asociadas	Referencia
CDH13	Superfamilia de cadherinas. Glicoproteína de adhesión célula-célula. Regulador negativo del crecimiento de axones, protege a las células endoteliales vasculares de la apoptosis debida al estrés oxidativo (resistencia a la aterosclerosis). Gen hipermetilado en muchos cánceres.	(315)
SMYD3	Histona metiltransferasa que funciona en complejos de ARN polimerasa II.	(315)
ANKS1B	Proteína de múltiples dominios que se expresa predominantemente en el cerebro y los testículos. Interactúa con beta amiloide. Desarrollo normal del cerebro y en la patogénesis de la enfermedad de Alzheimer. Sobre expresada en leucemia linfocítica aguda de células pre-B asociada con la translocación t (1; 19).	(315)
DAB1	Desarrollo del cerebro, dirigiendo la migración de las neuronas corticales. Transductor de señal que interactúa con las vías de la proteína quinasa para regular el posicionamiento neuronal en el cerebro en desarrollo.	(315)
DLG2	Familia del guanilato quinasa asociada a la membrana (MAGUK). Interactúa en sitios postsinápticos para formar un andamio multimérico para la agrupación de receptores, canales iónicos y proteínas de señalización asociadas.	(315)
FRMD4 A	Dominio FERM que regula la polaridad de las células epiteliales. Remodelación de las uniones adherentes y la formación de cables lineales de actina. Polimorfismos en este gen están asociados con la enfermedad de Alzheimer y también con la dependencia de la nicotina.	(315)
KCNIP4	Interactúan con los canales de potasio (Kv) dependientes de voltaje (KCNIP). Regulan la excitabilidad neuronal en respuesta a variaciones de calcio intracelular.	(315)
CTNNA2	Variante de transcripción. Enlazador entre los receptores de adhesión de cadherina y el citoesqueleto para regular la adhesión, diferenciación del sistema nervioso y la plasticidad morfológica de las sinapsis y la laminación cerebelosa e hipocampal durante el desarrollo.	(315)
AGBL4	Metalocarboxipeptidasa que media la desglutamilación de proteínas diana.	(315)
PRKCE	Proteínas quinasas C específicas de serina y treonina activadas por calcio y diacilglicerol. Activación de los canales neuronales, la apoptosis, cardio protección de la isquemia, la respuesta al choque térmico, la ansiedad y la exocitosis de la insulina.	(315)
SYN3	La sinapsinas codifican fosfoproteínas neuronales que se asocian con la superficie citoplásmica de las vesículas sinápticas. Papel en la sinaptogénesis y modulación de la liberación de neurotransmisores. Mutación asociada a enfermedades neuropsiquiátricas.	(315)
ARHGA P15	RHO GTP asas. Regulan diversos procesos biológicos.	(315)
PACRG	Gen de parkina. complejo molecular con chaperonas, incluidas las proteínas de choque térmico 70 y 90, y componentes de chaperonina. Mutaciones en la enfermedad de Parkinson juvenil autosómica recesiva, susceptibilidad a lepra, componente de los cuerpos de Lewy en pacientes con enfermedad de Parkinson.	(315)
C9orf92	Variante de transcripción 2 es un ARN no codificante.	(315)

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

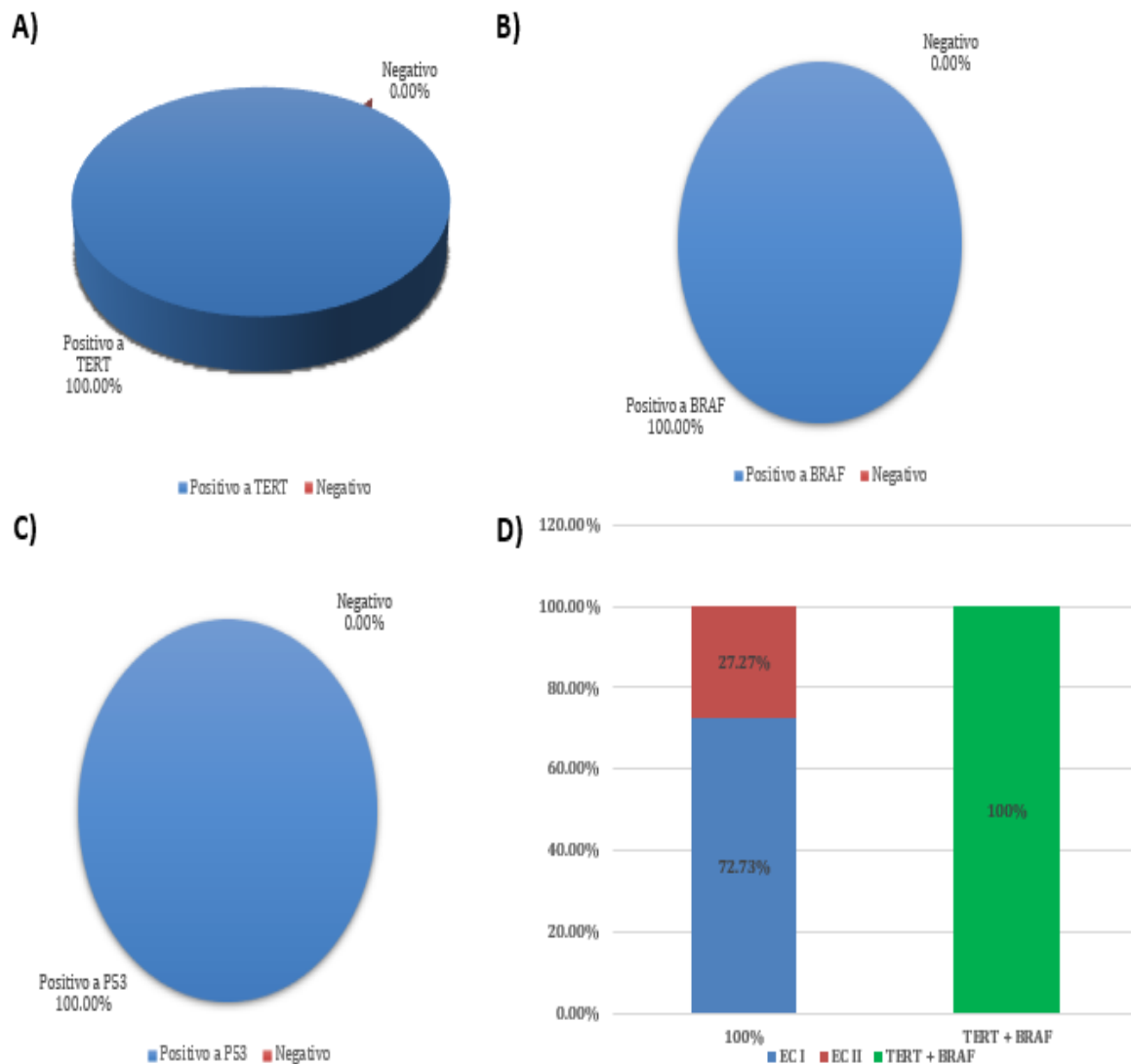


Figura 16. Expresión de TERT, BRAF y P53 en el Cáncer Diferenciado de Tiroides. Se objetiva que en el 100% de los casos estudiados de pacientes portadores de CDT expresarán TERT (A), BRAF (B) y P53 (C), además de la coexistencia de TERT y BRAF independiente del estadio clínico de la enfermedad. (D). Fuente propia.

**PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE
TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020**

Tabla 9 (parte 1 de 2)

**Expresión de P53 y otros genes asociados a P53 encontrados comunes en los 11 casos
de Cáncer Diferenciado de Tiroides**

Gen	Nombre	Sinónimos	Localización	Referencia
TP53	Proteína tumoral p53	P53; BCC7; LFS1; BMFS5; TRP53	17p13.1	(315)
WRAP53	Repetición WD que contiene antisentido para TP53	DKCB3; TCAB1; WDR79	17p13.1	(315)
TP53BP2	Proteína tumoral 2 proteína de unión a p53	BBP; 53BP2; ASPP2; P53BP2; PPP1R13A	1q41	(315)
TP53BP1	Proteína de unión p53 de la proteína tumoral 1	p202; 53BP1; TDRD30; p53BP1	15q15.3	(315)
TP53I11	Proteína tumoral proteína inducible p53 11	PIG11	11p11.2	(315)
NOP53	Factor de biogénesis del ribosoma NOP53	PICT1; PICT-1; GLTSCR2	19q13.33	(315)
USP53	Peptidasa 53 específica de ubiquitina	Q70EK8; UBP53	4q26	(315)
CFAP53	Proteína 53 asociada a flagelos y cilios	HTX6; CCDC11; Q68DA5; Q8WVQ5; Q9P2J7	18q21.1	(315)

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

Tabla 9 (parte 2 de 2)

Expresión de P53 y otros genes asociados a P53 encontrados comunes en los 11 casos de Cáncer Diferenciado de Tiroides

Gen	Función y patologías asociadas	Referencia
TP53	Codifica una proteína supresora de tumores que contiene dominios de activación transcripcional, unión al ADN y oligomerización. Induciendo la detención del ciclo celular, apoptosis, senescencia, reparación del ADN o cambios en el metabolismo.	(315)
WRAP53	Codifica un componente para síntesis de telómeros y función de telomerasa. Interactúa con disquerina, TERT y TERC. Regula niveles de ARNm de p53 endógeno y la inducción adicional de la proteína p53 dirigiéndose a la región no traducida 5' del ARNm de p53.	(315)
TP53BP2	Este gen codifica un miembro de la familia ASPP (proteína estimulante de la apoptosis de p53) de proteínas que interactúan con p53. Regula la apoptosis y el crecimiento celular con interacciones con otras moléculas reguladoras, incluidos miembros de la familia p53.	(315)
TP53BP1	Codifica una proteína de reparación de rotura de doble hebra del ADN, promoción de la señalización del punto de control después del daño del ADN, andamio para el reclutamiento de proteínas de respuesta al daño del ADN en los telómeros no protegidos.	(315)
TP53I11	Funciona como un mediador para equilibrar la activación de AKT y AMPK para adaptar las células a diferentes contextos celulares, como la unión y separación de la matriz extracelular.	(315)
NOP53	Funciona como un sensor nucleolar que regula la activación de p53/TP53 en respuesta a la perturbación de la biogénesis del ribosoma, daño del ADN y otras condiciones de estrés.	(315)
USP53	Proteólisis dependiente de ubiquitina-proteasoma. Promueve la apoptosis e inhibe la glucólisis en el adenocarcinoma de pulmón a través de la señalización FKBP51-AKT1. Enfermedades incluyen colestasis y colestasis intrahepática familiar progresiva.	(315)
CFAP53	Este gen pertenece a la familia CFAP53. Se encontró que se expresaba diferencialmente por las células ciliadas de la epidermis de la rana y en los fibroblastos de la piel humana. Las mutaciones en este gen están asociadas con heterotaxi-6 visceral, que implica a este gen en la determinación del patrón asimétrico de izquierda a derecha.	(315)

En la tabla 9 (parte 1 y 2) se muestra la expresión de P53 y otros genes expresados asociados a P53 encontrados comunes en los 11 casos de pacientes portadores de CDT de la población estudiada. Además, se describen características de los mismos, como sinónimos de su nomenclatura, localización cromosómica, función y sus referencias bibliográficas.

**PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE
TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020**

Tabla 10

**Traslocaciones tipo fusión encontradas en 3 casos de los 11 pacientes con Cáncer
Diferenciado de Tiroides**

Fusión	Variante histopatológica	Genes	Localización	Referencia
TDG - TMEM132B	Papilar folicular	Timina-ADN glicosilasa	Cromosoma 12 q23.3-q24.31	(315)
		Proteína transmembrana 132B		
MEIS1-ITSN2	Papilar clásico	Meis homeobox 1	Cromosoma 2 p14- p23.3	(315)
		Intersectina 2		
KCNE1B - FP671120.4	Papilar clásico	Subfamilia de canales regulados por voltaje de potasio E subunidad reguladora 1B	Cromosoma 21 p11.2- p11.2	(315)
		RNA ribosomal no codificante		

En la tabla 10 se muestra la presencia de las 3 traslocaciones tipo fusión que fueron encontradas en 3 casos de los 11 pacientes con el diagnóstico de CDT. Así mismo, se describen los genes componentes de cada una, variante histopatológica de cada tumor, localización cromosómica y referencia bibliográfica.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

Tabla 11

Genes componentes de las traslocaciones tipo fusión encontradas en 3 casos de los 11 pacientes con Cáncer Diferenciado de Tiroides

Gen	Función y patologías asociadas	Referencia
TDG	Elimina restos de timina de: desajustes G/T hidrolizando el enlace carbono-nitrógeno entre la columna vertebral de azúcar-fosfato del ADN y la timina desaparecida, emparejamientos incorrectos de C/T y T/T, elimina el uracilo y el 5-bromouracilo de los emparejamientos incorrectos con guanina. Defensa celular contra la mutación genética causada por la desaminación espontánea de 5-metilcitosina y citosina.	(315)
TMEM132B	Asociado con la pérdida auditiva no sindrómica, el trastorno de pánico y el cáncer. Función de adhesión celular para la familia TMEM132, conectando el medio extracelular con el citoesqueleto de actina intracelular.	(315)
MEIS1	Los genes homeobox, representada por los genes HOX, juegan un papel crucial en el desarrollo normal. Además, varias homeoproteínas están involucradas en la neoplasia. Este gen codifica una proteína homeobox que pertenece a la familia TALE ('extensión de bucle de tres aminoácidos') de proteínas que contienen homeodominio. Asociado a leucemia mieloide.	(315)
ITSN2	Codifica una proteína citoplásmica que contiene dominios SH3. Proteínas implicadas en la endocitosis mediada por clatrina. Regula la formación de vesículas recubiertas de clatrina. Induce la endocitosis del receptor de antígeno de células T (TCR).	(315)
KCNE1B	El producto del gen KCNE1, minK, forma un complejo estable con HERG (KCNH2; 152427) y esta heteromultimerización regula el rectificador retardado cardíaco de activación rápida, control de la frecuencia y el ritmo cardíaco, formando el canal de potasio I (Ks) cardíaco. Relacionado a cáncer de útero, gástrico, astroglioma, leucemia mielógena, cáncer de tiroides.	(315)
FP671120.4	ARN largos no codificantes (LncRNA), regula la expresión génica por medio de la modificación de la cromatina, la activación transcripcional y la interferencia transcripcional en la aparición y desarrollo de tumores.	(315)

En la tabla 11 se muestra la función de los genes componentes de las 3 traslocaciones tipo fusión encontradas en 3 casos de los 11 pacientes con CDT estudiados. Así mismo, se muestra la referencia bibliográfica.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

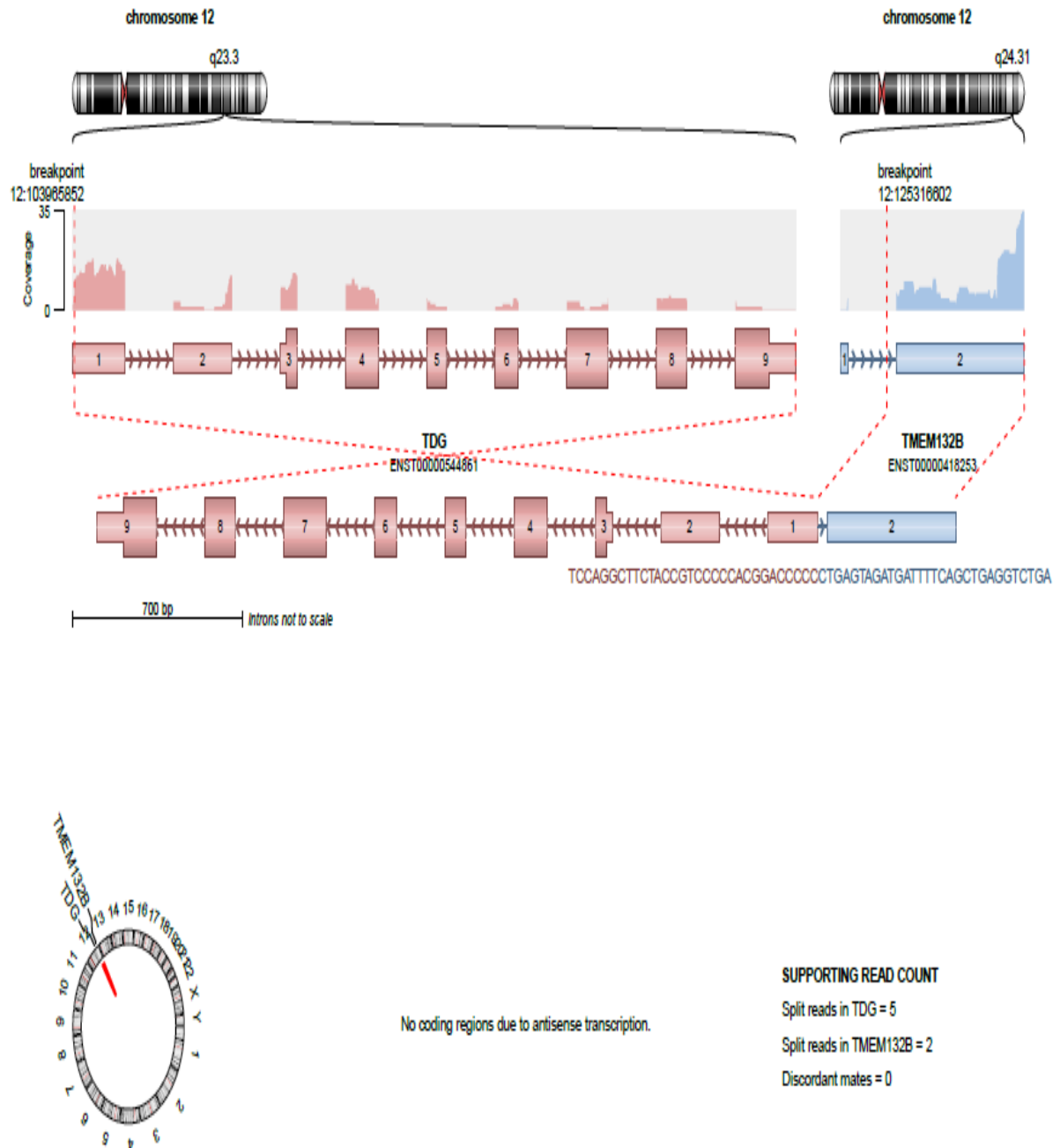


Figura 17. Traslocación/fusión TDG-TMEM132B de un caso de Cáncer Diferenciado de Tiroides. Encontrada en un caso de CDT variante papilar folicular, la cual no presenta regiones codificantes debido a ser una transcripción antisentido. Fuente análisis bioinformático Admera health.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

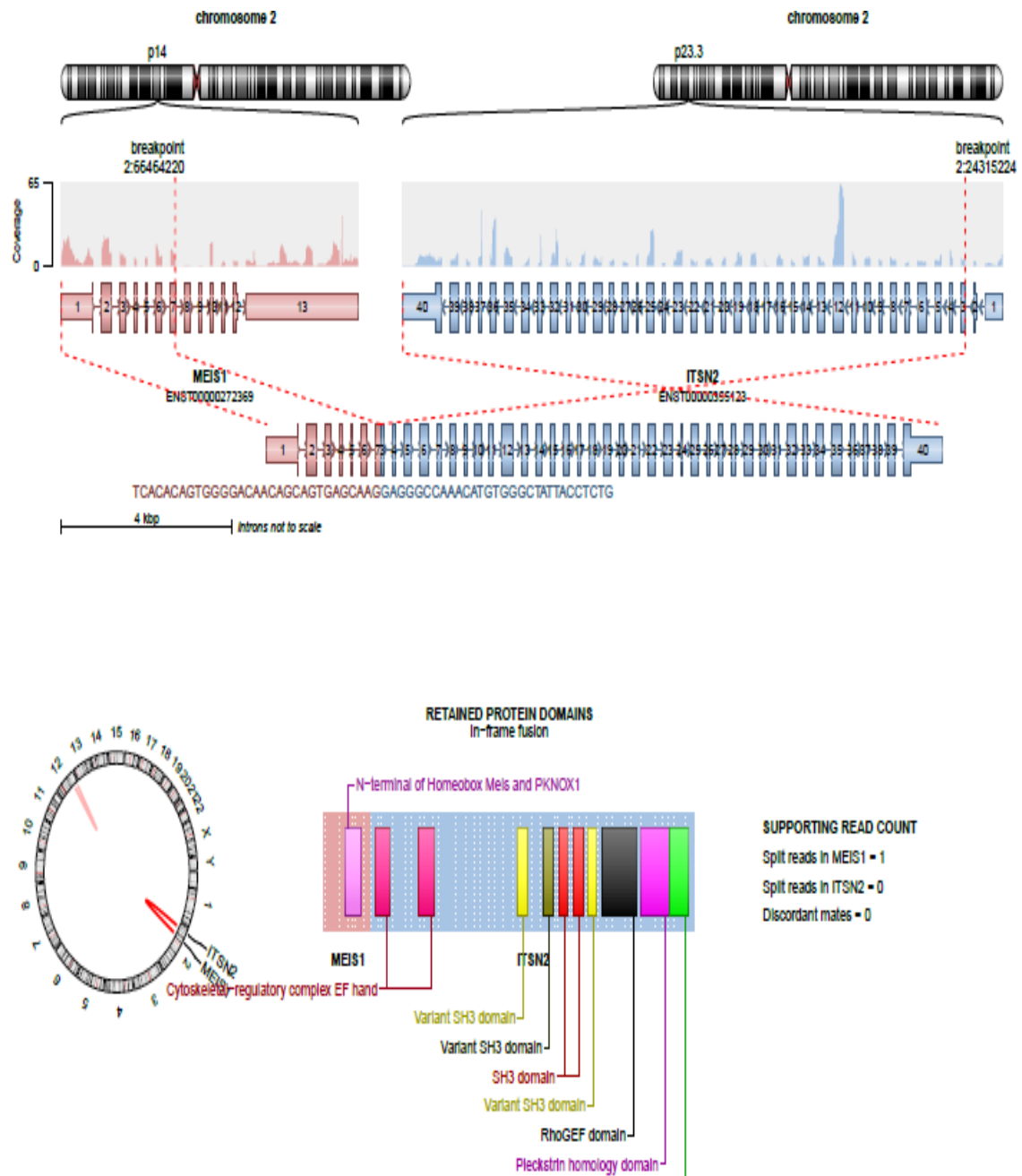
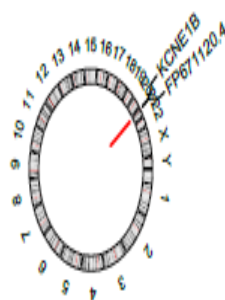
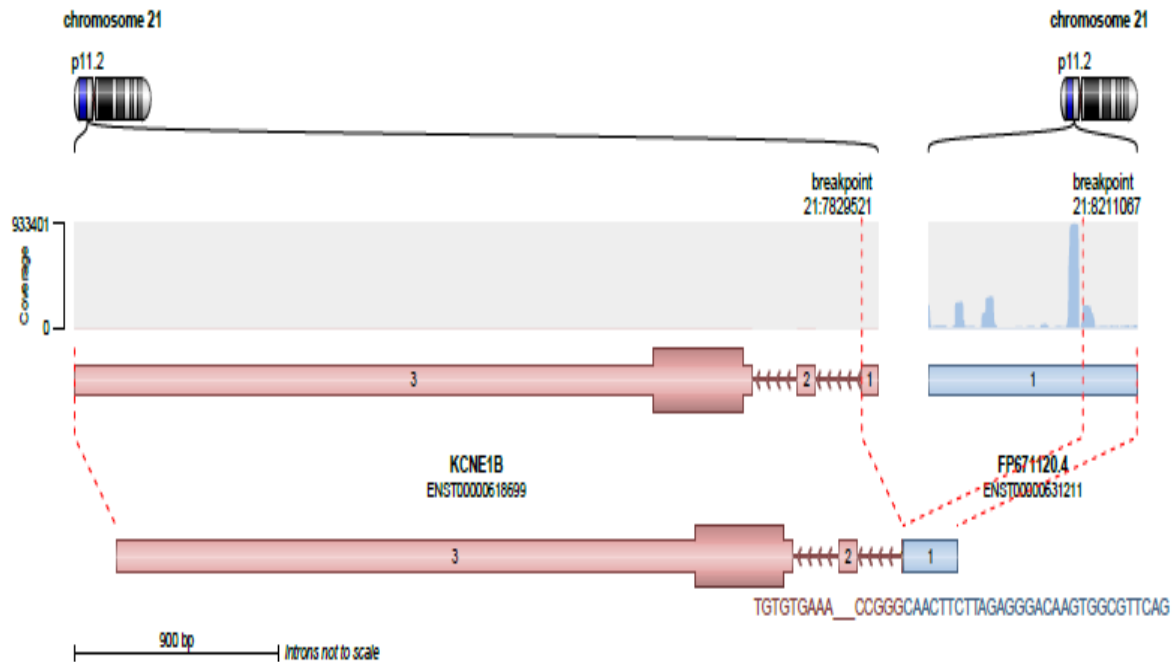


Figura 18. Traslocación/fusión MEIS1-ITSN2 de un caso de Cáncer Diferenciado de Tiroides. Encontrada en un caso de CDT variante papilar clásico, la cual presenta dominios de proteínas retenidos de ambos genes. Fuente análisis bioinformático Admera health.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020



No coding regions due to antisense transcription.

SUPPORTING READ COUNT

Split reads in KCNE1B = 0
Split reads in FP671120.4 = 4
Discordant mates = 0

Figura 19. Traslocación/fusión KCNE1B - FP671120.4 de un caso de Cáncer diferenciado de tiroides. Encontrada en un caso de CDT variante papilar clásico, la cual no presenta regiones codificantes debido a ser una transcripción antisentido. Fuente análisis bioinformático Admera health.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

Tabla 12

Ubicación de las mutaciones puntuales BRAF comunes en los 11 casos de Cáncer Diferenciado de Tiroides

BRAF	
Mutación	Referencia bibliográfica
Nombre: rs889134706 Tipo de variante: SNV Alelos: C > T, G Cromosoma: 7:140924572 (GRCh38) Exón 1 BRAF: Variante missense Consecuencia funcional: No reportada	(583)
Tipo de variante: SNV Alelos: T > A Cromosoma: 7: 140850164 (GRCh38) Exón 2	Mutación de novo
Tipo de variante: SNV Alelos: T > A Cromosoma: 7: 140800452 (GRCh38) Exón 7	Mutación de novo
Tipo de variante: SNV Alelos: T > A, C, G Cromosoma: 7: 140794401 (GRCh38) Exón 8	Mutación de novo
Tipo de variante: SNV Alelos: A > T Cromosoma: 7: 140777028 (GRCh38) Exón 14	Mutación de novo
Tipo de variante: SNV Alelos: A > T Cromosoma: 7: 140754195 (GRCh38) Exón 15	Mutación de novo

En la tabla 12 se muestra las mutaciones de BRAF encontradas en los 11 pacientes del estudio, tomando como referencia al banco genético humano de secuencia de ADN “Consortio de referencia del genoma Human Build 38” (GRCh38.p13). De las 6 mutaciones encontradas, la mutación de BRAF rs113488022 variante missense, ya se encuentra descrita en la literatura; las 5 mutaciones restantes son de novo. Se muestra la referencia bibliográfica.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

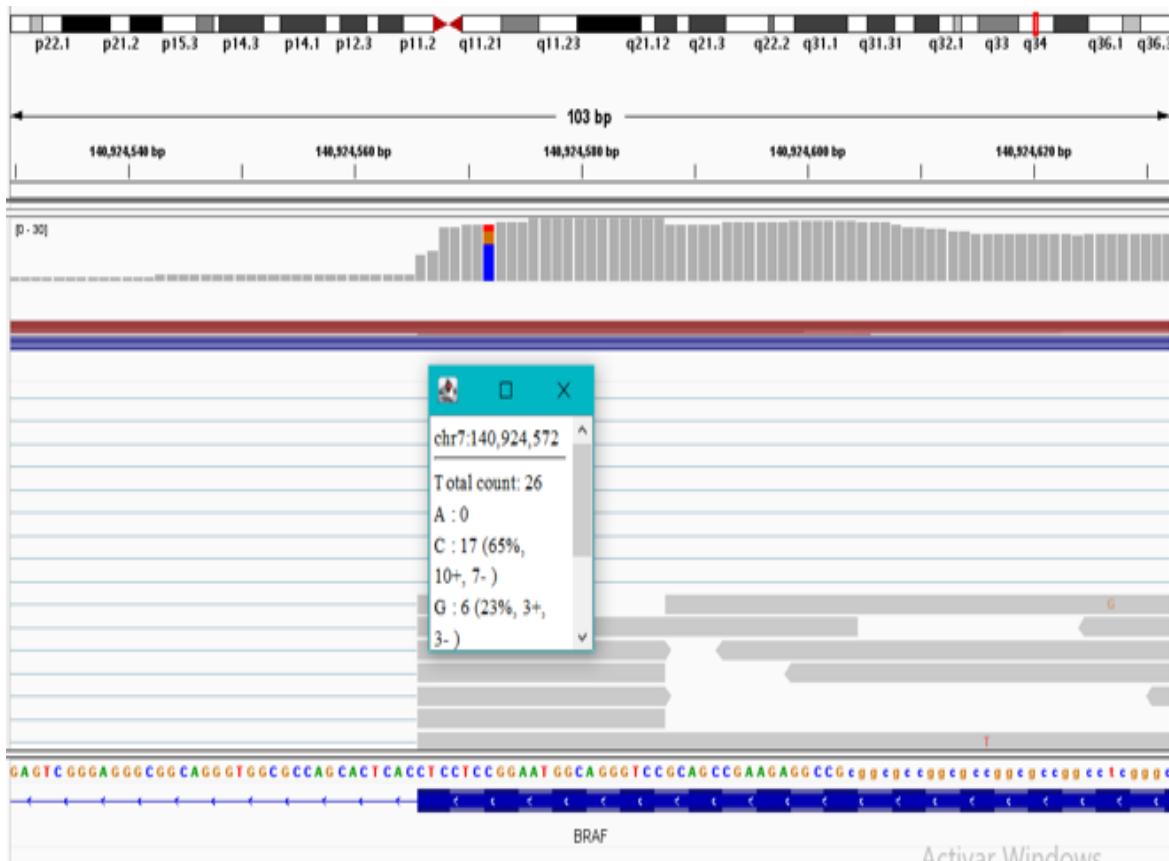


Figura 20. Mutación puntual BRAF Cromosoma 7: 140924572 C > T en el Cáncer Diferenciado de Tiroides. Se muestra la pista de cobertura predeterminada para un archivo de alineación BAM utilizando el software IGV (Integrative Genomics Viewer) de la combinación de los 11 pacientes del estudio. Si un nucleótido difiere de la secuencia de referencia en más del 20 % de las lecturas ponderadas de calidad, IGV colorea la barra en proporción al recuento de lecturas de cada base (A, C, G, T). Se encontró el polimorfismo de nucleótido único (SNP) BRAF 7: 140924572, SNV (variación en la secuencia de ADN que ocurre cuando se altera un solo nucleótido) C > T, G. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

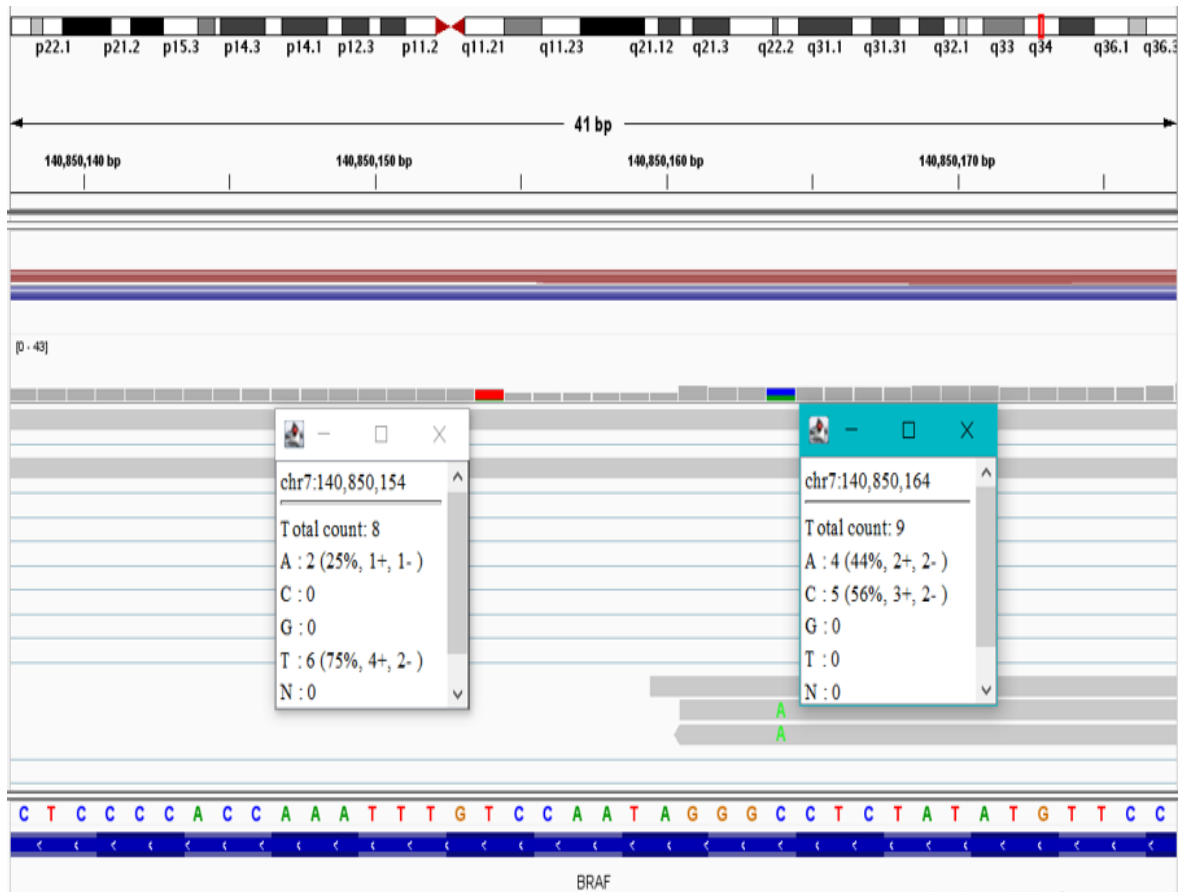


Figura 21. Mutaciones puntuales BRAF Cromosoma 7: 140850154 T > A; 140850164 T>A en el Cáncer Diferenciado de Tiroides. Se muestra la pista de cobertura predeterminada para un archivo de alineación BAM utilizando el software IGV (Integrative Genomics Viewer) de la combinación de los 11 pacientes del estudio. Si un nucleótido difiere de la secuencia de referencia en más del 20 % de las lecturas ponderadas de calidad, IGV colorea la barra en proporción al recuento de lecturas de cada base (A, C, G, T). Se encontraron los polimorfismos de nucleótido único (SNPs) BRAF 7: 140850154, SNV (variación en la secuencia de ADN que ocurre cuando se altera un solo nucleótido) T > A y BRAF 7: 140850164, SNV T > A. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

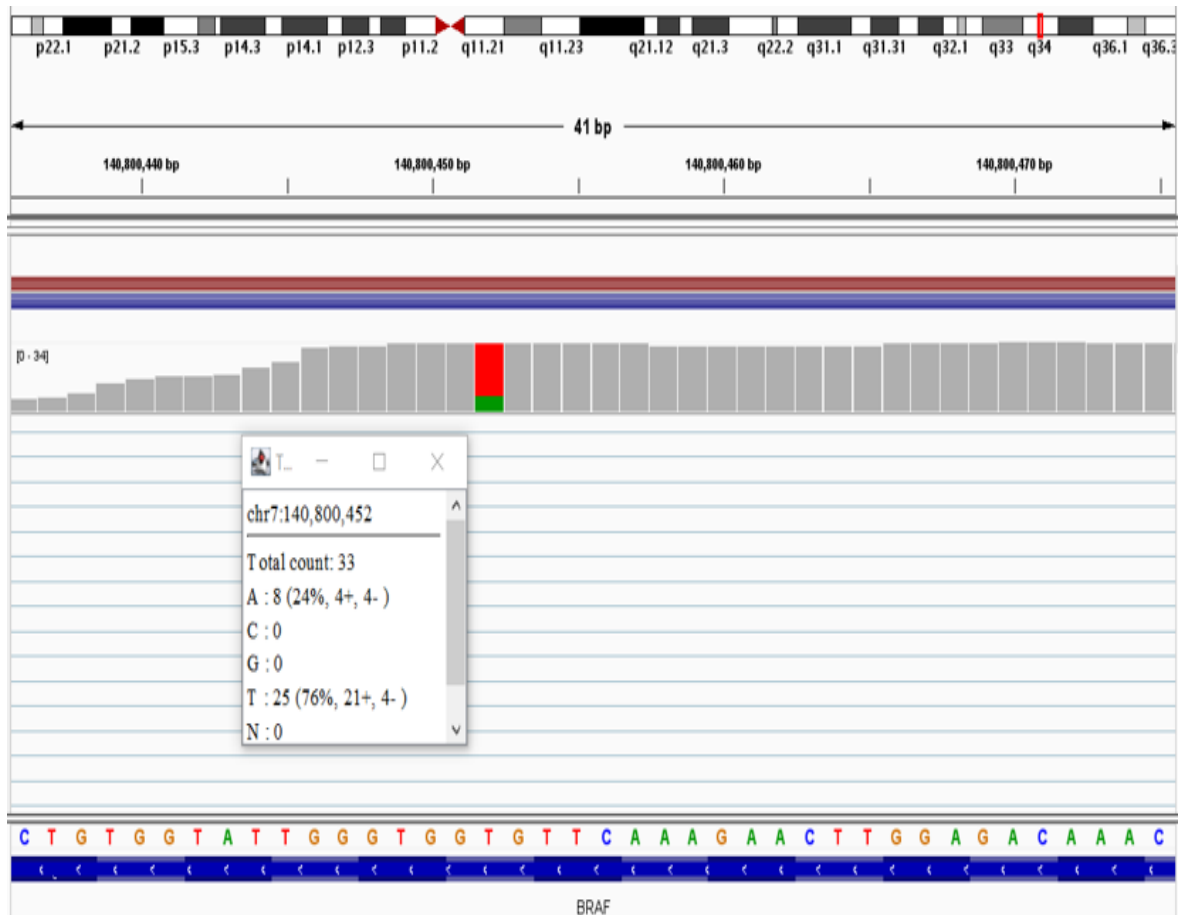


Figura 22. Mutaciones puntuales BRAF Cromosoma 7: 140800452 T > A en el Cáncer Diferenciado de Tiroides. Se muestra la pista de cobertura predeterminada para un archivo de alineación BAM utilizando el software IGV (Integrative Genomics Viewer) de la combinación de los 11 pacientes del estudio. Si un nucleótido difiere de la secuencia de referencia en más del 20 % de las lecturas ponderadas de calidad, IGV colorea la barra en proporción al recuento de lecturas de cada base (A, C, G, T). Se encontró el polimorfismo de nucleótido único (SNP) BRAF 7: 140800452, SNV (variación en la secuencia de ADN que ocurre cuando se altera un solo nucleótido) T > A. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

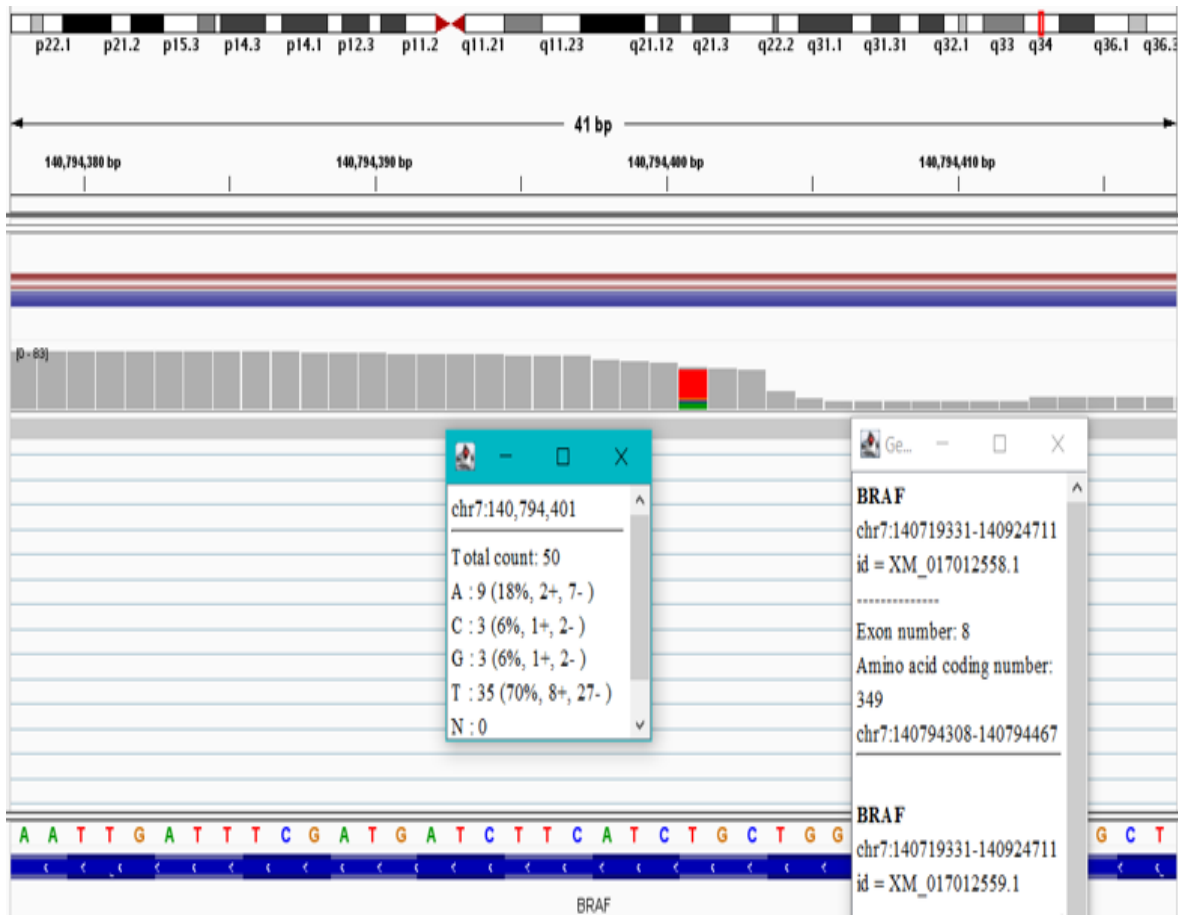


Figura 23. Mutación puntual BRAF Cromosoma 7: 140794401 T > A, C, G en el Cáncer Diferenciado de Tiroides. Se muestra la pista de cobertura predeterminada para un archivo de alineación BAM utilizando el software IGV (Integrative Genomics Viewer) de la combinación de los 11 pacientes del estudio. Si un nucleótido difiere de la secuencia de referencia en más del 20 % de las lecturas ponderadas de calidad, IGV colorea la barra en proporción al recuento de lecturas de cada base (A, C, G, T). Se encontró el polimorfismo de nucleótido único (SNP) BRAF 7: 140794401, SNV (variación en la secuencia de ADN que ocurre cuando se altera un solo nucleótido) T > A, C, G. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

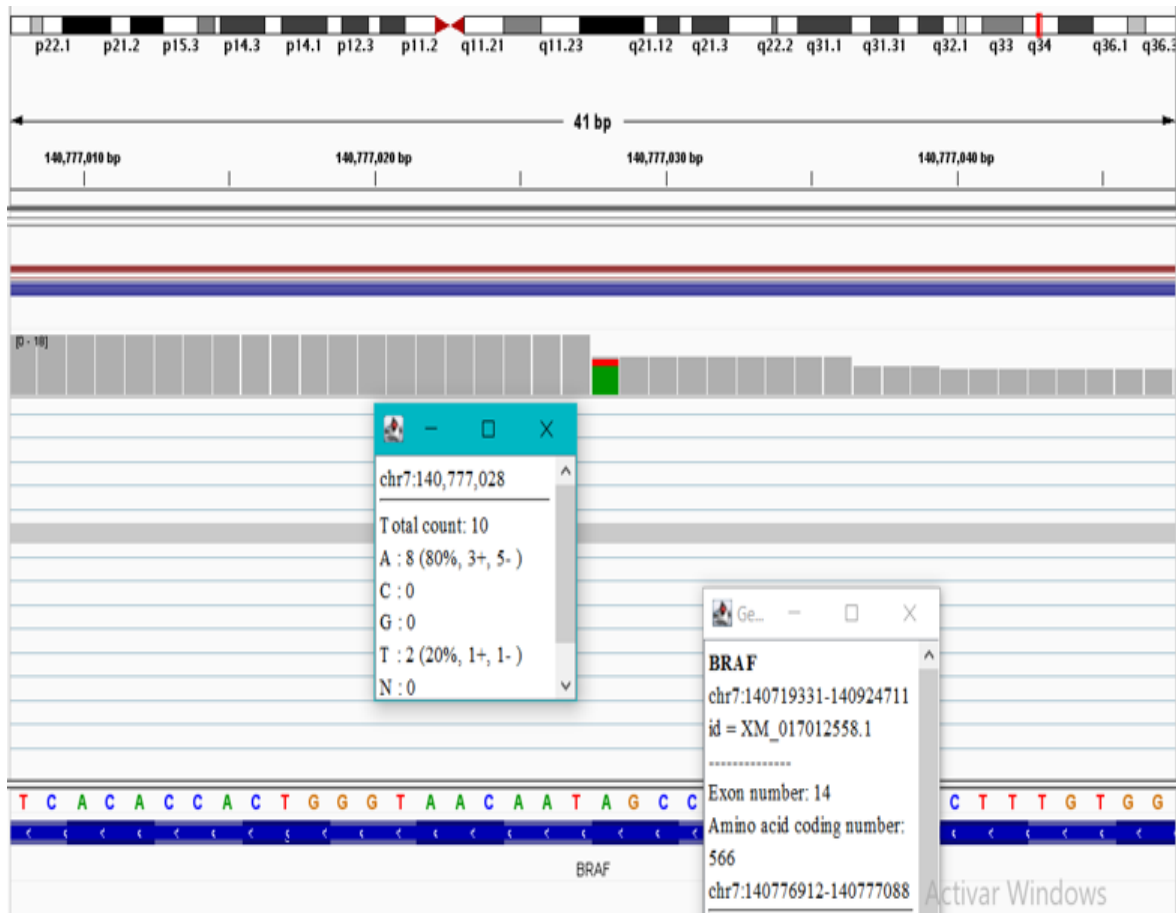


Figura 24. Mutación puntual BRAF Cromosoma 7: 140777028 A > T en el Cáncer Diferenciado de Tiroides. Se muestra la pista de cobertura predeterminada para un archivo de alineación BAM utilizando el software IGV (Integrative Genomics Viewer) de la combinación de los 11 pacientes del estudio. Si un nucleótido difiere de la secuencia de referencia en más del 20 % de las lecturas ponderadas de calidad, IGV colorea la barra en proporción al recuento de lecturas de cada base (A, C, G, T). Se encontró el polimorfismo de nucleótido único (SNP) BRAF 7: 140777028, SNV (variación en la secuencia de ADN que ocurre cuando se altera un solo nucleótido) A > T. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

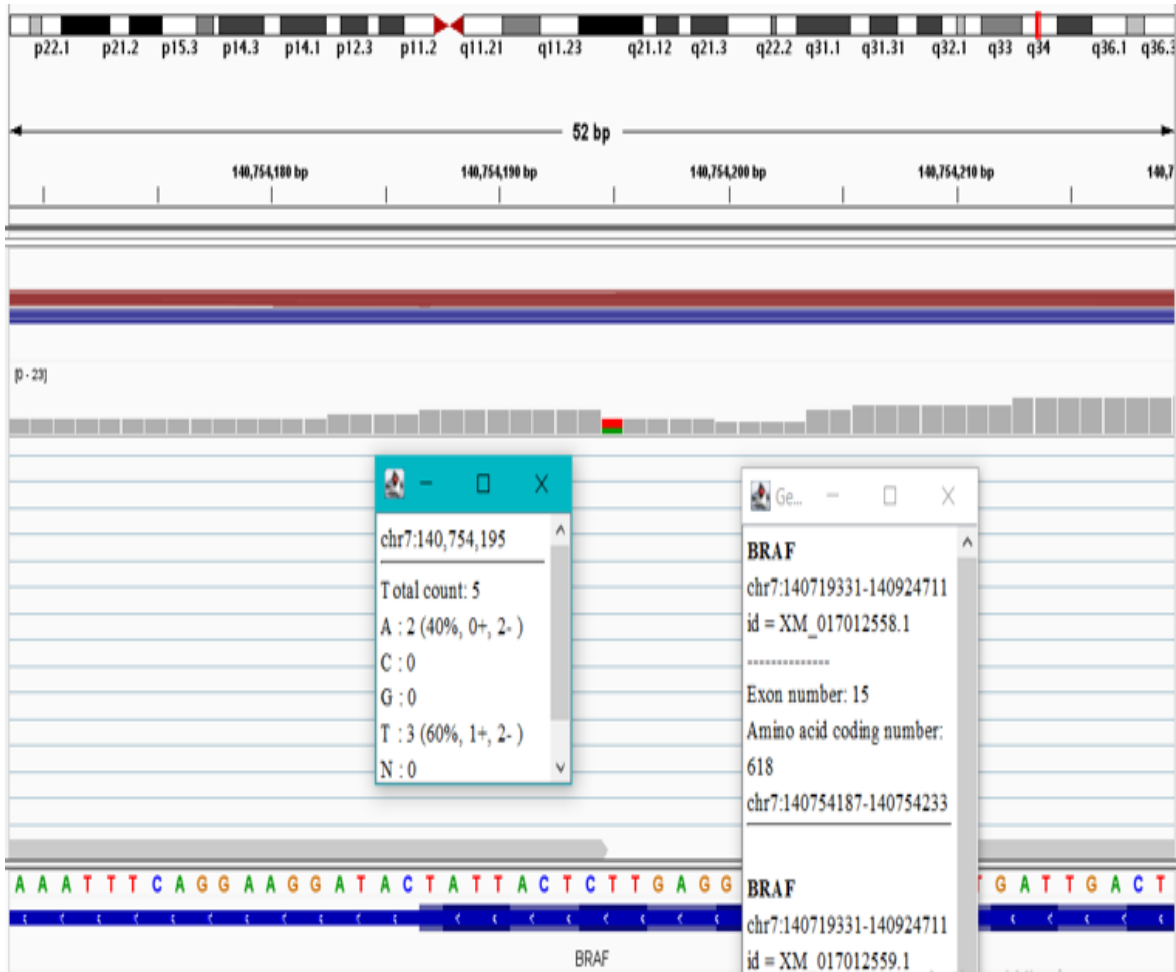


Figura 25. Mutación puntual BRAF Cromosoma 7: 140754195 A > T en el Cáncer Diferenciado de Tiroides. Se muestra la pista de cobertura predeterminada para un archivo de alineación BAM utilizando el software IGV (Integrative Genomics Viewer) de la combinación de los 11 pacientes del estudio. Si un nucleótido difiere de la secuencia de referencia en más del 20 % de las lecturas ponderadas de calidad, IGV colorea la barra en proporción al recuento de lecturas de cada base (A, C, G, T). Se encontró el polimorfismo de nucleótido único (SNP) BRAF 7: 140754195, SNV (variación en la secuencia de ADN que ocurre cuando se altera un solo nucleótido) A > T. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

Tabla 13 (parte 1 de 3)

Mutaciones puntuales TERT en el Cáncer Diferenciado de Tiroides

TERT	
Mutación	Referencia bibliográfica
Nombre: rs1751161082 Tipo de variante: SNV Alelos: T > C Cromosoma: 5:1293763 (GRCh38) Exón 2 TERT: Variante missense Consecuencia funcional: No reportada	(592)
Tipo de variante: SNV Alelos: C > G Cromosoma: 5:1293765 (GRCh38) Exón 2	Mutación de novo
Tipo de variante: SNV Alelos: C > G Cromosoma: 5:1293766 (GRCh38) Exón 2	Mutación de novo
Tipo de variante: SNV Alelos: C > A Cromosoma: 5:1293910 (GRCh38) Exón 2	Mutación de novo
Tipo de variante: SNV Alelos: A > G Cromosoma: 5:1293913 (GRCh38) Exón 2	Mutación de novo
Tipo de variante: SNV Alelos: C > A Cromosoma: 5:1293914 (GRCh37) Exón 2	Mutación de novo
Tipo de variante: SNV Alelos: G > T Cromosoma: 5:1294136 (GRCh38) Exón 2	Mutación de novo
Tipo de variante: SNV Alelos: A > C Cromosoma: 5:1294238 (GRCh38) Exón 2	Mutación de novo

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

Tabla 13 (parte 2 de 3)

Mutaciones puntuales TERT en el Cáncer Diferenciado de Tiroides

TERT	
Mutación	Referencia bibliográfica
Tipo de variante: SNV Alelos: A > C Cromosoma: 5:1294313 (GRCh38) Exón 2	Mutación de novo
Tipo de variante: SNV Alelos: T > G Cromosoma: 5:1294315 (GRCh38) Exón 2	Mutación de novo
Tipo de variante: SNV Alelos: C > T Cromosoma: 5:1294430 (GRCh38) Exón 2	Mutación de novo
Tipo de variante: SNV Alelos: G > A Cromosoma: 5:1294479 (GRCh38) Exón 2	Mutación de novo
Tipo de variante: SNV Alelos: T > C Cromosoma: 5:1280318 (GRCh38) Exón 4	Mutación de novo
Tipo de variante: SNV Alelos: T > A Cromosoma: 5:1278733 (GRCh38) Exón 6	Mutación de novo
Nombre: rs1749094617 Tipo de variante: SNV Alelos: T > C Cromosoma: 5:1272217 (GRCh38) Exón 7 TERT: Variante missense Consecuencia funcional: No reportada	(593)
Tipo de variante: SNV Alelos: G > T Cromosoma: 5: 1272218 (GRCh38) Exón 7	Mutación de novo

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

Tabla 13 (parte 3 de 3)

Mutaciones puntuales TERT en el Cáncer Diferenciado de Tiroides

TERT	
Mutación	Referencia bibliográfica
Tipo de variante: SNV Alelos: T > C Cromosoma: 5: 1272221 (GRCh38) Exón 7	Mutación de novo
Tipo de variante: SNV Alelos: T > C Cromosoma: 5:1264568 (GRCh38) Exón 11	Mutación de novo
Tipo de variante: SNV Alelos: A > C Cromosoma: 5: 1264493 (GRCh38) Exón 11	Mutación de novo
Tipo de variante: SNV Alelos: A > G Cromosoma: 5:1264446 (GRCh38) Exón 11	Mutación de novo
Nombre: rs33954691 Tipo de variante: SNV Alelos: G > A Cromosoma: 5:1255405 (GRCh38) Exón 13 TERT: Variante sinónima Consecuencia funcional: No reportada	(594)

En la tabla 13 (parte 1, 2 y 3) se muestra las mutaciones de TERT encontradas en los 11 pacientes del estudio, tomando como referencia al banco genético humano de secuencia de ADN “Consortio de referencia del genoma Human Build 38” (GRCh38.p13). De las 21 mutaciones encontradas, las mutaciones de TERT “rs1751161082 variante missense”, “rs1749094617 variante missense” y “rs33954691 variante sinónima”, ya se encontraban descritas en la literatura; las 18 mutaciones restantes son de novo. Se muestra la referencia bibliográfica.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

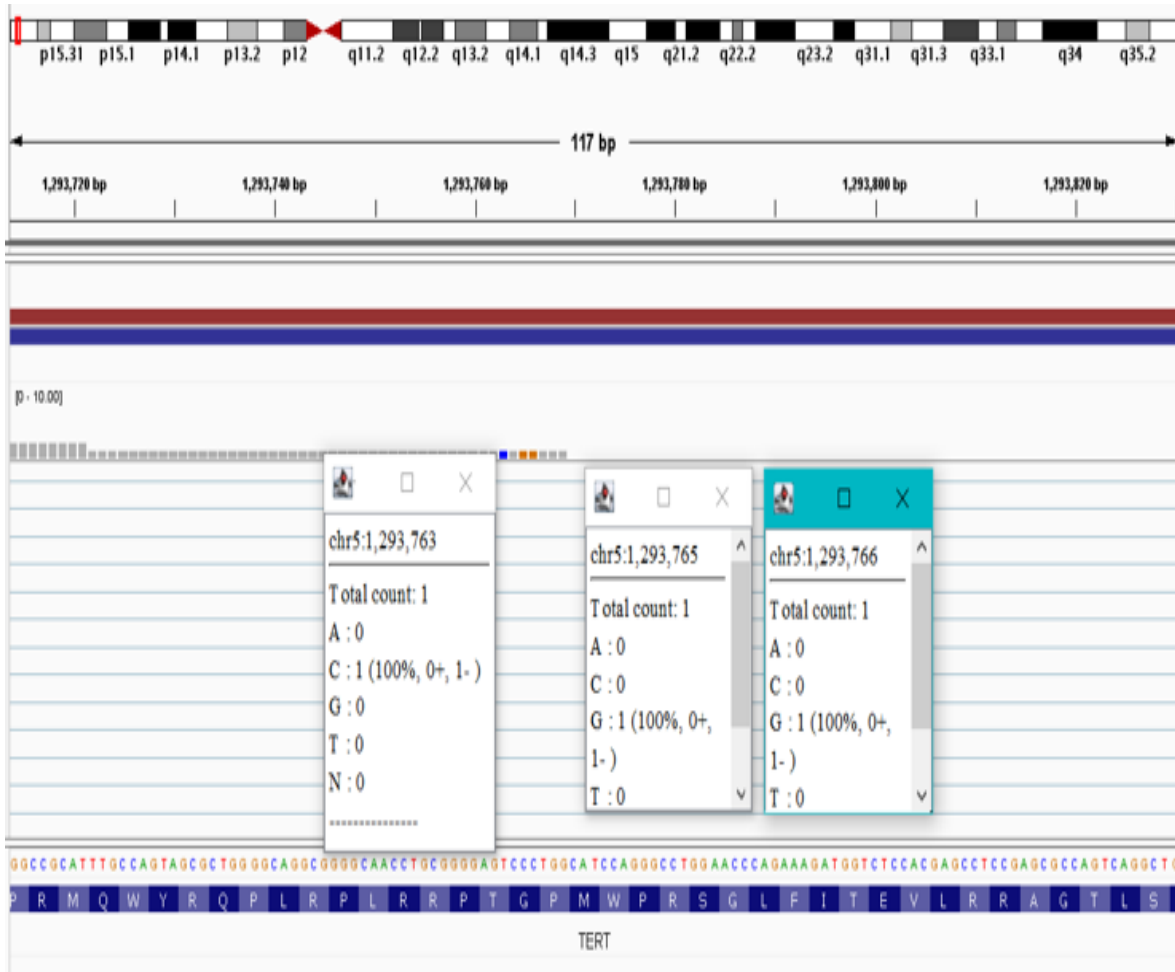


Figura 26. Mutaciones puntuales TERT Cromosoma 5: 1293763 T > C, 1293765 C > G, 1293766 C > G en el Cáncer Diferenciado de Tiroides. Se muestra la pista de cobertura predeterminada para un archivo de alineación BAM utilizando el software IGV (Integrative Genomics Viewer) de la combinación de los 11 pacientes del estudio. Si un nucleótido difiere de la secuencia de referencia en más del 20 % de las lecturas ponderadas de calidad, IGV colorea la barra en proporción al recuento de lecturas de cada base (A, C, G, T). Se encontraron los polimorfismos de nucleótido único (SNPs) TERT Cromosoma 5: 1293763, SNV (variación en la secuencia de ADN que ocurre cuando se altera un solo nucleótido) T > C, 1293765 SNV C > G y 1293766 SNV C > G. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

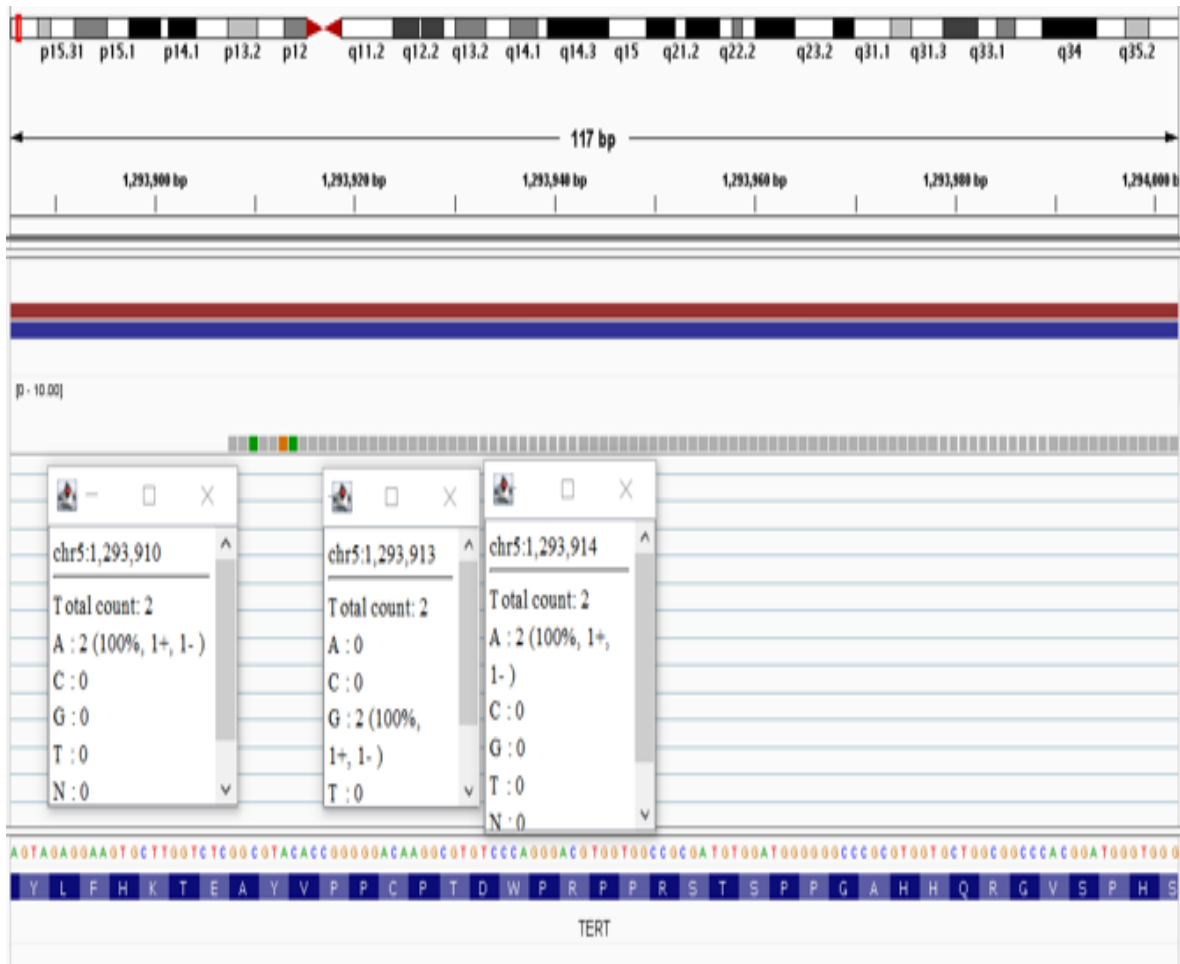


Figura 27. Mutaciones puntuales TERT Cromosoma: 5:1293910 C > A, 1293913 A > G, 1293914 C > A en el Cáncer Diferenciado de Tiroides. Se muestra la pista de cobertura predeterminada para un archivo de alineación BAM utilizando el software IGV (Integrative Genomics Viewer) de la combinación de los 11 pacientes del estudio. Si un nucleótido difiere de la secuencia de referencia en más del 20 % de las lecturas ponderadas de calidad, IGV colorea la barra en proporción al recuento de lecturas de cada base (A, C, G, T). Se encontraron los polimorfismos de nucleótido único (SNPs) TERT Cromosoma 5: 1293910, SNV (variación en la secuencia de ADN que ocurre cuando se altera un solo nucleótido) C > A, 1293913 SNV A > G y 1293914 SNV C > A. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

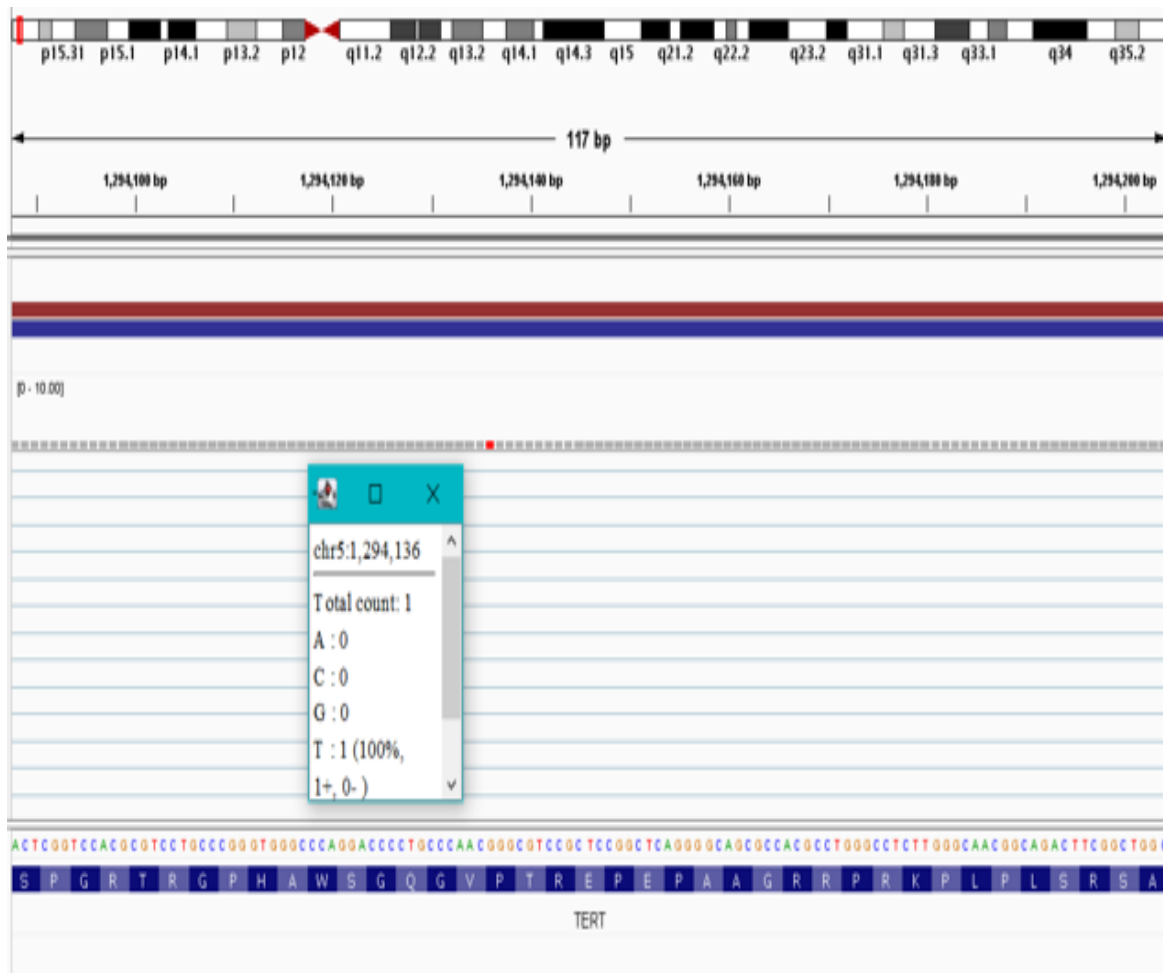


Figura 28. Mutaciones puntuales TERT Cromosoma 5: 1294136 G > T en el Cáncer Diferenciado de Tiroides. En la figura 36 se muestra la pista de cobertura predeterminada para un archivo de alineación BAM utilizando el software IGV (Integrative Genomics Viewer) de la combinación de los 11 pacientes del estudio. Si un nucleótido difiere de la secuencia de referencia en más del 20 % de las lecturas ponderadas de calidad, IGV colorea la barra en proporción al recuento de lecturas de cada base (A, C, G, T). Se encontró el polimorfismo de nucleótido único (SNPs) TERT Cromosoma 5: 1294136, SNV (variación en la secuencia de ADN que ocurre cuando se altera un solo nucleótido) G > T. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

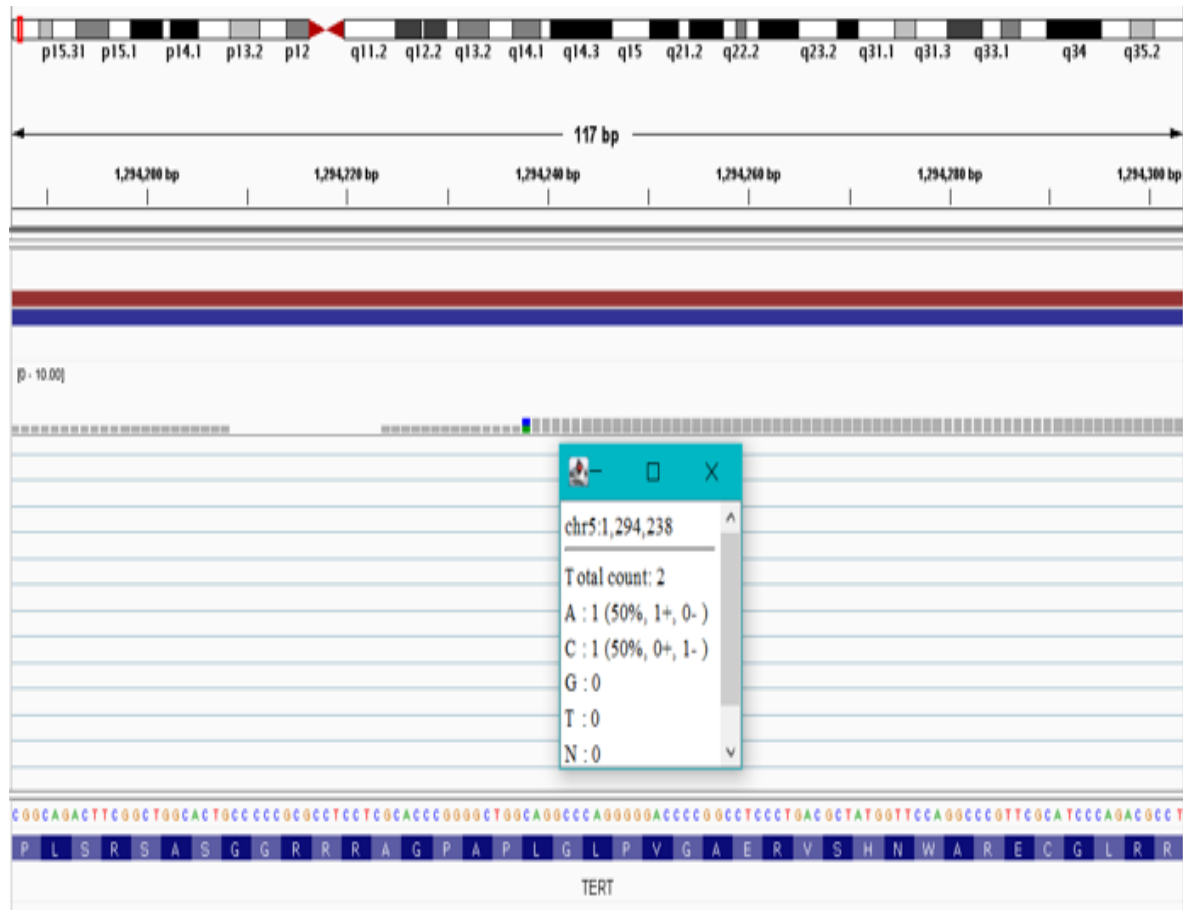


Figura 29. Mutaciones puntuales TERT Cromosoma 5: 1294238 A > C en el Cáncer Diferenciado de Tiroides. Se muestra la pista de cobertura predeterminada para un archivo de alineación BAM utilizando el software IGV (Integrative Genomics Viewer) de la combinación de los 11 pacientes del estudio. Si un nucleótido difiere de la secuencia de referencia en más del 20 % de las lecturas ponderadas de calidad, IGV colorea la barra en proporción al recuento de lecturas de cada base (A, C, G, T). Se encontró el polimorfismo de nucleótido único (SNPs) TERT Cromosoma 5: 1294238, SNV (variación en la secuencia de ADN que ocurre cuando se altera un solo nucleótido) A > C. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

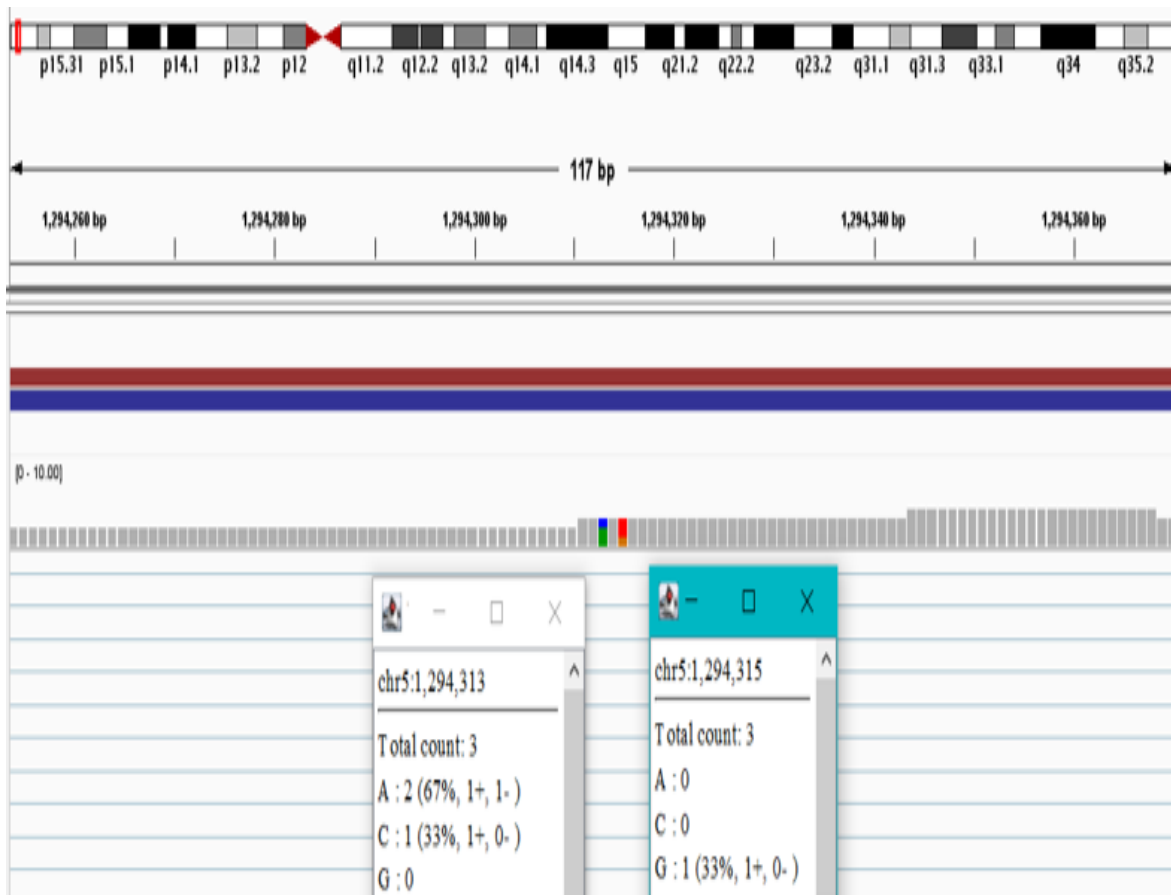


Figura 30. Mutaciones puntuales TERT Cromosoma 5: 1294313 A > C, 1294315 T > G en el Cáncer Diferenciado de Tiroides. Se muestra la pista de cobertura predeterminada para un archivo de alineación BAM utilizando el software IGV (Integrative Genomics Viewer) de la combinación de los 11 pacientes del estudio. Si un nucleótido difiere de la secuencia de referencia en más del 20 % de las lecturas ponderadas de calidad, IGV colorea la barra en proporción al recuento de lecturas de cada base (A, C, G, T). Se encontraron los polimorfismos de nucleótido único (SNPs) TERT Cromosoma 5: 1294313, SNV (variación en la secuencia de ADN que ocurre cuando se altera un solo nucleótido) A > C y 1294315 SNV T > G. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

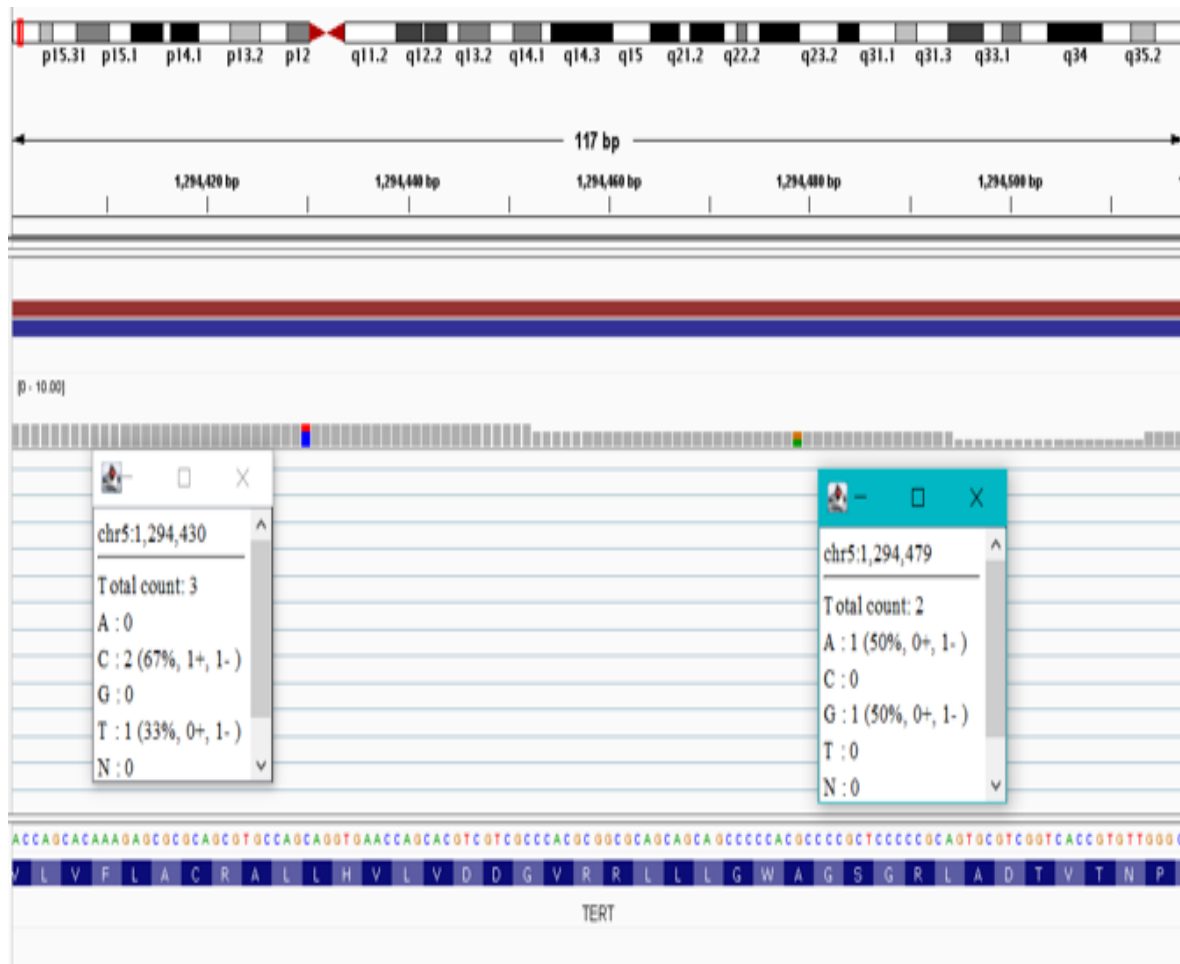


Figura 31. Mutaciones puntuales TERT Cromosoma 5: 1294430 C > T, 1294479 G > A en el Cáncer Diferenciado de Tiroides. Se muestra la pista de cobertura predeterminada para un archivo de alineación BAM utilizando el software IGV (Integrative Genomics Viewer) de la combinación de los 11 pacientes del estudio. Si un nucleótido difiere de la secuencia de referencia en más del 20 % de las lecturas ponderadas de calidad, IGV colorea la barra en proporción al recuento de lecturas de cada base (A, C, G, T). Se encontraron los polimorfismos de nucleótido único (SNPs) TERT Cromosoma 5: 1294430, SNV (variación en la secuencia de ADN que ocurre cuando se altera un solo nucleótido) C > T y 1294479 SNV G > A. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

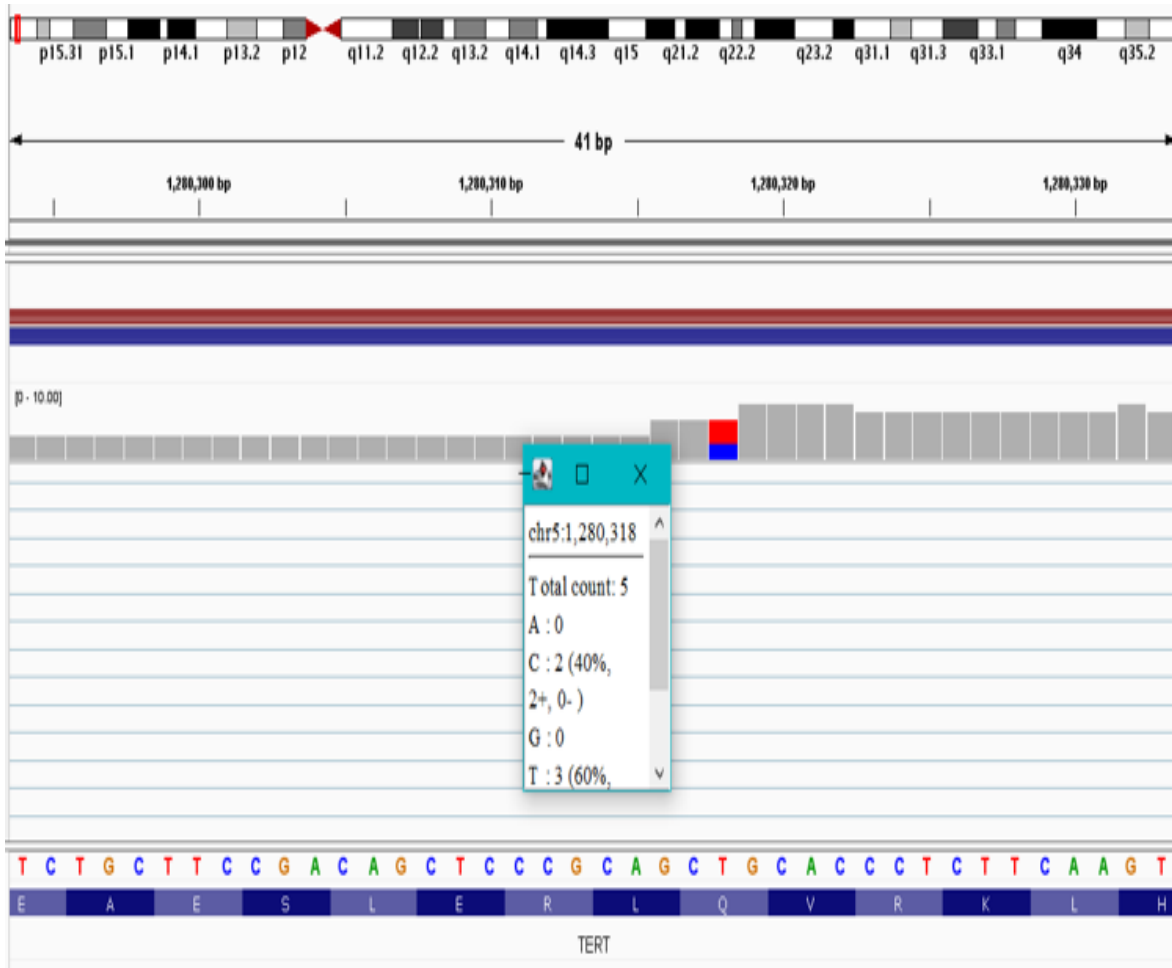


Figura 32. Mutacion puntual TERT Cromosoma 5: 1280318 T > C en el Cáncer Diferenciado de Tiroides. En la figura 40 se muestra la pista de cobertura predeterminada para un archivo de alineación BAM utilizando el software IGV (Integrative Genomics Viewer) de la combinación de los 11 pacientes del estudio. Si un nucleótido difiere de la secuencia de referencia en más del 20 % de las lecturas ponderadas de calidad, IGV colorea la barra en proporción al recuento de lecturas de cada base (A, C, G, T). Se encontró el polimorfismo de nucleótido único (SNP) TERT Cromosoma 5: 1280318, SNV (variación en la secuencia de ADN que ocurre cuando se altera un solo nucleótido) T > C. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

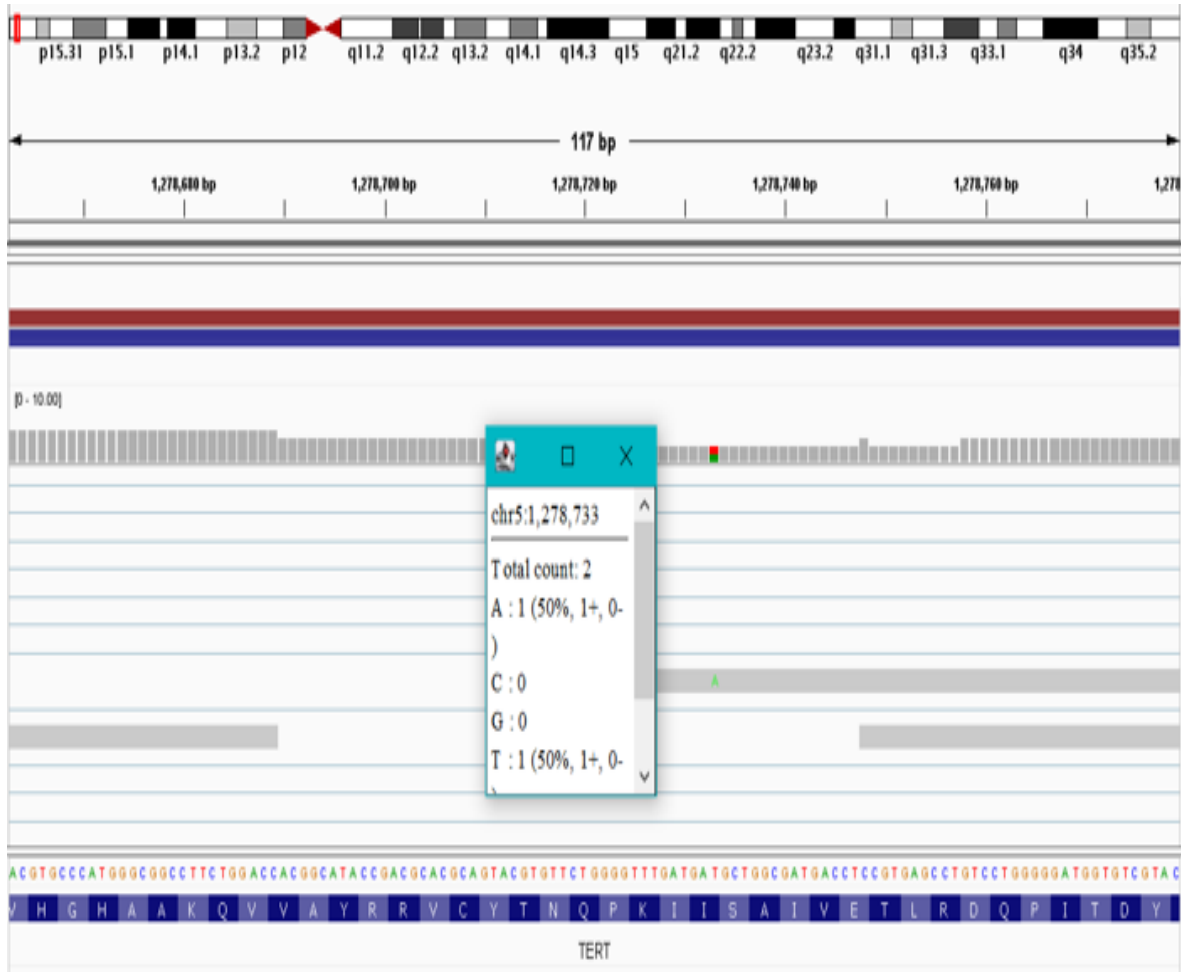


Figura 33. Mutacion puntual TERT Cromosoma 5: 1278733 T > A en el Cáncer Diferenciado de Tiroides. Se muestra la pista de cobertura predeterminada para un archivo de alineación BAM utilizando el software IGV (Integrative Genomics Viewer) de la combinación de los 11 pacientes del estudio. Si un nucleótido difiere de la secuencia de referencia en más del 20 % de las lecturas ponderadas de calidad, IGV colorea la barra en proporción al recuento de lecturas de cada base (A, C, G, T). Se encontró el polimorfismo de nucleótido único (SNP) TERT Cromosoma 5: 1278733, SNV (variación en la secuencia de ADN que ocurre cuando se altera un solo nucleótido) T > A. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

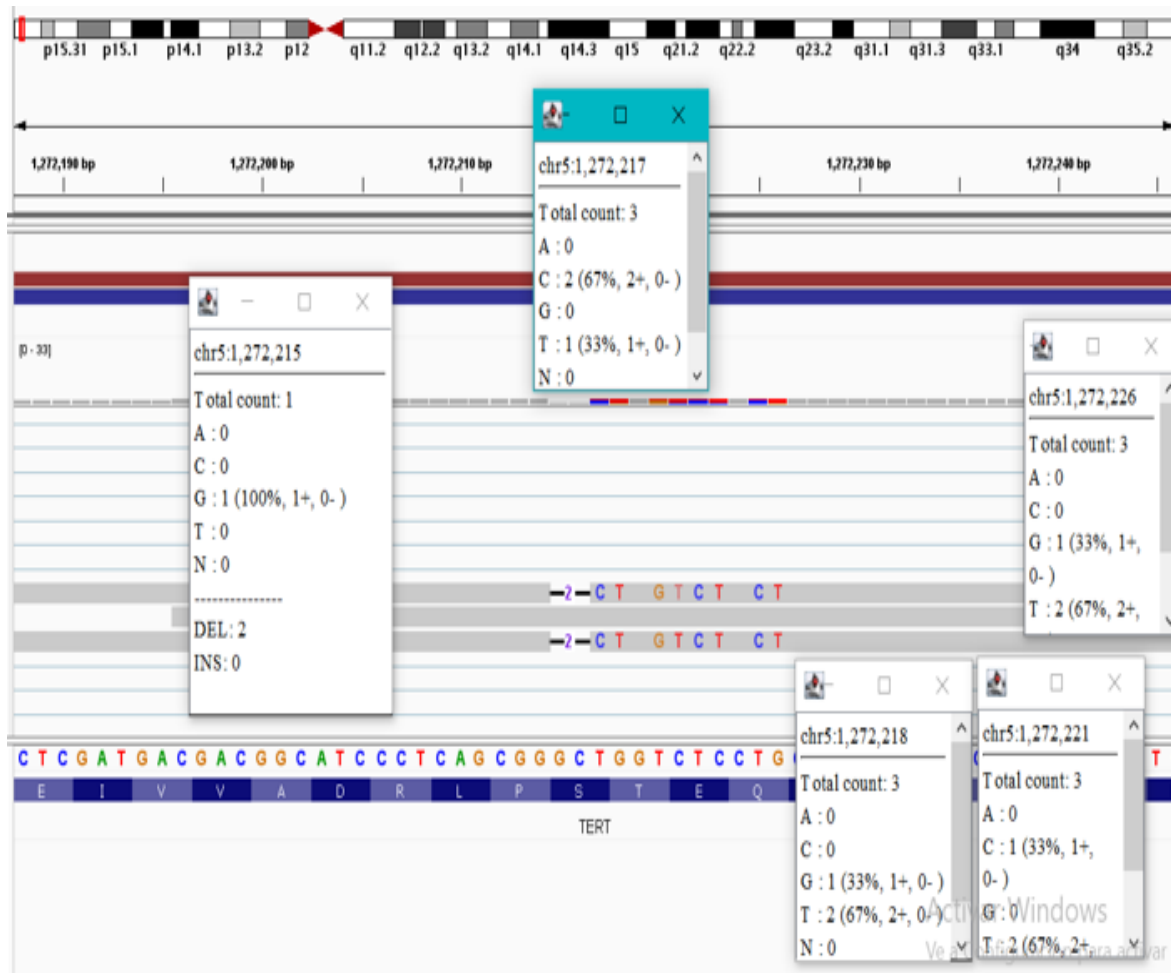


Figura 34. Mutaciones puntuales TERT Cromosoma 5: 1272217 T > C, 1272218 G > T, 1272221 T > C en el Cáncer Diferenciado de Tiroides. Se muestra la pista de cobertura predeterminada para un archivo de alineación BAM utilizando el software IGV (Integrative Genomics Viewer) de la combinación de los 11 pacientes del estudio. Si un nucleótido difiere de la secuencia de referencia en más del 20 % de las lecturas ponderadas de calidad, IGV colorea la barra en proporción al recuento de lecturas de cada base (A, C, G, T). Se encontraron los polimorfismos de nucleótido único (SNPs) TERT Cromosoma 5: 1272217, SNV (variación en la secuencia de ADN que ocurre cuando se altera un solo nucleótido) T > C, 1272218 SNV G > T y 1272221 SNV T > C. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

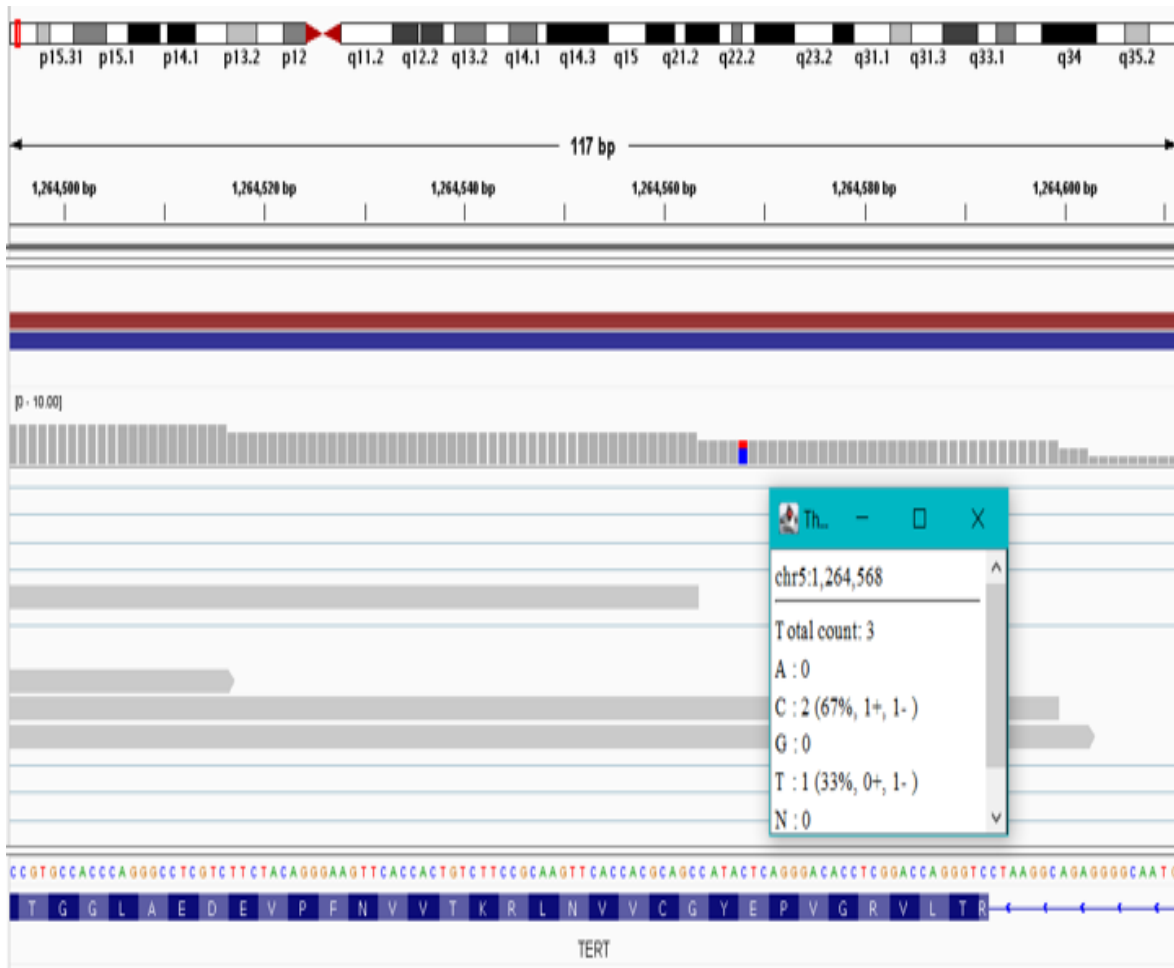


Figura 35. Mutacion puntual TERT Cromosoma 5: 1264568 T > C en el Cáncer Diferenciado de Tiroides. Se muestra la pista de cobertura predeterminada para un archivo de alineación BAM utilizando el software IGV (Integrative Genomics Viewer) de la combinación de los 11 pacientes del estudio. Si un nucleótido difiere de la secuencia de referencia en más del 20 % de las lecturas ponderadas de calidad, IGV colorea la barra en proporción al recuento de lecturas de cada base (A, C, G, T). Se encontró el polimorfismo de nucleótido único (SNP) TERT Cromosoma 5: 1264568, SNV (variación en la secuencia de ADN que ocurre cuando se altera un solo nucleótido) T > C. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

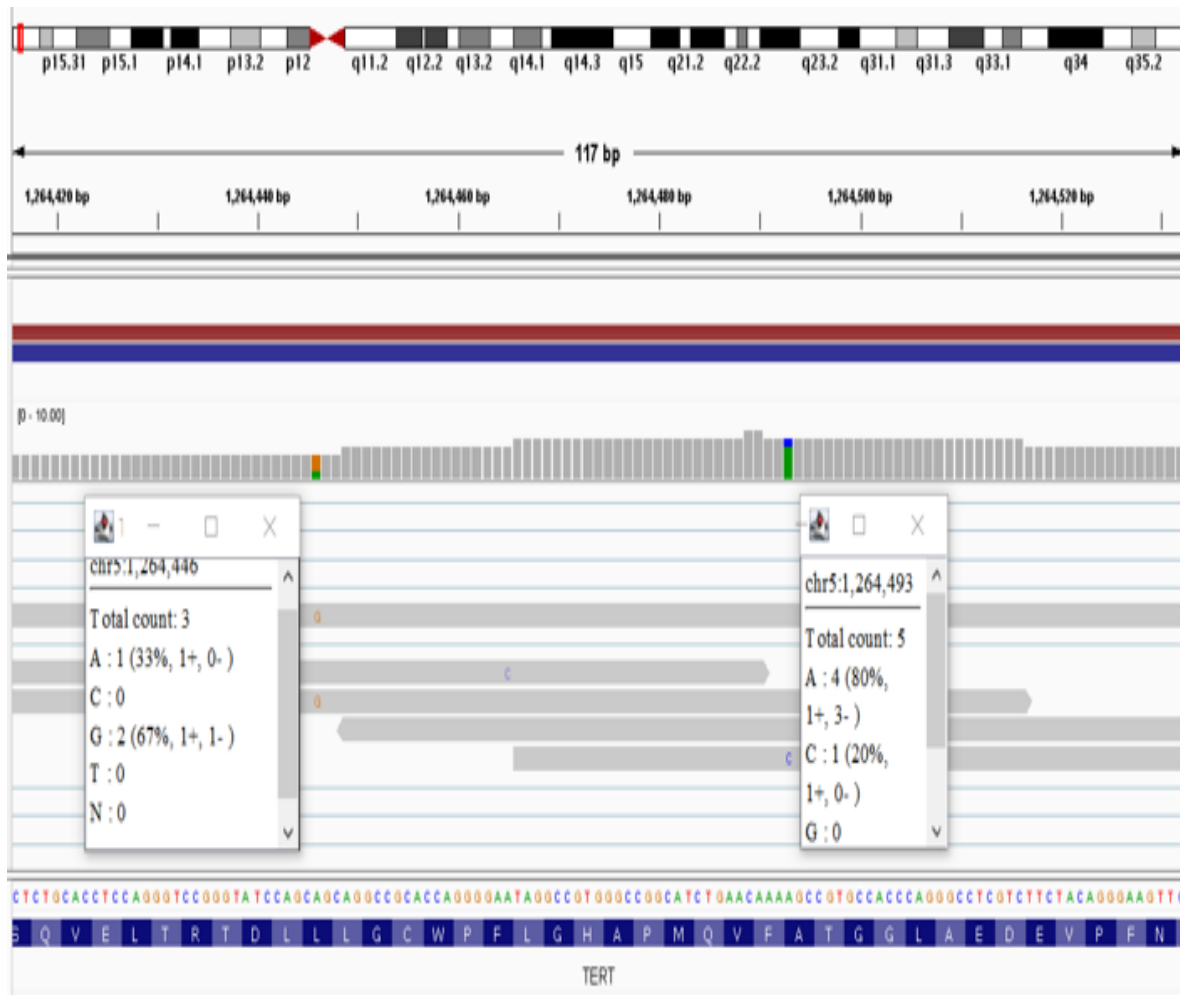


Figura 36. Mutaciones puntuales TERT Cromosoma 5: 1264446 A > G, 1264493 A > C en el Cáncer Diferenciado de Tiroides. Se muestra la pista de cobertura predeterminada para un archivo de alineación BAM utilizando el software IGV (Integrative Genomics Viewer) de la combinación de los 11 pacientes del estudio. Si un nucleótido difiere de la secuencia de referencia en más del 20 % de las lecturas ponderadas de calidad, IGV colorea la barra en proporción al recuento de lecturas de cada base (A, C, G, T). Se encontraron los polimorfismos de nucleótido único (SNPs) TERT Cromosoma 5: 1264446, SNV (variación en la secuencia de ADN que ocurre cuando se altera un solo nucleótido) A > G y 1264493 SNV A > C. Fuente propia.

PERFIL DE EXPRESIÓN GÉNICA POR RNA-SEQ DEL CÁNCER DIFERENCIADO DE TIROIDES DIAGNOSTICADO EN EL HNCASE ESSALUD, AREQUIPA 2019-2020

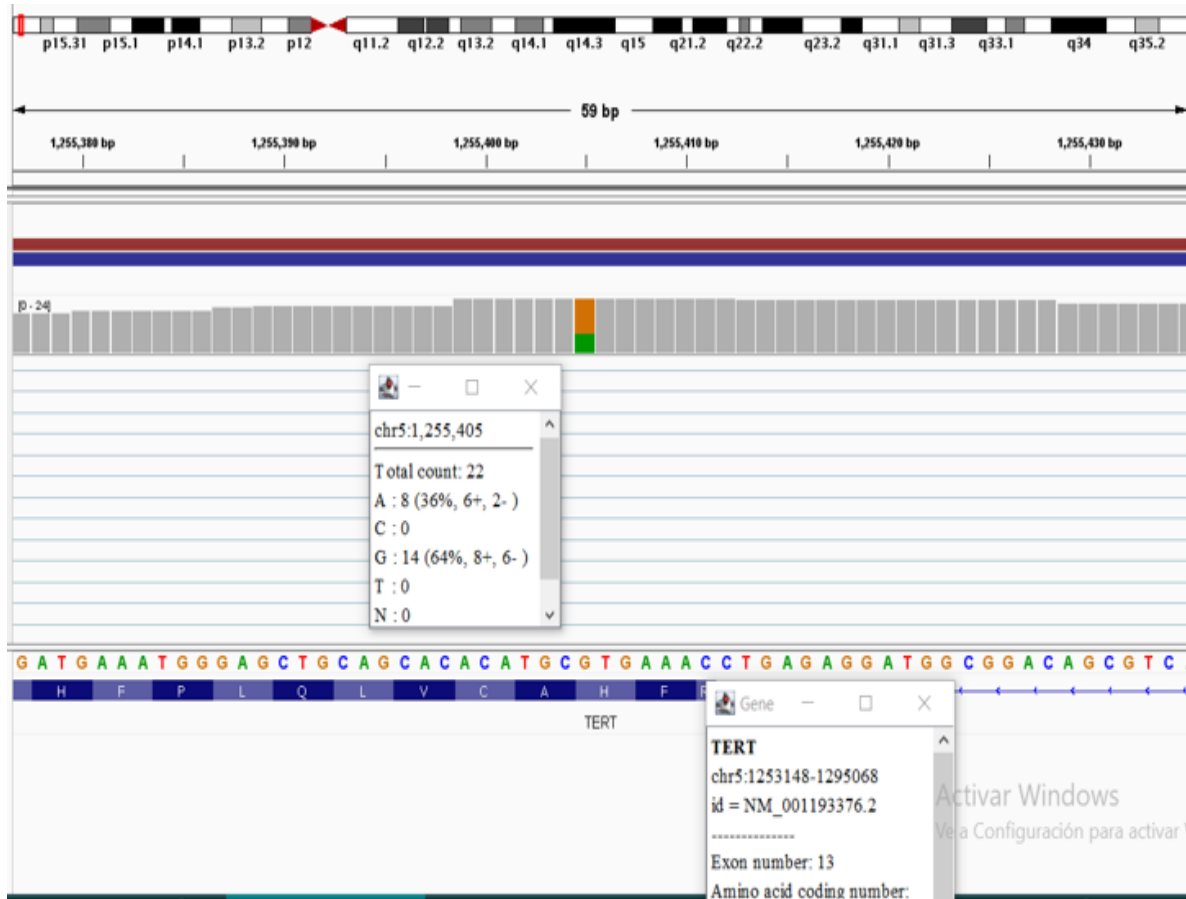


Figura 37. Mutacion puntual TERT Cromosoma 5: 1255405 G > A en el Cáncer Diferenciado de Tiroides. Se muestra la pista de cobertura predeterminada para un archivo de alineación BAM utilizando el software IGV (Integrative Genomics Viewer) de la combinación de los 11 pacientes del estudio. Si un nucleótido difiere de la secuencia de referencia en más del 20 % de las lecturas ponderadas de calidad, IGV colorea la barra en proporción al recuento de lecturas de cada base (A, C, G, T). Se encontró el polimorfismo de nucleótido único (SNP) TERT Cromosoma 5: 1255405, SNV (variación en la secuencia de ADN que ocurre cuando se altera un solo nucleótido) G > A. Fuente propia.

DISCUSIÓN Y COMENTARIOS

Se realizó el presente estudio con el objetivo de conocer el perfil de expresión génica por la técnica RNA-Seq de células tumorales de pacientes diagnosticados con cáncer diferenciado de tiroides en el HNCASE EsSalud Arequipa, 2019-2020.

Se llevó a cabo esta investigación debido al incremento de la incidencia de nuevos casos de cáncer diferenciado de tiroides a nivel mundial ⁽²⁹¹⁻²⁹⁴⁾ y en nuestro país ⁽²⁹⁵⁾, siendo una de las hipótesis, el incremento de la presencia de mutaciones, como la mutación BRAF^{V600E}, en la biología molecular de esta patología ^(296,297).

Así mismo, aunque la mortalidad de esta enfermedad no se ha visto incrementada ⁽²⁹¹⁻²⁹⁵⁾, se ha encontrado que la recurrencia, agresividad y mortalidad aumentan con la presencia y coexistencia de nuevas mutaciones del cáncer diferenciado de tiroides ^(298,299), por lo que, una de las estrategias para combatir esta enfermedad es la caracterización molecular individualizada tanto para el diagnóstico y terapias blanco ⁽³⁰⁰⁻³⁰²⁾ de los pacientes portadores de esta enfermedad, información que puede ser ofertada por la técnica RNA-Seq como lo demuestra la literatura ^(303, 304); lo que justifica la realización de esta investigación en pro de la búsqueda de un perfil de mutaciones autóctono del cáncer diferenciado de tiroides de los pacientes asegurados de la macro región sur del Perú que se atienden en el HNCASE Arequipa.

Para realizar el presente estudio, previa aprobación por el comité de ética de investigación del Hospital HNCASE EsSalud Arequipa, se revisaron las historias clínicas de los 11 casos de los pacientes operados de cáncer diferenciado de tiroides por el Servicio de Cirugía de Cabeza y Cuello en el periodo 2019-2020, cuyas muestras tumorales fueron obtenidas del banco de tejidos del Servicio de Anatomía Patológica, inmediatamente posterior a la cirugía,

previa corroboración y evaluación diagnóstica del anatomopatólogo, siendo trasladadas en los primeros 30 minutos en un tubo estéril eppendorf con hielo seco, almacenadas en cadena de frío al laboratorio de biología molecular de la Universidad Católica de Santa María Arequipa, donde fueron procesadas para aislamiento de ADN, y posteriormente enviadas al laboratorio NOVOGEN y Admera Health Biopharma Servicesen de USA para secuenciamiento mediante técnica RNA-Seq, además de una segunda opinión de la lectura bioinformática en el laboratorio del Instituto Pasteur de París, Francia; con los datos obtenidos se pudo recolectar la información en formato BAM, sobre el perfil de expresión génica y también mutacional con ayuda del softawre IGV. Así mismo, de la revisión de las historias clínicas se obtuvieron datos epidemiológicos, clínicos e histopatológicos.

En la **Figura 13** se aprecian las características de los pacientes con el diagnóstico de cáncer diferenciado de tiroides durante el periodo de estudio. El 63.64% fueron mujeres y 36.36% varones, con edad predominante entre los 40 y 49 años en 45,55% de casos.

Resultados que son equiparables a los descritos por U. Córdoba y T. Sáenz, en el capítulo de Epidemiología del Cáncer de Tiroides, del tratado “Avances y Perspectivas en el Cáncer Diferenciado de Tiroides” del 2018 ⁽³⁰⁵⁾, en el que mencionan una mayor frecuencia de presentación de esta enfermedad en el sexo femenino a razón de 6.1 mujeres por cada 1.9 varones, así como una mayor frecuencia entre los 45 y 54 años de edad en población de México y los EE.UU.

La **Figura 14** muestra las características citológicas, histopatológicas y sonográficas de los pacientes con el diagnóstico de Cáncer Difeenciado de Tiroides (CDT) durante el periodo de estudio. Se encontró una predominancia de la variante papilar en un 81.82%, frente a un 18.18% de variante folicular.

Resultados equiparables a los descritos por Mohammad Hossein Khosravi en su trabajo “Thyroid Cancers: Considerations, Classifications, and Managements” ⁽³⁰⁶⁾, donde nos muestra un 95% de frecuencias de CDT (papilar, folicular y células de Hürthle), distribuidos en un 70-90% con la variante papilar y la folicular en un 10-15%. También resaltar que se presentaron 2 casos de variantes histopatológicas agresivas del carcinoma papilar de tiroides, según la 4ta clasificación de la WHO, una de ellas la variante columnar, y la otra la variante Fascicitis like, ambas en un porcentaje del 9.09%; estos hallazgos discrepan con lo descrito en la literatura, por citar a Kennichi Kakudo y colaboradores, en su trabajo “The new 4th edition World Health Organization classification for thyroid tumors, Asian perspectives” ⁽³⁰⁷⁾, donde se narra una frecuencia de presentación de 0.15 a 0.2% para la variante Columnar y de 0.06% para la variante Fascitis like; esta discrepancia podría estar en relación a la escasa población de nuestro estudio, sin embargo, es de suma importancia resaltar que un 18.18% de nuestros pacientes, correspondió a variantes histopatológicas agresivas, lo que podría indicar una mayor frecuencia de presentación de estas variantes, en nuestra población y que debería de alertar y propugnar una mayor investigación al respecto.

Predominó la categoría Bethesda VI con un 63.64% frente a un 36.36% para Bethesda V, la escala TI-RADS 5 se presentó en el 100% de casos, el 72.73% fueron tumores unifocales, siendo bien diferenciados (G1) en el 81.82%. Se encontró tiroiditis crónica asociada al cáncer diferenciado de tiroides, en el 18.18% de casos.

Como características preoperatorias, la categoría Bethesda V y VI fueron las más prevalentes en nuestra investigación, esto en razón a que ambas, sumadas a otros hallazgos clínicos e imagenológicos indican el tratamiento quirúrgico del paciente con nódulo tiroideo, con corroboración histopatológica de CDT en los 11 casos, lo que demuestra altos niveles de sensibilidad y especificidad, tal como nos describe Ismael Mora-Guzmán y colaboradores en

su trabajo titulado “Efficiency of the Bethesda System for Thyroid Cytopathology”⁽³⁰⁸⁾, quienes encuentran valores de sensibilidad 98,6%, especificidad 97,6%, valor predictivo positivo 93,5%, valor predictivo negativo 99,5% y precisión diagnóstica global 97,9% para evaluación de malignidad. El Score preoperatorio ecográfico que se estudio fue el TI-RADS, en el que se muestra que en la totalidad de casos se diagnosticó categoría TI-RADS V, la cual corresponde a una alta predicción de cáncer, lo que se corroboró en los resultados histopatológicos, estos datos son similares a los descritos en la American Thyroid Association, donde Mendes GF y colaboradores en el trabajo “Fine needle aspiration biopsy of thyroid nodule smaller than 1.0 cm: accuracy of TIRADS classification system in more than 1000 nodules”⁽³⁰⁹⁾, encontraron presentación de cáncer de tiroides alrededor del 86% en nódulos con categoría TI-RADS V. En nuestra población se encontró que el 18.18% de los casos de CDT se asociaron a Tiroiditis crónica de Hashimoto, resultados similares a los que reportan Molnar y colaboradores en su trabajo Thyroid Carcinoma Coexisting with Hashimoto's Thyroiditis: Clinicopathological and Molecular Characteristics Clue up Pathogenesis⁽³¹⁰⁾, resaltando que esta condición se encontró un 16.4%.

El estadiaje anatomopatológico del CDT durante el periodo de estudio correspondió a los estadios I y II, con un 72.73% y 27.27% respectivamente.

En lo que se refiere al estadiaje anatomopatológico, la totalidad de los casos fueron diagnosticados en estadios clínicos tempranos y de buen pronóstico (grado I y grado II), esto en relación al incremento del diagnóstico precoz de esta enfermedad con el uso y accesibilidad a la ecografía tiroidea, tal como lo mencionado la literatura, como Benjamín Schmidt and Louise Davies en el capítulo 1, The Rising Incidence of Thyroid Cancer: Contributions from Healthcare Practice and Biologic Risk Factors, del libro “Management of Differentiated Thyroid Cancer”⁽³¹¹⁾, quienes comunican el sobre diagnóstico en relación

a detección preclínica por estudios de imagen en los sistemas de salud como la ecografía, lo que conduce a tratamientos oportunos y con mayor tasa de curación.

La **Figura 15 (parte 1-3)**, muestra los 76 genes con mayor frecuencia de expresión, comunes a los 11 casos del CDT de la población en estudio. En las **Tablas 7 y 8**, se exponen los 30 primeros de estos genes anteriormente mencionados, por criterio jerárquico, donde se detallan datos como su denominación, simbología, ubicación cromosómica, y función, encontrando que todos ellos codifican de proteínas; además se describen algunas patologías asociadas a las alteraciones de estos genes.

En el presente estudio el método de secuenciación de RNA Seq NGS, nos ha permitido identificar los genes con mayor frecuencia de transcripción, comunes y repetitivos en los 11 pacientes estudiados; siendo los 30 primeros genes con mayor frecuencia de expresión el PCB2, LRMDA, WWOX, RPS29, RBFOX1, FHIT, NTM, RAD51B, PRKN2, MACROD2, ASIC2, TBC1D5, GPC5, EXOC4, IMMP2L, CAMTA1, CDH13, SMYD3, ANKS1B, DAB1, DLG2, FRMDA4A, KCNIP4, CTNNA2. AGLB4, PRKCE, SYN3, ARHGAP15, PACRG, C9orf92; todos ellos ya descritos en la literatura, expresados en tejido tiroideo sano ^(314, 315) y cuyas alteraciones patogénicas participan en la oncogénesis del CDT ⁽³¹²⁻³¹⁵⁾, como lo menciona Jaime Miguel Pita y colaboradores en su trabajo “Cell cycle deregulation and TP53 and RAS mutations are major events in poorly differentiated and undifferentiated thyroid carcinomas” ⁽³¹⁴⁾, quienes también describen la participación de algunas de las mutaciones de estos genes descritos tales como “RAS, BRAF, TP53, CTNNB1 (β -catenina), PIK3CA, CDKN1A, CDKN1B, CDKN2A, CDKN2B, CDKN2C, AXIN1, PTEN”, en las variantes pobremente diferenciadas e indiferenciadas del cáncer de tiroides ⁽³¹⁶⁻³¹⁸⁾, información de mucha importancia puesto que nos da a entender las bases moleculares comunes que pueden tener neoplasias con distinta variedad y agresividad histopatológica, y

que se podrían idear targets terapéuticos precozmente en el tratamiento de esta neoplasia ⁽³¹⁹⁾. Muchos de estos genes descritos se relacionan con factores de transcripción, proteínas supresoras tumorales, inductores de apoptosis, proteínas del citoesqueleto, proteínas de adhesión celular, vías de proteína quinasa C, reparadores del DNA, proteínas chaperonas y canales de voltaje iónicos de la membrana celular ^(312, 315), cuyas alteraciones patológicas como mutaciones son mecanismos que han sido descritos en la carcinogénesis tiroidea y de otras neoplasias ⁽³²¹⁻³²⁸⁾. Llama la atención haber encontrado sobreexpresión de genes de proteínas relacionadas al desarrollo neuronal vinculadas al cáncer de tiroides como NTM ^(314, 315), ANKS1B ^(314, 315), DAB1 ^(314, 315), SYN3 ⁽³¹⁵⁾; este hallazgo podría estar relacionado al desarrollo embrionario de la glándula tiroidea que conduce a la presencia de células derivadas del neuroectodermo como las “células C” o parafoliculares ^(329, 330), y que ya se han venido reportando en la literatura la coexistencia de tumores tiroideos mixtos con presencia de células foliculares tiroideas y parafoliculares C ⁽³³¹⁻³³³⁾, y que en estudios transgénicos se han inducido hiperplasia de células C y tumores de células C en ratones ⁽³¹⁵⁾; uno de estos estudios transgénicos mencionados utilizó la isoforma larga de RET (con una mutación 634) ligada a un promotor de calcitonina ⁽³³⁴⁾, recordando que RET es un proto-oncogén involucrado en la génesis del cáncer medular de tiroides derivado de las células C. Curiosamente, los folículos anormales que contienen tiroglobulina que se cree que son de origen ultimo branquial ocurrieron en dos líneas fundadoras, acompañadas en una de tumores, solo carcinomas papilares en la mayoría de los ratones, pero una minoría mostró tumores tanto papilares como medulares ^(331, 334). Estos hallazgos permiten abrir un abanico de posibilidades, que ameritan una continua investigación en la capacidad totipotencial de células madres de múltiples tejidos.

En la **Figura 16** se muestra la expresión de TERT, BRAF y P53 en el Cáncer Diferenciado de Tiroides. Se puede objetivar la presencia de la expresión del gen TERT, BRAF y P53, así

como la coexistencia de TERT y BRAF en el 100% de los casos de pacientes portadores de cáncer diferenciado de tiroides.

La transcriptasa inversa de la telomerasa humana (TERT) es un gen importante involucrado en el mantenimiento de la integridad cromosómica y la estabilidad del genoma y codifica la subunidad catalítica de la transcriptasa inversa de la enzima telomerasa ⁽³³⁵⁾. En un reciente metaanálisis Jing Yang y colaboradores, en su trabajo “Association between TERT promoter mutations and clinical behaviors in differentiated thyroid carcinoma: a systematic review and meta-analysis”, demostraron que las frecuencias promedio de mutaciones del promotor TERT en carcinomas diferenciados de tiroides (CDT), papilar (CPT) y folicular (CFT) fueron 10,9%, 10,6% y 15,1%, respectivamente ⁽³³⁵⁾. Otras referencias actuales reportan una prevalencia entre el 7.5 y 25% ⁽³³⁶⁻³³⁹⁾. Estas cifras se ven incrementadas en las variantes de carcinoma de células de Hürthle (CCH) y las variantes pobremente diferenciado e indiferenciado o anaplásico (CTPD/CTA) con valores de 25% y 86% respectivamente ^(340, 341). Datos que no correlacionan con los encontrados en nuestro estudio, donde encontramos que la totalidad de pacientes expresaban TERT y tenían mutaciones de TERT, esto probablemente por lo poca cantidad de pacientes; sin embargo es importante recalcar que la gran mayoría de trabajos calcula los valores de prevalencia de esta mutación pesquizando 2 variantes de TERT desde que estas fueron reportadas por Liu y colaboradores en su trabajo “Highly prevalent TERT promoter mutations in aggressive thyroid cancers” ⁽³³⁶⁾ en el 2013, las que corresponden a chr5: 1, 295228 C > T (C228T) y chr5: 1, 295250 C > T (C250T), que representan cambios de nucleótidos de -124 C > T y -146 C > T del codón de inicio de la traducción de ATG en el sitio del promotor del gen TERT ⁽³⁴²⁾, mutaciones que no se han encontrado en nuestro estudio, lo que podría conducir a subestimar la verdadera prevalencia de otras mutaciones de TERT por falta de la búsqueda de otras variantes mutacionales. Además, se ha encontrado asociación significativa entre las mutaciones del promotor TERT

y la edad de los pacientes adultos mayores, según lo reportado por Liu T. y colaboradores en su trabajo The age- and shorter telomere-dependent TERT promoter mutation in follicular thyroid cell-derived carcinomas ⁽³⁴³⁾, que describen diferencias de edad de $59,2 \pm 15,5$ frente a $44,9 \pm 15,6$ años en pacientes con mutación TERT positivos frente a negativos ($P < 0,001$) ⁽³⁴³⁾. Curiosamente, el sexo masculino de los pacientes también se asocia con mutaciones del promotor TERT, siendo el 37,4% masculino en pacientes con mutación positiva frente al 24,8% en pacientes con mutación negativa, $P < 0,001$ ⁽³⁴³⁾. Estos hallazgos correlacionan con nuestro estudio donde el 90.90% de pacientes en ambos sexos se encontró entre los 40 y 79 años de edad, con predominio de la edad en el 75% de varones entre los 60 y 79 años.

El gen BRAF, esta ubicado en el cromosoma 7, es el gen mutado con mayor frecuencia en los cánceres de tiroides, lo que resulta en una potente activación de la vía MAPK (proteína quinasa activada por mitógenos) ⁽³⁴⁴⁾. El hotspot mutacional más común en BRAF es T1799A en el exón 15 (chr7: 140453136 A>T), lo que confiere un glutamato a la valina como sustitución en el aminoácido 600 (V600E) en la proteína BRAF. BRAF^{V600E} es la alteración genética más común en CPT, presenta una alta prevalencia en el CPT clásico y la variante de células altas, aunque generalmente es raro en la variante folicular ⁽³⁴⁴⁾. La prevalencia general de mutaciones BRAF se encuentra en alrededor del 45% (rango del 27,3% al 87,1%) ^(344, 345), con una prevalencia significativamente mayor en Asia, especialmente Corea, en relación con los países occidentales ⁽³⁴⁶⁻³⁴⁸⁾. Resultados muy cercanos a los encontrados en nuestro estudio, donde el 100% de pacientes presentaban la expresión y mutación de BRAF, en relación al gran porcentaje de la variante histopatológica papilar y de células altas de nuestra población. Así mismo, esta gran prevalencia encontrada podría corresponder a la escasa población de estudio; sin embargo, es importante mencionar que la gran mayoría de estudios pesquisan la mutación BRAF^{V600E}, lo que podría subestimar la real prevalencia de la mutación de BRAF al no enrolar otras variantes mutacionales. La literatura también

menciona incremento de mutaciones de la vía de MAPK, donde se incluye a BRAF por estímulo de radiación ultravioleta (UV), en países con elevados índices de exposición solar como Colombia en estudios de Melanoma ^(349, 350), recordando que nuestra de ciudad de Arequipa, tenemos unas de las mayores dosis de radiación UV, lo que sería motivo de investigación en relación a la presencia de BRAF en todos los casos del estudio. De la misma forma que ocurre con TERT, se ha descrito mayor prevalencia de la mutación BRAF y la edad, por citar a Changjiao Yan. y colaboradores en su investigación “Relationship between *BRAF* V600E and clinical features in papillary thyroid carcinoma” ⁽³⁵¹⁾, quienes encontraron diferencia significativa entre grupos de edades de 39.29 ± 11.44 versus 43.88 ± 10.77 con un valor p de 0.002, mas no encontraron diferencias entre sexos con un valor p de 0.195 ⁽³⁵¹⁾, datos que correlacionan con nuestro estudio donde el 90.90% de pacientes en ambos sexos se encontró entre los 40 y 79 años de edad.

Se describe la coexistencia de la expresión de TERT y BRAF en el 100% de los casos de pacientes portadores de cáncer diferenciado de tiroides, así mismo se muestra la relación entre el estadiaje clínico y estas mutaciones, donde el 72.73% de casos corresponde al estadiaje clínico I y el 27.27% restante al estadiaje clínico II, la totalidad de casos independiente del estadiaje clínico presentó la coexistencia de TERT y BRAF.

En nuestro trabajo se encontró que todos los pacientes presentaron expresión y mutaciones de TERT y BRAF simultáneamente, esta coexistencia se ha descrito en la literatura por Haoyu Ren y colaboradores en su trabajo “Co-existence of *BRAF*^{V600E} and *TERT* promoter mutations in papillary thyroid carcinoma is associated with tumor aggressiveness, but not with lymph node metastasis” ⁽³⁵²⁾, concluyendo que esta asociación predispone a una mayor agresividad clínico-patológica del cáncer papilar de tiroides, como la edad avanzada, tamaño tumoral más grande, extensión extratiroidea y estadio tumoral avanzado, pero no con

metástasis en los ganglios linfáticos centrales, metástasis en los ganglios linfáticos laterales, número de metástasis en los ganglios linfáticos > 5 y número de ganglios linfáticos afectados/extraídos ⁽³⁵²⁾. Otros metaanálisis corroboran la mayor relación entre esta asociación y la mayor agresividad de presentación de esta neoplasia ⁽³⁵³⁻³⁵⁸⁾. Estos resultados de la literatura discrepan parcialmente de los encontrados en nuestro trabajo, los cuales probablemente se relacionen a la escasa cantidad de pacientes; sin embargo, recalcar nuevamente, que las revisiones científicas antes descritas investigan la coexistencia de las mutaciones TERT C228T y C250T, con la BRAF^{V600E}, las que no fueron encontradas en nuestro estudio), lo cual puede conllevar a subestimar la coexistencia de otras variantes mutacionales de TERT y BRAF en el cáncer diferenciado de tiroides; además el 100% de nuestra población de estudio estuvo constituida por pacientes en estadios clínicos I y II, lo que conlleva a lesiones tumorales de pequeño tamaño y sin extensión extratiroidea, lo que podría estar en concordancia con la presencia de otras variantes mutacionales de TERT y BRAF no agresivas.

Se encontró que P53 se expresó en el 100% de los casos de pacientes portadores de cáncer diferenciado de tiroides.

Está bien documentado que el inicio y la progresión del carcinoma de tiroides se produce a través de la acumulación gradual de múltiples alteraciones genéticas ⁽³⁵⁹⁾. Una de las alteraciones moleculares fundamentales que discriminan al cáncer de tiroides anaplásico (CTA) de los cánceres diferenciados de tiroides (CDT) es la inactivación del gen supresor de tumores P53 ⁽³⁵⁹⁾. Las mutaciones de P53 son frecuentes en los tumores tiroideos indiferenciados (50-80% en los CTA) ^(360, 361). Sin embargo, varios hallazgos indican que las alteraciones en la secuencia de P53 también pueden desempeñar un papel en las primeras etapas de la cancerogénesis tiroidea ⁽³⁵⁹⁾. De hecho, recientemente se han encontrado

mutaciones de P53 en hasta el 40% de los CPT y en el 22% de los cánceres foliculares de tiroides (CFT) oncocíticos ^(362, 363). Por lo general, las mutaciones puntuales de P53 se localizan en la región entre los exones 5 y 8 ⁽³⁶⁴⁾. Otras referencias bibliográficas informan que se producían mutaciones del gen P53 en el 40-62% de los carcinomas indiferenciados de tiroides y en el 0-25% de los carcinomas bien diferenciados de tiroides ^(365, 366). Algunos autores sugieren que la sobreexpresión de p53, determinada por inmunohistoquímica, se atribuía a la mutación de un gen p53 en hasta el 95% de los casos del cáncer papilar de tiroides ⁽³⁶⁷⁾. Estos datos concordantes con los encontrados en nuestra investigación, donde en estadios clínicos iniciales de la enfermedad como el I y II de nuestra población expresan P53 y tienen mutaciones de P53, en la totalidad de los casos. En lo que respecta a la coexistencia de P53 con TERT y/o BRAF no se han observado correlaciones significativas entre la mutación BRAF^{V600E} o la sobreexpresión de la proteína p53 y las características clínico-patológicas de los pacientes portadores de CDT ⁽³⁶⁸⁾, así mismo es de destacar que un carcinoma papilar que alberga una mutación del promotor TERT tiene un mayor riesgo de transformación anaplásica independiente de la presencia de P53 u otras mutaciones ⁽³⁶⁹⁾.

En las **Tabla 9 (parte 1 y 2)** se presentará la expresión de P53 y otros genes relacionados a P53, junto a su ubicación cromosómica, simbología y funciones; encontrados en los 11 pacientes con cáncer diferenciado de tiroides del presente estudio.

El método de secuenciación de RNA Seq NGS, nos permitió identificar la expresión de P53 y otras asociadas a esta, encontrando 8 genes ejemplares comunes en los 11 pacientes del estudio, que son WRAP53, TP53BP2, TP53BP1, TP53I11, NOP53, TP53, USP53, CFAP53; dentro de las funciones se encuentran la de *telomerasa*, inducción de la *apoptosis*, *reparación de las cadenas de DNA*, activación de las *vías AKT* (grupo de enzimas de la familia serina-treonina que participan en varios procesos relacionados con el crecimiento y la supervivencia

celular, ayudan a transferir las señales en el interior de las células, también se llama proteína cinasa B) ⁽³⁷⁰⁾ y *AMPK* (enzima con actividad de varios procesos fisiológicos esenciales, incluida la síntesis de proteínas, la apoptosis y la autofagia) ⁽³⁷¹⁾, *activación de P53* tras daños de material ribosomal o del DNA, *detención de ciclo celular* y apoptosis, *inhibición de la glicolisis* con activación de apoptosis y alteración de los cilios y *fibroblastos*, respectivamente; las mutaciones de las mismas originan alteraciones de las proteínas que cumplen las funciones antes señaladas, y que al alterarse su funcionamiento son parte de la oncogénesis de múltiples neoplasias ⁽³⁷²⁻³⁸¹⁾, como la del CDT ^(314, 315).

En las **Tablas 10 y 11** se muestra la presencia de las traslocaciones y sus genes, componentes de cada una, junto a su ubicación cromosómica, simbología y funciones; de 03 pacientes con cáncer diferenciado de tiroides del presente estudio (01 caso portador de CDT variante papilar folicular y 02 casos con CDT variante papilar clásico), así mismo las **Figuras 17, 18 y 19** nos representan esquemática a estas traslocaciones.

Nuevamente el método de secuenciación de RNA Seq NGS, nos permitió identificar traslocaciones de genes en el presente estudio, encontrando 3 fusiones que son TDG - TMEM132B, MEIS1-ITSN2, KCNE1B - FP671120.4; se ha descrito que se requiere la proteína TDG para la reparación y desmetilación del ADN, desempeñando un papel en la prevención de tumores, actuando como un coactivador de P53 ^(382,383), que se une al promotor de TDG y regula la transcripción de TDG ⁽³⁸⁴⁾, por lo tanto se comporta como un supresor tumoral ⁽³⁸⁵⁾, relacionándolo con neoplasias como cáncer colorrectal ⁽³⁸⁶⁾, mieloma múltiple ⁽³⁸⁷⁾, adenocarcinoma pancreático ⁽³⁸⁸⁾, cáncer de piel no melanoma ⁽³⁸⁹⁾, carcinoma gástrico ⁽³⁹⁰⁾, carcinoma esofágico ⁽³⁹¹⁾ y melanoma ⁽³⁹²⁾, recientemente Dan Li y colaboradores reportan en su trabajo The *HOTAIRM1/miR-107/TDG* axis regulates papillary thyroid cancer cell proliferation and invasión, que la activación de TDG por la sobre expresión del ARN

mieloide 1 intergénico antisentido HOX (HOTAIRM1), el cual es una ARN de cadena larga no codificante, que actúa potenciando la actividad supresora tumoral, cuya disminución de este eje juega un rol importante en la proliferación e invasión del CPT ⁽³⁹²⁾, información que podría correlacionarse con la encontrada en nuestro estudio donde la traslocación antisentido ⁽³⁹³⁾ TDG-TMEM132B, puede originar ARNm que codifiquen proteína TDG alterada y ello favorecer la oncogénesis ^(314, 315). En lo que respecta a TMEM132B, pertenece al grupo de la familia de TMEM132 las que son moléculas de adhesión de la superfamilia de dominios de inmunoglobulina (IG) expresadas en el sistema nervioso central (SNC) que permiten una conexión clave putativa entre la matriz extracelular y el citoesqueleto celular basado en actina, con funciones importantes en la regulación de cambios en la morfología de las células neuronales. motilidad y migración ⁽³⁹⁴⁾, siendo relacionada a neoplasias como el cáncer de mama donde se produce hipermetilación del gen TMEM132C perdiendo la función de supresión tumoral ⁽³⁹⁵⁾, el cáncer de ovario donde TMEM132D donde se describe que la asociación de la expresión de este gen con la infiltración de linfocitos T CD8 otorga un pronóstico favorable y mayor supervivencia en el cáncer de ovario en estadio temprano ⁽³⁹⁶⁾, así mismo Weibo Xu y colaboradores en su investigación Identification of Key Functional Gene Signatures Indicative of Dedifferentiation in Papillary Thyroid Cancer informan la asociación de la expresión de TMEM132B y el cáncer papilar de tiroides (CPT) ⁽³⁹⁷⁾, información que podría correlacionar con nuestros hallazgos, ya que la traslocación antisentido TDG-TMEM132B, podría ocasionar la pérdida de función de TMEM132B y favorecer la oncogénesis ^(314, 315). Estudios recientes han demostrado que MEIS1 funciona para mantener la diferenciación celular en tejidos adultos ⁽³⁹⁸⁾, las proteínas MEIS (MEIS 1 MEIS 2, MEIS 3) se han asociado históricamente con tumorigénesis, metástasis e invasión en el cáncer ⁽³⁹⁹⁾, así mismo MEIS y las proteínas PBX-HOX asociadas pueden actuar como supresores de tumores u oncogenes en diferentes entornos celulares ⁽³⁹⁹⁾ y curiosamente, la

sobreexpresión de proteínas MEIS puede causar apoptosis dependiente de caspasa en algunas células ⁽³⁹⁹⁾, por lo que se le han encontrado asociadas a neoplasias tanto sólidas como hematológicas como cáncer de vejiga (MEIS 1-2) ^(399, 400), mama (MEIS 1) ^(399, 401), esófago, gástrico y colorrectal (MEIS1) ^(399, 402), glioblastoma, glioma y neuroblastoma (MEIS 1-3) ^(399, 403), renal (MEIS 1) ^(339, 404), leucemia y linfoma (MEIS 1) ^(339, 405), pulmón (MEIS 1) ^(339, 406), de piel como melanoma (MEIS 1-2) ^(339, 407), cavidad oral (MEIS 1) ^(339, 408), ginecológico (MEIS 1) ^(339, 407), pancreático (MEIS 1-2) ^(339, 408), próstata (MEIS 1-2) ^(339, 409), sarcoma (MEIS 1) ^(339, 410), timoma (MEIS 1-3) ^(339, 411) y cáncer de tiroides donde MEIS2 tiene participación en la tumorigénesis tiroidea (carcinoma papilar, variante oncocítica del carcinoma folicular, tumores de potencial maligno incierto, adenoma folicular y carcinoma anaplásico) ^(339, 412-415), que según Vriens y sus colaboradores en su trabajo Clinical and molecular features of papillary thyroid cancer in adolescents and young adults, indican que MEIS 2 está regulado positivamente en la oncogénesis del cáncer de tiroides ⁽⁴¹⁶⁾, estos resultados similares a los de nuestro estudio en los que la fusión de MEIS1-ITSN2, preserva los dominios de codificación de proteínas, preservando su función la que podría estar relacionada a la oncogénesis de esta neoplasia. ITSN1 contribuye a las propiedades tumorigénicas de varios cánceres humanos, incluidos el neuroblastoma, glioblastoma ^(417, 418), carcinomas de páncreas, liposarcomas y tumor de Wilm ⁽⁴¹⁹⁾, lo que sugiere un papel más amplio de este andamio en los cánceres humanos ⁽⁴²⁰⁾; ITSN2 también se ha relacionado con el cáncer; sin embargo, a diferencia de ITSN1, los niveles elevados de ITSN2 en pacientes con cáncer de mama se asocian con una supervivencia libre de enfermedad más prolongada de los pacientes ⁽⁴²¹⁾, lo que sugiere que ITSN2 puede desempeñar un papel protector o supresor de tumores en la tumorigénesis ⁽⁴²⁰⁾, Federica Panebianco y colaboradores han descrito en su trabajo Characterization of Thyroid Cancer Driven by Known and Novel ALK Fusions ⁽⁴²²⁾ han descrito por primera vez tres fusiones de las cuales una de ellas

corresponde a ALK-ITSN2 (ALK: kinasa de linfoma anaplásico), encontrada en el carcinoma papilar de tiroides y que les permitió orientar el diagnóstico entre un resultado incierto y de neoplasia tiroides por biopsia aspiración con aguja fina (BAAF), así mismo en nuestro trabajo encontramos la fusión MEIS1-ITSN2 no descrita hasta el momento, que podría también ser de ayuda en estos casos de biopsias inciertas. KCNE1 (originalmente llamada Mink) pertenece a una familia de cinco proteínas, incluida KCNE2 (originalmente llamada MiRP1), que, como KCNE1, puede regular a KCNQ1 y otras subunidades α como hERG, dotando a menudo de propiedades funcionales únicas ⁽⁴²³⁻⁴²⁵⁾. KCNQ1 es único entre las subunidades α del canal K⁺ controlado por voltaje en que, por ensamblaje con KCNE2 o KCNE3, puede formar constitutivamente canales activos de "fuga" de K⁺ ⁽⁴²⁶⁾. Las subunidades del canal de potasio KCNQ1 y KCNE2 forman un canal de K⁺ tiroideo estimulado por la hormona estimulante de la tiroides, constitutivamente activo, necesario para la biosíntesis normal de la hormona tiroidea ⁽⁴²⁶⁾. Mutaciones heredadas en los genes que codifican KCNQ1, hERG, KCNE1 y KCNE2 son todos asociados con arritmias cardíacas potencialmente mortales, que incluyen Síndrome QT ⁽⁴²⁷⁻⁴²⁹⁾ y puede tener un papel en la fibrilación auricular ⁽⁴³⁰⁻⁴³³⁾. KCNE se ha relacionado a múltiples neoplasias como el cáncer de útero, gástrico, astroglioma, leucemia mielógena ⁽³⁸⁵⁾; Nuria Comes y colaboradores en su investigación *Involvement of potassium channels in the progression of cancer to a more malignant phenotype*, describen la participación de estos canales en el control del potencial de membrana y la excitabilidad celular además de diferentes funciones celulares como regulación del volumen celular, proliferación, migración celular, angiogénesis así como apoptosis ⁽⁴³⁴⁾, detectando complejos hERG-KCNE en ambos escamosos y líneas celulares de cáncer de endometrio ⁽⁴³⁵⁾, así mismo Aline Hébrant y colaboradores describen en su trabajo de investigación *mRNA Expression in Papillary and Anaplastic Thyroid Carcinoma: Molecular Anatomy of a Killing Switch* ⁽⁴³⁶⁾, describen la ausencia total de ARNm de TSHR

(receptor de TSH), Tg (tiroglobulina) y TPO (tiroperoxidasa) en CTA (cáncer de tiroides anaplásico) ^(436, 437), lo que se corresponde bien con la ausencia total de Tg en el suero de los pacientes con CTA y la falta de respuesta de CTA a la TSH, una característica que diferencia a la CAT del CPT (cáncer papilar de tiroides), relatan además dentro de otros genes, la baja expresión de KCNQ1 como participe de la desdiferenciación del CPT a un CTA ^(436, 438), datos que podrían equiparse a nuestros resultados, donde encontramos que la fusión de KCNE1B - FP671120.4, da lugar a una transcripción sin sentido, perdiendo la zona de codificación de las proteínas de ambos genes. En lo que respecta a FP671120.4, pertenece al grupo de ARN no codificantes (ncRNA), los que se definen como transcripciones de nucleótidos con poca o sin capacidad de codificación de proteínas, que se han asociado con varios tumores malignos, que su desregulación puede conducir a cambios en la función de los oncogenes o supresores de tumores, mostrando su papel en la tumorigénesis y que su regulación de la expresión génica puede ocurrir en diferentes pasos, en etapas epigenética, transcripcional, postranscripcional y otras ⁽⁴³⁹⁾, así Anna Franca Cavaliere y colaboradores, en su trabajo Towards Personalized Medicine: Non-Coding RNAs and Endometrial Cancer ⁽⁴⁴⁰⁾, indican que la sobre expresión de FP671120.4 se comporta como un oncogén y se asocia como un factor de mal pronóstico en el cáncer de endometrio ⁽⁴⁴¹⁾; otra literatura donde Hossein Tabatabaeian y colaboradores en su publicación científica Non-Coding RNAs: Uncharted Mediators of Thyroid Cancer Pathogenesis ⁽⁴⁴²⁾, indica que los ARN no codificantes son moléculas emergentes con diversas capacidades para iniciar y promover el cáncer de tiroides tras la desregulación durante el desarrollo del cáncer de tiroides o la adquisición de resistencia terapéutica, y pueden desempeñar un papel clave en el fracaso del tratamiento y el mal pronóstico de los pacientes con esta neoplasia ⁽⁴⁴²⁾, identificando a SNHG3 (gen huésped 3 de ARN nucleolar pequeño), PAR5 (Región de Prader Willi/Angelman ARN5), CASC9 (Susceptibilidad al cáncer 9); DLX6-AS1 (Homeobox 6-

Antisentido 1), NEAT1 (Transcripción 1 del ensamblaje de paraspeckle nuclear), UNC5B-AS1 (Receptor B-Antisentido 1 Unc-5 Netrin), XIST (transcripción específica X-inactiva), MALAT1 (Transcripción 1 de adenocarcinoma de pulmón asociado a metástasis), HCP5 (complejo HLA P5), PTCSC3 (Candidato 3 de susceptibilidad al carcinoma de tiroides), PI3K (fosfatidilinositol 4,5bisfosfato 3 quinasa), mTOR (objetivo mecanicista de la rapamicina quinasa), EZH2 (potenciador de Zeste Homolog 2), ADAM9 (Dominio 9 de metalopeptidasa de ADAM), EGFR (Receptor del factor de crecimiento epidérmico), UPF1 (ARN Helicasa y ATPasa de UPF1), ER- β (receptor de estrógeno beta), SPAG9 (Antígeno 9 asociado a la permanente), BDNF (Factor neurotrófico derivado del cerebro), CDH6 (Cadherina 6), STAT3 (Transductor de señal y activador de la transcripción 3), ST6GAL2 (alfa-2, 6-sialiltransferasa 2), SCAI (Supresor de la invasión de células cancerosas) ⁽⁴⁴²⁾ como molecular de ARN no codificante involucrados en el cáncer tiroideo tanto de la variante papilar, folicular y anaplásica ⁽⁴⁴²⁾. De la revisión en la literatura no se han encontrado trabajos científicos que investiguen a FP671120.4 con el cáncer de tiroides, por lo que resulta relevante mencionar que en nuestro trabajo la fusión encontrada de este ocasiona una traslocación antisentido, no codificante de proteínas, asociada a esta neoplasia. Es importante mencionar a Valentina D y colaboradores quienes hacen una revisión actualizada en su trabajo Gene Fusions in Thyroid Cancer ⁽⁴⁴³⁾, de las fusiones de genes encontradas en los diversos subtipos histológicos del cáncer tiroideo, donde nos informan que desempeñan un papel importante tanto en la etiología como en la patogenia del cáncer y se consideran potenciales marcadores de diagnóstico y pronóstico y posibles dianas terapéuticas, el espectro y la prevalencia de las fusiones de genes en el cáncer tiroideo varía desde casos únicos hasta el 80%, según el tipo específico de cáncer ⁽⁴⁴³⁾, haciendo referencia a fusiones de genes en el cáncer papilar de tiroides (*RET*, *NTRK3*, *NTRK1*, *ALK*, *BRAF*, *PAX8-PPARG*, *THADA*) ⁽⁴⁴³⁾, cáncer folicular de tiroides (*PAX8-PPARG*, *CREB3L2-PPARG*, *DERL-*

COX6C)⁽⁴⁴³⁾, cáncer de tiroides pobremente diferenciado (*CCDC6-RET*, *NCOA4-RET*, *STRN-ALK*, *EML4-ALK*, *CCDC149-ALK*, *PAX8-PPARG*, *SCRIB-BRAF*, *ETV6-NTRK3*, *PDCD10-RET*, *TFG-RET*)⁽⁴⁴³⁾, cáncer anaplásico de tiroides (*NCOA4-RET*, *STRN-ALK*, *NUT-BRD4*, *KIAA1549-BRAF*)⁽⁴⁴³⁾ y el cáncer medular de tiroides (*MYH13-RET*, *EML4-ALK*, *GFPT1-ALK*)⁽⁴⁴³⁾. Como se evidencia no se han descrito las tres fusiones que hemos encontrado en nuestro estudio, analizando a las mismas, conllevan la participación de genes implicados en la patogenia del cáncer tiroideo, por lo que son de relevancia para futuras investigaciones de la biología molecular del cáncer de tiroides.

En el **Tabla 12 y Figuras 20-25**, se evidencian esquemáticamente la ubicación cromosómica de las mutaciones puntuales BRAF encontradas en la población de pacientes de nuestro estudio portadores de cáncer diferenciado de tiroides.

Se encontraron 6 mutaciones SNV (variante de un solo nucleótido) de BRAF tomando como referencia la base de datos del ensamblaje del genoma humano GRCh38.p7 (versión 7 del Consorcio de investigación del genoma humano 38)⁽⁴⁴⁴⁾, que a continuación se menciona: SNV C>T, G del Cromosoma 7:140924572 en el exón 1, la cual ya fue descrita en la literatura⁽⁴⁴⁵⁾, las restantes de novo no encontradas en la literatura como SNV T>A cromosoma7: 140850164 en el exón 2, SNV T>A cromosoma 7: 140800452 en el exón 7, SNV T>A, C, G cromosoma 7: 140794401 en el exón 8, SNV A>T cromosoma 7: 140777028 en el exón 14 y SNV A>T cromosoma 7: 140754195 en el exón 15; no se logro identificar las ubicaciones de la mutación BRAF^{V600E} (cromosoma 7: 140753335-140753336)⁽⁴⁴⁶⁾. Como ya se había mencionado, el hotspot mutacional más común en BRAF es T1799A en el exón 15, lo que confiere un glutamato a la valina como sustitución en el aminoácido 600 (V600E) en la proteína BRAF. BRAF^{V600E} es la alteración genética más común en CPT, presenta una alta prevalencia en el CPT clásico y la variante de células altas, aunque generalmente es raro en

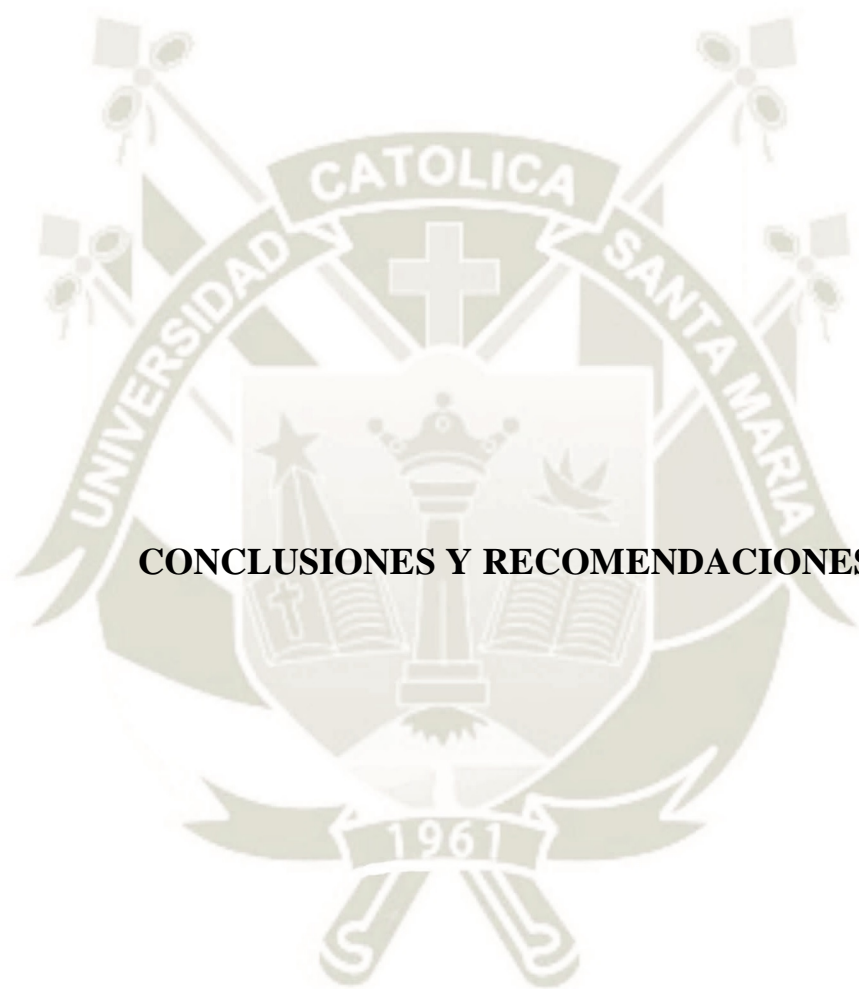
la variante folicular ⁽³⁴⁴⁾, y su localización genómica según la base de datos del GRCh38 se encuentra en el cromosoma 7: 140753335-140753336 ⁽⁴⁴⁶⁾; la presencia de BRAF^{V600E} ha demostrado una mayor agresividad y peor pronóstico en el cáncer de tiroides ⁽⁴⁴⁷⁾, siendo uno de los mecanismos la activación de genes posteriores de la vía MAPK, lo que puede conducir a una ligera regulación a la baja del simportador de yoduro de sodio (NIS), lo que provoca una localización inexacta de NIS en la membrana celular ^(448, 449), lo que conduce a una disminución en la capacidad de las células tumorales para captar yodo ^(450, 451) por lo que durante la terapia adyuvante con I-131 que se basa en la captación de yodo radiactivo por las células tumorales, la presencia de BRAF^{V600E} puede conducir a una disminución en el efecto terapéutico, lo que resulta en un mal pronóstico ⁽⁴⁵²⁾; en relación a nuestros hallazgos, correlaciona el no haber encontrado las mutaciones de BRAF agresivas con los estadios iniciales I y II de los pacientes de nuestra investigación (estadios de enfermedad localizada), y que sería importante medir el impacto clínico de las mutaciones de novo en el seguimiento de la evolución de esta enfermedad oncológica.

En la **Tabla 13 (parte 1-3) y Figuras 26-37**, se muestran esquemáticamente la ubicación cromosómica de las mutaciones puntuales TERT encontradas comunes en los 11 pacientes de la población de pacientes de nuestro estudio portadores de cáncer diferenciado de tiroides.

Se encontraron 21 mutaciones SNV (variante de un solo nucleótido) de TERT tomando como referencia la base de datos del ensamblaje del genoma humano GRCh38.p13 (versión 13 del Consorcio de investigación del genoma humano 38) ⁽⁴⁴⁴⁾, que a continuación se menciona: SNV T>C del cromosoma 5:1293763 en el exón 2 ⁽⁴⁵³⁾, SNV T>C del cromosoma 5:1272217 en el exón 7 ⁽⁴⁵⁴⁾ y SNV G>A del cromosoma 5:1255405 en el exón 13 ⁽⁴⁵⁵⁾, las cuales ya fueron descritas en la literatura, las restantes de novo no encontradas en la literatura como SNV C>G del cromosoma 5: 1293765 en el exón 2, SNV C>G del cromosoma 5:

1293766 en el exón 2, SNV C>A del cromosoma 5: 1293910 en el exón 2, SNV A>G del cromosoma 5: 1293913 en el exón 2, SNV C>A del cromosoma 5: 1293914 en el exón 2, SNV G>T del cromosoma 5: 1294136 en el exón 2, SNV A>C del cromosoma 5: 1294238 en el exón 2, SNV A>C del cromosoma 5: 1294313 en el exón 2, SNV T>G del cromosoma 5: 1294315 en el exón 2, SNV C>T del cromosoma 5: 1294430 en el exón 2, SNV G>A del cromosoma 5: 1294479 en el exón 2, SNV T>C del cromosoma: 5:1280318 en el exón 4, SNV T>A del cromosoma 5: 1278733 en el exón 6, SNV G>T del cromosoma 5: 1272218 en el exón 7, SNV T>C del cromosoma 5: 1272221 en el exón 7, SNV T>C del cromosoma 5: 1264568 en el exón 11, SNV A>C del cromosoma 5: 1264493 en el exón 11 y SNV A>G del cromosoma 5: 1264446 en el exón 11. En nuestro estudio no se encontraron las mutaciones 1295228 C>T y 1295250 C>T (denominadas C228T y C250T respectivamente), correspondientes a -124 C>T y -146 C>T del sitio de inicio de la traducción en el promotor de la transcriptasa inversa de telomerasa (TERT)^(456, 457), las cuales son las más estudiadas en el cáncer de tiroides y están relacionadas a la agresividad y pronóstico de la enfermedad^(458, 459), esto hallazgos podrían estar en relación a las variantes histopatológicas bien diferenciadas y de reciente diagnóstico de la población de nuestro trabajo, aun así, sería de mucho valor el seguimiento clínico cercano para el impacto de las mutaciones de novo encontradas en la respuesta al tratamiento y evolución de nuestros pacientes.

Finalmente, el conocimiento del conjunto de genes del cáncer de tiroides y su correlación con subgrupos clínicos de pacientes portadores de CDT, nos brindan información sobre la tumorigénesis del cáncer de tiroides y abren vías para desarrollar biomarcadores de pronóstico y terapias dirigidas a genes impulsores en el cáncer de tiroides, evidencia ya demostrada en múltiples publicaciones⁽⁴⁶⁰⁻⁴⁶³⁾.



CONCLUSIONES Y RECOMENDACIONES

CONCLUSIONES

Primera. Las muestras del cáncer diferenciado de tiroides analizadas comparten características epidemiológicas, clínicas, sonográficas e histopatológicas propias de esta patología.

Segunda. La expresión génica de las biopsias analizadas muestra un perfil de expresión estándar descrito por la literatura científica para este tipo de cáncer.

Tercera. Se han encontrado 5 nuevas mutaciones en el gen BRAF: SNV T>A cromosoma 7:140850164 exón 2, SNV T>A cromosoma 7:140800452 exón 7, SNV T>A, C, G cromosoma 7:140794401 exón 8, SNV A>T cromosoma 7:140777028 exón 14 y SNV A>T cromosoma 7:140754195 exón 15.

Cuarta. Se han encontrado 18 nuevas mutaciones en el gen TERT: SNV C>G cromosoma 5:1293765 exón 2, SNV C>G cromosoma 5:1293766 exón 2, SNV C>A cromosoma 5:1293910 exón 2, SNV A>G cromosoma 5:1293913 exón 2, SNV C>A cromosoma 5:1293914 exón 2, SNV G>T cromosoma 5:1294136 exón 2, SNV A>C cromosoma 5:1294238 exón 2, SNV A>C cromosoma 5:1294313 exón 2, SNV T>G cromosoma 5:1294315 exón 2, SNV C>T cromosoma 5:1294430 exón 2, SNV G>A cromosoma 5:1294479 exón 2, SNV T>C cromosoma 5:1280318 exón 4, SNV T>A cromosoma 5:1278733 exón 6, SNV G>T cromosoma 5:1272218 exón 7, SNV T>C cromosoma 5:1272221 exón 7, SNV T>C cromosoma 5:1264568 exón 11, SNV A>C cromosoma 5:1264493 exón 11 y SNV A>G cromosoma 5:1264446 exón 11.

Quinta. La búsqueda de proteínas de fusión se hizo con el software “Arriba (versión 1.2.0)”, encontrándose 3 tipos: TDG-TMEM132B, MEIS1-ITSN2, KCNE1B-FP671120.4, que por primera vez son descritas en esta tesis.

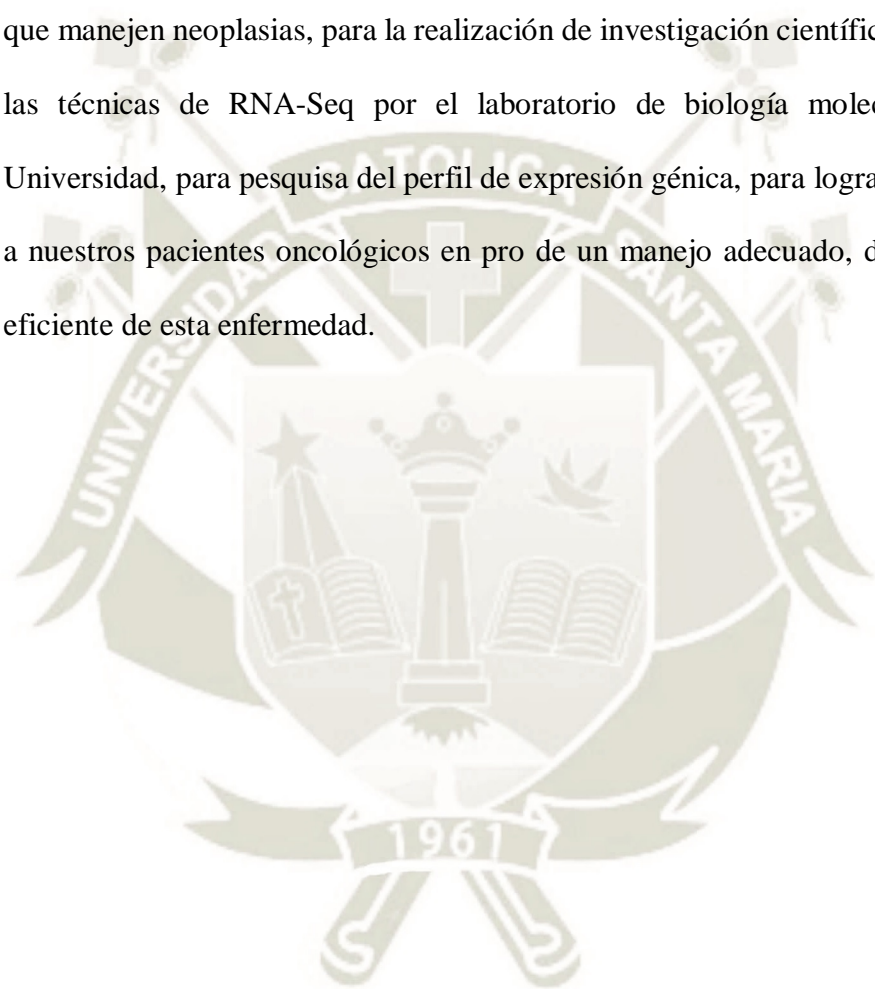


RECOMENDACIONES

- 1) Se debería permitir la accesibilidad al uso de técnicas de RNA-Seq en el sistema sanitario de nuestro país y utilizar el perfil de expresión génica durante la evaluación inicial y seguimiento del cáncer diferenciado de tiroides para estimar el pronóstico y respuesta al tratamiento precozmente, tal como ya se viene realizando en otros países a nivel mundial, en el manejo de pacientes portadores de esta neoplasia en nuestra población de Arequipa.
- 2) Se debería hacer investigación científica para el estudio genómico en búsqueda de mutaciones en el ADN de múltiples neoplasias con técnicas de RNA-Seq para la pesquisa de targets terapéuticos y pronósticos, según los hallazgos de perfil de expresión génica autóctono de cada población, para un tratamiento individualizado del cáncer.
- 3) Las nuevas mutaciones encontradas así como las proteínas de fusión deben de ser estudiadas y analizadas desde un punto de vista funcional y encontrar su relación funcional con el cancer y presentarlas como candidatos de target para diagnosticos y blancos de nuevas terapias.

PROPUESTA

- 1) Se debería de proponer y poner en marcha un convenio de prestaciones mutuas entre la Universidad Católica de Santa María, Ministerio de Salud, EsSalud, Fuerzas armadas e instituciones privadas de Arequipa que involucre a todas las instituciones que manejen neoplasias, para la realización de investigación científica y aplicación de las técnicas de RNA-Seq por el laboratorio de biología molecular de nuestra Universidad, para pesquisa del perfil de expresión génica, para lograr la accesibilidad a nuestros pacientes oncológicos en pro de un manejo adecuado, dirigido, precoz y eficiente de esta enfermedad.



REFERENCIA

1. GLOBOCAN <https://gco.iarc.fr/>.
2. Goodarzi E, et al. Epidemiology, incidence and mortality of thyroid cancer and their relationship with the human development index in the world: An ecology study in 2018. *WCRJ* 2018; 5 (2): e1091.
3. Fabián Pitoia. Cáncer de Tiroides en Búsqueda del Tratamiento Individualizado. *Medicina (Buenos Aires)* 2012; 72: 503-513.
4. Pérez B. Avances y Perspectivas en el Cáncer Diferenciado de Tiroides. México: Permanyer; 2018.
5. Atamari-Anahui N, et al. Tendencia nacional de la prevalencia y mortalidad por cáncer de tiroides con datos del Ministerio de Salud de Perú. *Medwave* 2019; 19(4): e7631.
6. Stacey L, et al. Global variation in the pattern of differentiated thyroid cancer. *The American Journal of Surgery* (2010) 200, 462–466.
7. Haytham B, et al. Invasiveness and Metastatic Aggressiveness in Small Differentiated Thyroid Cancers: Demography of Small Papillary Thyroid Carcinomas in the Swedish Population. *World J Surg* (2020) 44:461–468.
8. Sanam Lathief, et al. Advances and practical use of the molecular markers for thyroid cancer. *Advances in Cellular and Molecular Otolaryngology* 2016; 4: 33948.
9. Bryan R. Haugen, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid*; Volume 26, Number 1, 2016.
10. George E. Naoum, et al. Novel targeted therapies and immunotherapy for advanced thyroid cancers. *Molecular Cancer*. 2018; 17:51.
11. Li-Bo Yang, et al. The Clinicopathological Features of BRAF Mutated Papillary Thyroid Cancers in Chinese Patients. *Int J Endocrinol*. 2015; 2015: 642046.
12. Cha YJ, Koo JS. Next-generation sequencing in thyroid cancer. *J Transl Med*. 2016; 14(1):322. Published 2016 Nov 21.
13. He, J., Tian, Z., Yao, X. *et al*. A novel RNA sequencing-based risk score model to predict papillary thyroid carcinoma recurrence. *Clin Exp Metastasis* **37**, 257–267 (2020).
14. Romina L, et al. Uso de Inhibidores de Tirosina Quinasa en Carcinoma Medular de Tiroides. Experiencia Institucional. *Oncología Clínica* 2017; 22: 24-27.

15. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000; 100: 57-70.
16. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011; 144: 646-674.
17. Hofstra RM, Landsvater RM, Ceccherini I, Stulp RP, Stelwagen T, Luo Y, et al. A mutation in the RET proto-oncogene associated with multiple endocrine neoplasia type 2B and sporadic medullary thyroid carcinoma. *Nature* 1994; 367: 375-376.
18. Cohen Y, Xing M, Mambo E, Guo Z, Wu G, Trink B, et al. BRAF mutation in papillary thyroid carcinoma. *J Natl Cancer Inst*. 2003; 95: 625-627.
19. Xing M. Molecular pathogenesis and mechanisms of thyroid cancer. *Nat Rev Cancer*. 2013; 13: 184-199.
20. Xing M. BRAF mutation in thyroid cancer. *Endocr Relat Cancer*. 2005; 12: 245-262.
21. Xing M, Alzahrani AS, Carson KA, Viola D, Elisei R, Bendlova B, et al. Association between BRAF V600E mutation and mortality in patients with papillary thyroid cancer. *JAMA*. 2013; 309: 1493-1501.
22. Xing M, Westra WH, Tufano RP, Cohen Y, Rosenbaum E, Rhoden KJ, et al. BRAF mutation predicts a poorer clinical prognosis for papillary thyroid cancer. *J Clin Endocrinol Metab*. 2005; 90: 6373-6379.
23. Garcia-Rostan G, Zhao H, Camp RL, Pollan M, Herrero A, Pardo J, et al. Ras mutations are associated with aggressive tumor phenotypes and poor prognosis in thyroid cancer. *J Clin Oncol*. 2003; 21: 3226-3235.
24. Liu Z, Hou P, Ji M, Guan H, Studeman K, Jensen K, et al. Highly prevalent genetic alterations in receptor tyrosine kinases and phosphatidylinositol 3-kinase/akt and mitogen-activated protein kinase pathways in anaplastic and follicular thyroid cancers. *J Clin Endocrinol Metab*. 2008; 93: 3106-3116.
25. Gustafson S, Zbuk KM, Scacheri C, Eng C. Cowden syndrome. *Semin Oncol*. 2007; 34: 428-434.
26. Alvarez-Núñez F, Bussaglia E, Mauricio D, Ybarra J, Vilar M, Lerma E, et al. PTEN promoter methylation in sporadic thyroid carcinomas. *Thyroid*. 2006; 16: 17-23.
27. Hou P, Ji M, Xing M. Association of PTEN gene methylation with genetic alterations in the phosphatidylinositol 3-kinase/AKT signaling pathway in thyroid tumors. *Cancer*. 2008; 113: 2440-2447.

28. Schagdarsurengin U, Gimm O, Dralle H, Hoang-Vu C, Dammann R. CpG island methylation of tumor-related promoters occurs preferentially in undifferentiated carcinoma. *Thyroid*. 2006; 16: 633-642.
29. Murugan AK, Bojdani E, Xing M. Identification and functional characterization of isocitrate dehydrogenase 1 (IDH1) mutations in thyroid cancer. *Biochem Biophys Res Commun*. 2010; 393: 555-559.
30. Garcia-Rostan G, Tallini G, Herrero A, D'Aquila TG, Carcangiu ML, Rimm DL. Frequent mutation and nuclear localization of beta-catenin in anaplastic thyroid carcinoma. *Cancer Res*. 1999; 59: 1811-1815.
31. Donghi R, Longoni A, Pilotti S, Michieli P, Della Porta G, Pierotti MA. Gene p53 mutations are restricted to poorly differentiated and undifferentiated carcinomas of the thyroid gland. *J Clin Invest*. 1993;91: 1753-1760.
32. Fagin JA, Matsuo K, Karmakar A, Chen DL, Tang SH, Koeffler HP. High prevalence of mutations of the p53 gene in poorly differentiated human thyroid carcinomas. *J Clin Invest*. 1993; 91: 179-184.
33. Musholt PB, Musholt TJ, Morgenstern SC, Worm K, Sheu SY, Schmid KW. Follicular histotypes of oncocytic thyroid carcinomas do not carry mutations of the BRAF hot-spot. *World J Surg*. 2008; 32: 722-728.
34. Máximo V, Botelho T, Capela J, Soares P, Lima J, Taveira A, et al. Somatic and germline mutation in GRIM-19, a dual function gene involved in mitochondrial metabolism and cell death, is linked to mitochondrionrich (Hurthle cell) tumours of the thyroid. *Br J Cancer*. 2005; 92: 1892-1898.
35. Bounacer A, Wicker R, Caillou B, Cailleux AF, Sarasin A, Schlumberger M, et al. High prevalence of activating ret proto-oncogene rearrangements, in thyroid tumors from patients who had received external radiation. *Oncogene*. 1997; 15: 1263-1273.
36. Rabes HM, Demidchik EP, Sidorow JD, Lengfelder E, Beimfohr C, Hoelzel D, et al. Pattern of radiation-induced RET and NTRK1 rearrangements in 191 post-chernobyl papillary thyroid carcinomas: biological, phenotypic, and clinical implications. *Clin Cancer Res* 2000;6: 1093-1103.
37. Grieco M, Santoro M, Berlingieri MT, Melillo RM, Donghi R, Bongarzone I, et al. PTC is a novel rearranged form of the ret proto-oncogene and is frequently detected in vivo in human thyroid papillary carcinomas. *Cell* 1990; 60: 557-563.

38. Santoro M, Dathan NA, Berlingieri MT, Bongarzone I, Paulin C, Grieco M, et al. Molecular characterization of RET/PTC3; a novel rearranged version of the RET proto-oncogene in a human thyroid papillary carcinoma. *Oncogene* 1994; 9: 509-516.
39. Trovisco V, Soares P, Preto A, Castro P, Máximo V, Sobrinho-Simões M. Molecular genetics of papillary thyroid carcinoma: great expectations. *Arq Bras Endocrinol Metabol.* 2007; 51: 643-653.
40. Placzkowski KA, Reddi HV, Grebe SK, Eberhardt NL, McIver B. The Role of the PAX8/PPARgamma Fusion Oncogene in Thyroid Cancer. *PPAR Res* 2008; 2008: 672829.
41. Adjadj E, Schlumberger M, de Vathaire F. Germ-line DNA polymorphisms and susceptibility to differentiated thyroid cancer. *Lancet Oncol.* 2009; 10: 181-190.
42. Canbay E, Dokmetas S, Canbay EI, Sen M, Bardakci F. Higher glutathione transferase GSTM1 0/0 genotype frequency in Young thyroid carcinoma patients. *Curr Med Res Opin.* 2003; 19: 102-106.
43. Granja F, Morari J, Morari EC, Correa LA, Assumpção LV, Ward LS. Proline homozygosity in codon 72 of p53 is a factor of susceptibility for thyroid cancer. *Cancer Lett.* 2004; 210: 151-157.
44. Morari EC, Leite JL, Granja F, da Assumpção LV, Ward LS. The null genotype of glutathione s-transferase M1 and T1 locus increases the risk for thyroid cancer. *Cancer Epidemiol Biomarkers Prev.* 2002; 11:1485-1488.
45. Gudmundsson J, Sulem P, Gudbjartsson DF, Jonasson JG, Sigurdsson A, Bergthorsson JT, et al. Common variants on 9q22.33 and 14q13.3 predispose to thyroid cancer in European populations. *Nat Genet.* 2009; 41: 460-464.
46. Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell.* 2009; 136: 215-233.
47. Esquela-Kerscher A, Slack FJ. Oncomirs - microRNAs with a role in cancer. *Nat Rev Cancer.* 2006; 6: 259-269.
48. De la Chapelle A, Jazdzewski K. MicroRNAs in thyroid cancer. *J Clin Endocrinol Metab.* 2011; 96: 3326-3336.
49. Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, Teague J, Woffendin H, Garnett MJ, Bottomley W, Davis N, Dicks E, Ewing R, et al. Mutations of the BRAF gene in human cancer. *Nature.* 2002; 417:949-954.

50. Dhomen N, Marais R. New insight into BRAF mutations in cancer. *Curr Opin Genet Dev.* 2007; 17:31–39.
51. Caronia LM, Phay JE, Shah MH. Role of BRAF in thyroid oncogenesis. *Clin Cancer Res.* 2011; 17:7511–7517.
52. Khatami F, Tavangar SM. A Review of Driver Genetic Alterations in Thyroid Cancers. *Iran J Pathol.* 2018; 13(2):125-135.
53. Ciampi R, Knauf JA, Kerler R, Gandhi M, Zhu Z, Nikiforova MN, Rabes HM, Fagin JA, Nikiforov YE. Oncogenic AKAP9-BRAF fusion is a novel mechanism of MAPK pathway activation in thyroid cancer. *J Clin Invest.* 2005; 115:94–101.
54. Wan PT, Garnett MJ, Roe SM, Lee S, Niculescu-Duvaz D, Good VM, Jones CM, Marshall CJ, Springer CJ, Barford D, Marais R, and Cancer Genome Project. Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. *Cell.* 2004; 116:855–867.
55. Pollock PM, Harper UL, Hansen KS, Yudt LM, Stark M, Robbins CM, Moses TY, Hostetter G, Wagner U, Kakareka J, Salem G, Pohida T, Heenan P, et al. High frequency of BRAF mutations in nevi. *Nat Genet.* 2003; 33:19–20.
56. Poynter JN, Elder JT, Fullen DR, Nair RP, Soengas MS, Johnson TM, Redman B, Thomas NE, Gruber SB. BRAF and NRAS mutations in melanoma and melanocytic nevi. *Melanoma Res.* 2006; 16:267–273.
57. Uhlen M, Zhang C, Lee S, Sjöstedt E, Fagerberg L, Bidkhori G, Benfeitas R, Arif M, Liu Z, Edfors F, Sanli K, von Feilitzen K, Oksvold P, et al. A pathology atlas of the human cancer transcriptome. *Science.* 2017; 357: eaan2507.
58. Yan C, Huang ML, Li X, Wang T, Ling R. Relationship between BRAFV600E and clinical features in papillary thyroid carcinoma. *Endocr Connect.* 2019; 8:988–96.
59. Yoo SK, Song YS, Lee EK, Hwang J, Kim HH, Jung G, Kim YA, Kim SJ, Cho SW, Won JK, Chung EJ, Shin JY, Lee KE, et al. Integrative analysis of genomic and transcriptomic characteristics associated with progression of aggressive thyroid cancer. *Nat Commun.* 2019; 10:2764.
60. Lorentzen HF. Targeted therapy for malignant melanoma. *Curr Opin Pharmacol.* 2019; 46:116–121.
61. Ducreux M, Chamseddine A, Laurent-Puig P, Smolenschi C, Hollebecque A, Dartigues P, Samallin E, Boige V, Malka D, Gelli M. Molecular targeted therapy of BRAF-mutant colorectal cancer. *Ther Adv Med Oncol.* 2019; 11:1–15.

62. Bernicker EH, Allen TC, Cagle PT. Update on emerging biomarkers in lung cancer. *J Thorac Dis.* 2019; 11: S81–S88.
63. Yip L, Nikiforova MN, Yoo JY, McCoy KL, Stang MT, Armstrong MJ, Nicholson KJ, Ohori NP, Coyne C, Hodak SP, Ferris RL, LeBeau SO, Nikiforov YE, Carty SE. Tumor genotype determines phenotype and disease-related outcomes in thyroid cancer: a study of 1510 patients. *Ann Surg.* 2015; 262:519–25.
64. Croce L, Coperchini F, Magri F, Chiovato L, Rotondi M. The multifaceted anti-cancer effects of BRAF-inhibitors. *Oncotarget.* 2019; 10(61):6623-6640.
65. Landa I, Ibrahimasic T, Boucai L et al. Genomic and transcriptomic hallmarks of poorly differentiated and anaplastic thyroid cancers. *J Clin Invest* 2016; 126: 1052–66.
66. Agrawal N, Akbani R, Aksoy BA et al. For the cancer genome atlas research network. Integrated genomic characterization of papillary thyroid carcinoma. *Cell* 2014; 159: 676–90.
67. Kakudo, K., Bychkov, A., Bai, Y., Li, Y., Liu, Z., & Jung, C. K. (2018). The new 4th edition World Health Organization classification for thyroid tumors, Asian perspectives. *Pathology international*, 68(12), 641–664.
68. Cech TR, Steitz JA. The noncoding RNA revolution trashing old rules to forge new ones. *Cell.* 2014; 157(1):77–94.
69. Netto, George & Kaul, Karen. (2019). *Genomic Applications in Pathology* (2^{da} ed). Springer, Cham. <https://doi.org/10.1007/978-3-319-96830-8>.
70. Oszolak F, Milos PM. RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet.* 2011; 12(2):87–98.
71. Byron SA, Van Keuren-Jensen KR, Engelthaler DM, Carpten JD, Craig DW. Translating RNA sequencing into clinical diagnostics: opportunities and challenges. *Nat Rev Genet.* 2016; 17(5):257–71.
72. Roychowdhury S, Chinnaiyan AM. Translating cancer genomes and transcriptomes for precision oncology. *CA Cancer J Clin.* 2016; 66(1):75–88.
73. Rodriguez SA, Impey SD, Pelz C, Enestvedt B, Bakis G, Owens M, et al. RNA sequencing distinguishes benign from malignant pancreatic lesions sampled by EUS-guided FNA. *Gastrointest Endosc.* 2016; 84(2):252–8.

74. Li X, Quigg RJ, Zhou J, Gu W, Nagesh Rao P, Reed EF. Clinical utility of microarrays: current status, existing challenges and future outlook. *Curr Genomics*. 2008; 9(7):466–74.
75. Fedorowicz G, Guerrero S, Wu TD, Modrusan Z. Microarray analysis of RNA extracted from formalin-fixed, paraffin-embedded and matched fresh-frozen ovarian adenocarcinomas. *BMC Med Genet*. 2009; 2:23.
76. Coudry RA, Meireles SI, Stoyanova R, Cooper HS, Carpino A, Wang X, et al. Successful application of microarray technology to microdissected formalin-fixed, paraffin-embedded tissue. *J Mol Diagn*. 2007; 9(1):70–9.
77. Malone JH, Oliver B. Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biol*. 2011; 9:34.
78. Jaluria P, Konstantopoulos K, Betenbaugh M, Shiloach J. A perspective on microarrays: current applications, pitfalls, and potential uses. *Microb Cell Factories*. 2007; 6:4.
79. Dallas PB, Gottardo NG, Firth MJ, Beesley AH, Hoffmann K, Terry PA, et al. Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR -- how well do they correlate? *BMC Genomics*. 2005; 6:59.
80. Yuen T, Wurmbach E, Pfeffer RL, Ebersole BJ, Sealfon SC. Accuracy and calibration of commercial oligonucleotide and custom cDNA microarrays. *Nucleic Acids Res*. 2002;30(10): e48.
81. Fan X, Lobenhofer EK, Chen M, Shi W, Huang J, Luo J, et al. Consistency of predictive signature genes and classifiers generated using different microarray platforms. *Pharmacogenomics J*. 2010; 10(4):247–57.
82. Dumur CI, Fuller CE, Blevins TL, Schaum JC, Wilkinson DS, Garrett CT, et al. Clinical verification of the performance of the pathwork tissue of origin test: utility and limitations. *Am J Clin Pathol*. 2011; 136(6):924–33.
83. Cardoso F, van't Veer LJ, Bogaerts J, Slaets L, Viale G, Delaloge S, et al. 70-gene signature as an aid to treatment decisions in early-stage breast cancer. *N Engl J Med*. 2016; 375(8):717–29.
84. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*. 2000; 403(6769):503–11.

85. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature*. 2000; 406(6797):747–52.
86. Consortium M, Shi L, Reid LH, Jones WD, Shippy R, Warrington JA, et al. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol*. 2006; 24(9):1151–6.
87. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, et al. Minimum information about a microarray experiment (MIAME) - toward standards for microarray data. *Nat Genet*. 2001; 29(4):365–71.
88. Ma C, Lyons-Weiler M, Liang W, LaFramboise W, Gilbertson JR, Becich MJ, et al. In vitro transcription amplification and labeling methods contribute to the variability of gene expression profiling with DNA microarrays. *J Mol Diagn*. 2006; 8(2):183–92.
89. Schulze A, Downward J. Navigating gene expression using microarrays—a technology review. *Nat Cell Biol*. 2001; 3(8): E190–5.
90. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*. 2009; 10(1):57–63.
91. Kulkarni MM. Digital multiplexed gene expression analysis using the NanoString nCounter system. *Curr Protoc Mol Biol*. 2011; 94; Chapter 25: Unit25B 10.
92. Geiss GK, Bumgarner RE, Birditt B, Dahl T, Dowidar N, Dunaway DL, et al. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol*. 2008; 26(3):317–25.
93. Richard AC, Lyons PA, Peters JE, Biasci D, Flint SM, Lee JC, et al. Comparison of gene expression microarray data with count-based RNA measurements informs microarray interpretation. *BMC Genomics*. 2014; 15:649.
94. Malkov VA, Serikawa KA, Balantac N, Watters J, Geiss G, Mashadi-Hosseini A, et al. Multiplexed measurements of gene signatures in different analytes using the Nanostring nCounter Assay System. *BMC Res Notes*. 2009; 2:80.
95. Nielsen T, Wallden B, Schaper C, Ferree S, Liu S, Gao D, et al. Analytical validation of the PAM50- based Prosigna Breast Cancer Prognostic Gene Signature Assay and nCounter Analysis System using formalin-fixed paraffin-embedded breast tumor specimens. *BMC Cancer*. 2014; 14:177.

96. Prokopec SD, Watson JD, Waggott DM, Smith AB, Wu AH, Okey AB, et al. Systematic evaluation of medium-throughput mRNA abundance platforms. *RNA*. 2013; 19(1):51–62.
97. Reis PP, Waldron L, Goswami RS, Xu W, Xuan Y, Perez-Ordóñez B, et al. mRNA transcript quantification in archival samples using multiplexed, color-coded probes. *BMC Biotechnol*. 2011; 11:46.
98. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*. 2008; 5(7):621–8.
99. Consortium SM-I. A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. *Nat Biotechnol*. 2014; 32(9):903–14.
100. Cieslik M, Chinnaiyan AM. Cancer transcriptome profiling at the juncture of clinical translation. *Nat Rev Genet*. 2018; 19(2):93–109.
101. Feng H, Zhang X, Zhang C. mRIN for direct assessment of genome-wide and gene-specific mRNA integrity from large-scale RNA-sequencing data. *Nat Commun*. 2015; 6:7816.
102. Morlan JD, Qu K, Sinicropi DV. Selective depletion of rRNA enables whole transcriptome profiling of archival fixed tissue. *PLoS One*. 2012; 7(8): e42882.
103. Zhao W, He X, Hoadley KA, Parker JS, Hayes DN, Perou CM. Comparison of RNA-Seq by poly (A) capture, ribosomal RNA depletion, and DNA microarray for expression profiling. *BMC Genomics*. 2014; 15:419.
104. O’Neil D, Glowatz H, Schlumpberger M. Ribosomal RNA depletion for efficient use of RNA-seq capacity. *Curr Protoc Mol Biol*. 2013; 103; Chapter 4: Unit 4 19.
105. Li S, Tighe SW, Nicolet CM, Grove D, Levy S, Farmerie W, et al. Multi-platform assessment of transcriptome profiling using RNA-seq in the ABRF next-generation sequencing study. *Nat Biotechnol*. 2014; 32(9):915–25.
106. Sultan M, Amstislavskiy V, Risch T, Schuette M, Dokel S, Ralser M, et al. Influence of RNA extraction methods and library selection schemes on RNAseq data. *BMC Genomics*. 2014; 15:675.
107. Cieslik M, Chugh R, Wu YM, Wu M, Brennan C, Lonigro R, et al. The use of exome capture RNA-seq for highly degraded RNA with application to clinical cancer sequencing. *Genome Res*. 2015; 25(9):1372–81.

108. Li W, Turner A, Aggarwal P, Matter A, Storvick E, Arnett DK, et al. Comprehensive evaluation of AmpliSeq transcriptome, a novel targeted whole transcriptome RNA sequencing methodology for global gene expression analysis. *BMC Genomics*. 2015; 16:1069.
109. Zheng Z, Liebers M, Zhelyazkova B, Cao Y, Panditi D, Lynch KD, et al. Anchored multiplex PCR for targeted next-generation sequencing. *Nat Med*. 2014; 20(12):1479–84.
110. Levin JZ, Berger MF, Adiconis X, Rogov P, Melnikov A, Fennell T, et al. Targeted nextgeneration sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts. *Genome Biol*. 2009; 10(10): R115.
111. Nikiforov YE, Yip L, Nikiforova MN. New strategies in diagnosing cancer in thyroid nodules: impact of molecular markers. *Clin Cancer Res*. 2013; 19(9):2283–8.
112. National Human Genome Research Institute. Clinical Sequencing Evidence-Generating Research (CSER2) 2017 [updated 08/08/2017]. Available from: <https://www.genome.gov/27546194/clinicalsequencing-exploratory-research-cser/>.
113. Kivioja T, Vaharautio A, Karlsson K, Bonke M, Enge M, Linnarsson S, et al. Counting absolute numbers of molecules using unique molecular identifiers. *Nat Methods*. 2011; 9(1):72–4.
114. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et al. A survey of best practices for RNA-seq data analysis. *Genome Biol*. 2016; 17:13.
115. Martin JA, Wang Z. Next-generation transcriptome assembly. *Nat Rev Genet*. 2011; 12(10):671–82.
116. Babraham Bioinformatics. FastQC 2010 [updated 01/10/2018]. Available from: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
117. Hannon GJ. FASTX-Toolkit 2009 [updated 02/02/2010]. Available from: http://hannonlab.cshl.edu/fastx_toolkit/index.html.
118. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014; 30(15):2114–20.
119. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*. 2009; 25(9):1105–11.
120. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013; 29(1):15–21.

121. Engstrom PG, Steijger T, Sipos B, Grant GR, Kahles A, Ratsch G, et al. Systematic evaluation of spliced alignment programs for RNA-seq data. *Nat Methods*. 2013; 10(12):1185–91.
122. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010; 28(5):511–5.
123. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011; 12:323.
124. Martin J, Bruno VM, Fang Z, Meng X, Blow M, Zhang T, et al. Rnnotator: an automated de novo transcriptome assembly pipeline from stranded RNA-Seq reads. *BMC Genomics*. 2010; 11: 663.
125. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 2011;29(7):644–52.
126. Xie Y, Wu G, Tang J, Luo R, Patterson J, Liu S, et al. SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. *Bioinformatics*. 2014; 30(12):1660–6.
127. Robertson G, Schein J, Chiu R, Corbett R, Field M, Jackman SD, et al. De novo assembly and analysis of RNA-seq data. *Nat Methods*. 2010; 7(11):909–12.
128. Iyer MK, Chinnaiyan AM. RNA-Seq unleashed. *Nat Biotechnol*. 2011; 29(7):599–600.
129. Broad Institute. Picard [Available from: [http:// broadinstitute.github.io/picard/](http://broadinstitute.github.io/picard/)].
130. DeLuca DS, Levin JZ, Sivachenko A, Fennell T, Nazaire MD, Williams C, et al. RNA-SeQC: RNAseq metrics for quality control and process optimization. *Bioinformatics*. 2012; 28(11):1530–2.
131. Wang L, Wang S, Li W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics*. 2012; 28(16):2184–5.
132. T Hoen PA, Friedlander MR, Almlöf J, Sammeth M, Pulyakhina I, Anvar SY, et al. Reproducibility of highthroughput mRNA and small RNA sequencing across laboratories. *Nat Biotechnol*. 2013; 31(11):1015–22.
133. Sims D, Sudbery I, Ilott NE, Heger A, Ponting CP. Sequencing depth and coverage: key considerations in genomic analyses. *Nat Rev Genet*. 2014; 15(2):121–32.

134. Nazarov PV, Muller A, Kaoma T, Nicot N, Maximo C, Birembaut P, et al. RNA sequencing and transcriptome arrays analyses show opposing results for alternative splicing in patient derived samples. *BMC Genomics*. 2017; 18(1):443.
135. Hardwick SA, Deveson IW, Mercer TR. Reference standards for next-generation sequencing. *Nat Rev Genet*. 2017; 18(8):473–84.
136. National Center for Biotechnology Information. Sequence read archive [Available from: <https://www.ncbi.nlm.nih.gov/sra>].
137. National Center for Biotechnology Information. Gene expression omnibus [Available from: <https://www.ncbi.nlm.nih.gov/geo/>].
138. Jiang L, Schlesinger F, Davis CA, Zhang Y, Li R, Salit M, et al. Synthetic spike-in standards for RNAseq experiments. *Genome Res*. 2011;21(9):1543–5.
139. Baker SC, Bauer SR, Beyer RP, Brenton JD, Bromley B, Burrill J, et al. The External RNA Controls Consortium: a progress report. *Nat Methods*. 2005; 2(10):731–4.
140. Hardwick SA, Chen WY, Wong T, Deveson IW, Blackburn J, Andersen SB, et al. Spliced synthetic genes as internal controls in RNA sequencing experiments. *Nat Methods*. 2016; 13(9):792–8.
141. Lutzmayer S, Enugutti B, Nodine MD. Novel small RNA spike-in oligonucleotides enable absolute normalization of small RNA-Seq data. *Sci Rep*. 2017; 7(1):5913.
142. Lukas Paul PK, Horner G, Ante M, Hollaender I, Alexander S, Reda T. SIRVs: Spike-In RNA Variants as external isoform controls in RNA-sequencing. *bioRxiv*. Posted October 13, 2016. <https://doi.org/10.1101/080747>.
143. National Institute of Standards and Technology. ERCC 2.0: developing a new suite of RNA controls 2017 [updated 02/17/2017]. Available from: <https://www.nist.gov/programs-projects/ercc-20-developing-new-suite-rna-controls>.
144. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010; 26(1):139–40.
145. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014; 15(12):550.
146. Tarazona S, Furio-Tari P, Turra D, Pietro AD, Nueda MJ, Ferrer A, et al. Data quality aware analysis of differential expression in RNA-seq with NOISeq R/ Bioc package. *Nucleic Acids Res*. 2015; 43(21): e140.

147. Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol.* 2013; 31(1):46–53.
148. Wang W, Qin Z, Feng Z, Wang X, Zhang X. Identifying differentially spliced genes from two groups of RNA-seq samples. *Gene.* 2013; 518(1):164–70.
149. Shen S, Park JW, Lu ZX, Lin L, Henry MD, Wu YN, et al. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc Natl Acad Sci U S A.* 2014; 111(51): E5593–601.
150. Steijger T, Abril JF, Engstrom PG, Kokocinski F, Consortium R, Hubbard TJ, et al. Assessment of transcript reconstruction methods for RNA-seq. *Nat Methods.* 2013; 10(12):1177–84.
151. Bradford JR, Hey Y, Yates T, Li Y, Pepper SD, Miller CJ. A comparison of massively parallel nucleotide sequencing with oligonucleotide microarrays for global transcription profiling. *BMC Genomics.* 2010; 11:282.
152. Rehrauer H, Opitz L, Tan G, Sieverling L, Schlapbach R. Blind spots of quantitative RNA-seq: the limits for assessing abundance, differential expression, and isoform switching. *BMC Bioinformatics.* 2013; 14:370.
153. Sahraeian SME, Mohiyuddin M, Sebra R, Tilgner H, Afshar PT, Au KF, et al. Gaining comprehensive biological insight into the transcriptome by performing a broad-spectrum RNA-seq analysis. *Nat Commun.* 2017;8(1):59.
154. Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Van der Auwera GA, et al. Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv.* Posted November 14, 2017. <https://doi.org/10.1101/201178>.
155. Piskol R, Ramaswami G, Li JB. Reliable identification of genomic variants from RNA-seq data. *Am J Hum Genet.* 2013; 93(4):641–51.
156. Schram AM, Chang MT, Jonsson P, Drilon A. Fusions in solid tumours: diagnostic strategies, targeted therapy, and acquired resistance. *Nat Rev Clin Oncol.* 2017; 14(12):735–48.
157. McPherson A, Hormozdiari F, Zayed A, Giuliany R, Ha G, Sun MG, et al. deFuse: an algorithm for gene fusion discovery in tumor RNA-Seq data. *PLoS Comput Biol.* 2011; 7(5): e1001138.
158. Kim D, Salzberg SL. TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. *Genome Biol.* 2011; 12(8): R72.

159. Li Y, Chien J, Smith DI, Ma J. FusionHunter: identifying fusion transcripts in cancer using paired-end RNA-seq. *Bioinformatics*. 2011; 27(12):1708–10.
160. Torres-Garcia W, Zheng S, Sivachenko A, Vegesna R, Wang Q, Yao R, et al. PRADA: pipeline for RNA sequencing data analysis. *Bioinformatics*. 2014; 30(15):2224–6.
161. Jia W, Qiu K, He M, Song P, Zhou Q, Zhou F, et al. SOAPfuse: an algorithm for identifying fusion transcripts from paired-end RNA-Seq data. *Genome Biol*. 2013; 14(2): R12.
162. Beccuti M, Carrara M, Cordero F, Lazzarato F, Donatelli S, Nadalin F, et al. Chimera: a Bioconductor package for secondary analysis of fusion products. *Bioinformatics*. 2014; 30(24):3556–7.
163. Kumar S, Vo AD, Qin F, Li H. Comparative assessment of methods for the fusion transcripts detection from RNA-Seq data. *Sci Rep*. 2016; 6:21597.
164. Davila JI, Fadra NM, Wang X, McDonald AM, Nair AA, Crusan BR, et al. Impact of RNA degradation on fusion detection by RNA-seq. *BMC Genomics*. 2016; 17(1):814.
165. Rogers TM, Arnau GM, Ryland GL, Huang S, Lira ME, Emmanuel Y, et al. Multiplexed transcriptome analysis to detect ALK, ROS1 and RET rearrangements in lung cancer. *Sci Rep*. 2017; 7:42259.
166. Roy S, Coldren C, Karunamurthy A, Kip NS, Klee EW, Lincoln SE, et al. Standards and guidelines for validating next-generation sequencing bioinformatics pipelines: a joint recommendation of the Association for Molecular Pathology and the College of American Pathologists. *J Mol Diagn*. 2018; 20(1):4–27.
167. Monzon FA, Lyons-Weiler M, Buturovic LJ, Rigl CT, Henner WD, Sciulli C, et al. Multicenter validation of a 1,550-gene expression profile for identification of tumor tissue of origin. *J Clin Oncol*. 2009; 27(15):2503–8.
168. Prat A, Ellis MJ, Perou CM. Practical implications of gene-expression-based assays for breast oncologists. *Nat Rev Clin Oncol*. 2011; 9(1):48–57.
169. Baehner FL, Lee M, Demeure MJ, Bussey KJ, Kiefer JA, Barrett MT. Genomic signatures of cancer: basis for individualized risk assessment, selective staging and therapy. *J Surg Oncol*. 2011; 103(6):563–73.
170. Chibon F. Cancer gene expression signatures – the rise and fall? *Eur J Cancer*. 2013; 49(8):2000–9.

171. Pham MX, Teuteberg JJ, Kfoury AG, Starling RC, Deng MC, Cappola TP, et al. Gene-expression profiling for rejection surveillance after cardiac transplantation. *N Engl J Med*. 2010; 362(20):1890–900.
172. Cooper-Knock J, Kirby J, Ferraiuolo L, Heath PR, Rattray M, Shaw PJ. Gene expression profiling in human neurodegenerative disease. *Nat Rev Neurol*. 2012; 8(9):518–30.
173. Drew JE. Cellular defense system gene expression profiling of human whole blood: opportunities to predict health benefits in response to diet. *Adv Nutr*. 2012; 3(4):499–505.
174. Shih B, Watson S, Bayat A. Whole genome and global expression profiling of Dupuytren's disease: systematic review of current findings and future perspectives. *Ann Rheum Dis*. 2012;71(9):1440–7.
175. Meldrum C, Doyle MA, Tothill RW. Next-generation sequencing for cancer diagnostics: a practical perspective. *Clin Biochem Rev*. 2011; 32(4):177–95.
176. Saunders CJ, Miller NA, Soden SE, Dinwiddie DL, Noll A, Alnadi NA, et al. Rapid wholegenome sequencing for genetic disease diagnosis in neonatal intensive care units. *Sci Transl Med*. 2012; 4(154):154ra35.
177. Peters TL, Kumar V, Polikepahad S, Lin FY, Sarabia SF, Liang Y, et al. BCOR-CCNB3 fusions are frequent in undifferentiated sarcomas of male children. *Mod Pathol*. 2015; 28(4):575–86.
178. Vendrell JA, Taviaux S, Beganton B, Godreuil S, Audran P, Grand D, et al. Detection of known and novel ALK fusion transcripts in lung cancer patients using next-generation sequencing approaches. *Sci Rep*. 2017; 7(1):12510.
179. Harris MH, DuBois SG, Glade Bender JL, Kim A, Crompton BD, Parker E, et al. Multicenter feasibility study of tumor molecular profiling to inform therapeutic decisions in advanced pediatric solid tumors: the individualized cancer therapy (iCat) study. *JAMA Oncol*. 2016; 2:608.
180. Mody RJ, Wu YM, Lonigro RJ, Cao X, Roychowdhury S, Vats P, et al. Integrative clinical sequencing in the management of refractory or relapsed cancer in youth. *JAMA*. 2015; 314(9):913–25.
181. Robinson DR, Wu YM, Lonigro RJ, Vats P, Cobain E, Everett J, et al. Integrative clinical genomics of metastatic cancer. *Nature*. 2017; 548(7667):297–303.

182. National Cancer Institute. NCI-MATCH [Available from: <https://www.cancer.gov/about-cancer/treatment/clinical-trials/nci-supported/nci-match>].
183. Clinical Sequencing Evidence-Gathering Research Consortium. KidsCanSeq [Available from: <https://cser-consortium.org/projects/27>].
184. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science*. 2005; 310(5748):644–8.
185. Santoro M, Melillo RM, Fusco A. RET/PTC activation in papillary thyroid carcinoma: European Journal of Endocrinology Prize Lecture. *Eur J Endocrinol*. 2006; 155(5):645–53.
186. Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature*. 2007; 448(7153):561–6.
187. Maher CA, Kumar-Sinha C, Cao X, KalyanaSundaram S, Han B, Jing X, et al. Transcriptome sequencing to detect gene fusions in cancer. *Nature*. 2009; 458(7234):97–101
188. Edgren H, Murumagi A, Kangaspeska S, Nicorici D, Hongisto V, Kleivi K, et al. Identification of fusion genes in breast cancer by paired-end RNAsequencing. *Genome Biol*. 2011; 12(1): R6.
189. Robinson DR, Kalyana-Sundaram S, Wu YM, Shankar S, Cao X, Ateeq B, et al. Functionally recurrent rearrangements of the MAST kinase and notch gene families in breast cancer. *Nat Med*. 2011;17(12):1646–51.
190. Seshagiri S, Stawiski EW, Durinck S, Modrusan Z, Storm EE, Conboy CB, et al. Recurrent R-spondin fusions in colon cancer. *Nature*. 2012; 488(7413):660–4.
191. Wu YM, Su F, Kalyana-Sundaram S, Khazanov N, Ateeq B, Cao X, et al. Identification of targetable FGFR gene fusions in diverse cancers. *Cancer Discov*. 2013; 3(6):636–47.
192. Lindeman NI, Cagle PT, Beasley MB, Chitale DA, Dacic S, Giaccone G, et al. Molecular testing guideline for selection of lung cancer patients for EGFR and ALK tyrosine kinase inhibitors: guideline from the College of American Pathologists, International Association for the Study of Lung Cancer, and Association for Molecular Pathology. *J Mol Diagn*. 2013; 15(4):415–53.

193. Kwak EL, Bang YJ, Camidge DR, Shaw AT, Solomon B, Maki RG, et al. Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N Engl J Med.* 2010; 363(18):1693–703.
194. Nikiforova MN, Mercurio S, Wald AI, Barbi de Moura M, Callenberg K, Santana-Santos L, et al. Analytical performance of the ThyroSeq v3 genomic classifier for cancer diagnosis in thyroid nodules. *Cancer.* 2018; 124:1682.
195. Monzon FA, Koen TJ. Diagnosis of metastatic neoplasms: molecular approaches for identification of tissue of origin. *Arch Pathol Lab Med.* 2010; 134(2):216–24.
196. Erlander MG, Ma XJ, Kesty NC, Bao L, Salunga R, Schnabel CA. Performance and clinical evaluation of the 92-gene real-time PCR assay for tumor classification. *J Mol Diagn.* 2011; 13(5):493–503.
197. Varadhachary G. New strategies for carcinoma of unknown primary: the role of tissue-of-origin molecular profiling. *Clin Cancer Res.* 2013; 19(15):4027–33.
198. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med.* 2004; 351(27):2817–26.
199. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, et al. Prospective validation of a 21-gene expression assay in breast cancer. *N Engl J Med.* 2015; 373(21):2005–14.
200. Silvestri GA, Vachani A, Whitney D, Elashoff M, Porta Smith K, Ferguson JS, et al. A bronchial genomic classifier for the diagnostic evaluation of lung cancer. *N Engl J Med.* 2015; 373(3):243–51.
201. Alexander EK, Kennedy GC, Baloch ZW, Cibas ES, Chudova D, Diggans J, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. *N Engl J Med.* 2012; 367(8):705–15.
202. Gerami P, Cook RW, Wilkinson J, Russell MC, Dhillon N, Amaria RN, et al. Development of a prognostic genetic signature to predict the metastatic risk associated with cutaneous melanoma. *Clin Cancer Res.* 2015; 21(1):175–83.
203. Berger AC, Davidson RS, Poitras JK, Chabra I, Hope R, Brackeen A, et al. Clinical impact of a 31-gene expression profile test for cutaneous melanoma in 156 prospectively and consecutively tested patients. *Curr Med Res Opin.* 2016; 32(9):1599–604.

204. Zager JS, Gastman BR, Leachman S, Gonzalez RC, Fleming MD, Ferris LK, et al. Performance of a prognostic 31-gene expression profile in an independent cohort of 523 cutaneous melanoma patients. *BMC Cancer*. 2018; 18:130.
205. Spratt DE, Zhang J, Santiago-Jimenez M, Dess RT, Davis JW, Den RB, et al. Development and validation of a novel integrated clinical-genomic risk group classification for localized prostate cancer. *J Clin Oncol*. 2017; 36:581: JCO2017742940.
206. Sinicropi D, Qu K, Collin F, Crager M, Liu ML, Pelham RJ, et al. Whole transcriptome RNA-Seq analysis of breast cancer recurrence risk using formalin-fixed paraffin-embedded tumor tissue. *PLoS One*. 2012; 7(7): e40092.
207. Mardis ER. Applying next-generation sequencing to pancreatic cancer treatment. *Nat Rev Gastroenterol Hepatol*. 2012; 9(8):477–86.
208. Downing JR, Wilson RK, Zhang J, Mardis ER, Pui CH, Ding L, et al. The pediatric cancer genome project. *Nat Genet*. 2012; 44(6):619–22.
209. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011; 144(5):646–74.
210. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000; 100(1):57–70.
211. Van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C. Ten years of next-generation sequencing technology. *Trends Genet*. 2014; 30(9):418–26.
212. Blum A, Wang P, Zenklusen JC. SnapShot: TCGAanalyzed tumors. *Cell*. 2018; 173(2):530.
213. Giordano TJ. Genomic hallmarks of thyroid neoplasia. *Annu Rev Pathol*. 2018; 13:141–62.
214. Cancer Genome Atlas Research N. Integrated genomic characterization of papillary thyroid carcinoma. *Cell*. 2014; 159(3):676–90.
215. Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, et al. Comprehensive characterization of cancer driver genes and mutations. *Cell*. 2018; 173(2):371–85.e18.
216. Kunstman JW, Juhlin CC, Goh G, Brown TC, Stenman A, Healy JM, et al. Characterization of the mutational landscape of anaplastic thyroid cancer via whole-exome sequencing. *Hum Mol Genet*. 2015; 24(8):2318–29.

217. Karunamurthy A, Panebianco F, Hsiao S J, Vorhauer J, Nikiforova MN, Chiosea S, et al. Prevalence and phenotypic correlations of EIF1AX mutations in thyroid nodules. *Endocr Relat Cancer*. 2016; 23(4):295–301.
218. Landa I, Ibrahimasic T, Boucai L, Sinha R, Knauf JA, Shah RH, et al. Genomic and transcriptomic hallmarks of poorly differentiated and anaplastic thyroid cancers. *J Clin Invest*. 2016; 126(3):1052–66.
219. Yoo SK, Lee S, Kim SJ, Jee HG, Kim BA, Cho H, et al. Comprehensive analysis of the transcriptional and mutational landscape of follicular and papillary thyroid cancers. *PLoS Genet*. 2016; 12(8): e1006239.
220. Martin-Marcos P, Zhou F, Karunasiri C, Zhang F, Dong J, Nanda J, et al. eIF1A residues implicated in cancer stabilize translation preinitiation complexes and favor suboptimal initiation sites in yeast. *elife*. 2017; 6.
221. Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature*. 2014; 505(7484):495–501.
222. Wreesmann VB, Ghossein RA, Hezel M, Banerjee D, Shaha AR, Tuttle RM, et al. Follicular variant of papillary thyroid carcinoma: genome-wide appraisal of a controversial entity. *Genes Chromosomes Cancer*. 2004; 40(4):355–64.
223. Taylor AM, Shih J, Ha G, Gao GF, Zhang X, Berger AC, et al. Genomic and functional approaches to understanding cancer aneuploidy. *Cancer Cell*. 2018; 33(4):676–89. e3.
224. Vandin F, Clay P, Upfal E, Raphael BJ. Discovery of mutated subnetworks associated with clinical data in cancer. *Pac Symp Biocomput*. 2012:55–66.
225. Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, et al. COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Res*. 2017; 45(D1): D777–D83.
226. Eszlinger M, Krohn K, Hauptmann S, Dralle H, Giordano TJ, Paschke R. Perspectives for improved and more accurate classification of thyroid epithelial tumors. *J Clin Endocrinol Metab*. 2008; 93(9):3286–94.
227. Nikiforov YE, Carty SE, Chiosea SI, Coyne C, Duvvuri U, Ferris RL, et al. Highly accurate diagnosis of cancer in thyroid nodules with follicular neoplasm/suspicious for a follicular neoplasm cytology by ThyroSeq v2 next-generation sequencing assay. *Cancer*. 2014; 120(23):3627–34.

228. Wylie D, Beaudenon-Huibregtse S, Haynes BC, Giordano TJ, Labourier E. Molecular classification of thyroid lesions by combined testing for miRNA gene expression and somatic gene alterations. *J Pathol Clin Res*. 2016; 2(2):93–103.
229. Fagin JA, Wells SA Jr. Biologic and clinical perspectives on thyroid cancer. *N Engl J Med*. 2016; 375(11):1054–67.
230. Kidger AM, Keyse SM. The regulation of oncogenic Ras/ERK signalling by dual-specificity mitogen activated protein kinase phosphatases (MKPs). *Semin Cell Dev Biol*. 2016; 50:125–32.
231. Afkhami M, Karunamurthy A, Chiosea S, Nikiforova MN, Seethala R, Nikiforov YE, et al. Histopathologic and clinical characterization of thyroid tumors carrying the BRAF (K601E) mutation. *Thyroid*. 2016; 26(2):242–7.
232. Pozdeyev N, Gay L, Sokol ES, Hartmaier RJ, Deaver KE, Davis SN, et al. Genetic analysis of 779 advanced differentiated and anaplastic thyroid cancers. *Clin Cancer Res*. 2018; 24:3059.
233. Cheng DT, Mitchell TN, Zehir A, Shah RH, Benayed R, Syed A, et al. Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): a hybridization capture-based nextgeneration sequencing clinical assay for solid tumor molecular oncology. *J Mol Diagn*. 2015; 17(3):251–64.
234. Frampton GM, Fichtenholtz A, Otto GA, Wang K, Downing SR, He J, et al. Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. *Nat Biotechnol*. 2013; 31(11):1023–31.
235. Prasad ML, Vyas M, Horne MJ, Virk RK, Morotti R, Liu Z, et al. NTRK fusion oncogenes in pediatric papillary thyroid carcinoma in Northeast United States. *Cancer*. 2016; 122(7):1097–107.
236. Hardee S, Prasad ML, Hui P, Dinauer CA, Morotti RA. Pathologic characteristics, natural history, and prognostic implications of BRAF (V600E) mutation in pediatric papillary thyroid carcinoma. *Pediatr Dev Pathol*. 2017; 20(3):206–12.
237. Vanden Borre P, Schrock AB, Anderson PM, Morris JC 3rd, Heilmann AM, Holmes O, et al. Pediatric, adolescent, and young adult thyroid carcinoma harbors frequent and diverse targetable genomic alterations, including kinase fusions. *Oncologist*. 2017; 22(3):255–63.

238. Hoadley KA, Yau C, Hinoue T, Wolf DM, Lazar AJ, Drill E, et al. Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. *Cell*. 2018; 173(2):291–304. e6.
239. Ding L, Bailey MH, Porta-Pardo E, Thorsson V, Colaprico A, Bertrand D, et al. Perspective on oncogenic processes at the end of the beginning of cancer genomics. *Cell*. 2018; 173(2):305–20. e10.
240. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The immune landscape of cancer. *Immunity*. 2018; 48(4):812–30. e14.
241. Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, et al. Oncogenic signaling pathways in the cancer genome atlas. *Cell*. 2018; 173(2):321–37. e10.
242. Nikiforova MN, Lynch RA, Biddinger PW, Alexander EK, Dorn GW 2nd, Tallini G, et al. RAS point mutations and PAX8-PPAR gamma rearrangement in thyroid tumors: evidence for distinct molecular pathways in thyroid follicular carcinoma. *J Clin Endocrinol Metab*. 2003; 88(5):2318–26.
243. Harach HR, Soubeyran I, Brown A, Bonneau D, Longy M. Thyroid pathologic findings in patients with Cowden disease. *Ann Diagn Pathol*. 1999; 3(6):331–40.
244. Swierniak M, Pfeifer A, Stokowy T, Rusinek D, Chekan M, Lange D, et al. Somatic mutation profiling of follicular thyroid cancer by next generation sequencing. *Mol Cell Endocrinol*. 2016; 433:130–7.
245. Jung SH, Kim MS, Jung CK, Park HC, Kim SY, Liu J, et al. Mutational burdens and evolutionary ages of thyroid follicular adenoma are comparable to those of follicular carcinoma. *Oncotarget*. 2016; 7(43):69638–48.
246. Hemmer S, Wasenius VM, Knuutila S, Joensuu H, Franssila K. Comparison of benign and malignant follicular thyroid tumours by comparative genomic hybridization. *Br J Cancer*. 1998; 78(8):1012–7.
247. Roque L, Rodrigues R, Pinto A, Moura-Nunes V, Soares J. Chromosome imbalances in thyroid follicular neoplasms: a comparison between follicular adenomas and carcinomas. *Genes Chromosomes Cancer*. 2003; 36(3):292–302.
248. Garcia-Rendueles ME, Ricarte-Filho JC, Untch BR, Landa I, Knauf JA, Voza F, et al. NF2 loss promotes oncogenic RAS-induced thyroid cancers via YAP-dependent transactivation of RAS proteins and sensitizes them to MEK inhibition. *Cancer Discov*. 2015; 5(11):1178–93.

249. Liu T, Wang N, Cao J, Sofiadis A, Dinets A, Zedenius J, et al. The age- and shorter telomere-dependent TERT promoter mutation in follicular thyroid cell-derived carcinomas. *Oncogene*. 2014; 33(42):4978–84.
250. Melo M, da Rocha AG, Vinagre J, Batista R, Peixoto J, Tavares C, et al. TERT promoter mutations are a major indicator of poor outcome in differentiated thyroid carcinomas. *J Clin Endocrinol Metab*. 2014; 99(5): E754–65.
251. Kim TH, Kim YE, Ahn S, Kim JY, Ki CS, Oh YL, et al. TERT promoter mutations and long-term survival in patients with thyroid cancer. *Endocr Relat Cancer*. 2016; 23(10):813–23.
252. Liu R, Xing M. TERT promoter mutations in thyroid cancer. *Endocr Relat Cancer*. 2016; 23(3):R143–55.
253. Song J, Yang Z. Case report: whole exome sequencing of circulating cell-free tumor DNA in a follicular thyroid carcinoma patient with lung and bone metastases. *J Circ Biomark*. 2018; 7:1849454418763725.
254. Carcangiu ML. Hurthle cell carcinoma: clinicopathological and biological aspects. *Tumori*. 2003; 89(5):529–32.
255. Satoh M, Yagawa K. Electron microscopic study on mitochondria in Hurthle cell adenoma of thyroid. *Acta Pathol Jpn*. 1981; 31(6):1079–87.
256. Gasparre G, Porcelli AM, Bonora E, Pennisi LF, Toller M, Iommarini L, et al. Disruptive mitochondrial DNA mutations in complex I subunits are markers of oncogenic phenotype in thyroid tumors. *Proc Natl Acad Sci U S A*. 2007; 104(21):9001–6.
257. Evangelisti C, de Biase D, Kurelac I, Ceccarelli C, Prokisch H, Meitinger T, et al. A mutation screening of oncogenes, tumor suppressor gene TP53 and nuclear encoded mitochondrial complex I genes in oncogenic thyroid tumors. *BMC Cancer*. 2015; 15:157.
258. Ganly I, Ricarte Filho J, Eng S, Ghossein R, Morris LG, Liang Y, et al. Genomic dissection of Hurthle cell carcinoma reveals a unique class of thyroid malignancy. *J Clin Endocrinol Metab*. 2013; 98(5): E962–72.
259. Corver WE, Ruano D, Weijers K, den Hartog WC, van Nieuwenhuizen MP, de Miranda N, et al. Genome haploidisation with chromosome 7 retention in oncogenic follicular thyroid carcinoma. *PLoS One*. 2012; 7(6):e38287.

260. Xu B, Ghossein R. Genomic landscape of poorly differentiated and anaplastic thyroid carcinoma. *Endocr Pathol.* 2016; 27(3):205–12.
261. Volante M, Rapa I, Gandhi M, Bussolati G, Giachino D, Papotti M, et al. RAS mutations are the predominant molecular alteration in poorly differentiated thyroid carcinomas and bear prognostic impact. *J Clin Endocrinol Metab.* 2009; 94(12):4735–41.
262. Pita JM, Figueiredo IF, Moura MM, Leite V, Cavaco BM. Cell cycle deregulation and TP53 and RAS mutations are major events in poorly differentiated and undifferentiated thyroid carcinomas. *J Clin Endocrinol Metab.* 2014; 99(3):E497–507.
263. Hiltzik D, Carlson DL, Tuttle RM, Chuai S, Ishill N, Shaha A, et al. Poorly differentiated thyroid carcinomas defined on the basis of mitosis and necrosis: a clinicopathologic study of 58 patients. *Cancer.* 2006; 106(6):1286–95.
264. Bonhomme B, Godbert Y, Perot G, Al Ghuzlan A, Bardet S, Belleanne G, et al. Molecular pathology of anaplastic thyroid carcinomas: a retrospective study of 144 cases. *Thyroid.* 2017; 27(5):682–92.
265. Kasthuber ER, Lowe SW. Putting p53 in context. *Cell.* 2017; 170(6):1062–78.
266. Eischen CM. Genome stability requires p53. *Cold Spring Harb Perspect Med.* 2016; 6(6): pii: a026096.
267. Tomasini R, Mak TW, Melino G. The impact of p53 and p73 on aneuploidy and cancer. *Trends Cell Biol.* 2008; 18(5):244–52.
268. Klemi PJ, Joensuu H, Eerola E. DNA aneuploidy in anaplastic carcinoma of the thyroid gland. *Am J Clin Pathol.* 1988; 89(2):154–9.
269. Pinto AE, Silva G, Banito A, Leite V, Soares J. Aneuploidy and high S-phase as biomarkers of poor clinical outcome in poorly differentiated and anaplastic thyroid carcinoma. *Oncol Rep.* 2008; 20(4):913–9.
270. Baldini E, Sorrenti S, Tartaglia F, Catania A, Palmieri A, Pironi D, et al. New perspectives in the diagnosis of thyroid follicular lesions. *Int J Surg.* 2017; 41(Suppl 1): S7–S12.
271. Cibas ES. Fine-needle aspiration in the workup of thyroid nodules. *Otolaryngol Clin N Am.* 2010; 43(2):257–71. vii–viii.
272. Nishino M, Nikiforova M. Update on molecular testing for cytologically indeterminate thyroid nodules. *Arch Pathol Lab Med.* 2018; 142(4):446–57.

273. Vargas-Salas S, Martinez JR, Urra S, Dominguez JM, Mena N, Uslar T, et al. Genetic testing for indeterminate thyroid cytology: review and meta-analysis. *Endocr Relat Cancer*. 2018;25(3): R163–R77.
274. Nikiforova MN, Wald AI, Roy S, Durso MB, Nikiforov YE. Targeted next-generation sequencing panel (ThyroSeq) for detection of mutations in thyroid cancer. *J Clin Endocrinol Metab*. 2013; 98(11): E1852–60.
275. Nikiforov YE, Carty SE, Chiosea SI, Coyne C, Duvvuri U, Ferris RL, et al. Impact of the multi-gene ThyroSeq next-generation sequencing assay on cancer diagnosis in thyroid nodules with atypia of undetermined significance/follicular lesion of undetermined significance cytology. *Thyroid*. 2015; 25(11):1217–23.
276. Nikiforova MN, Mercurio S, Wald AI, Barbi de Moura M, Callenberg K, Santana-Santos L, et al. Analytical performance of the ThyroSeq v3 genomic classifier for cancer diagnosis in thyroid nodules. *Cancer*. 2018; 124(8):1682–90.
277. Jug RC, Datto MB, Jiang XS. Molecular testing for indeterminate thyroid nodules: performance of the Afirma gene expression classifier and ThyroSeq panel. *Cancer Cytopathol*. 2018; <https://doi.org/10.1002/cncy.21993>.
278. Kloos RT. Molecular profiling of thyroid nodules: current role for the Afirma gene expression classifier on clinical decision making. *Mol Imaging Radionucl Ther*. 2017; 26(Suppl 1):36–49.
279. Duh QY, Busaidy NL, Rahilly-Tierney C, Gharib H, Randolph G. A systematic review of the methods of diagnostic accuracy studies of the Afirma gene expression classifier. *Thyroid*. 2017; 27(10):1215–22.
280. Aragon Han P, Olson MT, Fazeli R, Prescott JD, Pai SI, Schneider EB, et al. The impact of molecular testing on the surgical management of patients with thyroid nodules. *Ann Surg Oncol*. 2014; 21(6):1862–9.
281. Sacks WL, Bose S, Zumsteg ZS, Wong R, Shiao SL, Braunstein GD, et al. Impact of Afirma gene expression classifier on cytopathology diagnosis and rate of thyroidectomy. *Cancer Cytopathol*. 2016; 124(10):722–8.
282. Guo HQ, Zhao H, Zhang ZH, Zhu YL, Xiao T, Pan QJ. Impact of molecular testing in the diagnosis of thyroid fine needle aspiration cytology: data from mainland China. *Dis Markers*. 2014; 2014:912182.

283. Nicholson KJ, Yip L. An update on the status of molecular testing for the indeterminate thyroid nodule and risk stratification of differentiated thyroid cancer. *Curr Opin Oncol*. 2018; 30(1):8–15.
284. Bisarro Dos Reis M, Barros-Filho MC, Marchi FA, Beltrami CM, Kuasne H, Pinto CAL, et al. Prognostic classifier based on genome-wide DNA methylation profiling in well-differentiated thyroid tumors. *J Clin Endocrinol Metab*. 2017; 102(11):4089–99.
285. Chen H, Luthra R, Routbort MJ, Patel KP, Cabanillas ME, Broaddus RR, Williams MD. Molecular Profile of Advanced Thyroid Carcinomas by Next-Generation Sequencing: Characterizing Tumors Beyond Diagnosis for Targeted Therapy. *Mol Cancer Ther*. 2018 Jul; 17(7):1575-1584. doi: 10.1158/1535-7163.MCT-17-0871. Epub 2018 Apr 25. PMID: 29695638.
286. Fang Y, Ma X, Zeng J, Jin Y, Hu Y, Wang J, Liu R, Cao C. The Profile of Genetic Mutations in Papillary Thyroid Cancer Detected by Whole Exome Sequencing. *Cell Physiol Biochem*. 2018; 50(1):169-178. doi: 10.1159/000493966. Epub 2018 Oct 2. PMID: 30278442.
287. Agliata, A, De Martino, M, Brigida Ferraro, M and Rosario Guarracino, M. (2014). IGV-plus: A Java Software for the Analysis and Visualization of Next-Generation Sequencing Data. *Vogiatzis (1ra Ed), Dynamics of Information Systems Computational and Mathematical Challenges* (pp.149-160). Springer. DOI 10.1007/978-3-319-10046-3__8.
288. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform*. 2013 Mar;14(2):178-92. doi: 10.1093/bib/bbs017. Epub 2012 Apr 19. PMID: 22517427; PMCID: PMC3603213.
289. Uhrig S, Ellermann J, Walther T, Burkhardt P, Fröhlich M, Hutter B, Toprak UH, Neumann O, Stenzinger A, Scholl C, Fröhling S, Brors B. Accurate and efficient detection of gene fusions from RNA sequencing data. *Genome Res*. 2021 Mar;31(3):448-460. doi: 10.1101/gr.257246.119. Epub 2021 Jan 13. PMID: 33441414; PMCID: PMC7919457.
290. Vega D. Secuenciamiento masivo (RNA-seq) y bioinformática del transcriptoma de tejido graso de Gallus gallus procedentes de centros de venta de Lima [Tesis]. Lima:

- Universidad Nacional Mayor de San Marcos, Facultad de Farmacia y Bioquímica, Escuela Profesional de Ciencia de los Alimentos; 2019.
- <https://cybertesis.unmsm.edu.pe/handle/20.500.12672/11306?show=full>
291. Teodoriu L, Ungureanu MC, Leustean L, et al. Updated Incidence of Thyroid Cancer in the North East Region of Romania after 35 Years of Chernobyl Fallout. Is There a Link between?. *Diagnostics (Basel)*. 2021;11(5):907. Published 2021 May 19.
 292. Salazar-Vega, J., Ortiz-Prado, E., Solis-Pazmino, P. et al. Thyroid Cancer in Ecuador, a 16 years population-based analysis (2001–2016). *BMC Cancer* **19**, 294 (2019).
 293. Xie L, Wang S, Qian Y, Jia S, Wang J, Li L, Zhang W, Yu H, Bao P, Qian B. Increasing Gap Between Thyroid Cancer Incidence and Mortality in Urban Shanghai, China: An Analysis Spanning 43 Years. *Endocr Pract*. 2021 Nov;27(11):1100-1107.
 294. Pellegriti G, Frasca F, Regalbuto C, Squatrito S, Vigneri R. Worldwide increasing incidence of thyroid cancer: update on epidemiology and risk factors. *J Cancer Epidemiol*. 2013;2013:965212.
 295. Atamari-Anahui N, et al. Tendencia nacional de la prevalencia y mortalidad por cáncer de tiroides con datos del Ministerio de Salud de Perú. *Medwave* 2019; 19(4):e7631.
 296. Chan Kwon Jung, Mark P. Little, Jay H. Lubin, Alina V. Brenner, Samuel A. Wells, Alice J. Sigurdson, Yuri E. Nikiforov, The Increase in Thyroid Cancer Incidence During the Last Four Decades Is Accompanied by a High Frequency of *BRAF* Mutations and a Sharp Increase in *RAS* Mutations, *The Journal of Clinical Endocrinology & Metabolism*, Volume 99, Issue 2, 1 February 2014, Pages E276–E285.
 297. Kowalska A, Walczyk A, Kowalik A, Pałyga I, Trybek T, Kopczyński J, Kajor M, Chrapek M, Pięciak L, Chłopek M, Gózdź S, Kamiński G. Increase in Papillary Thyroid Cancer Incidence Is Accompanied by Changes in the Frequency of the BRAF V600E Mutation: A Single-Institution Study. *Thyroid*. 2016 Apr;26(4):543-51.
 298. Jin M, Song DE, Ahn J, Song E, Lee Y-M, Sung T-Y, Kim TY, Kim WB, Shong YK, Jeon MJ, Kim WG. Genetic Profiles of Aggressive Variants of Papillary Thyroid Carcinomas. *Cancers*. 2021; 13(4):892.
 299. Liu R, Bishop J, Zhu G, Zhang T, Ladenson PW, Xing M. Mortality Risk Stratification by Combining BRAF V600E and TERT Promoter Mutations in Papillary Thyroid Cancer: Genetic Duet of BRAF and TERT Promoter Mutations in Thyroid Cancer Mortality. *JAMA Oncol*. 2017 Feb 1;3(2):202-208.
 300. Jin Y, Liu B, Younis MH, Huang G, Liu J, Cai W, Wei W. Next-Generation Molecular Imaging of Thyroid Cancer. *Cancers*. 2021; 13(13):3188.

301. Al-Jundi M, Thakur S, Gubbi S, Klubo-Gwiezdzinska J. Novel Targeted Therapies for Metastatic Thyroid Cancer-A Comprehensive Review. *Cancers (Basel)*. 2020;12(8):2104. Published 2020 Jul 29.
302. Maria E Cabanillas, Mabel Ryder, Camilo Jimenez, Targeted Therapy for Advanced Thyroid Cancer: Kinase Inhibitors and Beyond, *Endocrine Reviews*, Volume 40, Issue 6, December 2019, Pages 1573–1604.
303. Luo H, Xia X, Kim GD, Liu Y, Xue Z, Zhang L, Shu Y, Yang T, Chen Y, Zhang S, Chen H, Zhang W, Li R, Tang H, Dong B, Fu X, Cheng W, Zhang W, Yang L, Peng Y, Dai L, Hu H, Jiang Y, Gong C, Hu Y, Zhu J, Li Z, Caulin C, Wei T, Park J, Xu H. Characterizing dedifferentiation of thyroid cancer by integrated analysis. *Sci Adv*. 2021 Jul 28;7(31):eabf3657.
304. Han J, Chen M, Wang Y, et al. Identification of Biomarkers Based on Differentially Expressed Genes in Papillary Thyroid Carcinoma. *Sci Rep*. 2018;8(1):9912. Published 2018 Jul 2.
305. Pérez B. Avances y Perspectivas en el Cáncer Diferenciado de Tiroides. México: Permanyer; 2018.
306. Khosravi M, ET AL. Thyroid Cancers: Considerations, Classifications, and Managements. INTECHOPEN.
307. Kakudo K, Bychkov A, Bai Y, Li Y, Liu Z, Jung CK. The new 4th edition World Health Organization classification for thyroid tumors, Asian perspectives. *Pathol Int*. 2018; 68(12):641-664.
308. Mora-Guzmán I, Muñoz de Nova JL, Marín-Campos C, et al. Efficiency of the Bethesda System for Thyroid Cytopathology. Rendimiento del sistema Bethesda en el diagnóstico citopatológico del nódulo tiroideo. *Cir Esp*. 2018; 96(6):363-368.
309. Mendes GF, Garcia MR, Falsarella PM, et al. Fine needle aspiration biopsy of thyroid nodule smaller than 1.0 cm: accuracy of TIRADS classification system in more than 1000 nodules. *Br J Radiol*. 2018; 91(1083):20170642.
310. Molnár C, Molnár S, Bedekovics J, et al. Thyroid Carcinoma Coexisting with Hashimoto's Thyroiditis: Clinicopathological and Molecular Characteristics Clue up Pathogenesis. *Pathol Oncol Res*. 2019; 25(3):1191-1197.
311. Anne T. Mancino (Redactor), Lawrence T. Kim. Management of Differentiated Thyroid Cancer. Usa: springer; 2017.
312. <https://www.ncbi.nlm.nih.gov/gene/84105>
313. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL570>

314. Pita JM, Figueiredo IF, Moura MM, Leite V et al. Cell cycle deregulation and TP53 and RAS mutations are major events in poorly differentiated and undifferentiated thyroid carcinomas. *J Clin Endocrinol Metab* 2014 Mar;99(3):E497-507.
315. <https://www.genecards.org/>.
316. Kunstman JW, Juhlin CC, Goh G, Brown TC, Stenman A, Healy JM, Rubinstein JC, Choi M, Kiss N, Nelson-Williams C, Mane S, Rimm DL, Prasad ML, Höög A, Zedenius J, Larsson C, Korah R, Lifton RP, Carling T. Characterization of the mutational landscape of anaplastic thyroid cancer via whole-exome sequencing. *Hum Mol Genet.* 2015 Apr 15;24(8):2318-29.
317. Fang Y, Ma X, Zeng J, Jin Y, Hu Y, Wang J, Liu R, Cao C. The Profile of Genetic Mutations in Papillary Thyroid Cancer Detected by Whole Exome Sequencing. *Cell Physiol Biochem.* 2018;50(1):169-178. doi: 10.1159/000493966. Epub 2018 Oct 2. PMID: 30278442.
318. Giordano TJ, Kuick R, Thomas DG, Misek DE, Vinco M, Sanders D, Zhu Z, Ciampi R, Roh M, Shedden K, Gauger P, Doherty G, Thompson NW, Hanash S, Koenig RJ, Nikiforov YE. Molecular classification of papillary thyroid carcinoma: distinct BRAF, RAS, and RET/PTC mutation-specific gene expression profiles discovered by DNA microarray analysis. *Oncogene.* 2005 Oct 6;24(44):6646-56. doi: 10.1038/sj.onc.1208822. PMID: 16007166.
319. Gerber TS, Schad A, Hartmann N, Springer E, Zechner U, Musholt TJ. Targeted next-generation sequencing of cancer genes in poorly differentiated thyroid cancer. *Endocr Connect.* 2018;7(1):47-55. doi:10.1530/EC-17-0290.
320. Khatami F, Larijani B, Nikfar S, Hasanzad M, Fendereski K, Tavangar SM. Personalized treatment options for thyroid cancer: current perspectives. *Pharmacogenomics Pers Med.* 2019;12:235-245.
321. Xing M. Molecular pathogenesis and mechanisms of thyroid cancer. *Nat Rev Cancer.* 2013;13(3):184-199.
322. Prete A, Borges de Souza P, Censi S, Muzza M, Nucci N, Sponziello M. Update on Fundamental Mechanisms of Thyroid Cancer. *Front Endocrinol (Lausanne).* 2020;11:102. Published 2020 Mar 13.
323. Ensieh Farahani, Hira K. Patra, Jaganmohan R. Jangamreddy, Iran Rashedi, Marta Kawalec, Rama K. Rao Pariti, Petros Batakis, Emilia Wiechec, Cell adhesion molecules and their relation to (cancer) cell stemness, *Carcinogenesis*, Volume 35, Issue 4, April 2014, Pages 747–759.
324. Kloeber JA, Lou Z. Critical DNA damaging pathways in tumorigenesis. *Semin Cancer Biol.* 2021 Apr 24:S1044-579X(21)00113-9.
325. Chen X, Duan N, Zhang C, Zhang W. Survivin and Tumorigenesis: Molecular Mechanisms and Therapeutic Strategies. *J Cancer* 2016; 7(3):314-323.

326. Rao VR, Perez-Neut M, Kaja S, Gentile S. Voltage-gated ion channels in cancer cell proliferation. *Cancers (Basel)*. 2015;7(2):849-875. Published 2015 May 22.
327. Chernet BT, Levin M. Transmembrane voltage potential is an essential cellular parameter for the detection and control of tumor development in a *Xenopus* model. *Dis Model Mech*. 2013 May;6(3):595-607.
328. Wawrzynow B, Zylicz A, Zylicz M. Chaperoning the guardian of the genome. The two-faced role of molecular chaperones in p53 tumor suppressor action. *Biochim Biophys Acta Rev Cancer*. 2018 Apr;1869(2):161-174.
329. Johansson E, Andersson L, Örnros J, et al. Revising the embryonic origin of thyroid C cells in mice and humans. *Development*. 2015;142(20):3519-3528.
330. Nilsson M, Fagman H. Mechanisms of thyroid development and dysgenesis: an analysis based on developmental stages and concurrent embryonic anatomy. *Curr Top Dev Biol*. 2013;106:123-70.
331. Nilsson M, Williams D. On the Origin of Cells and Derivation of Thyroid Cancer: C Cell Story Revisited. *Eur Thyroid J*. 2016;5(2):79-93.
332. Preto A, Cameselle-Teijeiro J, Moldes-Boullosa J, Soares P, Cameselle-Teijeiro J, Silva P, et al. Telomerase expression and proliferative activity suggest a stem cell role for thyroid solid cell nests. *Mod Pathol*. 2004;17:819–826.
333. Harmon BG, Kelley LC. Immunohistochemistry of ultimobranchial thyroid carcinomas in seven slaughtered cows and one bull. *J Vet Diagn Invest*. 2001;13:101–105.
334. Reynolds L, Jones K, Winton DJ, Cranston A, Houghton C, Howard L, Ponder BA, Smith DP. C-cell and thyroid epithelial tumours and altered follicular development in transgenic mice expressing the long isoform of MEN 2A RET. *Oncogene*. 2001;20:3986–3994.
335. Yang J, Gong Y, Yan S, Chen H, Qin S, Gong R. Association between TERT promoter mutations and clinical behaviors in differentiated thyroid carcinoma: a systematic review and meta-analysis. *Endocrine*. 2020;67(1):44-57.
336. Liu X, Bishop J, Shan Y, Pai S, Liu D, Murugan AK, Sun H, El-Naggar AK, Xing M. Highly prevalent TERT promoter mutations in aggressive thyroid cancers. *Endocr Relat Cancer*. 2013;20:603–610.
337. Liu T, Wang N, Cao J, Sofiadis A, Dinets A, Zedenius J, Larsson C, Xu D. The age- and shorter telomere-dependent TERT promoter mutation in follicular thyroid cell-derived carcinomas. *Oncogene*. 2014;33:4978–4984.
338. Liu R, Xing M. TERT promoter mutations in thyroid cancer. *Endocr Relat Cancer*. 2016;23:R143–R155.

339. Marques IJ, Moura MM, Cabrera R, Pinto AE, Simoes-Pereira J, Santos C, Menezes FD, Montezuma D, Henrique R, Rodrigues Teixeira M, et al. Identification of somatic TERT promoter mutations in familial nonmedullary thyroid carcinomas. *Clin Endocrinol (Oxf)* 2017;87:394–399.
340. Panebianco F, Nikitski AV, Nikiforova MN, Nikiforov YE. Spectrum of TERT promoter mutations and mechanisms of activation in thyroid cancer. *Cancer Med*. 2019;8(13):5831-5839.
341. Alzahrani AS, Alsaadi R, Murugan AK, Sadiq BB. TERT promoter mutations in thyroid cancer. *Horm Cancer*. 2016;7:165–177.
342. Choi YS, Choi SW, Yi JW. Prospective Analysis of TERT Promoter Mutations in Papillary Thyroid Carcinoma at a Single Institution. *J Clin Med*. 2021;10(10):2179. Published 2021 May 18.
343. Liu T, Wang N, Cao J, Sofiadis A, Dinets A, Zedenius J, Larsson C, Xu D. The age- and shorter telomere-dependent TERT promoter mutation in follicular thyroid cell-derived carcinomas. *Oncogene*. 2014 Oct 16;33(42):4978-84.
344. Song YS, Lim JA, Park YJ. Mutation Profile of Well-Differentiated Thyroid Cancer in Asians. *Endocrinol Metab (Seoul)*. 2015 Sep;30(3):252-62.
345. Yan C, Huang M, Li X, Wang T, Ling R. Relationship between BRAF V600E and clinical features in papillary thyroid carcinoma. *Endocr Connect*. 2019 Jul;8(7):988-996. doi: 10.1530/EC-19-0246. PMID: 31252408; PMCID: PMC6652244.
346. Xing M, Alzahrani AS, Carson KA, Viola D, Elisei R, Bendlova B, et al. Association between BRAF V600E mutation and mortality in patients with papillary thyroid cancer. *JAMA* 2013;309:1493-501.
347. Ito Y, Yoshida H, Maruo R, Morita S, Takano T, Hirokawa M, et al. BRAF mutation in papillary thyroid carcinoma in a Japanese population: its lack of correlation with high-risk clinicopathological features and disease-free survival of patients. *Endocr J* 2009;56:89-97.
348. Musholt TJ, Fottner C, Weber MM, Eichhorn W, Pohlenz J, Musholt PB, et al. Detection of papillary thyroid carcinoma by analysis of BRAF and RET/PTC1 mutations in fineneedle aspiration biopsies of thyroid nodules. *World J Surg* 2010;34:2595-603.

349. Shain AH, Joseph NM, Yu R, et al. Genomic and Transcriptomic Analysis Reveals Incremental Disruption of Key Signaling Pathways during Melanoma Evolution. *Cancer Cell*. 2018;34(1):45-55.e4.
350. Gutiérrez-Castañeda LD, Gamboa M, Nova JA, Pulido L, Tovar-Parra JD. Mutations in the BRAF, NRAS, and C-KIT Genes of Patients Diagnosed with Melanoma in Colombia Population. *Biomed Res Int*. 2020;2020:2046947. Published 2020 Jul 22.
351. Yan C, Huang M, Li X, Wang T, Ling R. Relationship between BRAF V600E and clinical features in papillary thyroid carcinoma. *Endocr Connect*. 2019;8(7):988-996. doi:10.1530/EC-19-0246.
352. Ren H, Shen Y, Hu D, et al. Co-existence of BRAFV600E and TERT promoter mutations in papillary thyroid carcinoma is associated with tumor aggressiveness, but not with lymph node metastasis. *Cancer Manag Res*. 2018;10:1005-1013. Published 2018 May 3.
353. Xing M, Liu R, Liu X, et al. BRAF V600E and TERT promoter mutations cooperatively identify the most aggressive papillary thyroid cancer with highest recurrence. *J Clin Oncol*. 2014;32(25):2718–2726.
354. Lee SE, Hwang TS, Choi YL, et al. Prognostic significance of TERT promoter mutations in papillary thyroid carcinomas in a BRAF(V600E) mutation-prevalent population. *Thyroid*. 2016;26(7):901–910.
355. Jin L. BRAF and TERT promoter mutations in the aggressiveness of papillary thyroid carcinoma: a study of 653 patients. *Oncotarget*. 2016;7(14):18346–18355.
356. Asmann YW, Sun J, Zhang J, et al. BRAF V600E and TERT promoter mutations in papillary thyroid carcinoma in Chinese patients. *PLoS One*. 2016;11(4):e0153319.
357. Vuong HG, Altibi AMA, Duong UNP, Hassell L. Prognostic implication of BRAF and TERT promoter mutation combination in papillary thyroid carcinoma-A meta-analysis. *Clin Endocrinol (Oxf)* 2017;87(5):411–417.
358. Moon S, Song YS, Kim YA, et al. Effects of coexistent BRAFV600E and TERT promoter mutations on poor clinical outcomes in papillary thyroid cancer: a meta-analysis. *Thyroid*. 2017;27(5):651–660.
359. Manzella L, Stella S, Pennisi MS, et al. New Insights in Thyroid Cancer and p53 Family Proteins. *Int J Mol Sci*. 2017;18(6):1325. Published 2017 Jun 21.

360. Liu L, Li D, Chen Z, Yang J, Ma Y, Cai H, Shan C, Lv Z, Zhang X. Wild-Type P53 Induces Sodium/Iodide Symporter Expression Allowing Radioiodide Therapy in Anaplastic Thyroid Cancer. *Cell Physiol Biochem*. 2017;43(3):905-914.
361. Quiros, R.M.; Ding, H.G.; Gattuso, P.; Prinz, R.A.; Xu, X. Evidence that one subset of anaplastic thyroid carcinomas are derived from papillary carcinomas due to BRAF and p53 mutations. *Cancer* 2005, *103*, 2261–2268.
362. Cancer Genome Atlas Research Network. Integrated genomic characterization of papillary thyroid carcinoma. *Cell* 2014, *159*, 676–690.
363. Dralle, H.; Machens, A.; Basa, J.; Fatourehchi, V.; Franceschi, S.; Hay, I.D.; Nikiforov, Y.E.; Pacini, F.; Pasiaka, J.L.; Sherman, S.I. Follicular cell-derived thyroid cancer. *Nat. Rev. Dis. Primers* 2015, *1*, 15077.
364. Farid, N.R. P53 mutations in thyroid carcinoma: Tidings from an old foe. *J. Endocrinol. Investig.* 2001, *24*, 536–545.
365. Morita N, Ikeda Y and Takami H: Clinical significance of p53 protein expression in papillary thyroid carcinoma. *World J Surg.* 32:2617–2622. 2008.
366. Parameswaran R, Brooks S and Sadler GP: Molecular pathogenesis of follicular cell derived thyroid cancers. *Int J Surg.* 8:186–193. 2010.
367. Zafon C, Obiols G, Castellvi J, Tallada N, Baena JA, Simo R, et al: Clinical significance of RET/PTC and p53 protein expression in sporadic papillary thyroid carcinoma. *Histopathology.* 50:225–231. 2007.
368. Shin MK, Kim JW, Min SK, et al. Associations of the BRAF (V600E) mutation and p53 protein expression with clinicopathological features of papillary thyroid carcinomas patients. *Oncol Lett.* 2015;10(3):1882-1888.
369. Oishi, N., Kondo, T., Ebina, A. *et al.* Molecular alterations of coexisting thyroid papillary carcinoma and anaplastic carcinoma: identification of *TERT* mutation as an independent risk factor for transformation. *Mod Pathol* 30, 1527–1537 (2017).
370. Mundi PS, Sachdev J, McCourt C, Kalinsky K. AKT in cancer: new molecular insights and advances in drug development. *Br J Clin Pharmacol.* 2016;82(4):943-956.
371. Cheng J, Shuai X, Gao J, Cai M, Wang G, Tao K. Prognostic significance of AMPK in human malignancies: A meta-analysis. *Oncotarget.* 2016;7(46):75739-75748.
372. Chiba K, Lorbeer FK, Shain AH, McSwiggen DT, Schruf E, Oh A, Ryu J, Darzacq X, Bastian BC, Hockemeyer D. Mutations in the promoter of the telomerase

- gene *TERT* contribute to tumorigenesis by a two-step mechanism. *Science*. 2017 Sep 29;357(6358):1416-1420.
373. Tang, D., Lotze, M., Kang, R. *et al.* Apoptosis promotes early tumorigenesis. *Oncogene* 30, 1851–1854 (2011).
374. De Almeida LC, Calil FA, Machado-Neto JA, Costa-Lotuf LV. DNA damaging agents and DNA repair: From carcinogenesis to cancer therapy. *Cancer Genet*. 2021 Apr;252-253:6-24.
375. Yang, X., Liu, Y., Li, M. *et al.* Predictive and preventive significance of AMPK activation on hepatocarcinogenesis in patients with liver cirrhosis. *Cell Death Dis* 9, 264 (2018).
376. Guo J, Cheng J, Zheng N, Zhang X, Dai X, Zhang L, Hu C, Wu X, Jiang Q, Wu D, Okada H, Pandolfi PP, Wei W. Copper Promotes Tumorigenesis by Activating the PDK1-AKT Oncogenic Pathway in a Copper Transporter 1 Dependent Manner. *Adv Sci (Weinh)*. 2021 Sep;8(18):e2004303.
377. Yue X, Zhao Y, Xu Y, Zheng M, Feng Z, Hu W. Mutant p53 in Cancer: Accumulation, Gain-of-Function, and Therapy. *J Mol Biol*. 2017;429(11):1595-1606.
378. Viner-Breuer, R., Yilmaz, A., Benvenisty, N. *et al.* The essentiality landscape of cell cycle related genes in human pluripotent and cancer cells. *Cell Div* 14, 15 (2019).
379. Yu L, Chen X, Wang L, Chen S. Oncogenic virus-induced aerobic glycolysis and tumorigenesis. *J Cancer*. 2018;9(20):3699-3706. Published 2018 Sep 8.
380. Mao H, Sun Y. Primary cilium and its role in tumorigenesis. *Zhejiang Da Xue Xue Bao Yi Xue Ban*. 2021 Apr 25;50(2):245-260.
381. Alexander J, Cukierman E. Cancer associated fibroblast: Mediators of tumorigenesis. *Matrix Biol*. 2020 Sep;91-92:19-34.
382. Kim E.J. Um S.J. Thymine-DNA glycosylase interacts with and functions as a coactivator of p53 family proteins. *Biochem. Biophys. Res. Commun*. 2008; 377: 838-842.
383. Sun W. Yang J. Functional mechanisms for human tumor suppressors. *J. Cancer*. 2010; 1: 136-140.
384. da Costa N.M. Hautefeuille A. Cros M.P. Melendez M.E. Waters T. Swann P. Hainaut P. Pinto L.F. Transcriptional regulation of thymine DNA glycosylase (TDG) by the tumor suppressor protein p53. *Cell Cycle*. 2012; 11: 4570-4578.
385. Vasovcak P. Krepelova A. Menigatti M. Puchmajerova A. Skapa P. Augustinakova A. Amann G. Wernstedt A. Jiricny J. Marra G. Wimmer K. Unique mutational profile associated

- with a loss of TDG expression in the rectal cancer of a patient with a constitutional PMS2 deficiency. *DNA Repair*. 2012; 11: 616-623.
386. Sjolund A.B. Senejani A.G. Sweasy J.B. MBD4 and TDG. Multifaceted DNA glycosylases with ever expanding biological roles. *Mutat. Res*. 2012; 743–744: 12-25.
387. Reddy Parine N, Alanazi IO, Shaik JP, Aldhaian S, Aljebreen AM, Alharbi O, Almadi MA, Azzam NA, Alanazi M. TDG Gene Polymorphisms and Their Possible Association with Colorectal Cancer: A Case Control Study. *J Oncol*. 2019 May 23;2019:7091815.
388. T. Yatsuoka, T. Furukawa, T. Abe et al., “Genomic analysis of the thymine-DNA glycosylase (TDG) gene on 12q22–q24. 1 in human pancreatic ductal adenocarcinoma,” *International Journal of Pancreatology*, vol. 25, no. 2, pp. 97–102, 1999.
389. I. Ruczinski, T. J. Jorgensen, Y. Y. Shugart et al., “A population-based study of dna repair gene variants in relation to non-melanoma skin cancer as a marker of a cancer-prone phenotype,” *Carcinogenesis*, vol. 33, no. 9, pp. 1692–1698, 2012.
390. W.-Q. Li, N. Hu, P. L. Hyland et al., “Genetic variants in DNA repair pathway genes and risk of esophageal squamous cell carcinoma and gastric adenocarcinoma in a Chinese population,” *Carcinogenesis*, vol. 34, no. 7, pp. 1536–1542, 2013.
391. Mancuso, P., Tricarico, R., Bhattacharjee, V. et al. Thymine DNA glycosylase as a novel target for melanoma. *Oncogene* **38**, 3710–3728 (2019).
392. Li, D., Chai, L., Yu, X. et al. The *HOTAIRM1/miR-107/TDG* axis regulates papillary thyroid cancer cell proliferation and invasion. *Cell Death Dis* **11**, 227 (2020).
393. Balbin OA, Malik R, Dhanasekaran SM, et al. The landscape of antisense gene expression in human cancers. *Genome Res*. 2015;25(7):1068-1079.
394. Luis Sanchez-Pulido, Chris P Ponting, TMEM132: an ancient architecture of cohesin and immunoglobulin domains define a new family of neural adhesion molecules, *Bioinformatics*, Volume 34, Issue 5, 01 March 2018, Pages 721–724.
395. de Almeida BP, Apolónio JD, Binnie A, Castelo-Branco P. Roadmap of DNA methylation in breast cancer identifies novel prognostic biomarkers. *BMC Cancer*. 2019;19(1):219. Published 2019 Mar 12.
396. Karapetsas A, Giannakakis A, Dangaj D, Lanitis E, Kynigopoulos S, Lambropoulou M, Tanyi JL, Galanis A, Kakolyris S, Trypsianis G, Coukos G, Sandaltzopoulos R. Overexpression of GPC6 and TMEM132D in Early Stage Ovarian Cancer Correlates with CD8+ T-Lymphocyte Infiltration and Increased Patient Survival. *Biomed Res Int*. 2015;2015:712438.
397. Xu W, Li C, Ma B, et al. Identification of Key Functional Gene Signatures Indicative of Dedifferentiation in Papillary Thyroid Cancer. *Front Oncol*. 2021;11:641851. Published 2021 Apr 28.

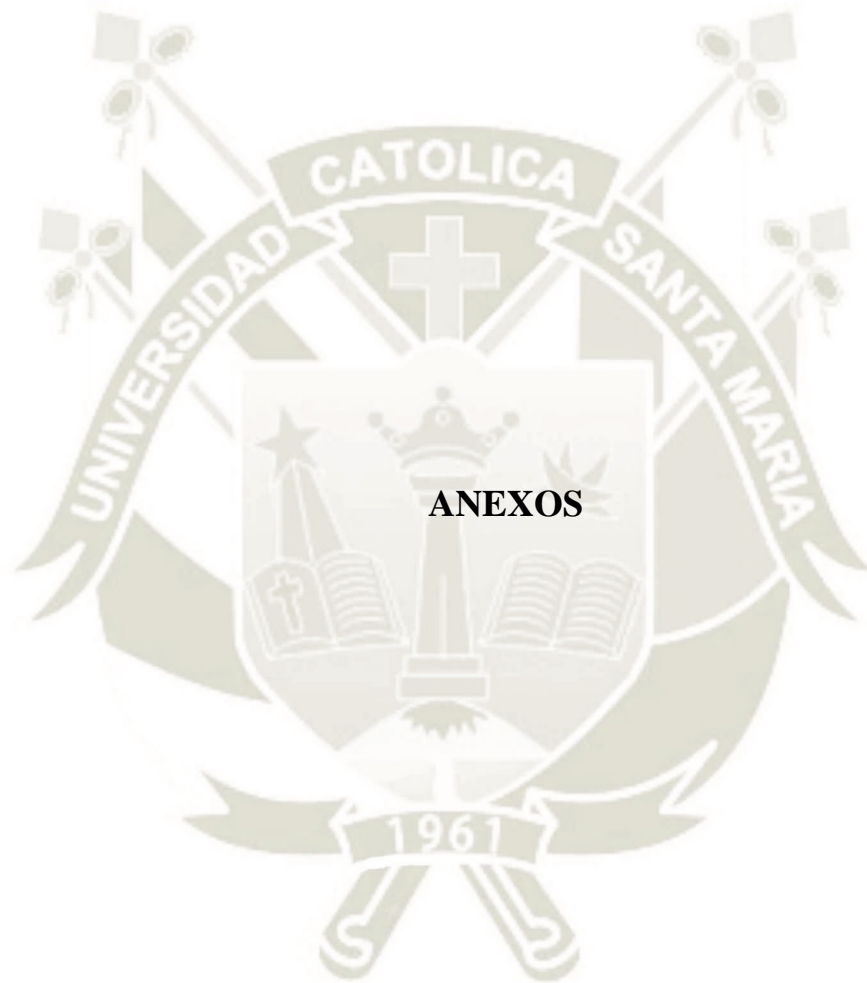
398. Wang Y, Yin L, Hillgartner FB. The homeodomain proteins PBX and MEIS1 are accessory factors that enhance thyroid hormone regulation of the malic enzyme gene in hepatocytes. *J Biol Chem.* 2001 Jun 29;276(26):23838-48.
399. GİrgİN B, KaradaĖ-Alpaslan M, KocabaŞ F. Oncogenic and tumor suppressor function of MEIS and associated factors. *Turk J Biol.* 2020;44(6):328-355. Published 2020 Dec 14.
400. Kitchen MO, Bryan RT, Haworth KE, Emes RD, Luscombe C et al. (2015). Methylation of HOXA9 and ISL1 predicts patient outcome in high-grade non-invasive bladder cancer. *PLoS One* 10: e0137003.
401. Dard A, Reboulet J, Jia Y, Bleicher F, Duffraisse M et al. (2018). Human HOX proteins use diverse and context-dependent motifs to interact with TALE class cofactors. *Cell Reports* 22: 3058-3071.
402. Moghbeli M, Rad A, Farshchian M, Taghehchian N, Gholamin M et al. (2016). Correlation between Meis1 and Msi1 in esophageal squamous cell carcinoma. *Journal of Gastrointestinal Cancer* 47: 273-277.
403. Xu X, Bao Z, Liu Y, Jiang K, Zhi T et al. (2018). PBX3/MEK/ERK1/2/LIN28/let-7b positive feedback loop enhances mesenchymal phenotype to promote glioblastoma migration and invasion. *Journal of Experimental & Clinical Cancer Research* 37: 158.
404. Zhu J, Cui L, Xu A, Yin X, Li F et al. (2017). MEIS1 inhibits clear cell renal cell carcinoma cells proliferation and in vitro invasion or migration. *BMC Cancer* 17: 176.
405. Mohr S, Doebele C, Comoglio F, Berg T, Beck J et al. (2017). Hoxa9 and Meis1 cooperatively induce addiction to syk signaling by suppressing miR-146a in acute myeloid leukemia. *Cancer Cell* 31: 549-562 e511.
406. Liu N, Liu Z, Zhang W, Li Y, Cao J et al. (2018). MicroRNA433 reduces cell proliferation and invasion in nonsmall cell lung cancer via directly targeting E2F transcription factor 3. *Molecular Medicine Reports* 18: 1155-1164.
407. Chen G, Xie Y (2018). miR-495 inhibits proliferation, migration, and invasion and induces apoptosis via inhibiting PBX3 in melanoma cells. *OncoTargets and Therapy* 11: 1909-1920.
408. Platais C, Radhakrishnan R, Niklander Ebensperger S, Morgan R, Lambert DW et al. (2018). Targeting HOX-PBX interactions causes death in oral potentially malignant and squamous carcinoma cells but not normal oral keratinocytes. *BMC Cancer* 18: 723.
409. Thomsen CB, Andersen RF, Steffensen KD, Adimi P, Jakobsen A (2019). Delta tocotrienol in recurrent ovarian cancer. A phase II trial. *Pharmacological Research* 141: 392-396.
410. Qiu Z, Wang X, Shi Y, Da M (2019). miR-129-5p suppresses proliferation, migration, and induces apoptosis in pancreatic cancer cells by targeting PBX3. *Acta Biochimica et Biophysica Sinica* 51: 997-1007.

411. Bhanvadia RR, VanOpstall C, Brechka H, Barashi NS, Gillard M et al. (2018). MEIS1 and MEIS2 expression and prostate cancer progression: a role for HOXB13 binding partners in metastatic disease. *Clinical Cancer Research* 24: 3668-3680.
412. Lin L, Huang M, Shi X, Mayakonda A, Hu K et al. (2019). Superenhancer-associated MEIS1 promotes transcriptional dysregulation in Ewing sarcoma in co-operation with EWSFLI1. *Nucleic Acids Research* 47: 1255-1267.
413. Jia Y, Bleicher F, Merabet S (2018). A systematic survey of HOX and TALE expression profiling in human cancers. *The International Journal of Developmental Biology* 62: 865-876.
414. Jacques C, Guillotin D, Fontaine JF, Franc B, Mirebeau-Prunier D et al. (2013). DNA microarray and miRNA analyses reinforce the classification of follicular thyroid tumors. *The Journal of Clinical Endocrinology and Metabolism* 98: E981-989.
415. Takahashi Y, Hamada J, Murakawa K, Takada M, Tada M et al. (2004). Expression profiles of 39 HOX genes in normal human adult organs and anaplastic thyroid cancer cell lines by quantitative real-time RT-PCR system. *Experimental Cell Research* 293:144-153.
416. Wang Y, Hardin H, Chu YH, Esbona K, Zhang R et al. (2019a). Long non-coding RNA Expression in Anaplastic Thyroid Carcinomas. *Endocrine Pathology* 30: 262-269.
417. Vriens MR, Moses W, Weng J, Peng M, Griffin A et al. (2011). Clinical and molecular features of papillary thyroid cancer in adolescents and young adults. *Cancer* 117: 259-267.
418. Gu, H. Zhang, F. Qin, X. Liu, W. Li, L. Fu, G. Ying, B. Li, M. Zhang, Y. Ma Intersectin1-S, a multidomain adapter protein, is essential for malignant glioma proliferation *Glia*, 63 (2015), pp. 1595-1605.
419. A. Russo, J.P. O'Bryan Intersectin 1 is required for neuroblastoma tumorigénesis *Oncogene*, 31 (2012), pp. 4828-4834.
420. 559. Herrero-Garcia E, O'Bryan JP. Intersectin scaffold proteins and their role in cell signaling and endocytosis. *Biochim Biophys Acta Mol Cell Res.* 2017 Jan;1864(1):23-30.
421. K. Specht, N. Harbeck, J. Smida, K. Annecke, U. Reich, J. Naehrig, R. Langer, J. Mages, R. Busch, E. Kruse, L. Klein-Hitpass, M. Schmitt, M. Kiechle, H. Hoefler. Expression profiling identifies genes that predict recurrence of breast cancer after adjuvant CMF-based chemotherapy *Breast Cancer Res. Treat.*, 118 (2009), pp. 45-56.
422. Panebianco F, Nikitski AV, Nikiforova MN, et al. Characterization of thyroid cancer driven by known and novel ALK fusions. *Endocr Relat Cancer.* 2019;26(11):803-814.
423. McDonald, T.V. et al. A minK-HERG complex regulates the cardiac potassium current IKr. *Nature* 388, 289–292 (1997).
424. Abbott, G.W. et al. MiRP1 forms IKr potassium channels with HERG and is associated with cardiac arrhythmia. *Cell* 97, 175–187 (1999).

425. McCrossan, Z.A. & Abbott, G.W. The MinK-related peptides. *Neuropharmacology* 47, 787–821 (2004).
426. Roepke, T., King, E., Reyna-Neyra, A. *et al.* *Kcne2* deletion uncovers its crucial role in thyroid hormone biosynthesis. *Nat Med* 15, 1186–1194 (2009).
427. Splawski, I., Tristani-Firouzi, M., Lehmann, M.H., Sanguinetti, M.C. & Keating, M.T. Mutations in the hminK gene cause long QT syndrome and suppress IKs function. *Nat. Genet.* 17, 338–340 (1997).
428. Tyson, J. *et al.* IsK and KvLQT1: mutation in either of the two subunits of the slow component of the delayed rectifier potassium channel can cause Jervell and LangeNielsen syndrome. *Hum. Mol. Genet.* 6, 2179–2185 (1997).
429. Curran, M.E. *et al.* A molecular basis for cardiac arrhythmia: HERG mutations cause long QT syndrome. *Cell* 80, 795–803 (1995).
430. Chen, Y.H. *et al.* KCNQ1 gain-of-function mutation in familial atrial fibrillation. *Science* 299, 251–254 (2003).
431. Ehrlich, J.R., Zicha, S., Coutu, P., Hebert, T.E. & Nattel, S. Atrial fibrillation associated minK38G/S polymorphism modulates delayed rectifier current and membrane localization. *Cardiovasc. Res.* 67, 520–528 (2005).
432. Temple, J. *et al.* Atrial fibrillation in KCNE1-null mice. *Circ. Res.* 97, 62–69 (2005).
433. Yang, Y. *et al.* Identification of a KCNE2 gain-of-function mutation in patients with familial atrial fibrillation. *Am. J. Hum. Genet.* 75, 899–905 (2004).
434. Comes N, Serrano-Albarrás A, Capera J, Serrano-Novillo C, Condom E, Ramón Y Cajal S, Ferreres JC, Felipe A. Involvement of potassium channels in the progression of cancer to a more malignant phenotype. *Biochim Biophys Acta.* 2015 Oct;1848(10 Pt B):2477-92.
435. T. Suzuki, K. Takimoto, Selective expression of HERG and Kv2 channels influences proliferation of uterine cancer cells, *Int. J. Oncol.* 25 (2004) 153–159.
436. Barhanin, J., Lesage, F., Guillemare, E., Fink, M., Lazdunski, M., Romey, G. K(v)LQT1 and IsK (minK) proteins associate to form the I(Ks) cardiac potassium current. *Nature* 384: 78–80, 1996.
437. Brabant G, Maenhaut C, Kohrle J, Scheumann G, Dralle H, *et al.* (1991) Human thyrotropin receptor gene: expression in thyroid tumors and correlation to markers of thyroid differentiation and dedifferentiation. *Mol. Cell Endocrinol.* 82, R7–12.
438. Frohlich H, Boini KM, Seebohm G, Strutz-Seebohm N, Ureche ON, *et al.* (2011) Hypothyroidism of gene-targeted mice lacking Kcnq1. *Pflugers Arch.*, 461, 45–52.
439. Vallone, C.; Rigon, G.; Gulia, C.; Baffa, A.; Votino, R.; Morosetti, G.; Zaami, S.; Briganti, V.; Catania, F.; Gaffi, M.; *et al.* Non-Coding RNAs and Endometrial Cancer. *Genes* 2018, 9, 187.

440. Cavaliere AF, Perelli F, Zaami S, Piergentili R, Mattei A, Vizzielli G, Scambia G, Straface G, Restaino S, Signore F. Towards Personalized Medicine: Non-Coding RNAs and Endometrial Cancer. *Healthcare (Basel)*. 2021 Jul 30;9(8):965.
441. Wang Y, Liu Y, Guan Y, et al. Integrated analysis of immune-related genes in endometrial carcinoma. *Cancer Cell Int*. 2020;20:477. Published 2020 Oct 2.
442. Tabatabaeian H, Peiling Yang S, Tay Y. Non-Coding RNAs: Uncharted Mediators of Thyroid Cancer Pathogenesis. *Cancers (Basel)*. 2020;12(11):3264. Published 2020 Nov 4.
443. Yakushina VD, Lerner LV, Lavrov AV. Gene Fusions in Thyroid Cancer. *Thyroid*. 2018 Feb;28(2):158-167.
444. <https://www.ncbi.nlm.nih.gov/snp/?term=BRAF+7%3A+140924572>.
445. https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.33/.
446. <https://www.ncbi.nlm.nih.gov/clinvar/variation/375939/>.
447. Silver, J.A., Bogatchenko, M., Pusztaszeri, M. *et al*. *BRAF V600E* mutation is associated with aggressive features in papillary thyroid carcinomas ≤ 1.5 cm. *J of Otolaryngol - Head & Neck Surg* **50**, 63 (2021).
448. Xia M, Zhang Y, Jin K, Lu Z, Zeng Z, Xiong W. Communication between mitochondria and other organelles: a brand-new perspective on mitochondria in cancer. *Cell Biosci*. 2019; 9():27.
449. Wang H, Zhou Y, Oyang L, Han Y, Xia L, Lin J, Tang Y, Su M, Tan S, Tian Y, Chen X, Luo X, Liang J, Rao S, Wang Y, Xiong W, Zeng Z, Wang H, Li G, Liao Q. LPLUNC1 stabilises PHB1 by counteracting TRIM21-mediated ubiquitination to inhibit NF- κ B activity in nasopharyngeal carcinoma. *Oncogene*. 2019 Jun; 38(25):5062-5075.
450. Fagin JA, Wells SA Jr. Biologic and Clinical Perspectives on Thyroid Cancer. *N Engl J Med*. 2016 Sep 15; 375(11):1054-67.
451. Riesco-Eizaguirre G, Gutiérrez-Martínez P, García-Cabezas MA, Nistal M, Santisteban P. The oncogene BRAF V600E is associated with a high risk of recurrence and less differentiated papillary thyroid carcinoma due to the impairment of Na⁺/I⁻ targeting to the membrane. *Endocr Relat Cancer*. 2006 Mar; 13(1):257-69.
452. Ge J, Wang J, Wang H, et al. The *BRAF* V600E mutation is a predictor of the effect of radioiodine therapy in papillary thyroid cancer. *J Cancer*. 2020;11(4):932-939. Published 2020 Jan 1.
453. <https://www.ncbi.nlm.nih.gov/snp/?term=TERT+5%3A+1293763>.
454. <https://www.ncbi.nlm.nih.gov/snp/?term=TERT+5%3A+1272217>.
455. <https://www.ncbi.nlm.nih.gov/snp/?term=TERT+5%3A+1255405>.

456. Bae JS, Kim Y, Jeon S, et al. Clinical utility of *TERT* promoter mutations and ALK rearrangement in thyroid cancer patients with a high prevalence of the *BRAF* V600E mutation. *Diagn Pathol* 2016; 11: 21.
457. Jung CK, Kim Y, Jeon S, Jo K, Lee S, Bae JS. Clinical utility of *EZH1* mutations in the diagnosis of follicular-patterned thyroid tumors. *Hum Pathol* 2018; 81: 9–17.
458. Pani F, Macerola E, Basolo F, Boi F, Scartozzi M, Mariotti S. Aggressive differentiated thyroid cancer with multiple metastases and NRAS and TERT promoter mutations: A case report. *Oncol Lett.* 2017 Aug;14(2):2186-2190.
459. Vuong HG, Duong UN, Altibi AM, Ngo HT, Pham TQ, Tran HM, Gandolfi G, Hassell L. A meta-analysis of prognostic roles of molecular markers in papillary thyroid carcinoma. *Endocr Connect.* 2017 Apr;6(3):R8-R17.
460. Xu Q, Song A, Xie Q. The Integrated Analyses of Driver Genes Identify Key Biomarkers in Thyroid Cancer. *Technol Cancer Res Treat.* 2020 Jan-Dec;19:1533033820940440.
461. Xia F, Jiang B, Chen Y, Du X, Peng Y, Wang W, Wang Z, Li X. Prediction of novel target genes and pathways involved in tall cell variant papillary thyroid carcinoma. *Medicine (Baltimore).* 2018 Dec;97(51):e13802.
462. Ibrahimasic T, Xu B, Landa I, Dogan S, Middha S, Seshan V, Deraje S, Carlson DL, Migliacci J, Knauf JA, Untch B, Berger MF, Morris L, Tuttle RM, Chan T, Fagin JA, Ghossein R, Ganly I. Genomic Alterations in Fatal Forms of Non-Anaplastic Thyroid Cancer: Identification of *MED12* and *RBM10* as Novel Thyroid Cancer Genes Associated with Tumor Virulence. *Clin Cancer Res.* 2017 Oct 1;23(19):5970-5980.
463. Erinjeri NJ, Nicolson NG, Deyholos C, Korah R, Carling T. Whole-Exome Sequencing Identifies Two Discrete Druggable Signaling Pathways in Follicular Thyroid Cancer. *J Am Coll Surg.* 2018 Jun;226(6):950-959.e5.



Anexo 1: Ficha de Recolección de Datos

1.- Presencia de células cancerígenas del cáncer diferenciado de tiroides

Si () No ()

2.- Secuenciación RNA-Seq:

2.1.- Grado de expresión de proteínas mutadas:

Nombre proteína	Alta expresión	
-----	Moderada expresión	
	Baja expresión	

2.2.- Tipo de mutación cromosómica:

Código de mutación	Delección	
-----	Duplicación	
	Inversión	
	Translocación	

3.- Resultado histopatológico:

3.1.- Variante histopatológica

.....

3.2. Estadiaje histopatológico:

Cáncer diferenciado de Tiroides:

< 55 años	Cualquier T	Cualquier N	M0	Estadio I	
	Cualquier T	Cualquier N	M1	Estadio II	
> = 55 años	T1	N0 / NX	M0	Estadio I	
	T1	N1	M0	Estadio II	
	T2	N0 / NX	M0	Estadio I	
	T2	N1	M0	Estadio II	
	T3a / T3b	Cualquier N	M0	Estadio II	
	T4a	Cualquier N	M0	Estadio III	
	T4b	Cualquier N	M0	Estadio IVA	
Cualquier T	Cualquier N	M1	Estadio IVB		